ECOLOGICAL ENVIRONMENT QUALITY COMPREHENSIVE EVALUATION MODEL BASED ON GREY CLUSTERING

LI Yan ling
School of Information Science & Engineering, Shandong Agricultural University,
Taian, Shandong, P.R.China 271018

KEY WORDS: ecological environment quality comprehensive evaluation, e-government, grey clustering, whitening function

ABSTRACT:

Ecological environment quality comprehensive evaluation evaluate different cells' ecological environment quality, on the basis of region eco-environment investigation , the indexes and math methods selected according to the region, to discover the problems and put forward comprehensive father measures. In electronic government spatial decision-making assistant systems, Estimate the region's eco-environment quality not only provides science evidence for the departments environment management and decision making of the government. It is also very important for making sustainable development of the region.
At present, region eco-environment quality comprehensive evaluation is just underway, and is often handworked. The computation process is also complex. As for evaluation method, fuzzy comprehensive evaluation method is usually used. These methods have their advantages as well as disadvantages.
In System Theory and Control Theory, colors are often used to present the degree of knowledge and understandability of investigator. "Black" expresses complete lack of information, "white" expresses full information, "grey" expresses insufficiency inadequacy information. If there is insufficiency or indeterminate information in a system, then the system is called has grey character. The system with grey character is called grey system. Strictly speaking, the external word known and changed by people is neither "black" nor "white", instead, it is "grey". Simply we ignore some uncertain factors when we deal with actual problems, so we regard some grey systems as white systems to know and dispose. So, grey systems are ubiquitous.
In this article, the author introduces the principle and general process of grey clustering, including confirming clustering whitening number, the evaluation indexes' whitening functions' eigenvalue, dimensionless procession of data, decision of whitening functions, get the weight values of the indexes for different sorts, get clustering coefficients and make sure the sorts of the clustering objects. Thereinto, whitening functions have very important impact on the clustering results. This article adopts the usually used triangle whitening functions. Then, the evaluation index system of the experimental area-ZunYi county of Guizhou province is built, obtains data such as weights and whitening functions. At last, eco-environment quality evaluation of the area is carried out with this method. The results are analyzed and proved nicety, impersonal and is identical with the fact.
This method can be used for eco-environment quality evaluation of any region.

1 FOREWORD

Ecological environment quality comprehensive evaluation evaluate different cells' ecological environment quality, on the basis of region eco-environment investigation , the indexes and math methods selected according to the region, to discover the problems and put forward comprehensive father measures. In electronic government spatial decision-making assistant systems, Estimate the region's eco-environment quality not only provides science evidence for the departments environment management and decision making of the government. It is also very important for making sustainable development of the region. [1]

At present, region eco-environment quality comprehensive evaluation is just underway, and is often handworked. The computation process is also complex. As for evaluation method, fuzzy comprehensive evaluation method is usually used. These methods have their advantages as well as disadvantages. [2][3]For example, in the process of fuzzy comprehensive evaluation, it is important to confirm weights and select subordinate functions, while they are liable to human factors.*

According to grey system theory, the paper builds eco-environment quality comprehensive evaluation model by grey clustering method, and states the general evaluation process combining the example of Zunyi county of Guizhou province.

* Li Yanling(1980-), E-mail: liyanlingsd@yahoo.com.cn

2 PRINCIPLE AND STEPS OF GREY CLUSTERING

2.1 Introduction to grey system

In system theory and control theory, colors are often used to represent the degree of what investigators know about the system and its inner information. "Black" expresses complete lack of information, "white" expresses full information, "grey" expresses insufficiency inadequacy information. And, numbers lack of information are grey numbers, elements lack of information are grey
elements, relations lack of information are grey relations. If there is insufficiency or indeterminate information in a system, then the system is called has grey character. The system with grey character is called grey system. In system science, systems with physical antetype are called latent systems such as circuit system, God's image system; while systems without physical antetype are called nonessential systems such as society system, economy system, ecology system, environment system etc, accordingly grey system is distinguished as nonessential grey system and latent grey system.

Strictly speaking, the external word known and changed by people is neither "black" nor "white", instead, it is "grey". Simply we ignore some uncertain factors when we deal with actual problems, so we regard some grey systems as white systems to know and dispose. So, grey systems are ubiquitous. [4]

2.2 Principle and steps of grey clustering method

Grey clustering is a content of grey system theory. it is a method based on whitening functions of grey numbers. Grey clustering analysis method is grey statistical method; it manages whitening value (measure value or analysis data) of clustering objects (evaluation objects) to different clustering indexes (evaluation indexes) by N grey kinds (evaluation degrees), and judge the clustering object's grey kind.[4]

Treat n evaluation samples as clustering objects k（k=1，2，…，n）, p evaluation criterion(such as higher, high, middle, low etc), as clustering kinds j(j=1，2，…，p），m evaluation indexes as clustering indexes i(i=1，2，…，m）. Suppose d_{ki} is whitening value of sample k to index i , f_{ij} is whitening function of index i to evaluation degree j . So, the steps are: [5][6]

1) giving clustering whitening value d_{ki}.
2) confirming eigenvalue λ_{ij} of evaluation indexes by classification standards of the indexes.
3) no dimension process of data.
4) confirming whitening function $f_{ij}(x)$ of grey clusters, in general, there are three types:

① more than type（$\lambda_{ij}(1) < \lambda_{ij} < \infty$）

$$f_{ij}(x)=\begin{cases} 1 & x\in[\lambda_{ij},\infty) \\ \dfrac{x-\lambda_{ij}(1)}{\lambda_{ij}-\lambda_{ij}(1)} & x\in[\lambda_{ij}(1),\lambda_{ij}] \\ 0 & x\notin[\lambda_{ij}(1),\infty) \end{cases}$$

② interzone type（$\lambda_{ij}(1) \le \lambda_{ij} \le \lambda_{ij}(2)$）

$$f_{ij}(x)=\begin{cases} \dfrac{x-\lambda_{ij}(1)}{\lambda_{ij}-\lambda_{ij}(1)} & x\in[\lambda_{ij}(1),\lambda_{ij}] \\ \dfrac{\lambda_{ij}(2)-x}{\lambda_{ij}(2)-\lambda_{ij}} & x\in[\lambda_{ij},\lambda_{ij}(2)] \\ 0 & x\notin[\lambda_{ij}(1),\lambda_{ij}(2)] \end{cases}$$

③ less than type（$0 < \lambda_{ij} < \lambda_{ij}(2)$）

$$f_{ij}(x)=\begin{cases} 1 & x\in[0,\lambda_{ij}] \\ \dfrac{\lambda_{ij}(2)-x}{\lambda_{ij}(2)-\lambda_{ij}} & x\in[\lambda_{ij},\lambda_{ij}(2)] \\ 0 & x\notin[0,\lambda_{ij}(2)] \end{cases}$$

5) getting weights of the indexes in different kinds η_{ij}

$$\eta_{ij}=\frac{\lambda_{ij}}{\sum_{i=1}^{m}\lambda_{ij}}$$

6) getting clustering coefficient α_{ij} $\alpha_{ij}=\sum_{i=1}^{m} f_{ij}(d_{ki})\eta_{ij}$

and construct clustering vector α_k $\alpha_k=\{\alpha_{k1},\alpha_{k2},...,\alpha_{kp}\}$

7) confirming evaluation quality degree, performing clustering analysis

get the max element α_{kj} in α_k $\alpha_{kj}=\max_j\{\alpha_{kj}\}$

then the sample k for evaluation belongs to j degree.

3 APPLICATION EXAMPLE

3.1 Evaluation index system

At present, there is no uniform eco-environment quality evaluation index system either home or abroad. This is because the actual eco-environment quality is very different, so it is very difficult to build uniform index system to solve eco-environment quality evaluation for different areas. So, according to *eco function division provisional regulations* issued by SEPA in May 15th, 2002, we build index system by the principles as follows:

1) selecting environment factor collection which determine eco-environment quality and population, society, economy, factor collection which affect environment quality.
2) selecting factors which can reflect main aspects, change characters and results of eco-environment quality, that is, environment result variable factor collection.
3) The evaluation factors can be analyzed separately or synthetically.
4) The evaluation factors have the characteristic of reflecting the dynamic change of eco-environment quality.
5) Data is easy to obtain, understand and number.

According to natural environment characteristic of Zunyi county, main eco-environment problems and population, society, economy status, we select nine elements which are representative, clear, and easy to Statistic, totally 20 factors, and form index system of eco-environment quality evaluation(as listed in the following table)

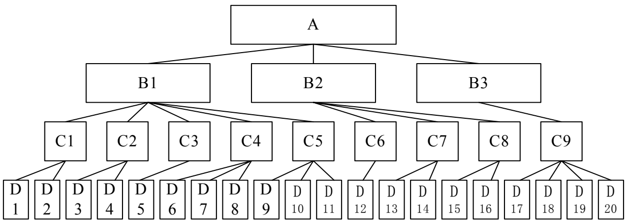
Factor collections	Elements	Factors
Environment and resource factor collection	physiognomy	Percentage of precipitous fields area (%)
		Percentage of karst fields area (%)
	climate	Active accumulate temperature higher than 10℃(℃)
		Average precipitation by year(mm)
	Water resource	Water possession per person (m^3/person)
	infield	Infield area per person (mu/ person)
		Infield percentage (assart ratio)(%)
		Percentage of precipitous fields more than 25 degree area (%)
	Plants resource and conservations	Forest coverage (%)
		Grassland area percentage (%)
		Conservations area percentage (%)
population, society, economy factor collection	population	population density (person/Km^2)
	society	Farmers' average net income (Yuan)
		commissariat output per person (Kilogram)
	economy	GDP per person(Yuan)
		aerarian input per person (Yuan)
environment result variable factor collection	Ecology destroy	water and soil run off area percentage(%)
		Desert area percentage (%)
		Percentage of plants area destroyed by small coal mine (%)

		Flood and drought generant ratio (%)

Table 1. The evaluation index system

All factors data is in unit of villages and towns.

The above factor collections, elements and factors are interrelated and restricted, and form a complex system of multi layers. Arranging and streaking the index system with AHP (Analytical Hierarchy Process) method, we divide Zunyi county eco-environment quality comprehensive evaluation index system into 4 layers: target layer, constraint layer, element layer, factor layer. The hierarchical configuration is:



In this diagram, A- target layer, represents the eco-environment quality comprehensive evaluation results of the villages and towns of Zunyi county.

B-constraint layer, represents the main element collection which restrict and affect eco-environment quality. As to Zunyi county, there are three factor collections, that is : environment and resource factor collection (B1), population, society, economy factor collection (B2), environment result variable factor collection (B3).

C-element layer, the above three factor collections reflect eco-environment quality by variety environment elements. Regards indexes' concision and data available, we select physiognomy (C1), climate (C2), water resource (C3), infield (C4), plants resource and conservations (C5), population (C6), society (C7), economy (C8), ecology destroy (C9) nine elements.

D-factor layer, is the details and list of the evaluation indexes, the factors. At the range of the above 9 elements, we select the indexes which can reflect eco-environment quality and its character in detail, considering both natural background (environmental and resource character) of eco-environment, and the aftereffect of environment after human activities as well as factors of social economy which have direct or indirect effect on eco-environment, to make the evaluation more all-around.

3.2 Confirming evaluation criterion

In this article, all grade and mark criterion are made by consulting related environment criterion, industry criterion and design criterion, as well as the grade and mark criterion of Guizhou province. The detailed criterion is listed in the following table:

index	Grade and mark		
	excellent	good	bad
Percentage of precipitous fields area (%)	<20	[20，30)	≥50
Percentage of karst fields area (%)	<60	[60，70)	≥90
Active accumulate temperature higher than 10℃(℃)	≥5000	[4600，5000)	<4000
Average precipitation by year(mm)	≥1100	[1050，1100)	<900
Water possession per person (m^3/person)	≥4000	[3000，4000)	<1000
Infield area per person (mu/ person)	≥2.5	[2.0,2.5)	<1.0
Infield percentage (assart ratio)(%)	≥40	[35,40)	<25
Percentage of precipitous fields more than 25 degree area (%)	<2.0	[2.0,4.0)	≥8.0
Forest coverage (%)	>50	(40,50]	≤20
Grassland area percentage (%)	≥15.0	[10.0,15.0)	<1.0
Conservations area percentage (%)	≥30	[20,30)	<5
population density (person/Km^2)	≤200	(200,300]	>500
Farmers' average net income (Yuan)	≥2500	[2200,2500)	<1600
commissariat output per person (Kilogram)	≥500	[400,500)	<200
GDP per person(Yuan)	≥5000	[4000,5000)	<2000
aerarian input per person (Yuan)	≥140	[120,140)	<80
water and soil run off area percentage(%)	≤35	(35,40]	>50
Desert area percentage (%)	≤10	(10,15]	>25
Percentage of plants area destroyed by small coal mine (%)	≤0.0001	(0.0001,0.0002]	>0.0004
Flood and drought generant ratio (%)	≤15	(15,20]	>30

Table 2. Grade criteria

3.3 Grey clustering analysis model of eco-environment comprehensive evaluation

3.3.1 **Clustering whitening value d_{ki} :** d_{ki} represents real value of clustering object i about clustering index j , i=1，2，…，35，j=1，2，…，20. Whitening number matrix is a 35×20 matrix made up of the indexes' original value, here is some data:

User id	Percentage of precipitous fields area	…	Flood and drought generant ratio
52032101	11.5205		15
52032103	14.9809		29
52032105	21.0377		34
52032109	24.7906		22
52032111	13.7584		17

Table 3. Part of whitening value

3.3.2 **The evaluation indexes' whitening function eigenvalue** λ_{ij} **:** It is confirmed according to the grade criteria in 3.2.

index	5	4			1
		lower limit	Middle limit	upper limit	
Percentage of precipitous fields area (%)	<0.4	0.4	0.5	0.6	≥1
Percentage of karst fields area (%)	<0.67	0.67	0.725	0.78	≥1
Active accumulate temperature higher than 10 ℃(℃)	≥1	0.92	0.96	1	<0.8
Average precipitation by year(mm)	≥1	0.95	0.975	1	<0.82
Water possession per person (m^3/person)	≥1	0.75	0.875	1	<0.25
Infield area per person (mu/ person)	≥1	0.8	0.9	1	<0.4
Infield percentage (assart ratio)(%)	≥1	0.875	0.9375	1	<0.625
Percentage of precipitous fields more than 25 degree area (%)	<0.25	0.25	0.375	0.5	≥1
Forest coverage (%)	>1	0.8	0.9	1	≤0.4
Grassland area percentage (%)	≥1	0.67	0.835	1	<0.067
Conservations area percentage (%)	≥1	0.67	0.835	1	<0.167
population density (person/Km^2)	≤0.4	0.4	0.5	0.6	>1
Farmers' average net income (Yuan)	≥1	0.88	0.94	1	<0.64
commissariat output per person (Kilogram)	≥1	0.8	0.9	1	<0.4
GDP per person(Yuan)	≥1	0.8	0.9	1	<0.4
aerarian input per person (Yuan)	≥1	0.86	0.93	1	<0.57
water and soil run off area percentage(%)	≤0.7	0.7	0.75	0.8	>1
Desert area percentage (%)	≤0.4	0.4	0.5	0.6	>1
Percentage of plants area destroyed by small coal mine (%)	≤0.25	0.25	0.375	0.5	>1
Flood and drought generant ratio (%)	≤0.5	0.5	0.585	0.67	>1

Table 4. Part of whitening function eigenvalue value

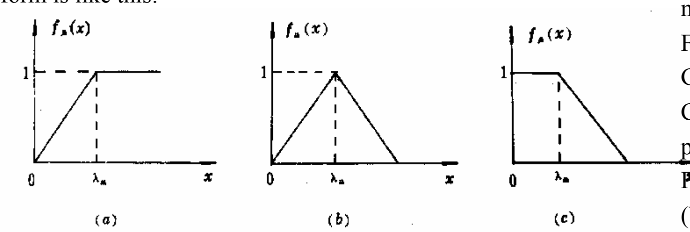
3.3.3 **No dimension process of data:** The dimension of the indexes is different from each other, so they need no dimension process in order to make them comparable. We divide corresponding data by $\max\max\{d_{ki},\lambda_{ij}\}$, here is some data:

User id	Percentage of precipitous fields area	…	Flood and drought generant ratio
52032101	0.23041		0.5
52032103	0.299618		0.967
52032105	0.420754		1.133
52032109	0.495812		0.733

Table 5. Part of data after no dimension process

3.3.4 **Confirming whitening function:** Grey cluster whitening function have very important impact on clustering results. In this article, the threshold value λ_{ij} of whitening functions is confirmed according to the clustering indexes. Thereby avoid errors even mistakes estimated by experience anciently. Whitening function is independent of whitening number.

The usually used triangle whitening function is adopted, whose form is like this:



Whitening function can be got according to formula in 2.2 and table in 3.3.2. For example, whitening function of percentage of precipitous fields area is :

$$f_{11}(d_{11})=\begin{cases}1 & d_{11}\in[1,\infty)\\ \dfrac{d_{11}-0.9}{1-0.9} & d_{11}\in[0.9,1]\\ 0 & d_{11}\notin[0.9,\infty)\end{cases}$$

$$f_{12}(d_{11})=\begin{cases}\dfrac{d_{11}-0.8}{0.9-0.8} & d_{11}\in[0.8,0.9]\\ \dfrac{1-d_{11}}{1-0.9} & d_{11}\in[0.9,1]\\ 0 & d_{11}\notin[0.8,1]\end{cases}$$

$$f_{13}(d_{11})=\begin{cases}\dfrac{d_{11}-0.6}{0.7-0.6} & d_{11}\in[0.6,0.7]\\ \dfrac{0.8-d_{11}}{0.8-0.7} & d_{11}\in[0.7,0.8]\\ 0 & d_{11}\notin[0.6,0.8]\end{cases}$$

$$f_{14}(d_{11})=\begin{cases}\dfrac{d_{11}-0.4}{0.5-0.4} & d_{11}\in[0.4,0.5]\\ \dfrac{0.6-d_{11}}{0.6-0.5} & d_{11}\in[0.5,0.6]\\ 0 & d_{11}\notin[0.4,0.6]\end{cases}$$

$$f_{15}(d_{11})=\begin{cases}1 & d_{11}\in[0,0.4]\\ \dfrac{0.5-d_{11}}{0.5-0.4} & d_{11}\in[0.4,0.5]\\ 0 & d_{11}\notin[0,0.5]\end{cases}$$

The rest may be deduced by analogy. Put the measured value of the evaluation cells, and get whitening function value.

3.3.5 Getting weights of the indexes in different kinds (demarcating clustering weight):

According to formula in 2.2, threshold values are confirmed first. Threshold value is usually point of intersection d_{ij} of f_{ij}=1 and $f_{ij}\leqslant 1$.

index	5	… 1
Percentage of precipitous fields area (%)	0.4	1
Percentage of karst fields area (%)	0.67	1
Active accumulate temperature higher than 10℃(℃)	1	0.8
Average precipitation by year(mm)	1	0.82
Water possession per person (m^3/person)	1	0.25
Infield area per person (mu/ person)	1	0.4
Infield percentage (assart ratio)(%)	1	0.625
Percentage of precipitous fields more than 25 degree area (%)	0.25	1
Forest coverage (%)	1	0.4
Grassland area percentage (%)	1	0.067
Conservations area percentage (%)	1	0.167
population density (person/Km^2)	0.4	1
Farmers' average net income (Yuan)	1	0.64
commissariat output per person (Kilogram)	1	0.4
GDP per person(Yuan)	1	0.4
aerarian input per person (Yuan)	1	0.57
water and soil run off area percentage(%)	0.7	1
Desert area percentage (%)	0.4	1
Percentage of plants area destroyed by small coal mine (%)	0.25	1
Flood and drought generant ratio (%)	0.5	1
$\Sigma\ \lambda_{ij}$	15.57	13.539

Table 6. Part of threshold value

So we can get the weight as below:

index	5	… 1
Percentage of precipitous fields area (%)	0.02569043	0.073861
Percentage of karst fields area (%)	0.04303147	0.073861
Active accumulate temperature higher than 10℃ (℃)	0.06422608	0.059089
Average precipitation by year(mm)	0.06422608	0.060566
Water possession per person (m^3/person)	0.06422608	0.018465
Infield area per person (mu/ person)	0.06422608	0.029544
Infield percentage (assart ratio)(%)	0.06422608	0.046163
Percentage of precipitous fields more than 25 degree area (%)	0.01605652	0.073861
Forest coverage (%)	0.06422608	0.029544

Grassland area percentage (%)	0.06422608	0.004949
Conservations area percentage (%)	0.06422608	0.012335
population density (person/Km^2)	0.02569043	0.073861
Farmers' average net income (Yuan)	0.06422608	0.047271
commissariat output per person (Kilogram)	0.06422608	0.029544
GDP per person(Yuan)	0.06422608	0.029544
aerarian input per person (Yuan)	0.06422608	0.042101
water and soil run off area percentage(%)	0.04495825	0.073861
Desert area percentage (%)	0.02569043	0.073861
Percentage of plants area destroyed by small coal mine (%)	0.01605652	0.073861
Flood and drought generant ratio (%)	0.03211304	0.073861

Table 7. Part of weight data

3.3.6 **Getting clustering coefficient:** For example, the 1 mark clustering coefficient of Nanbai town is computed like this:

$\alpha_{11} = \sum_{i=1}^{20} f_{i1}(d_{1i})\eta_{i1}$ =0+1×0.073860699+1×0.059088559+0+1×0.018465175+1×0.029544279+0+0+0+1×0.004948667+1×0.012334737+1×0.073860699+0+1×0.029544279+0+0+0+0+1×0.073860699+0=0.375507793。

The detailed data is listed below：

User id	5	…	1	result
52032109	0.2398		0.2022	5
52032119	0.2618		0.1505	5
52032163	0.3502		0.2332	5
52032161	0.3371		0.2055	5
52032137	0.2413		0.187	5
52032107	0.2477		0.2097	5
52032111	0.119		0.1426	4
52032113	0.1638		0.1332	4
52032145	0.0336		0.12	3
52032141	0.1039		0.1076	3
52032155	0.1156		0.1395	3
52032129	0.1156		0.2096	3
52032159	0.0713		0.0653	3
52032151	0.1048		0.076	3
52032127	0.0261		0.2662	1
52032103	0.0961		0.2099	1
52032105	0.0364		0.3456	1
52032102	0.1823		0.3532	1
52032165	0.0472		0.2409	1
52032115	0.1444		0.276	1
52032117	0.0899		0.2022	1
52032121	0.1923		0.2931	1
52032139	0.0815		0.1901	1
52032125	0.1013		0.3325	1
52032143	0.0846		0.2338	1
52032133	0.1068		0.4004	1
52032135	0.1599		0.2836	1
52032101	0.3216		0.3755	1
52032157	0.1156		0.2548	1
52032104	0.2039		0.3985	1
52032131	0.0034		0.3023	1
52032147	0.1164		0.3633	1
52032149	0.0818		0.281	1
52032153	0.0763		0.2298	1
52032123	0.1177		0.213	1

Table 8. Part of clustering coefficient and cluster results

3.3.7 **Confirming clustering object's degree:** Confirm clustering object's degree by the maximal clustering coefficient of all degrees. The object belongs to the degree whose clustering coefficient is maximal. The detailed results are in the above table.

3.4 Evaluation results analysis

The evaluation results are consistent with the index grade criteria used. In the case we study, grade criteria vary from 5 to 1, so are the evaluation results gained. The higher the score of an evaluation cell, the better its eco-environment quality. The results inosculate the actual condition well.

4 CONCLUSION

Considering eco-environment quality comprehensive evaluation process and model, developing eco-environment quality comprehensive evaluation system, can manage evaluation data scientifically. Making external and scientific evaluation on this basis can lighten work intensity, save time and energy, provide scientific gist for environment management and decision of government of all levels. But region eco-environment quality comprehensive evaluation is just underway, and this is only an attempt. The evaluation result is not absolute concept, but opposite concept. Instead, it represents opposite choiceness degree of the evaluation cells' eco-environment quality.

The clustering results of grey clustering analysis method have high resolving power, and make full use of the foregone information of the system. This method also overcomes the limitation that the results are influenced by some factor in single factor evaluation, and reflects the effect of every evaluation factor in the round. The results impersonally reflect condition of the evaluation cells.

References

Ma suting, Tang jie, Lin nianfeng. Synthetic assessment of eco-environment in the Western Jilin Province based on multi-sources GIS and RS spatial information. Resources Science, 2004,26(4): pp. 140-145.

Wang huiyong, Chen yuliang, Rui yongqin. Matter-elements analysis method of comprehensive evaluation for pavement condition. Journal of Traffic and Transportation Engineering, 2004, 4(2): pp. 6-7.

Shen zhenyao, Xie tongfang, Qian tianwei. Method comparison of fuzzy synthetic evaluation, grey cluster analysis and matter elements analysis in environment assessment. Pollution prevention and cure technology, 1997, 10(2): pp. 72-74.

Tang minkang. The grey classification and its application. Journal of Southern Institute of Metallurgy, 1995, 16(4): pp. 63-68.

Zou zhiyun, Li shuo. The grey clustering analysis of evaluation for transit comprehensive development level. Journal of Wuhan Transportation University, 2000, 24(1): pp. 38-41

Li qingfu, Hu qunfang, Liu wen, Ran ganglun. Asphalt pavement performance evaluation based on gray-clustering decision. Journal of Zhengzhou University(Engineering Science), 2003，24(2): pp. 44-47.