Evaluation of Water Quality using Grey Clustering

Changjun Zhu
College of Urban Construction
Hebei University of Engineering
Handan, 056038,China
christorf@126.com

Qinghua Liu
Huaihe Conservancy & Hydropower Development Company
Bengbu, 233001, China

***Abstract*—Evaluation of water environment quality plays an important role in environment science. Because of many factors affecting environment quality, it is a basic task to select rational pattern and make full use of limited information from monitors so as to describe environment quality objectively. In view of the deficiency of the traditional methods, based on the grey theory, a grey clustering model is established to evaluate water quality. The proposed model was applied to assess the water quality of 20 sections in Suzhou River. The evaluation result was compared with that of the traditional method and the reported results in the Suzhou River. It is indicated that the performance of the proposed model is practically feasible in the application of water quality assessment and its application is simple.**

Keywords-evaluation; water quality;grey clustering

I. INTRODUCTION

Water quality assessment is an important monitoring project. At present, there are many methods to evaluate the water quality ,such as Fuzzy Comprehensive Evaluation, Unascertained measure, Gray correlation analysis method, gray clustering method, integrated pollution index method[1][2]. Every method has his own characters. The characters of environment has complexity and comprehensive. Panfeng(2003) applied fuzzy comprehensive evaluation in evaluating water quality. Xu junjie (2002) recognized that the environment problem is multi-factors problems. These traditional methods can not solve the complicated nonlinear relationship between evaluations indicators and the grade of water quality. Because there are many factors affecting the water quality in river, and the factors is nonlinear relations with water quality. Traditional assessment can not meet the demands for evaluating precision, while artificial Neural networks are loosely based on the neural structure of the brain which provide the ability to learn from the input data they are given and then apply this to unknown data, in effect they can generalize and associate unknown data. The objectives of this study are to construct models for evaluating the water quality at Suzhou river using the grey clustering method.

II. PRINCIPLE THEORY

Grey clustering method is based on the whiten function of grey. The process of grey clustering can be described as follows:

(1) $i = 1,2,\cdots\cdots,n$ represents clustering sample which is also called surveying sections;

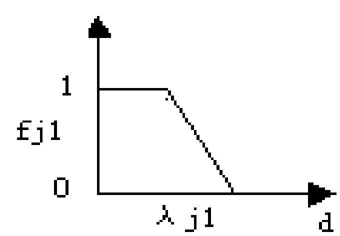
$j = 1,2\cdots\cdots m$ represents clustering index which is pollutant index in this paper;

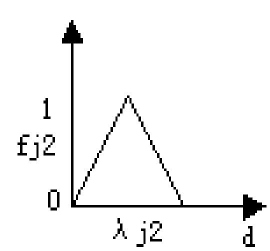
$k = 1,2,\cdots\cdots K$ represents gray classification

(2) According to surveying sample, whiten number $d_{i,j}$ ($i \in [1,2,\cdots\cdots,n]$ $j \in [1,2,\cdots\cdots,m]$) can be got. Sample matrix is as follows

$$D = (d_{i,j}) = \begin{vmatrix} d_{11} & d_{12} & \cdots\cdots & d_{1m} \\ d_{21} & d_{22} & \cdots\cdots & d_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ d_{n1} & d_{n2} & \cdots\cdots & d_{nm} \end{vmatrix}$$

(3) According to grading standards, whiten function $f_{jk} \in [0,1]$ can be made. The whiten function has three kinds of basic form. λ_{jk} is Threshold value.





978-0-7695-3543-2/09 $25.00 © 2009 IEEE
DOI 10.1109/WKDD.2009.28

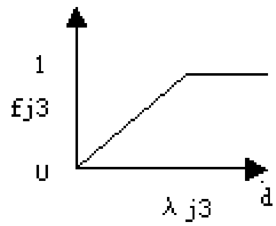


Figure 1. Whiten function f_{jk}

(4) Clustering weight $\eta_{jk} = \frac{\lambda_{jk}}{\sum_{j=1}^{m} \lambda_{jk}}$. When the dimension of clustering index is different, and the value of the indexes are different each other, zero dimension can be processed.

(5) Clustering coefficient σ_{ik} . $\sigma_{ik} = \sum_{j=1}^{m} f_{jk}(d_{ij}) \bullet \eta_{jk}$;

(6) Clustering vector

$$\sigma_c = \begin{vmatrix} \sigma_{11} & \sigma_{12} & \cdots\cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots\cdots & \sigma_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots\cdots & \sigma_{nk} \end{vmatrix}$$

(7) clustering

$$\sigma_{ik^*} = (\sigma_{ik})^{\max}$$

It represents that the clustering object belongs to gray classification k^*. in row vector $\sigma_i = (\sigma_{i1}, \sigma_{i2}, \cdots\cdots \sigma_{ik})$, the grey classification corresponding with the maximum clustering coefficient is the rank of water quality.

III. APLICATION

A. Study area

Suzhou is a famous city with much water, in which river port interlocks and numerously lake. water surface in whole city approximately is 3607 square kilometers, which approximately composes the total area 42.52%. The urban district water surface is approximately 24 square kilometers, composing the urban district area 20.15%. Suzhou altogether has more than 4000 bigger lakelet, the big lake swings has 87; Altogether has size rivers 20,000, total length 1457km. Outside the moat the Suzhou old city area circle will become a relatively independent region, spreads across the river course has formed “three horizontal three straight link” the urban river course network of rivers and lakes system with the city. The rivers height is equally 0.8-1.3m (the Woosung elevation), the water depth is equally 3 meters, the gradient is zero nearly. The river water excretes weakly, therefore the fluent speed of flow is very small, mean velocity of stream is 0.05m/s~0.1m/s.

B. Sample and Standard

In 2004, water quality is monitored in 20 sections(Table2). The sections are regarded as clustering object($i = 1,2,\cdots\cdots,20$),the pollutant index is clustering index($j = 1,2,3,4,5$), d_{ij} is sample value of jth pollutant index at ith section. the standard of evaluation is adopted as “surface water quality Standard”（GH2B1-2002）which can be seen in table1.

TABLE I. SURFACE WATER QUALITY STANDARD（GB3838-2002）

Rank	Pollutant index		
	CODmn	***BOD5***	***DO***
I	2	3	7.5
II	4	3	6
III	6	4	5
IV	10	6	3
V	15	10	2

TABLE II. OBSERVED DATA MG/L

Section	Pollutant index		
	CODmn	***BOD5***	***DO***
1	7.00	5.10	1.70
2	7.30	5.40	2.60
3	7.00	5.50	2.60
4	6.60	4.80	4.40
5	7.20	4.60	2.30
6	10.00	15.80	0.90
7	11.90	27.80	1.80
8	7.80	15.60	0.60
9	7.40	6.50	1.70
10	7.40	4.50	3.20
11	7.10	5.80	3.70
12	8.00	6.40	4.00
13	7.10	4.60	3.90
14	6.90	4.40	3.50
15	7.70	7.40	2.20
16	7.30	6.00	1.60
17	7.00	5.50	1.30
18	10.10	9.40	0.50
19	10.50	17.60	1.05
20	14.20	27.20	25.00

Before evaluating the water quality, the data should be preprocessed as followings

$$\frac{x}{\bar{x}}$$

C. Whiten function

According to the basic style of white function, the white function in this paper can be described as followings:

$$f_{j1}(d)=\begin{cases}1 & d\in[0,\lambda_{j1}]\\ \dfrac{\lambda_{j2}-d}{\lambda_{j2}-\lambda_{j1}} & d\in[\lambda_{j1},\lambda_{j2}]\\ 0 & d\in[\lambda_{j2},\infty)\end{cases}$$

$$f_{jk}(d)=\begin{cases}\dfrac{d-\lambda_{j,k-1}}{\lambda_{j,k}-\lambda_{j,k-1}} & d\in[\lambda_{j,k-1},\lambda_{j,k}]\\ \dfrac{\lambda_{j,k+1}-\lambda_{j,k}}{\lambda_{j,k+1}-\lambda_{j,k}} & d\in[\lambda_{j,k},\lambda_{j,k+1}]\\ 0 & d\in[\lambda_{j,k+1},\infty]\end{cases}$$

$$f_{j,k}(d)=\begin{cases}0 & d\in[0,\lambda_{j,k-1}]\\ \dfrac{d-\lambda_{j,k-1}}{\lambda_{j,k}-\lambda_{j,k-1}} & d\in[\lambda_{j,k-1},\lambda_{j,k}]\\ 1 & d\in[\lambda_{j,k},\infty)\end{cases}$$

D. Grey Clustering weight

According to $\eta_{jk}=\dfrac{\lambda_{jk}}{\sum_{j=1}^{m}\lambda_{jk}}$, the weight can be calculated.

TABLE III. WEIGHT OF GREY CLUSTERING

Index	Grey classification k				
	□	*II*	*III*	*IV*	*V*
DO	0.6592	0.575	0.343	0.181	0.07
BOD_5	0.1306	0.212	0.317	0.417	0.515
COD_{cr}	0.2102	0.273	0.340	0.402	0.415

E. Results

According to the above process, water quality in Suzhou city can be evaluated as followings. The result can be shown in table4.

TABLE IV. EVALUATION RESULT

Section	$\sigma_{ik}=\sum_{j=1}^{3}f_{jk}(d_{ij})\bullet\eta_{jk}$	**Rank**
1	（0.0910,0.0712,0.015,0.154,0.1013）	Ⅳ
2	（0.1011,0.1618,0.1123,0.345,0.125）	Ⅳ
3	（0.1022,0.0935,0.021,0.185,0.0873）	Ⅳ
4	（0.0789,0.1439,0.018,0.147,0.2246）	Ⅴ
5	（0.0495,0.1479,0.125,0.164,0）	Ⅳ
6	（0.1328,0.097,0.1162,0.1021,0.1308）	Ⅰ
7	（0.2499,0.0292,0.1330,0.1255, 0.1703）	Ⅰ
8	（0, 0.1298, 0.1049,0.1345,0.0737）	Ⅳ
9	（0.1002,0.1149,0.045,0.2389,0.0012）	Ⅳ
10	（0.0049,0.1377, 0.0802,0.14505,0.1020）	Ⅳ
11	（0.1071,0.1033, 0.097,0.2874,0.1165）	Ⅳ
12	（0.1250,0.1378, 0.0978,0.2497,0.0861）	Ⅳ
13	（0.0645,0.1554,0.1453,0.15539,0.1650）	Ⅴ
14	（0.0027,0.0885, 0.1147,0.1380,0.0673）	Ⅳ
15	（0.1405,0.0411,0.1875,0.2019,0.1162）	Ⅳ
16	（0.1276,0.1492,0.1230,0.1583,0.1079）	Ⅳ
17	（0.2855, 0.0645,0.1962, 0.0252, 0.1057）	Ⅰ
18	（0.0126,0.0296,0.1174,0.2213,0.1644）	Ⅳ
19	（0.0192, 0.1713,0.0866,0.2443,0）	Ⅳ
20	（0.1026,0.1944,0.2283,0.3177,0.0788）	Ⅳ

IV. CONCLUSIONS

Application of grey clustering to the water quality prediction is a novel research area. Based on the modeling results obtained in this study, the following conclusions can be drawn:

- The grey clustering method is proposed by virtue of the dynamic characteristics of water quality, which increased the evaluation precision. Therefore, the method is reliable and effective.
- The grey clustering which can overcome the disadvantages of single factors can reflect the water quality at present. This model can approach the reality.
- This method is not only suitable for the evaluation of the river water quality, but also suitable for other respects (such as groundwater quality, prediction of quality in the atmospheric environment, and so on).

REFERENCES

[1] XIANG Su-lin, Yang Liu-chun, Ran Quan.Application of Regression Analytical Method in Dynamic Prediction of Groundwater Pollution[J].Journal of Yangtze University,2006,3(1):40-42.

[2] Wang Hong-yu, Ma Fang, Yang Kai et al. A Study of Information Renewal GM(1,1) Model for Predicting Urban Medium and Long-term Water Demand[J].Engineering Journal of Wuhan University,2004,37(6):32-35.

[3] LIN Li-rong, JIN Meng-gui, WANG Xian-guo. Grey prediction of data sequence of the hardness in shallow groundwater in Luoyang[J]. Hydrogeology and Engineering Geology, 2006(1):41-45.

[4] Deng Ju-long. Basic Method of Gray System[M].Wuhan: Central China University of Science and Technology Press,1996.

[5] Wang Kai-zhang, Liu Fu-sheng, Sun Ming. The Application of Gray Model in Dawu Water Quality Prediction Water Resources Site[J].Journal of Shandong Agricultural University (Natural Science),2002,33(1):66-71.

[6] LIU Xian-feng , ZOU Ji-yan. Forecasting of Cell Capacitance Based on Grey Theory[J]. Journal of Dalian University of Techonology, 2005,45(5):630-632.

[7] LI Ru-zhong, Wang Jia-quan,Qian Jia-zhong. Prediction of River Water Quality Based on Gray Dynamic Model Group[J].Bulletin of Soil and Water Conservation,2002,22(4):10-13