For office use only	Team Control Number	For office use only
T1________________	**10086**	F1________________
T2________________		F2________________
T3________________	Problem Chosen	F3________________
T4________________	**B**	F4________________

Summary

In our paper, we focus on addressing problems of water quality forecast and the building of the lakes evaluation system. Then we apply them into Chao Hu. Lastly, we analyze the results, meanwhile, putting forward suggestions on the improvement of the land management.

First and foremost, this paper mainly places emphasis on the nitrogen and phosphorus input as the criterion of lake, then a model for predicting water quality can be divided into two sections. We formulate an export coefficient model accounting for factors influencing the output of N and P in order to estimate and forecast nitrogen and phosphorus load under the different land use. On the basis of obtaining N and P outputs, BP neural network model is built by combining with indictors as inputs like temperature, PH and so on to predict potentially-toxic algal blooms. To ground this model in reality, we incorporate 91 groups data collected from the websites for train and 19 groups data used for the simulation. The simulation results agreeing well with real situation indicate that the model is efficient and reliable.

Secondly, a strategy based on analytic hierarchy process is proposed to build lake evaluation model. We collect the values of seven indexes judging lakes through local knowledge and expertise. Then we use AHP to determine the weight of seven factors and finally evaluate Chao Hu successfully. We draw conclusion that Chao Hu lies inmiddle level.

Eventually, the sensitivity analysis of the evaluation model is carried out to ensure the utility. It is found that the model is shortage of insufficient stability by analyzing the sensitivity of environmental awareness, for lakes evaluation is a complicated system and cannot be described by simple indexes. So the model need to be improved.

Abstract

In our paper, we focus on addressing problems of water quality forecast and the building of the lakes evaluation system. Then we apply them into Chao Hu. Lastly, we analyze the results, meanwhile, putting forward suggestions on the improvement of the land management.

First and foremost, we build a model to predict the water quality and potentially-toxic algal blooms. We divide the model into two sections. The first is the Export Coefficient Model, which is used to estimate and forecast nitrogen and phosphorus load under the different land-use. On the basis of obtaining N and P outputs, BP neural network model is built by combining with indictors as inputs like temperature, PH and so on to predict potentially-toxic algal blooms. To ground this model in reality, we incorporate 91 groups data collected from the websites for train and 19 groups data used for the test. The data of Chao Hu are used for simulation. The simulation results agree well with real situation indicating that the model is efficient and reliable.

Secondly, we build the lake evaluation model based on analytic hierarchy process. We collect the values of seven indexes judging lakes through local knowledge and expertise. Then we use AHP to determine the weight of seven factors and finally evaluate Chao Hu successfully. We draw conclusion that Chao Hu lies in belonging to middle level.

Eventually, the sensitivity analysis of the evaluation model is carried out to ensure the utility. It is found that the model is shortage of insufficient stability by analyzing the sensitivity of environmental awareness, for lakes evaluation is a complicated system and cannot be described by simple indexes. So the model need to be improved.

Contents

1 Introduction

1.1 Present situation of Chao Hu

Chao Hu is located in the central An Hui Province, at north bank of the Yangtze River, whose basin area comes up to 13500 square kilometers. Watershed topography mainly consists of hills and plains. The key rivers surrounding the Chao Hu where the densities of river net are high take on centripetal selection distribution, what's more, lakes, reservoirs and paddy fields are widely dispersed. With the rapid growth of population and industrial production around Chao Hu catchment, water demand and wastewater discharge increase gradually, correspondingly rendering more sewage to be drained to lakes. Agricultural activities and unreasonable land use give birth to weight the load of nitrogen and phosphorus concentrations in water, inevitably, accelerating the lake ecosystem's deterioration. Especially, non-point source pollution problem is more prominent.

1.2 Previous research

The potential impact of no-point source pollution on Chao Hu ecosystem raises widely concern in the early 1990s, Zhang Zhi Yuan and Wang Pei Hua reported that no-point source on the average of TN load account for 49% of the total, and TP is about 40%, meanwhile, discussing the spatial distribution of pollutants in water according to the monitoring data of Chao Hu from 1986 to 1985. However, the initial research phase did not involve the major source of no-point source pollution, effects of different management scenarios on the lake and administrative console problems. Wang Zong Zhi exploited fuzzy clustering method to analyze the major source of several pollutants in the Chao Hu basin, thereby gaining the relationship between eleven land use patterns and nutrient outputs. Here, we focus on the need for solving the problem about linking different domains influencing pollutants input and also forecasting and evaluating the effects of different management scenarios on lakes.

1.3 Outline of our model

Firstly we divide the land into nine type in our paper. Then we build the Export Coefficient Model to get the load of N and P. Given the amount of N and P, we analyze the water quality. Next combining the meteorology and the other factors, we forecast the potentially-toxic algal blooms. Lastly, we building the evaluation model based on the AHP. What's more, we create an evaluation system. We can evaluate the Chao Hu with the evaluation model and evaluation system.

2 Problem Restatement and Analysis

2.1 Problem Restatement

Firstly, from the question we can know that the goods and services that lakes provide result from complex interactions between meteorology, hydrology, nutrient loads and in-lake processes. In addition, we know that Hydrology and nutrient loads are, in turn, influenced by socio-economic factors such as human habitation, water abstraction and land-management, within their catchments. The last, the question let us build models to link these different domains and forecast the effects of different management scenarios on lakes and evaluate the lake.

2.2 Problem Analysis

In order to build a model for forecasting the effects of different land management scenarios on lakes, we find that the reason for which land management can have great impacts on lakes is that different land use contributes different nitrogen and phosphorus concentrations which are the criterion of water quality to Chao Hu, apparently, we can formulate model based on the relationships between land management and nitrogen, phosphorus load to forecast the water quality. Besides land management, we cannot ignore the effects of temperature, hydrology and meteorology on the output of nitrogen and phosphorus.

Furthermore, nitrogen and phosphorus loads are closely related to potentially-toxic algal blooms. The number of algal blooms can be represented by chlorophyll, so the neural network can be built by chlorophyll as output and factors influencing potentially-toxic algal blooms reproduction as inputs to forecast the water quality indirectly.

Eventually, we could get data about nitrogen and phosphorus concentrations with chlorophyll, then a six-level evaluation system can be established combined with factors affecting the above data by AHP algorithm.

3 General Assumptions

- Assuming that data collected is accurate and reliable, and can reflect real conditions preferably.
- The different land use management scenarios were formulated in order to improve models' applicability, moreover, nutrient export coefficients from land are only subject to land use types, not its area and so on.
- Temperature, rainfall runoff and other factors which affect Nitrogen and Phosphorus concentrations in soil drained to Chao Hu should be at the same rate when we change the proportions of different land use.
- Under a certain scope, pollutants from agricultural fertilizer, septic tank, industrial plant and so on do not change with time and obey conservation law of mass.

4 Forecasting models about effects of different management scenarios on lakes

Different land use managements will have an impact on TN and TP load in lakes, moreover, water containing elevated level of N,P with proper temperature and other factors precipitate potentially-toxic algal blooms reproduction, to some extent, making water quality severely worse. Firstly, we establish a model to estimate nitrogen and phosphorus concentrations and forecasting model for water quality which lay a firm foundation for evaluation model.

4.1 Estimation of N,P concentrations and Export Coefficient Model

4.1.1 Introduction to models

Researchers put forward and apply export coefficient model in the process of studying the relationships between nutrient and lake eutrophication.[1]On the basis of considering land use management, we can determine export coefficients of different pollutants by combining the emission of no-point pollutants and the number and distribution of livestock. At the same time, the model also accounts for factors about nitrogen fixation in plants and air precipitation of nitrogen. The model is applicable in forecasting and evaluating no-point polluted catchment with big scale. The general expression of model is that:

$$L_i = \sum_{j=1}^{n} E_{ij} A_j + p \tag{1}$$

In the equation, i refers to the type of pollutant,j refers to the kind of nutrient source. L_i refers to the total loads of pollutant i. E_ij refers to the export coefficient of i pollutant exporting j nutrient source. A_j refers to the number of j nutrient source, p refers to the number of nutrient source due to raining.

Export coefficient models exploit principles of black box, avoiding the complicated progress which no-point pollution may happen, and it has a certain accuracy.

4.1.2 The methods of determining model's parameters

The key to solve problem is to determine a reasonable output coefficient when applying export coefficient model, on the basis of land use condition, the way to acquire export coefficient is field inspection and consulting literature. We will proceed them as follows:

Method	Field inspection	Consult literature
Operation	Monitor water quality And Calculate nutrient loads	Use previous research to get coefficient by consulting literature
Strength	High precision Reveal pollution's properties excellently	Simplicity Low costs
Weakness	Take long time High costs	Distinct characteristics Uncertainties

Figure 1: Symbols Used in This Section

4.1.3 Determine export coefficients based on hydrological materials

Area has the greatest impacts on the land use coefficient. under the premise of the law conservation of mass, we determine the different land export coefficients taking advantage of monitoring data, which is applicable in present situation being shortage of no-point pollution data.

- Classifying land use and selecting small watershed
  The land use is divided into n types according to research,then selecting m small watersheds,pollutants' inputs and outputs obey the law conservation of mass, we get

$$L_i = L_{ps} + L_{io} + \sum_{j=1}^{n} [E_{ij}(M_{ij})] A_j + p \tag{2}$$

- Solving parameter L_i

$$L_i = C_i \times \frac{Q}{k_i} \tag{3}$$

C_i refers to the average of i pollutant of monitoring concentrations a year.Q refers to annual flow.

- Geting parameter about lps and lio According to **Equation** (2) and **Equation** (3),we have

$$\frac{C_i \times Q}{k_i} = L_{ps} + L_{io} + \sum_{j=1}^{n} [E_{ij}(M_{ij})]A_j + p \tag{4}$$

L_{ps} refers to country life, L_{io} refers to the breeding of livestock

$$L_{ps} = \frac{C_k \times Q_k}{D_k \times k_i} \tag{5}$$

C_k refers to the average of i pollutant of monitoring concentrations. Q_k refers to total flows during dry season.D_k refers to the time of dry season.

$$L_{io} = \sum_{j=1}^{n} [E_{ij} M_{ij}]A_j \tag{6}$$

i refers to pollutant types, j refers to types of livestock,E_{ij} refers to the number of i th pollutant exporting to j_{th} pollution source in units. M_{ij} refers to the nutrient of i_{th} pollutant exporting to j_{th} pollution source.

- Getting solution by a set of simultaneous equations.
For pollutant i, then we have

$$\begin{cases} \dfrac{C_{i1} \times Q_1}{k_{i1}} = \dfrac{C_{k1} \times Q_{k1}}{D_{k1} \times k_{i1}} \times 365 + L_{io1} + \displaystyle\sum_{j=1}^{n} [E_{ij}(M_{ij})]A_{j1} + p_1 \\ \qquad \cdots \\ \dfrac{C_{it} \times Q_t}{k_{it}} = \dfrac{C_{kt} \times Q_{kt}}{D_{kt} \times k_{it}} \times 365 + L_{iot} + \displaystyle\sum_{j=1}^{n} [E_{ij}(M_{ij})]A_{jt} + p_t \\ \qquad \cdots \\ \dfrac{C_{im} \times Q_m}{k_{im}} = \dfrac{C_{km} \times Q_{km}}{D_{km} \times k_{im}} \times 365 + L_{iom} + \displaystyle\sum_{j=1}^{n} [E_{ij}(M_{ij})]A_{jm} + p_m \end{cases} \tag{7}$$

4.2 Forecasting model based on BP neural network

4.2.1 The selection of output and input factors

In view of particularity of studying object, we get 129 groups data processed which are from all kinds of web sites and papers. The reason why the potentially-toxic algal blooms reproduction can be predicted indirectly via forecasting the content of chlorophyll is that the total content of chlorophyll a is the most direct index to represent the quantity of algal blooms. After determining the output factor, reasonable selection of networks input factors is of great importance to apply BP neural network models accurately and guarantee the precision of the model, for input factor may exist noise factor related to outputs and redundant factor reflecting system information. Either of the two conditions will add complexity to model and difficulty to analyze problem so as to affect the predictive ability of model. The principle of sorting out input factors of neural network is to select environment factors of network structure which have nothing to with outputs.

At last, we regard N,P input totals as input factors and get rid of factor lacking data and noise factor and redundant factor.

4.2.2 Determine Network structure

Theoretical proof long before showed that, three BP layers can satisfy any complicated nonlinear function to fit approximation problem when every layer neurons adopt sigmoid function[2]. The conclusion can provide reference for ensuring algal blooms as three layers network structure, one input layer, one hidden layer and one output layer. Although selecting the number of hidden layer's neurons and activation function between layer and interlayer is regular, the results are even altogether different[3].

4.2.3 Select neurons of hidden layer

Through training neurons of different hidden layers, using mean squared normalized error performance function analyzes the numbers of neurons in hidden layer on the impacts of training effects, then we draw a conclusion from results that training effects fluctuate in a small scale with the increasement of neurons of hidden layer in number when they are more than three. Moreover, we can get the best result when the numbers of neurons are twenty, so we set up twenty neurons as hidden layer considering not many data.

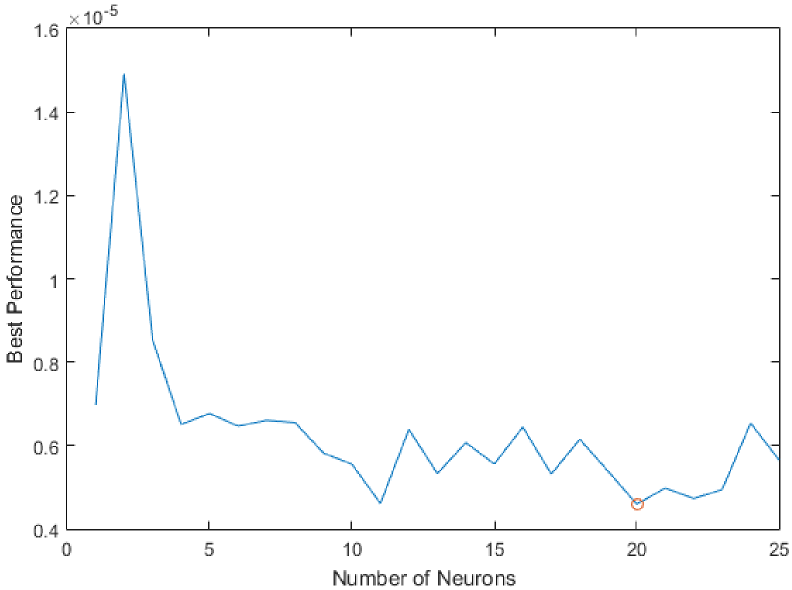

Figure 2: Best Hidden Neurons

4.2.4 Activation Function

We set up a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer

$$y = purelin(V \cdot tansig(W \cdot x + b_1) + b_2) \tag{8}$$

In the equation:x represents N,P concentrations and y represents the content of chlorophyll a ,W and b represent link weights between input layer and hidden layer. Y and b represent link weights between hidden layers and output layer.

5 Evaluation model of lakes based on AHP

Analytic hierarchy process method is a kind of qualitative and quantitative analysis of multiplicate objective decision, which is applicable in the problem hard to be solved. We sort out 7 indicators from the three aspects of ecological features, native functionality and social environment accounting for the real situation of lake ecosystem to apply in the AHP model.

5.1 Construction of AHP Model

The process of building evaluation model is as follows

- Build hierarchical structure
  We build hierarchy model by looking through materials, the result is as follows:

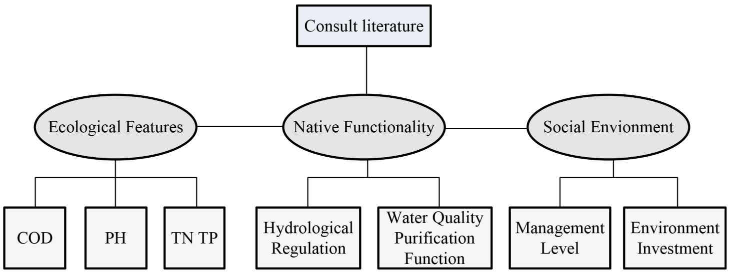

Figure 3: Construction of AHP Model

- Construct judgment matrix
  We can get corresponding judgment matrix in the way of 1-9 scales to evaluate by expertise and quantifying among factors, the results can be seen from the below figure

$$\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 1 & 7 & 5 \\ 1/7 & 1 & 1/5 \\ 1/5 & 5 & 1 \end{pmatrix} \end{array} \tag{9}$$

$$\begin{array}{c} \\ x_4 \\ x_5 \end{array} \begin{array}{c} \begin{array}{cc} x_4 & x_5 \end{array} \\ \begin{pmatrix} 1 & \frac{1}{3} \\ 3 & 1 \end{pmatrix} \end{array} \tag{10}$$

$$\begin{array}{c} \\ x_6 \\ x_7 \end{array} \begin{array}{c} \begin{array}{cc} x_6 & x_7 \end{array} \\ \begin{pmatrix} 1 & 3 \\ \frac{1}{3} & 1 \end{pmatrix} \end{array} \tag{11}$$

$$\begin{array}{c} \\ y_1 \\ y_2 \\ y_3 \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{pmatrix} 1 & 3 & 1/2 \\ 1/3 & 1 & 1/5 \\ 2 & 5 & 1 \end{pmatrix} \end{array} \tag{12}$$

x_1	x_2	x_3	x_4	x_5	x_6	x_7	y_1	y_2	y_3
TN TP	PH	COD	HYR	PUR	INV	MAN	ECO	NAR	SOC

Table 1: The description of symbols

- Description Consistency check of single hierarchical arrangement
  We define CI as Consistency index

$$CI = \frac{\lambda_{max} - n}{n-1} \tag{13}$$

$$CR = \frac{CI}{RI} \tag{14}$$

Generally, CR value is used to judge. when CR is less than 0.1, inverse symmetric matrices can be excepted. RI values can be seen in the following table

Table 2: The values of RI

n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.45	1.51	1.41

- Consistency check of total taxis of hierarchy
  Level A has m indexes like A1,A2,,,Am, and the weights of A total taxis of hierarchy are a1, am. Level B includes n indexes like B1,,,Bn. Eventually, we could get the weight of B's total taxis of hierarchy which can be seen in the below according to the A level single hierarchical arrangement.

Level A / *LevelB*	A_1	A_2	$\cdots$	A_m	*Weight*
	a_1	a_2	$\cdots$	a_m	
B_1	b_{11}	b_{11}	$\vdots$	b_{11}	$\sum_{j=1}^{m} a_j b_{1j}$
B_2	b_{21}	b_{22}	$\vdots$	b_{2m}	$\sum_{j=1}^{m} a_j b_{2j}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\sum_{j=1}^{m} a_j b_{n-1j}$
B_n	b_{n1}	b_{n2}	$\vdots$	b_{nm}	$\sum_{j=1}^{m} a_j b_{nj}$

Figure 4: Ranking table

We get consistency check of total taxis of hierarchy

$$CR = \frac{\sum_{j=1}^{m} a_j CI_j}{\sum_{j=1}^{m} a_j RI_j} \tag{15}$$

When $CR < 0.1$, the results of total taxis of hierarchy are satisfying

At last, we can get:

$$Z = -0.2172x_1 + 0.0231x_2 + 0.0685x_3 - 0.0272x_4 + 0.0818x_5 + 0.4365x_6 + 0.1455x_7 \tag{16}$$

5.2 The construction of evaluation system

We fetch optimal values from all indicators as the best condition for the evaluation model. We define this condition as the first level. On the contrary, the worst condition can be got and defined as six-level, so the condition can be divided into six levels.

6 Application of the Model

6.1 The application of evaluation model in Chao Hu

6.1.1 The construction of evaluation system on Chao Hu

We collect the best values of evaluation index as follow

Plugging values into model, we obtain the best evaluation value: 0.5

In the same way , the worst evaluation value is: 0.0

Evaluation system can be divided as shown below:

Range of Evaulation value	$0.5 \sim 0.4$	$0.4 \sim 0.3$	$0.3 \sim 0.2$	$0.2 \sim 0.1$	$0.1 \sim 0$
Level	1	2	3	4	5

Table 3: The criterion of evaluation

6.1.2 Determine evaluation grade on Chao Hu

The target data we collect are described in the following table:

Table 4: Symbols Used in This Section

Index	TN,TP	COD	PH	HYD	PUR	INV	MAN
Value	0.6	0.5	0.7	0.4	0.3	0.6	0.5

Then we get evaluation value 0.277, so it belongs to III.

6.1.3 Results Analysis

As we know from the results of AHP, the social environmental indicator's weight is biggest. The second is the ecological index. According to our evaluation system, the level of Chao Hu is . It must be caused by ecological indexes. Local government and environmental administration should take some actions to improve the status of the Chao Hu. For example , manage the land-use reasonably in the watershed and reducing the emissions of NP. From the point of land-use, increasing the area of the forest and grassland and reducing the area of plowland also can improve the status greatly.

6.2 The application of forecast model in Chao Hu

6.2.1 Water quality change prediction

Land use types and export coefficient of Chao Hu catchment are shown in the below:

Table 5: The loads of N and P in different land use

Types	Grass Land	Urban Land	Brushwood	Dry Land	Wasteland
TN	16.86	16.02	10.5	24.19	18.15
TP	3.2	4.54	2.33	2.71	3.25
Types	Dry Land	Residential Land	Wetland	Water	Paddy Field
TN	13.9	20.98	2.16	0	26.21
TP	2.64	4.53	0.9	0	3.28

As we can know from the table six, the load of TN of the unit area non-point source of various land-use is 0 26.21mg/L and the load of TN of the unit area non-point source of various land-use is 0 4.54mg/L. The paddy field and dry farm's load of unit area TN are the highest. The rural resident and urban land-use's load of unit area TP are the highest. The reason of the above situation might be the cause we give as following:

1. The daily sewage of rural-urban fringe zone and the area of plow are expanding in the Chao Hu watershed.

2. The scale of livestock farming increase.
   According to the above table seven, we can see that the TNTP discharge of the paddy field is the highest. Causing the result is because of the percentage of paddy field that is large. With the development of society, the area of urban land-use is increasing and the area of rural resident and dry farm are reducing.

6.2.2 The forecast of potentially-toxic algal blooms reproduction

We collect relevant data about Chao Hu from 129 groups, which are assigned subsequently, and 91 groups are used as train data, 19 groups as validation data and 19 groups as train data, The results converge reasonably well when data are trained at the 18 generation and the effects of curve is shown as follows:

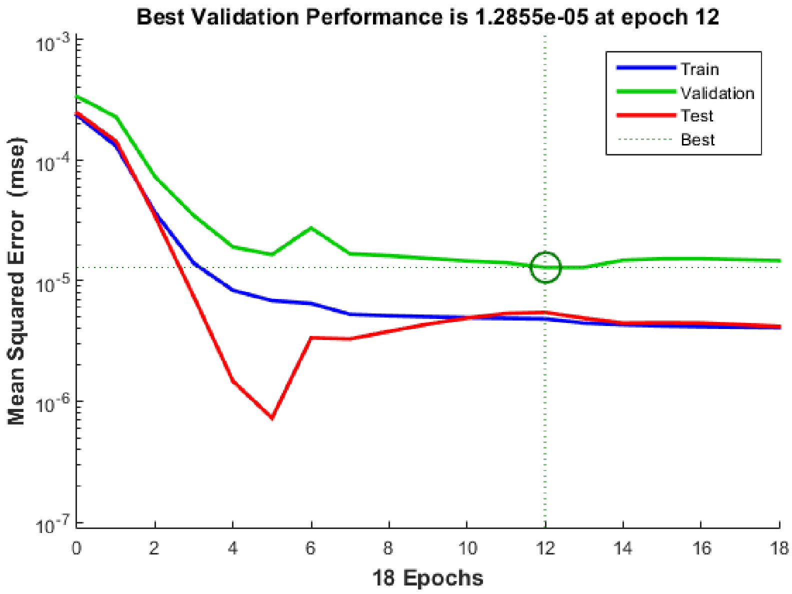

Figure 5: Training Performance

The correlation coefficient curve as follow:

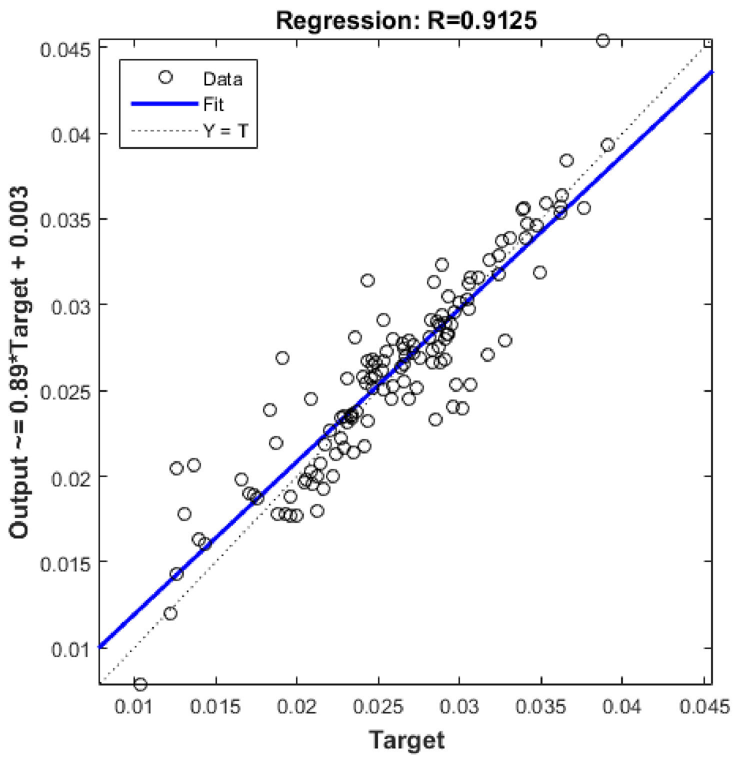

Figure 6: Regression

Analysis result, the correlation between the model trained with data and real data is up to 91%,reflecting that BP neutral network method is feasible to predict .the N,P concentrations have a great impact on the quantity of chlorophyll a and getting a better outcome.

That's to say, the growth of potentially-toxic algal blooms in water is affected by N,P nutrient load, by statistics analysis and prediction, potentially-toxic algal blooms will reproduce in water when nitrogen content exceed 2.5mg/L and nitrogen content is bigger than 0.3mol/L, furthermore, chlorophyll a content will increase with increasing N,P nutrient load.

7 The sensitivity analysis of the evaluation model

We only consider seven evaluating indexes to evaluate a lake in our evaluation model. Obviously, these indicators are incomplete. As for the other indicators, we assume that they are all at the optimal conditions. But actual situation is not always like what we assume. We test the sensitivity of the model with environmental awareness.

The environmental awareness belongs to the index system of social environment. The uncertainty of environmental awareness is large in the actual cases. We analyze the fluctuation of the result under different environmental awareness.

Sensitivity calculation results are as follows:

$$S(z, x_e) = \frac{dz}{dx_e} \cdot \frac{x_e}{z} = -3.06 \tag{17}$$

As we can see from the results, evaluation model is greatly influenced by environmental awareness. The reason for this result may be that the environmental awareness influences some indexes. What's more, stability of our model also need to be further improved.

8 Model evaluation

8.1 Strengths

The export coefficient model based on hydrology needs few parameters and its operation is simple. Not only does it have a comprehensive consideration but also certain accuracy.

We use the BP neural net model to forecast the Water quality and potentially-toxic algal blooms, avoiding analyzing the complex relationship between various factors.

Given the factors including ecologynature and society, the model of evaluation with the AHP can integrability evaluate a lake.

8.2 Weaknesses

Lack of Data Support: the data for the problem is hard to get. The data we collect are little for the model we build.

Estimated Parameters: Due to lack of data, some values used in the calculations had to be estimated.

Simplified Assumption: Simplifying assumptions had to be made in order to create a solvable model.

8.3 Future Work

Because the system of evaluation of the lake is very complex. Only seven indexes can not make it clear. Our evaluating model needs more indexes to be perfect. What's more, we should combine the AHP with the Fuzzy Algorithm rather than only the AHP to make the model of evaluation more reasonable. We should collect more data for the model to improve the stability of the model.

References

[1] Omernik J M. The influence of land use on stream nutrient levels[R]. New York: USEPA Corvallis, OR, 1976

[2] Nielsen H Komogorovs mapping neural network existence theorem[C]

Proceedings of the International Confereneeon Neural NetworksNew York:IEEE Press1987.1(3):1113

[3] Michele S,Lawrence W.Harding JR.Developing an empirical model of phytoplankton primary production:a neural network case study[J].Ecological Modeling,1999,120(59):213223.