

Application of Grey Clustering Method Based on Improved Analytic Hierarchy Process in Water Quality Evaluation

Jia Wang¹, Xu Wang^{2,a}, Xinhua Zhang¹, Haichen Li², Xiaohui Lei², Hao Wang² and Lixin Wang³

¹State Key Laboratory of Hydraulics and Mountain River Engineering, Sichuan University, Chengdu 610065, China;

²State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China;

³Xishan Coal Electricity Group Co., Ltd Taiyuan 030205, China

Abstract. To highlight the differences in water quality impacts of different indicators in water samples, this paper proposes a grey clustering method based on improved analytic hierarchy process to evaluate the quality of surface water. According to the pollution degree of different indicators in the water quality sample, the importance score is assigned, and the weight of different indicators is calculated by the analytic hierarchy process. The weight participates in the calculation of the grey clustering coefficient, and the evaluated water quality category considers the difference of the contribution rate of different pollutant indicators. The water quality samples of three water periods (Abundant, Normal, and Poor water flow periods) in the four sections of Qingshui River in Duyun City (Tea garden, Youhang, Yingpan and Jiadeng) were selected for evaluation, and the conventional grey clustering method and single factor method were used. The evaluation results were compared and analyzed. The evaluation results show that the improved grey clustering method is more scientific and reasonable and can provide a basis for water quality assessment and water environment management of water environment management departments.

1 Introduction

With the rapid development of urbanization in China, the water environment problem has become a constraint for development. Since the 18th National Congress of China, the Ministry of Water Resources has issued Document No. 1 for the implementation of the important spirit of strengthening the urbanization of ecological civilization, and the water environment has become a top priority [1]. The water environment is a complex system controlled by multiple factors and multiple indicators, and these influencing factors and hydrological indicators cause uncertainty due to limited or inaccurate information obtained through monitoring.

The grey clustering evaluation method can measure and discriminate the water quality under the condition of less sample and lack of information and has been widely used [2]. At present, many scholars [3-6] have used grey clustering method to evaluate the surface water quality, and some scholars have improved the method in terms of whitening weight function [7] and index transformation [8]. However, in the current research, the contribution rate of each water quality index is the same when calculating the clustering coefficient process, which does not reflect the difference of contribution of different water quality indicators. And the analytic hierarchy process can turn the decision-making problem of multi-index, multi-criteria and difficult to quantify into a multi-level single-objective problem. After determining

the quantitative relationship between the elements of the same level and the elements of the previous level through the pairwise comparison, the simple mathematical operations, get the weight of multiple indicators. Therefore, this paper combines the analytic hierarchy process to determine the weight of each pollutant, and the improvement of the grey clustering method can make the water quality evaluation result more reasonable and stricter than the conventional grey clustering method.

2 Improved grey clustering method

The improved grey clustering method determines different weights according to the measured concentration values of each pollutant index. The grey clustering not only considers the greyness of the water pollution degree, but also considers the degree of whitening of the system. At the same time, combined with the analytic hierarchy process, according to the influence of the measured values of different water quality indicators on the water quality category, determining the weight coefficient can make the work of water quality assessment more scientific and comprehensive.

According to the theory of grey clustering, suppose $n(i=1,2,\dots,n)$ water quality cluster samples, and $n(i=1,2,\dots,n)$ water quality clustering indicators (such as DO, BOD, $\text{NH}_3\text{-N}$, TP etc.). According to the national

^a Corresponding author: wangxu_04@126.com

project of surface water quality standard basic project standard limit and combined with the actual situation of each water body, the water quality is divided into p grade, p ($k = I, II, \dots, p$) grey scale.

2.1 Water whitening weight function

According to the category to which the clustering indicator belongs, different water whitening weight functions are constructed.

The first grey class of the j th indicator (class I water) whitening weight function is

$$f_{j1}(x) = \begin{cases} 1 & x \in [0, x_{j1}] \\ \frac{x_{j2} - x}{x_{j2} - x_{j1}} & x \in (x_{j1}, x_{j2}) \\ 0 & x \in [x_{j2}, \infty) \end{cases} \quad (1)$$

The whitening weight function of the k th class of the j th indicator (k class water, $k = I, II, \dots, p-1$) is

$$f_{jk}(x) = \begin{cases} \frac{x - x_{j,k-1}}{x_{jk} - x_{j,k-1}} & x \in (x_{j,k-1}, x_{jk}) \\ \frac{x_{j,k+1} - x}{x_{j,k+1} - x_{jk}} & x \in [x_{jk}, x_{j,k+1}) \\ 0 & x \notin (x_{j,k-1}, x_{j,k+1}) \end{cases} \quad (2)$$

The p -th grey class of the j th indicator (p -type water) whitening weight function is

$$f_{jp}(x) = \begin{cases} 0 & x \in [0, x_{j,p-1}] \\ \frac{x - x_{j,p-1}}{x_{jp} - x_{j,p-1}} & x \in (x_{j,p-1}, x_{jp}) \\ 1 & x \in [x_{jp}, \infty) \end{cases} \quad (3)$$

In the formula: x is the measured value of the water quality indicator, x_{jk} is the threshold of various whitening functions of various indicators, which is determined by reference to water quality standards at all levels in the country.

2.2 Cluster weight

The clustering weight is a measure of the weight of each indicator against the same grey class and is denoted by η_{kj} , which represents the weight of the k th grey class of the j th indicator. Because the units of various indicators are different, the absolute values differ greatly. Therefore, the grey classes are first normalized, and the methods are as follows:

For indicators with larger values, such as BOD, $\text{NH}_3\text{-N}$, TP, etc., the formula (4) is used:

$$r_{kj} = S_{kj} / \frac{1}{p} \sum_{k=1}^p S_{kj} \quad (4)$$

For indicators with a larger value, a milder pollution, such as DO, the equation (5) is used:

$$r_{kj} = S_{kj} / \frac{1}{p} \sum_{k=1}^p \frac{1}{S_{kj}} \quad (5)$$

In the equation: S_{kj} is the k th grey class (standard value) of the j th indicator; r_{kj} is dimensionless number of S_{kj} . The calculation of the clustering weight η_{kj} of the k th grey class of the j th indicator can be seen in the equation (6)

$$\eta_{kj} = \frac{r_{kj}}{\sum_{j=1}^m r_{kj}} \quad (6)$$

2.3 Analytic method for calculating measured water quality weights

In the multi-index comprehensive evaluation, the weighting value of the AHP method is commonly used. The index factor is judged according to its relative importance, which are giving 1~9 different importance scores. The larger the value, the more important it is. In the water quality assessment work, the concentration of each pollutant should be strictly controlled. The more serious the pollution is, the greater the important value is. This paper borrows the classification of surface water and water environmental quality standards to judge the pollution degree of the measured values of single indicators. The specific assignment is shown in Table 1. Then, according to the important values, the judgment matrix is constructed, the weight value of each index is calculated, and the consistency check of the judgment matrix is needed.

Table 1. Water quality monitoring data of different sections in Duyun City in 2013 for the analytic hierarchy process

standard	ClassI	ClassII	ClassIII	ClassIV	ClassV
weight value	1	3	5	7	9

2.4 Improved clustering coefficient

The clustering coefficient σ_{ik} of the k th grey class of the j th index is calculated as equation (7):

$$\sigma_{ik} = \sum_{j=1}^m \omega_{ij} f_{jk}(x_{ij}) \eta_{jk} \quad (7)$$

2.5 Clustering object

According to the principle of maximum membership, if $\sigma_{ik} = \max\{\sigma_{ik}\}$, The clustering object (measuring point) corresponding to σ_{ik}^* belongs to the grey class k^* . From this, the water quality level of each monitoring point can be determined, and then the water quality is polluted.

3 Case Analysis

3.1 Water quality samples and indicators

The improved grey clustering method is applied to the section of each water function zone of Qingshui River in Duyun City, Guizhou Province. In this paper, 12 water quality monitoring samples from 4 sections including Tea gardens, Yuohang, Yingpan and Jiadeng in the

Abundant, Normal, and Poor water flow periods were used for water quality evaluation. The water quality targets for each section are Class II, IV, III, and III. The water quality category standard refers to the Surface Water Environmental Quality Standard (GB3838-2002). In this paper, 5 main evaluation indicators were selected:

dissolved oxygen (DO), permanganate index (CODMn), five-day biochemical oxygen demand (BOD₅), ammonia nitrogen (NH₃-N), and total phosphorus (TP). The measured water quality data of each water quality monitoring section are shown in Table 2.

Table 2. Water quality monitoring data of different sections in Duyun City in 2013 mg/L

Section	Monitor time	DO	COD _{Mn}	BOD ₅	NH ₃ -N	TP
Chayuan	Poor water flow periods	8	1.15	2	0.06	0.01L*
	Abundant water flow periods	7	1.6	0.5	0.06	0.01L*
	Normal water flow periods	6	1.2	0.8	0.04	0.01L*
Youhang	Poor water flow periods	3	2.6	3.3	22.5	0.29
	Abundant water flow periods	6	1.9	2.1	1.95	0.07
	Normal water flow periods	7.6	1.3	2L*	0.39	0.04
Yingpan	Poor water flow periods	7.2	3.3	5	2.82	0.1L*
	Abundant water flow periods	6.4	1.8	0.6	1.10	0.02
	Normal water flow periods	7.4	1.2	0.5L*	0.122	0.06
Jiadeng	Poor water flow periods	7.7	2.26	2L*	0.632	0.06
	Abundant water flow periods	8.2	1.6	1.3	0.068	0.06
	Normal water flow periods	8.4	1.2	0.6	0.13	0.01L*

*L is the content of the substance less than this value, and is replaced by this value in the calculation.

3.2 Water quality evaluation

This paper takes the water quality evaluation of the Yingpan section as an example. The water quality samples of the remaining sections can be evaluated in the same way. First, construct a whitening weight function for the measured data of the section during the wet season. The whitening weight function values are shown in Table 3.

Table 3. The whitening weight function value of the Yingpan section during the flood season

grey class Index J	Class I	ClassII	ClassIII	Class IV	ClassV
DO	0.267	0.733	0	0	0
COD _{Mn}	1	0	0	0	0
BOD ₅	1	0	0	0	0
NH ₃ -N	0	0	0.920	0.080	0
TP	1	0	0	0	0

According to the measured water quality data, different scores are assigned to the pollutant index, a judgment matrix is constructed, and the judgment matrix is tested, that is, the maximum eigenvalue $\lambda_{max}=5$, consistency rate $CR=0$, which satisfies the condition $CR<0.1$. Therefore, the judgment matrix satisfies the consistency requirement. Then calculate the proportion of each indicator of the section during the wet season, the results are as follows.

Table 4. Judgment matrix and weight value of pollutants in the flood season of Yingpan section

A_i	DO	COD _{Mn}	BOD ₅	NH ₃ -N	TP	weight

DO	1	3	3	3/9	3	0.200
COD _{Mn}	1/3	1	1	1/9	1	0.067
BOD ₅	1/3	1	1	1/9	1	0.067
NH ₃ -N	9/3	9	9	1	9	0.600
TP	1/3	1	1	1/9	1	0.067

The surface water quality category is divided into 5 grades, which are 5 grey classes. According to formulas (4) ~ (6), the classification standard values of surface water quality are calculated, and the grey clustering weights of each index value are obtained. The results are shown in Table 5.

Table 5. Grey clustering weights η_{kj} for each indicator value

Index J	ClassI	ClassII	ClassIII	Class IV	ClassV
DO	0.314	0.230	0.175	0.187	0.193
COD _{Mn}	0.170	0.199	0.189	0.202	0.208
BOD ₅	0.363	0.212	0.180	0.173	0.198
NH ₃ -N	0.092	0.179	0.227	0.218	0.200
TP	0.062	0.180	0.229	0.220	0.202

After calculating the whitening weight function value, the cluster weight and the measured index value, the clustering coefficient is obtained according to formula (7). The specific results are shown in Table 6.

Table 6. Grey clustering coefficient of Yingpan section during flood season σ_{ik}

grey class	Class I	Class II	Class III	Class IV	Class V
Abundant water flow periods	0.063	0.028	0.122	0.010	0

of Yingpan					
---------------	--	--	--	--	--

According to the principle of maximum membership, in 2013 the flood season of the Yingpan section was classified as grey classes III, which is the third-grade water quality. The same method can be used to obtain the water quality of different sections in different water periods.

4 Case Analysis

The water quality of each section of the Qingshui River Basin in Duyun City in 2013 was evaluated according to the grey clustering evaluation based on the analytic hierarchy process. See Table 7 and the conventional grey clustering method and single factor method were used to calculate the water quality of the area. Evaluation, comparative analysis.

Table 7. Comparison of water quality evaluation results of different methods in different water periods of Qingshui River in Duyun City in 2013

Water samples		Evaluation results		
Section	Water periods	Improved grey clustering method	conventional grey clustering method	Single factor method
Chayuan	Poor water flow periods	ClassI	ClassI	ClassI
	Abundant water flow periods	ClassI	ClassI	ClassII
	Normal water flow periods	ClassI	ClassI	ClassII
Youhang	Poor water flow periods	Class IV	Class IV	worse than Class V
	Abundant water flow periods	ClassV	ClassI	ClassI
	Normal water flow periods	ClassI	ClassI	ClassII
Yingpan	Poor water flow periods	ClassV	ClassII	worse than Class V
	Abundant water flow periods	ClassIII	ClassI	Class IV
	Normal water flow periods	ClassI	ClassI	ClassII
Jiadeng	Poor water flow periods	ClassI	ClassI	ClassIII
	Abundant water flow periods	ClassI	ClassI	ClassII
	Normal water flow periods	ClassI	ClassI	ClassI

According to the water quality targets of each section, it can be seen from Table 7 that the improved grey clustering method is used to evaluate the water quality of the tea garden and Jiadeng section in the Abundant, Normal, and Poor water flow periods, which have reached the Class I water quality standard and the water quality is up to standard; The water quality during the dry period of the section and the dry section of the Yingpan section exceeded the standard, and the grey class V and the fifth grade water quality were more serious. The section of the oil line is the sewage control area, and the section of the Yingpan is the downstream buffer zone. Therefore, it is necessary to strengthen the sewage control of the two water function zones during the wet season and the dry season. The specific pollution sources need further analysis.

By comparison, the single-factor method evaluates water quality and the water quality level is lower; while the conventional grey clustering evaluation method has the best water quality. The improved grey clustering method and single factor method are basically consistent with the evaluation results of tea garden, oil line, Yingpan and Jiadeng sections in different water periods. Only in the samples of the Yingpan section during the flood season, the single factor method is not suitable for water quality assessment, which is contrary to the evaluation results of other methods.

It can be seen from the dry season data of the oil line section that the ammonia nitrogen index exceeds the standard and exceeds the water quality of the V class by

more than 11 times, which obviously limits the benign development of the water environment in the water function area. Nevertheless, it is unreasonable to evaluate the water quality sample by conventional cluster analysis method which result of evaluation is II class. The main reason is that when the grey clustering coefficient is calculated, the contribution rate of each index to water pollution is set equal, which weakens the influence of ammonia nitrogen on the overall water quality.

The results of single factor and improved grey clustering method for water quality compliance are different in the flood season of Yingpan section. The main standard of this water quality sample is ammonia nitrogen, and the super-class III standard is 10%, and the pollution is not serious. Since the one-factor method selects the single-indicator with the worst water quality to determine the water quality category, the water environment is more stringent and too protected, which is not conducive to the economic development of developing regions.

5 Conclusion and Recommendations

(1) The grey clustering method improved by the analytic hierarchy process is compared with the traditional grey clustering method. The method considers the difference of the influence of different indicators on the water quality in the water quality sample and gives different weights according to the pollution degree of different

indicators. The weight of major pollutants makes the water quality assessment work more comprehensive and scientific.

(2) The single factor method has a single-index “one-vote decision”. Its water quality category depends on the worst water quality index. The method is more stringent, and it protects the water environment too much, restricting the rapid economic development of relatively backward areas. However, this method can evaluate the main pollutants in the water sample and can be combined with the improved grey clustering method to analyze the water quality, which provides a more accurate theoretical basis for water quality management.

(3) According to the different types of water quality samples, this paper assigns different importance scores to water quality indicators. The weights of each water quality index are calculated by AHP, and the assignment method can be determined by Delphi method. This method needs to be verified in more watersheds.

Acknowledgements

This paper was supported by National Key R&D Program of China (2017YFC0404405), National Natural Science Foundation of China (51709276, 51579162), IWHR Research & Development Support Program (WR0145B762017). The authors also thank anonymous reviewers for their helpful comments and suggestions.

References

1. Tang Y, Peng X, Yang W, et al. Emerging Pollutants - Part I: Occurrence, Fate and Transport[J]. *Water Environment Research*, 2017, 89(10):1810-1828.
2. Zhan C, Ting G U, Yan P. Evaluation of Environmental Carrying Capacity Based on Grey Relational Analysis and Matter-Element Evaluation Model[J]. *Environmental Science & Technology*, 2017.
3. Yuan L X, YU-Yong. Application of Clustering Analysis Methods in Environmental Monitor (Water Analysis)[J]. *Environmental Science & Technology*, 2011.
4. Su-Ting M A, Hai-Ren Y E. Shanxi Reservoir Eutrophication Grey Clustering Evaluation and Control Countermeasures[J]. *Environmental Science & Technology*, 2012.
5. Pang B, Yu-Xia L I, Tong L. Application of Grey Clustering Method and Fuzzy Comprehensive Assessment Method to Assess Eutrophication Level of Water Quality[J]. *Environmental Science & Technology*, 2011, 34(11):185-188.
6. Zhou L F, Shi-Guo X U, Sun W G. Comprehensive evaluation on water environment quality of Zhalong wetland based on grey clustering method[J]. *Journal of Dalian University of Technology*, 2007.
7. Yan Z, Ting W, Xi X. Application of Improved Grey Clustering Relation Model in Groundwater Quality Assessment [J]. *Yellow River*, 2012,34(7):56-58.
8. Yicheng Z, Hui L. [J]. Application of grey clustering method based on index transformation value in water quality evaluation. *Gansu Water Resource and Hydropower Technology*, 2015,51(1):6-8,52.