Water Supply Optimization Strategy Using Qualitative Analysis Approaches

Guiyuan Yuan
Shandong University of
Science and Technology,
Qingdao 266590, China

Shanshan Li
Shandong University of
Science and Technology,
Qingdao 266590, China

Cong Liu*
Shandong University of
Science and Technology,
Qingdao 266590, China

Abstract: This paper aims to list the possible risk of water resources with fully considering of multiple factors, and analyse the factors influence on the shortage of water resources, and finally put forward optimization strategy and reasonable suggestions through a systematic qualitative analysis. More specifically, (1) establishing four models to measure water supply capacity: From the natural environment, social water, economic factors, and the new technology factors to create the model, (2) taking Beijing as a survey centre, to get the weight of each risk factor to draw the top 5 main risk factors, (3) building the equation by multiple linear regression, through the curve fitting analysis to predict water status in Beijing city for the past15 years, (4) making an intervention plan and modify model parameters on the basis of economy and environment, and (5) Based on the intervention plan, time variable is modified to predict whether Beijing will be affected by water shortage in the future.

Keywords: AHP; water resources; multiple linear regression; multivariate regression model

1. DESCRIPTION

Water resource on earth generally refers to the overall content of the hydrosphere[1,2]. Including the human control and direct irrigation, power generation, water supply, shipping, aquaculture and other purposes of surface water and groundwater, as well as rivers, lakes, wells, springs, tidal, harbor and aquaculture waters, etc. In a narrow sense, it refers to the amount of fresh water that can be recovered and renewed every year[3,4]. Water resources are important natural resources for the development of the national economy[5]. In many parts of the world, the demand for water has exceeded the level of the water resources, and there are many areas on the verge of the use of water resources is not balanced[6,7,8]. Therefore, we need to establish a set of effective water resource model to solve the world water problem. Throughout the modeling process, we fully consider the feasibility of our model for effectiveness, efficiency, and cost.

The problems to be solved inludes:

1) Develop a model that provides a measure of the ability of a region to provide clean water to meet the needs of its population. You may need to consider the dynamic nature of the factors that affect both supply and demand in your modeling process.

2) Using the UN water scarcity map pick one country or region where water is either heavily or moderately overloaded. Explain why and how water is scarce in that region. Make sure to explain both the social and environmental drivers by addressing physical and/or economic scarcity.

2. SYMBOL CONVENTIONS
2.1 Symbol Conventions
ZY: Total water consumption;

GY: industrial water consumption;

NY: agricultural water consumption;

SH: third industrial and domestic water consumption;

ZL: total water resources;

JS: precipitation;

DX: underground water;

DB: surface water content of;

ZS: reclaimed water;

λ : characteristic root;

E: cumulative contribution rate;

U: risk factors of water shortage;

V: hierarchical domain;

T: membership degree;

w^{*} : entropy weight;

B: comprehensive evaluation risk value;

S: *correlation* matrix;

R: fuzzy matrix.

3. ANALYSIS OF THE PROBLEM

For question 1, a four model was established to measure the water supply capacity, and a model was created to solve the task from four aspects: natural environment, social water, economic factors, and new technology.

For question 2, according to the selected regions of the economic environment variable is introduced to solve the model; Beijing, China as a water reference City, Beijing city's natural environment and climate, economic and population factors, water supply system, technical reserves into the model,

through social and environmental driving factors to explain reasons of water shortage in Beijing City.

4. MODELS

4.1 Water Resources Model

Beijing city as a research center, according to the economic environment of Beijing City, the introduction of the model to solve the environmental variables.

4.1.1 Analysis of the present situation

Will Beijing's natural environment, topography and climate precipitation, economic and population factors, water supply system, technical reserves, water pollution, water price, water purification ability is introduced into the model through social and environmental driving factors to explain the cause of water shortage in Beijing city. Then statistical results is the following table:

Table 1. Beijing water resources statistical results

project	2005	2006	2007	2008	2009	2010	2011	2012
Ground water	15.7	14.7	14.8	16.5	18.5	18.5	19.2	21.4
Surface water	7.8	5.3	6.1	8.2	7.6	6.0	7.6	12.8
Reclaimed water	2.5	2.0	2.1	2.0	2.6	3.6	5.0	2.1
Amount of water	26	22	23	26.7	28.2	28.1	31.8	36.3

4.1.2 Model Testing

Design water supply for y, the factors affecting the total water resources amount is the amount of surface water resources, total amount of groundwater resources and renewable water were recorded for x_1, x_2, x_3, establish three element linear regression model $= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$, among them ε is random error obeys normal distribution n $(0, \delta), \beta_0, \beta_1, \beta_2, \beta_3$ regression coefficients.

Using a 2005-2012 chronology data and on the type, then

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon (\alpha = 1, 2, 3 \dots, 8)$$

Let b_0, b_1, b_2, b_3 respectively, the parameters of $\beta_0, \beta_1, \beta_2, \beta_3$ of the estimated value, then get the regression equation $\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3$

$$\hat{y}_\alpha = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 (\alpha = 1, 2, \dots, 8)$$

$y_\alpha (\alpha = 1, 2 \dots, 8)$ as sample value; y goes alpha as the regression equation given value becomes the regression value,

called the $y_\alpha - \hat{y}_\alpha (\alpha = 1, 2 \dots, 8)$ as residual, it can be classified as the deviation of the sample value and the return value.

According to the least square method, the least square sum of residuals is achieved

$$Q = \sum_{\alpha=1} (y_\alpha - \hat{y}_\alpha)^2$$

According to the limit principle of calculus b_0, b_1, b_2, b_3 must be satisfied for the minimum

$$\frac{\delta Q}{\delta b_i} = 0 (i = 0, 1, 2, 3)$$

The normal equations in $\left(X^T X \right) \hat{B} = X^T Y$

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ 1 & x_{31} & x_{32} & x_{33} \\ 1 & x_{41} & x_{42} & x_{43} \\ 1 & x_{51} & x_{52} & x_{53} \\ 1 & x_{61} & x_{62} & x_{63} \\ 1 & x_{71} & x_{72} & x_{73} \\ 1 & x_{81} & x_{82} & x_{83} \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{pmatrix} \quad \hat{B} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Estimated value of solution:

$$\hat{B} = (X^T X)^{-1} (X^T Y) = \begin{pmatrix} -4.732 \\ 2.138 \\ 0.498 \\ 0.274 \end{pmatrix}$$

This is calculated by $\hat{B}_i$ regression equation through the regression equation as follows:

$$\hat{y} = -4.653 + 2.091 x_1 + 0.398 x_2 + 0.286 x_3$$

hold $(x_{1\alpha}, x_{2\alpha}, x_{3\alpha})$ and on the type of solution of each regression value $(\hat{y}_1, \hat{y}_2 \dots, \hat{y}_8)$

Calculate ,

$$\bar{y} = \frac{1}{8}\sum_{\alpha=1}^{8} y_\alpha \rightarrow U = \sum_{\alpha=1}^{8}(\hat{y}_\alpha - \bar{y})^2 = 181$$

Degree of freedom $f_u = p = 3$

$$Q = \sum_{\alpha=1}^{8}(\hat{y}_\alpha - y_\alpha)^2 = 28 \quad \text{Degree of freedom}$$

$$f_\alpha = n - p - 1 = 8 - 3 - 1 = 4$$

Total deviation square $S_{yy} = U + Q$

Verifying the significance of regression equation is equivalent to test hypothesis

$$H_0 : \beta_1 = 0, \beta_2 = 0, \beta_3 = 0.$$

For statistical measurement

$$F = \frac{U/f_u}{Q/f_Q} = \frac{U/P}{Q/n-P-1} = \frac{181/3}{30/4} = 8.67$$

Use $\alpha = 0.05$ search table with a critical value of $F_{0.05}(3,4) = 6.59$,easy to see the

8.67>6.59, and therefore refused to H_0.

Namely, the effect of regression equation was significant.

4.2 model of influence factors of water consumption

From the chart we can see that before 1994 due to policy, farming, the industrial structure of the problem data fluctuations, and in 1994 after can be regarded as a continuous curve, so we choose after 2003 data to carry on the fitting analysis, relational expression is obtained.
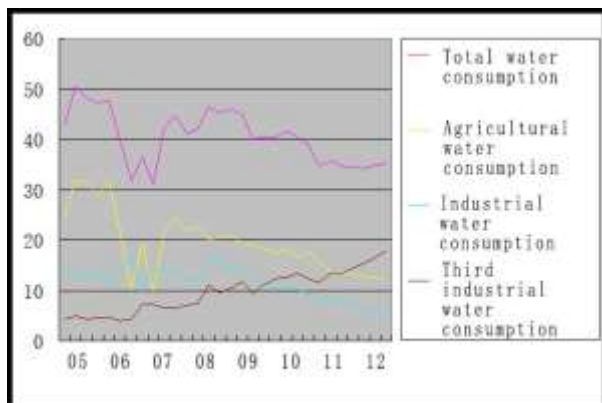

Figure. 1 the total water use agricultural water consumption, industrial water consumption, and the third industry and other water use fitting tables

Considering the continuous change of time, the curve of agricultural water consumption was obtained by fitting, as shown in Figure 2.
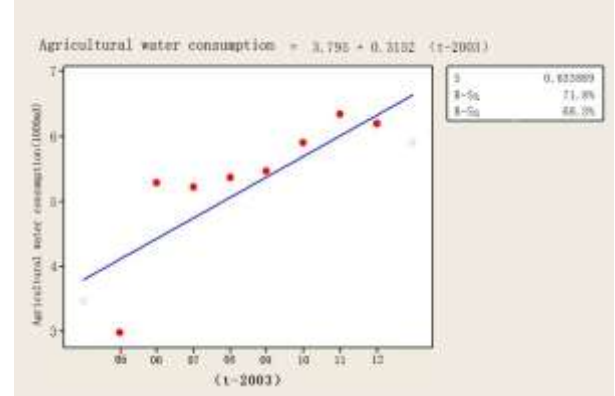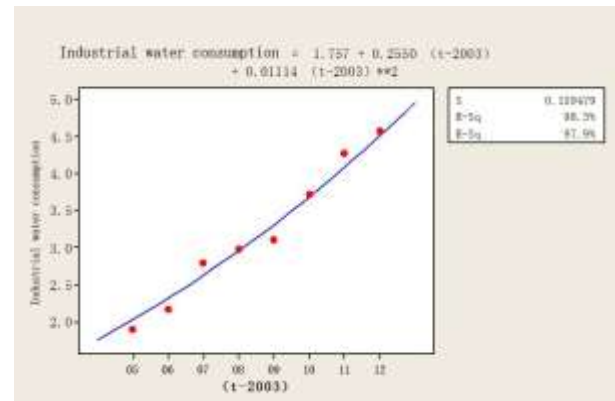

Figure. 2 fitting image of agricultural water

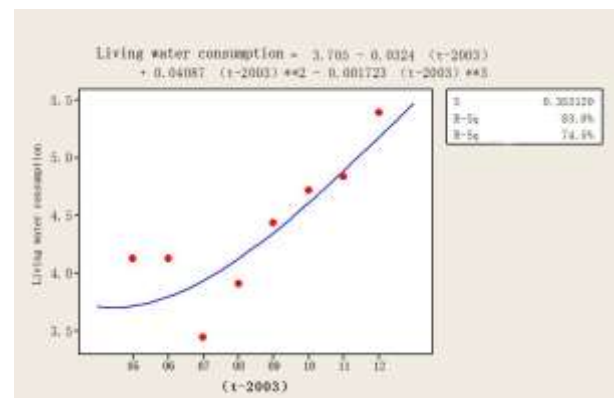Figure. 3 fitting image of industrial water
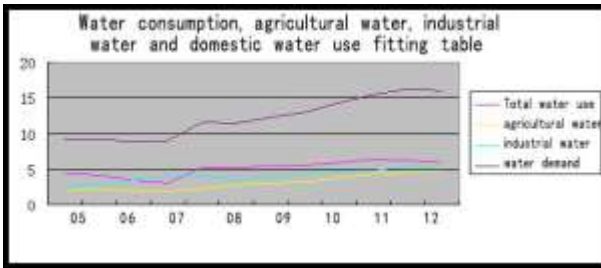

Figure. 4 the water of life

Figure. 5 Total water consumption, agricultural water consumption, industrial water consumption and other water use fitting table third

4.3 model of risk of factors of water shortage

4.3.1 Analysis on risk factors of water shortage

There are many factors that affect the water resources of Beijing City, the paper divides the risk factors from two aspects: water supply and water use.

Table 2. statistical analysis of risk factors of water resources shortage in Beijing city from 2005 to 2012

Risk factors	2005	2006	2007	2008	2009	2010	2011	2012
DX	15.7	14.7	14.8	16.5	18.5	18.5	19.2	21.4
DB	7.8	5.3	6.1	8.2	7.6	6.0	7.6	12.8
ZS	2.5	2.0	2.1	2.0	2.6	3.6	5.0	2.1
GY	9.2	7.5	8.4	7.7	6.8	6.2	5.8	5.2
NY	17.4	15.5	13.8	13.5	13.2	12.8	12.4	12.0
SH	12.0	10.8	13.0	12.8	13.4	13.7	13.9	14.7

To quantitatively analyze the influence degree of each risk factor of the status of water resources, the judgment matrix in AHP pairwise comparison of each risk factor, to determine the weights of various risk factors, which determine their influence on the total amount of water resources.

The establishment of the matrix A is as follows：

$$
\begin{array}{c|ccccccccc}
 & C1 & C2 & C3 & C4 & C5 & C6 & C7 & R & W \\
\hline
C1 & 1 & 3.0776 & 1.3884 & 8.1893 & 2.8837 & 4.7657 & 2.9678 & 24.2725 & 0.271337 \\
C2 & 0.3247 & 1 & 0.4511 & 2.6609 & 0.9370 & 1.5485 & 0.9643 & 7.8868 & 0.088164 \\
C3 & 0.7203 & 2.2167 & 1 & 5.8589 & 2.0771 & 3.4326 & 2.1376 & 17.4826 & 0.195434 \\
C4 & 0.1221 & 0.3758 & 0.1695 & 1 & 0.3521 & 0.5819 & 0.3624 & 18.2615 & 0.204141 \\
C5 & 0.2098 & 1.0672 & 0.4815 & 2.8398 & 1 & 1.6526 & 1.0291 & 8.2801 & 0.092561 \\
C6 & 0.2098 & 0.6458 & 0.2913 & 1.7184 & 0.6051 & 1 & 0.6227 & 5.0932 & 0.056935 \\
C7 & 0.3370 & 1.0370 & 0.4678 & 2.7594 & 0.9717 & 1.6058 & 1 & 8.1787 & 0.091427 \\
\end{array}
$$

among : $A = \left(a_{ij} \right),\ S = d_i - d_j, d_i = \sum_j c_{ij}$

$$ R_i = \sum_{j=1}^{7} a_{ij} \quad W_i = \dfrac{R_i}{\displaystyle\sum_{i=1}^{7} R_i} $$

W_i :The weight of the risk factor

$$ W_1 = 0.271337 \quad W_2 = 0.88164 \quad W_3 = 0.195434 $$
$$ W_4 = 0.204141 \quad W_5 = 0.092961 \quad W_6 = 0.056935 $$
$$ W_7 = 0.091527 $$

Through the calculation of 5 main risk factors, respectively, the weight of the relatively large surface water C4, C2 reclaimed water, C5 agricultural water, C6, industrial water, C7 third industry and domestic water.

4.3.2 Principal component regression

According to the ontology with fewer samples, variable dimension, multiple linear problem is serious by principal component regression model can in reducing the dimension of the variables at the same time eliminate variables highly correlated due to multicollinearity problems, so as to improve the accuracy of the regression model parameter estimation. At the same time, the establishment of the model also from the side to test the effectiveness of the above extraction of five risk factors.

A water shortage rate is Y, $Y = (U - V)/U$. U for the annual total water, V for the annual total water resources. In order to X_1, X_2, X_3, X_4, X_5 as explanatory variables, Y as the explanatory variables for multiple linear regression. Regression model can be written as

$$ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 $$

water resource scarcity risk factors were obtained from the principal component analysis:X_1,X_2,X_3,X_4,X_5, The regression equation obtained by SPSS is:

$$ Y = -0.024 X_1 + 0.1 X_2 - 0.124 X_3 - 0.15 X_4 - 0.09 X_5 + 0.271 $$

Goodness of fit test and F test of the regression equation: the test of goodness of fit, the model had an adjusted $R_2 = 0.898$, variance analysis of the significant level of less than 0.05. The test results show that the regression equation fitted well with the original data.

X_1 is known for industrial water comprehensive factors, X_2 for life and water environment comprehensive factors, X_5 for agricultural water integrated factors, by regression equations available, according to their contribution to the rate of water shortage of the rate of its importance to sort for:

Industrial water > living and environmental water > agricultural water.

It can be known that the regression model can well explain the relationship between water shortage rate and risk factors, and the significance of the index is consistent with the actual situation, which can be applied.

4.4 predicted and forecast model

4.4.1 Predicted total water consumption

Task three solution: water condition forecast for the next 15 years Environmental drivers: select Beijing City, according to the economic environment variables and time variables of Beijing city to solve the model. On the basis of population increase, scientific and technological progress and environmental change, the water status of Beijing city in the next 15 years is predicted.

Agricultural water consumption:

$$NY = 0.3152(t-2003)+3.795$$

Industrial water consumption:

$$GY =$$
$$0.01114(t-2003)2+0.2550(t-2003)+1.757$$

Third industry and other water use:

$$SH = -0.001723(t-2003)3+$$
$$0.04807(t-2003)2+0.0324(t-2003)+3.705$$

The total water use: $ZY = NY + GY + SH$

Predict the next 15 years, the total annual water consumption results is the following table:

Table 3. the total annual water consumption

Year	NY	GY	SH	Total
2016	7.10	17.00	12.33	23.53
2017	7.12	17.20	12.41	23.71
2018	7.14	17.52	12.25	23.87
2019	7.17	17.86	12.33	24.00
2020	7.25	17.69	12.45	25.23
2021	7.25	17.81	12.56	25.31
2022	7.56	17.73	12.67	25.28
2023	7.12	17.92	12.70	25.54
2024	7.67	17.93	12.71	25.57
2025	7.78	17.92	12.78	25.56
2026	7.81	17.87	12.85	25.63
2027	7.92	17.89	12.88	25.69
2028	7.93	17.92	12.92	25.77
2029	7.89	17.98	12.98	25.85
2030	7.90	18.12	13.23	16.23

4.4.2 Forecast total water resources

Based on our analysis on data from 2005 to 2012, surface water and groundwater in a certain range fluctuations, reclaimed water quantity rise in the basic trend, in view of this,

we have a bold prediction of four years of water resources in the future. According to the principle of Statistics found that the basic surface water and underground water fluctuations in a certain range, while the sewage treatment capacity because of the support of national policy it has been in a rising state, we take years of data for comparative analysis, speculated that the next four years surface water and underground water to maintain in the eight years ago, the average level of sewage treatment capacity for nearly three years on average, so the water resources amount in the next four years is speculated as follows table:

Table 4. water resources amount in the next four years

Year	Surface water quantity	Underground water	Sewage treatment capacity	Total water resources
2016	7.70	17.00	12.33	23.53
2017	7.70	17.20	12.41	23.71
2018	7.70	17.52	12.25	23.87
2019	7.70	17.86	12.33	24.00
2020	7.70	17.69	12.45	25.23
2021	7.71	17.81	12.56	25.31
2022	7.72	17.73	12.67	25.28
2023	7.72	17.92	12.70	25.54
2024	7.83	17.93	12.71	25.57
2025	7.78	17.92	12.78	25.56
2026	7.81	17.87	12.85	25.63
2027	7.92	17.89	12.88	25.69
2028	7.93	17.92	12.92	25.77
2029	7.89	17.98	12.98	25.85
2030	7.90	18.12	13.23	16.23

Description: the total amount of water resources can be solved by linear regression equation:

$$\hat{y} = -4.732 + 2.138x1 + 0.498x2 + 0.274x3.$$

5. WATER SHORTAGE DRIVEN FACTOR INTERVENTION PROGRAM

5.1 Holistic analysis

Intervention program from the natural environment, topography and climate, economic and population factors, water supply system, technical reserves, water pollution, water price, water purification ability of each model and the model parameters are taken into account, as far as possible according to the different situation of different places to draw the best interventions.

5.2 Model introduction

Since The index system is a reasonable abstraction and Simulation for the sustainable utilization of water resources, but the contribution of each index to the system is different, which can be used to a set of normalized weight W_1, W_2,... ,

W_n to describe. Traditional weight distribution commonly used Delphi expert consultation method and analytic hierarchy process (AHP). Among them, the Delphi method is not only a higher requirement for consulting experts, but also a large number of rounds of consulting work. AHP uses nine scale score comparison to judge the result, makes the person feel the operation difficulty, and the computation complexity, but also needs to carry on the consistency check. Under the premise of guaranteeing the scientific nature, in order to calculate the simple and convenient, this paper uses the improved three scale analytic hierarchy process, its computation steps like formula (1) ~ (3).

Subjective comparison matrix：$C = \begin{bmatrix} c_{ij} \end{bmatrix}_{n \times n}$,

Where n is the number of indicators, the element C_{ij} is expressed in terms of S_k, the relative importance of C_i relative to C_j,

Establishment of sensory judgment matrix:

$$S = \begin{bmatrix} s_{ij} \end{bmatrix}_{n \times n}, (2)$$

Among them, $S = d_i - d_j, d_i = \sum_j c_{ij}$

Calculation of objective judgment matrix:

$$R = \begin{bmatrix} r_{ij} \end{bmatrix}_{n \times n}, (3)$$

Among them,

$$r_{ij} = p^{(Sij/Sm)}$$

$$s_m = \max_{i,j} s_{ij} = \max_i(d_i) - \min_j(d_j)$$

P is defined as the scale extension of the user, such as p=3 or 7, this article takes 3. We can get the weight vector R, W2, by normalizing the arbitrary column of matrix [W1,... , Wn]T.

The criterion layer 4 indexes, through consultation and calculation, it is concluded that their weights. The present situation of water resources for 0.2765 and water resources utilization efficiency 0.1917, the pressure of the sustainable utilization of water resources for 0.1329 and ability of sustainable utilization of water resources is 0.3988. Then the index system in the next level of the indicators were calculated, the last layer of indicators in the index system of the weight of the factors (Table 5).

Table 5. w*ater resources sustainable use specific evaluation index weight*

target	Criterion layer	weight
General level of sustainable utilization of water resources	Current situation of water resources *(S)*	0.2765
	Water consumption *(E)*	0.1917
	Water resources sustainable utilization pressure *(P)*	0.1329
	Sustainable utilization of water resources(C)	0.3988

Events related to network news, forums, blog, microblogging and other Internet information resources, information capacity, high redundancy and susceptible to noise interference. To collect comprehensive information on these and filtering, on the one hand a huge workload and high labor costs, On the other hand due to the passage of time, some of the information on the Internet will disappear, to be covered by the new information, Therefore, quantitative study of this article only collect and use information representative corpus as samples. Although, in the foregoing chapter we have made under the existing study based on the results of the theory and use of the public emergency network news also from a public opinion which reflects the evolution of the development and changes of the network news coverage. However, Insufficient sample of the present study was to analyze the qualitative part brings some limitations.

Since public opinion guidance and management is a very practical work, specific to the implementation level. We need to deal with a large number of complex issues and deal with complicated things, due to the lack of implementation experience, guidance and management research in the public emergency network public opinion aspects, This article does not find significant innovations, suggestions and measures proposed are mainly to further enrich the existing research-based.

6. MODEL EVALUATION

6.1 Strengths

(1)interfere with the plan in the principal component analysis model based on, eliminating the precipitation, GDP, industrial water and other indicators related effect, reduce the workload of index selection, more objectively reflect the display relationship between risk factors and water shortage.

(2) in this paper, a comprehensive evaluation model of water resources based on entropy weight is established, which can avoid the problem that the pairwise contrast structure judgment matrix is easy to happen. The theory of entropy in the information theory was introduced from the data itself reflects the information disordering effect value to calculate the weight coefficients of the weight determination of the certain theory basis to reduce the subjectivity of evaluation and subjective factors influence and evaluation of the results.

(3) the establishment of a gray prediction model, to predict the risk of water shortage in Beijing 2011, 2012 years, and effectively quantify the risk of water shortages in the next two years in Beijing.

(4) all models of this paper, the whole process of clear thinking, simple calculation, the objective of the evaluation results.

6.2 Weaknesses

(1) the inevitable part of the principal component analysis model is to be lost.

(2) the grey model is used to predict, in view of the smoothness of the sequence is not good, the selection of the background value is more difficult.

(3) due to the environment and human impact of water resources is greater, so the results of the prediction error is inevitable.

(4) In the future, we prefer to integrate the optimization with some traditional data mining approaches [8], [9], [10] to find more effective solutions.

7. REFERENCES

[1] Gordon F L Florida west coast desalter negotiations-progress report #3 1998(03) Scott A S The work exchanger for SWRO energy recovery 1999(04).

[2] Gordon F L.Developer selected for 25MGD(94625m3/d) Florida west coast seawater desalting plant[J]. The international Desalination &Water Reuse, 1999, 9(1): 11~16 .

[3] Xiaojun Wang; Hui Zhao;Zhi Geng,Current groundwater exploitation situation and protection countermeasure in China, China Water Resources; 2010 13.

[4] Siqing Lin; Weirun Zhang, Cost of seawater desalination. J MEMBRANE SCIENCE AND TECHNOLOGY, Vol. 22 No.4: 1007-8924,2001 .

[5] L. Awerbuch, Integration of desalination and po wer: the challenge in the Middle East. Paper presented at Middle East Energy. Dubai (8–10 October 2001) .

[6] Ting Lu, Qi Gao, Xudong Wang, Cong Liu, "Modified Page Rank Model in Investigation of Criminal Gang", International Journal of Science and Engineering Applications, 4(3): 100-104, 2015.

[7] Baohua Liu, Xudong Wang, Qi Gao, Cong Liu, "A K-Means based Model towards Ebola Virus Prorogation Prediction", International Journal of Science and Engineering Applications, 4(3): 95-99, 2015.

[8] Libao Zhang, Faming Lu, An Liu, Pingping Guo and Cong Liu*, "Application of K-Means Clustering Algorithm for Classification of NBA Guards", International Journal of Science and Engineering Applications, 5(1), pp. 1-6, 2016.