

# A Water Quality Evaluation Algorithm Based on Factor Weighting Model

Tantan Man<sup>1, a</sup>, Huobin Tan<sup>1, b</sup> and Kebing Zhang<sup>1, c</sup>

<sup>1</sup>School of Software, Beihang University, Beijing 100191, China

<sup>a</sup>mantantan@buaa.edu.com, <sup>b</sup>thbin@buaa.edu.cn, <sup>c</sup>buaazkb@163.com

**Keywords:** water quality evaluation, factor weighting model, fuzzy mathematical model

**Abstract.** The water pollution level is affected by many factors that may have different pollution degree in different rivers. The factor weighting model combining the factor analysis model with the fuzzy mathematical model can have the advantages of both sides. The results based on the simulated pollution data set shows that this new algorithm based on the factor weighting model provides an effective water pollution evaluation. And this algorithm is thought to provide a common method for water quality evaluation.

## Introduction

As water pollution problem is becoming more serious every day, the water pollution treatment has received an universal attention and an effective water quality evaluation algorithm needs to be determined. Water quality evaluation algorithm is needed to give an accurate evaluation result of water pollution level according to different indicators like chemical oxygen demand (COD), nitrogen (N) and ammonia. It can provide basis for comprehensive control of pollution, the regional environmental planning and management.

Many scientific researchers have taken great effort to do research work dealing with the water quality evaluation methods. The nemerow index method [1] calculates the main pollutant but has slightly higher results. The synthetical index method reflect the outstanding pollution and the backward propagation (BP) neural network method [2] has difficulty in giving an accurate result when it comes to some abnormal cases.

This study intends to build an appropriate evaluation model of water quality based on processed data using factor weighting model. The water environment itself is controlled by multiple factors and it is complex and nonlinear fuzzy system [3]. The water quality evaluation algorithm based on factor weighting model can give pollution factors a dynamical weight according to their pollution degree and avoid the interference of huge data.

The data sets are simulated according to the data published from the environmental monitoring center in Yunnan province, China [4]. The expected result is to having a greater performance than the existing algorithm through the above data set.

## Factor Weighting Model

Currently, there are many models of water quality evaluation, such as the fuzzy mathematical model, the clustering model [3,5]. However, these conventional models of water quality evaluation do not solve the nonlinear relationship between the water quality observation index and the pollution grade. The factor analysis model can solve the non-linear relationship between the water quality observation index and the pollution grade, but it is too absolute for calculating the membership of single observation index. Experiments [6] show that the combination of two approximate mathematical models can have the advantages of both sides. We build our model by combining the factor analysis model with the fuzzy mathematical model to evaluate water quality more efficiently and precisely.

We named the model “factor weighting model”. Steps are as follows:

**Step1:** We combine factor membership degree with fuzzy membership degree to get a new factor weighting membership degree:

$$ly(x) = \left\{ \begin{array}{l} 0 \\ \left\{ \frac{(x-x_0)}{(x_1-x_0)} \mid \frac{(x_1-x)}{(x_1-x_0)} \right\} \\ 1 \end{array} \right\} (x < x_0, x_0 \leq x < x_1, x \geq x_1) \quad (1)$$

Where:  $x_0$  is the former value of water sample.

$x_1$  is the latter value of water sample.

$x$  is the current value of water sample.

**Step2:** Get factor membership degree matrix  $R$  of the water quality indicators:

We set  $W$  as the collection of water quality sample and  $L$  as the collection of indicator's pollution level. Establish the following formula:

$$\left\{ \begin{array}{l} W = \{A_1, A_2, \dots, A_n\} \\ L = \{L_1, L_2, \dots, L_m\} \end{array} \right\} \quad (2)$$

Where:  $A$  is the value of sample.

$n$  is the number of targets.

$m$  is the number of pollution levels.

Calculate the factor membership degree of individual indicators by equation (2). Then, we can get factor membership degree matrix  $R$  of the water quality indicators.

$$R = \begin{array}{ccc} A_1L_1 & \dots & A_1L_m \\ \dots & \dots & \dots \\ A_nL_1 & \dots & A_nL_m \end{array} \quad (3)$$

**Step3:** Calculate the comprehensive weight matrix  $B$  of water quality:

As water pollution is caused by a number of indicators, the concentration of different substances should have different influences on the weight of the comprehensive assessment. In theory, the greater the level of contamination from certain pollutants, the more severe the degree of water pollution. The weight of single indicators is calculated as the formula (4).

$$W_i = \frac{A_i}{\bar{L}_i} \quad \text{or} \quad W_i = \frac{A_i}{\frac{1}{m} \sum_{k=1}^m L_k} \quad (4)$$

Where:  $A_i$  is the value of the current sample.

$L_k$  is the single pollution level of indicator.

Unite the weights of all indicator by using the formula (5):

$$\overline{W}_i = W \frac{i}{\sum_{k=1}^n W_k} \quad (5)$$

Where:  $W_k$  is the Unified weight of single indicator.

Suppose the number of water quality indicators is  $n$ , we can get the comprehensive weight matrix  $B$  of water quality.

$$B = [\overline{W}_1, \overline{W}_2, \dots, \overline{W}_n] \quad (6)$$

**Step4:** Calculate water quality weight by the composite calculation of matrix  $B$  and matrix  $R$ , and get the result:

$$(BR)_{ij} = \max(\{\min(B_{i1}, R_{1j}), \dots, \min(B_{ik}, R_{kj}), \dots, \min(B_{in}, R_{nj})\}) \quad (7)$$

Where:  $B_{ij}$  is the element of the  $i$  row and  $j$  column in the matrix  $B$ .

$R_{ij}$  is the element of the  $i$  row and  $j$  column in the matrix  $R$ .

We can get the matrix  $Y$  of water quality weight by using the formula (7):

$$Y = [y_1, y_2, \dots, y_{m-1}, y_m] \quad (8)$$

In the matrix  $Y$ , each element corresponds to the weight of the water quality level. The pollution level corresponding to the element with the highest weight is the evaluation result of water quality.

## Empirical analysis

According to the "national surface water quality standards" in China [4], the water quality is

divided into five types which represent by I , II, III, IV, V, respectively. To better validate this model, we simulate the assessment criteria for water quality at all types. The Simulated pollution factors standard values of all types are shown in Table 1.

Table 1 Simulated quality standard of surface water

Types	I	II	III	IV	V
<b>COD</b>	20	40	60	80	100
<b>NH3-N</b>	10	15	30	70	100
<b>BOD5</b>	25	50	69	88	99
<b>TP</b>	12	23	56	77	89
<b>E. coli</b>	10	30	45	80	100

In order to visually validate this model, we made simulated samples for general water quality and heavily polluted water quality. The simulated data of pollution factors are shown in Table 2.

Table 2 The simulated measure data of pollution factors

Sample	COD	NH3-N	BOD5	TP	E. coli
<b>1</b>	10	15	8	7	9
<b>2</b>	10	20	10	15	1005
...					

We use the algorithm based on the factor weighting model to get the matrix of water quality evaluation of Table 3. The evaluation results are shown in Table 3. The final weight of sample 1 is “ $Y_1 = [0.014, 0.022, 0.023, 0, 0.974]$ ”. The weight matrix of sample 2 is “ $Y_2 = [0.014, 0.022, 0.023, 0, 0.974]$ ”. Each element corresponds to the weight of the water quality type that the samples belong to.

Table 3 Evaluation results of the factor weighting model

Sample	I	II	III	IV	V
<b>1</b>	0.343	0.519	0.113	0.02	0.013
<b>2</b>	0.014	0.022	0.023	0	0.947
...					

In order to classify the water quality visually, we convert the weight matrix of each sample to a pie chart. As the maximum weight of sample 1 belongs to Level II, the evaluation result of sample 1 is Level II. In the same way, the evaluation result of sample 2 is Level V.

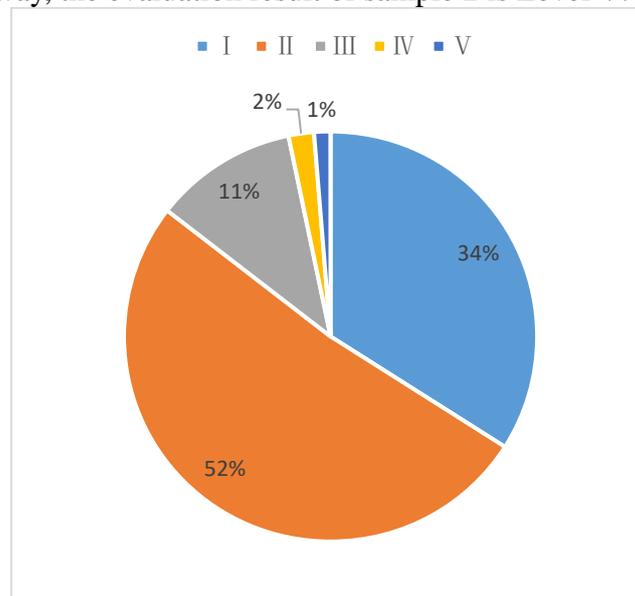


Fig. 1 Weights of all water quality level that sample 1 belong to

## Conclusion

In order to find a universal and efficient water quality evaluation algorithm, we combine the advantages of fuzzy and factor analysis models to establish an algorithm based on factor weighting model. The algorithm takes into account the inter relation of each index, and overcomes the disadvantages of the subjective weight. The factor model can not only sort and grade water samples, but also solve the complex nonlinear relationship between the evaluation factors and water quality levels. Experimental results show that the algorithm based on factor weighting is a general and effective method for water quality evaluation.

## References

- [1] Jie C, Qing L, Hui Q. Application of improved nemerow index method based on entropy weight for groundwater quality evaluation[J]. International Journal of Environmental Sciences, 2012, 2(3):1284-1290.
- [2] Miao Q, Zhao L, Gao A, et al. Optimization Design and Application of Water Quality Evaluation Model based on BP Neural Network[J]. 2010:1-5.
- [3] Wang D, Singh V P, Zhu Y. Hybrid fuzzy and optimal modeling for water quality evaluation[J]. Water Resources Research, 2007, 43(5):297-304.
- [4] Information on <http://www.ynem.com.cn/>.
- [5] Zhu C, Liu Q. Evaluation of Water Quality using Grey Clustering[C]// International Workshop on Knowledge Discovery & Data Mining. 2009:803-805.
- [6] Nováková P, Št'astná, M. Combination of the factor analysis and Fourier transformation during evaluation of the water quality.[J]. Acta Universitatis Agriculturae Et Silviculturae Mendelianae Brunensis, 2005:95-102.