

Optimization Design of Monitoring Sections of Qinghe River Basin

Xuejiao Feng^{1,a}, Xiang He^{1,b}, Jinxiang Fu^{1,c}, Fan Wang^{1,d}

¹Shenyang Jianzhu University, Hunnan District Hunnan Donglu Number
9, Shenyang, Liaoning, China: 110168

a.email: 1611916143@qq.com, b.email: 307805543@qq.com, c.email: fujinxiang@sina.com

d.email: 609341881@qq.com

Keywords: Qinghe River; Sections Optimization; K-means Cluster; Matter Element Analysis.

Abstract. K-means fuzzy cluster analysis and matter element analysis have been used to select optimize points of water quality monitoring in Qinghe River, taking 8 monitoring section data in June 2014. K-means fuzzy clustering analysis method is suitable for the processing of continuous data set, and matter element analysis method is for handling multi-objective data set. The two methods are combined to optimize single mathematical statistical method. As a result of optimization, 8 monitoring points can be reduced to 6 monitoring sites, according to sectional properties and actual situation.

Water quality monitoring sections are set up to meet the needs at a minimum cost and maximum efficiency to make the monitoring section (station) have the best overall function. The arrangement of environmental monitoring point has relation to the success of the work of environmental monitoring, and optimal layout is an important link in the manifestation of scientific environmental monitoring[1].

Presently, being applied for section optimization across the world, this method is mainly utilized to analyze whether the adjacent section needs to be deleted or not, according to similarity of the monitoring data. The common monitoring section optimization methods are Experience Formula Method [2], Fuzzy Clustering Analysis[3], Multi-Objective Decision Analysis Method[4], Principal Component Analysis[5], Matter-element Analysis Method[6], Genetic Algorithm[7], Dynamic Degree Method [8] etc.

This study is based on fuzzy clustering analysis method, with the k-means clustering analysis and the matter-element analysis being carried out on monitoring section optimization. K-means clustering method is suitable for processing continuous data and large data set. The matter element analysis method is an effective method for pollutant indexes to handle multiple incompatible problems.

1. Qinghe water quality data

Only one section called Qing Liao control the quality of Qinghe water in Tieling city. In this study, we monitored 8 sections of Tieling Qinghe River to get the data from June 15th to June 21st in 2014. With the use of above data, we propose a method of simulation, verify feasibility of the method, and provide a reference for the relevant research.

Qinghe is the first level tributaries of Liaohe. We monitored 8 section samples of Qinghe in Tieling region from June 15th to June 21st in 2014. And the data of monitoring sections are as follows:

Table 1-1 Qinghe River water quality monitoring data

Section name	COD (mg/L)	BOD (mg/L)	Total nitrogen concentration (mg/L)	Ammonia concentration (mg/L)	Petroleum concentration (mg/L)	Total phosphorous concentration (mg/L)
Big Gu Family Town	5.0	1.6	3.98	0.012	0.04	0.13
Big Three Families	5.0	1.5	4.21	0.104	0.08	0.13
Eight Trees	15.0	1.1	4.04	0.058	0.07	0.11
Qinghe Reservoir Inventory Mouth	10.1	1.6	5.71	0.063	0.07	0.07
Qinghe Reservoir Dispatch Mouth	16.3	2.9	5.36	0.016	0.07	0.03
Stone Tower	18.2	2.6	6.71	0.200	0.06	0.03
Two Stockaded Village (main)	17.0	5.0	6.98	0.417	0.06	0.06
Qing Liao	15.3	4.0	7.05	0.623	0.05	0.05

Qinghe river 8 monitoring sections are respectively called Big Gu Family Town, Big Three Families, Eight Trees, Qinghe Reservoir Inventory Mouth, Qinghe Reservoir Dispatch Mouth, Stone Tower, Two Stockaded Village (main) and Qing Liao.

2. Fuzzy clustering analysis method

2.1 K - means algorithm

K - means clustering analysis algorithm use two stage repeated cycle, and condition of ending the cycle is that there is no data elements being redistributed. Algorithm firstly select k points randomly as the initial clustering center, and then calculate the sample's distance to the center of the cluster. The sample is then attributed to the nearest clustering center ; If two adjacent clustering centers don't have any change after the calculation, it will be the end of sample adjustment, explaining that the clustering criterion function is convergent.

This paper used SPSS software clustering analysis module to do K means clustering analysis simulation, with utilization of corresponding function in SPSS calculation. And we utilize single factor variance analysis of SPSS to test the result of clustering.

2.2. The fuzzy clustering analysis method to simulate Qinghe river sections

In order to compare indicators for monitoring data which has correlation in some degree, we give standard pretreatment to raw statistics by SPSS software. The calculation method of SPSS data standardization is standard root extraction method, in which standard extraction use original data minus average data and then divide the average standard deviation .By using the software of SPSS , we select clustering number 4 and 10 times iterations to analyze, through the original data and experience .

Output results include initial clustering center and the distance between each point to the initial one which is calculated by Euclidean distance formula.

Table 2-1 Initial cluster center

Number	Cluster			
	1	2	3	4
1	-1.45090	-1.45090	.48051	1.02430
2	-.68171	-.75442	1.06346	.04545
3	-1.15389	-.97986	1.16903	.91177
4	-.79010	-.37384	1.97441	.06052
5	-1.75543	1.36533	-.97524	-.19505
6	1.28816	1.28816	-.62910	-1.10842

Table 2-2 Iteration historical recorda

Iteration	Changes within the cluster center			
	1	2	3	4
1	.000	1.306	.736	.794
2	.000	.000	.000	.000

Because the clustering center are in small changes or no changes ,it comes to achieve convergence.The absolute maximum coordinates of any center is changed at .000.The current iteration is 2,and the minimum distance to the center is 2.429.

Table 2-3 Cluster members

Case number	Cluster	Distance
1	1	.000
2	2	1.306
3	2	1.109
4	2	1.170
5	4	.794
6	4	.794
7	3	.736
8	3	.736

Table 2-4 Final cluster centers

Number	Cluster			
	1	2	3	4
1	-1.45090	-.50707	.63990	.84616
2	-.68171	-.82714	1.42704	.15452
3	-1.15389	-.64442	1.14254	.40103
4	-.79010	-.50505	1.50838	-.35574
5	-1.75543	.84521	-.58514	.19505
6	1.28816	.64907	-.50927	-1.10842

The analysis of variance is carried out from the result of clustering between categories distance based on the SPSS software. Variance analysis results show that corresponding statistics of P to the value of F was 0.01, which is significantly less than the level of $\alpha=0.05$, rejecting the null hypothesis. This comes to explain that under the significant level of 5%, there are obvious differences between three samples, and the clustering effect is good.

Table 2-5 Final cluster centers between the distance

Cluster	1	2	3	4
1		2.902	4.896	4.258
2	2.902		4.127	2.724
3	4.896	4.127		2.580
4	4.258	2.724	2.580	

Combined with actual optimization, results are as follows: Reserve 1, 4, 5, 8 section and delete 2, 3, 7 sections.

3. The matter-element analysis

3.1 The matter-element analysis mathematical principles

Matter element is to describe basic units of objects, which are composed of objects, characteristics and value constitution of objects. The objects are given the name of N , with the value about the characteristics of C being V , and matter-element can be expressed as three ordered tuple $R=[N,C,V]$. If multiple characteristics have corresponding value for N , it will be described as N dimension matter-element. And it is as follows

$$R = \begin{bmatrix} N & c_1 & v_1 \\ & c_2 & v_2 \\ & \vdots & \vdots \\ & c_n & v_n \end{bmatrix} \quad (1)$$

1) Comprehensive correlation functions for each point

Put the data into (2)(3), and linear correlation functions of single pollution index for each point are calculated.

$$K_a(X_{ij}) = \frac{X_{ij} - C_j}{C_j - A_j} \quad (2)$$

$$K_b(X_{ij}) = \frac{X_{ij} - C_j}{C_j - B_j} \quad (3)$$

X_{ij} : The measuring point I , j indicators to measure value

A_j : The minimum section J index (best value)

B_j : The maximum section J index (worst value)

C_j : The expected value of J index;

S_j : J measures of monitoring level limits the average;

x_j : J measures of monitoring the best ideal point and the ideal point average.

From the above results to calculate the comprehensive pollution index of all correlation functions:

$$w_j = \frac{x_j / s_j}{\sum_{j=1}^n (x_j / s_j)} \quad (4)$$

$$K_a(X_i) = \sum_{j=1}^m W_j \cdot K_a(X_{ij}) \quad (5)$$

$$K_b(X_i) = \sum_{j=1}^m W_j \bullet K_b(X_{ij}) \tag{6}$$

W_j :J pollution indicators weight

Using index exceeding bid method to calculate according to surface water environment quality standard , pollution indexes can be obtained at normalized weights.Water quality value are calculated along with all comprehensive correlation functions $K_a(X_i)$ and the calculated value $K_b(X_i)$.

2)Evaluation standard

The correlation function K (x) represents the evaluation of membership degree which the units comply with in standard range.When $K(x) \geq 1.0$, it represents the object evaluated being over standard limit value . And the greater the value is,the larger the development potential is . When $0 \leq K(x) \leq 1.0$, it represents evaluation objects are in accordance with standard requirements of degree of the objects,and the bigger the numerical value is , the closer it is to the standard limit.When $-1.0 \leq K(x) \leq 0$, it represents the object under evaluation does not accord with a standard request,but has the condition to transform into a standard object,and the bigger the value is , the easier it can transform.When $K(x) \leq -1.0$,it represents the evaluation object does not meet standard requirements,and does not have the ability to transform into a standard object either.

3.2 Section optimization and selection

Six indexes are used in the matter element analysis,which respectively are chemical oxygen demand COD, biochemical oxygen demand BOD, ammonia nitrogen in NH3-N, total nitrogen, total phosphorus and oil.Based on the matter element analysis theory, we formulate indicators of "best ideal point a", "the most ideal point B "and "point C mathematical expectation".Through the original data table,we can get the best point, the most ideal point and the mathematical expectation of the table below:

Table 3-1: Statistical results of water quality monitoring indicators

	COD	BOD	Total nitrogen	Ammonia	Petroleum	Total phosphorus
The best ideal point a	5.0	1.1	3.98	0.012	0.04	0.03
The worst ideal point b	18.2	5.0	7.05	0.623	0.08	0.13
Mathematical expectation point c	12.74	2.54	5.51	0.19	0.0625	0.076

Based on the matter element analysis theory, we utilize index exceeding method according to "surface water environment quality standard" (GB3838-2002) to get the pollution index normalized weights which are shown in the following table:

Table 3-2 Calculation of monitoring indicator weights w_j

Item	I	II	III	IV	V	s_j	x_j	x_j/s_j	w_j
COD	≤ 15	15	20	30	40	24	11.6	0.483333	0.061164
BOD	≤ 3	3	4	6	10	5.2	3.05	0.586538	0.074224
Total nitrogen	≤ 0.2	0.5	1.0	1.5	2.0	1.040	5.515	5.302885	0.671058
Ammonia	≤ 0.15	0.5	1.0	1.5	2.0	1.030	0.318	0.308738	0.03907
Petroleum	≤ 0.05	0.05	0.05	0.5	1.0	0.33	0.06	0.181818	0.023008
Total phosphorus	≤ 0.02	0.1	0.2	0.3	0.4	0.077	0.08	1.038961	0.131476

According to formula (4) , the value of W_j can be calculated which can help to prepare subsequent calculation of correlation function. At the same time, according to formula(2), (3) ,the value of $K_a(X_{ij})$ and $K_b(X_{ij})$ can be calculated.According to formula (4), (5), the value of water quality monitoring points for comprehensive correlation function $K_a(X_i)$ and $K_b(X_i)$ can be calculated.The following table 3-3 shows:

Table 3-3 Calculated of water quality monitoring point comprehensive correlation function $K_a(X_i)$ and $K_b(X_i)$

Monitoring point	$K_a(X_i)$	$K_b(X_i)$	Monitoring point	$K_a(X_i)$	$K_b(X_i)$
1	-0.68841	0.695934	5	-0.1811	0.132459
2	-0.53159	0.537838	6	0.440722	-0.47149
3	-0.62523	0.577955	7	0.806743	-0.74074
4	-0.01895	-0.01301	8	0.778875	-0.70312

According to the data in the table, we draw up diagrams combining with the meaning of correlation function. $K_a(X_i)$ is abscissa, and $K_b(i)$ is ordinate, which are as follows:

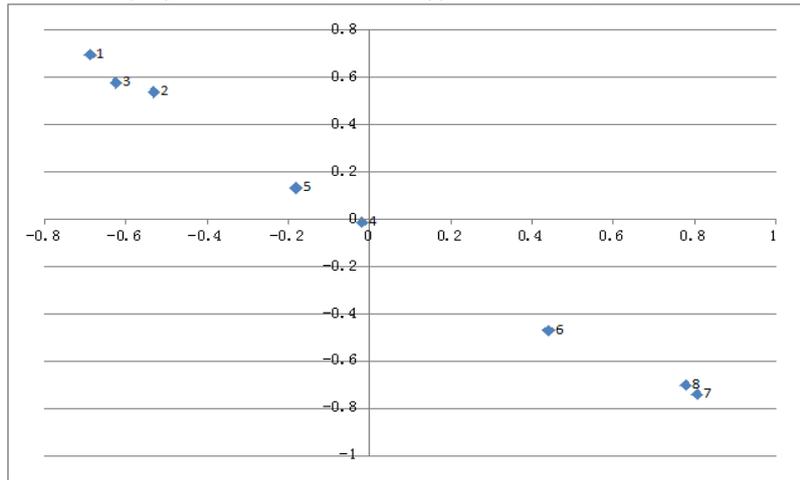


Fig 3-1: Scatter diagram of the monitoring point correlation function

Therefore, combined with the actual situation of matter element analysis, the Qinghe River's optimization result is: 1, 2, 4, 5, 6, 8 reserved, 3, 7 section removed.

4. The optimization results and discussion section

By the method of fuzzy cluster analysis and matter element analysis, the results of two methods for Qinghe River optimization section scheme can be seen as follows:

Table 4-1: Qinghe River surface water environment monitoring section optimization program

River	The section category	The existing section	Trade-off situation	Trade-off reason
Qinghe River	The background section	Big Gu Family Town	Keep	Qinghe source
	The contrast section	Big Three Families	Keep	Qinghe river contrast section
	The contrast section	Eight Trees	Cancel	Significantly correlated with three subsidiary fracture monitoring data or extremely significant correlation
	The contrast section	Qinghe Reservoir Inventory Mouth	Keep	Qinghe reservoir inventory mouth
	The control section	Qinghe Reservoir Dispatch Mouth	Keep	Qinghe reservoir dispatch mouth
	The control section	Stone Tower	Keep	Tieling industrial park upstream
	The control section	Two Stockaded Village (river)	Cancel	With two stockaded village (river) monitoring mouth repeatedly
	The cut section	Qing Liao	Keep	Qinghe in Liaohe cross section, the control section

Through the above study, it is realistic to reduce 8 monitoring sections to 6. Matter element analysis method combined with method of fuzzy cluster analysis on section optimization can make up for defects of single statistical method, and the results are more scientific and reliable. Moreover, this kind of method is also more suitable for multiple river basins.

Reference

- [1] Jiang Li & Ma Fei. Based on the water quality monitoring section of Matlab optimization setting research [J]. Journal of environmental protection in Inner Mongolia. 2006, 18 (3) : 48-50.
- [2] Dali Wang. River water quality monitoring control section set study [J]. Journal of environmental monitoring management and technology. 2001, 13 (004) : 41-42.
- [3] Yazhong Li, Xiaojie Hao. Shanxi fenhe river basin surface water environmental monitoring cross section optimization Settings [J]. Journal of north university (natural science edition), 2011, 32 (6).
- [4] Chilundo M, Kelderman P and Others. The Design of a water quality monitoring network for the Limpopo River Basin in Mozambique [J]. Journal of Physics and Chemistry of the Earth, Parts a/B/C. 2008, 33 (8) : 655-665.
- [5] Yunlong Ni. Monitoring points optimization mathematical model of water environment study [J]. Journal of environmental science and technology in jiangsu. 2007, 20 (2) : 58-60.
- [6] Xiaofeng Gong, Chunli Chen, Jin Zhao, etc. Poyang lake river water quality monitoring of lean optimization stationing [J]. Journal of lake science. 2006, 18 (005) : 545-549.
- [7] Wenbo Yang, Yin Fan, Hua Li, Tingting Liu. The matter-element analysis method in the application of water quality monitoring section optimization [J]. The people of the Yellow River. 2012 (11) : 82-84.
- [8] Telci I T, Nam K, Guan J, et al. The Optimal water quality monitoring network design for river systems [J]. Journal of environmental management, 2009, 90 (10) : 2987-2998.