

# An Optimized Model of Water Resources Assessment and Its Application

A-Ling Song

Institute of Sciences, PLA University of Science and  
Technology, Nanjing, 210007  
Nanjing City, Jiangsu Province, China

Rong-Ping Fan

Institute of Sciences, PLA University of Science and  
Technology, Nanjing, 210007  
Nanjing City, Jiangsu Province, China

Hai-Feng Liu

Institute of Sciences, PLA University of Science and  
Technology, Nanjing, 210007  
Nanjing City, Jiangsu Province, China  
e-mail: 25191908@qq.com

Shou-Sheng Liu

Institute of Sciences, PLA University of Science and  
Technology, Nanjing, 210007  
Nanjing City, Jiangsu Province, China

**Abstract-** In this paper, the classical WPI model is optimized by using principal component analysis, and a PCA based evaluation model is proposed. When the model is applied to the water resources data of Beijing area, the experiments have shown that the evaluation efficiency of WPI is significantly improved. And when used to evaluate WPI in different city and countryside, the optimized model has also achieved satisfying results. On this basis, this paper puts forward a prediction model for future water environment, which has predicted future water resource in the Shan-gan-ning district and also puts forward a set of intervention programs based on related analysis. The data shows that the model is effective in alleviate the water shortage.

**Keywords-** Water scarcity; WPI; Principal component analysis; Correlation analysis; Sensitivity analysis.

## I. INTRODUCTION

Water scarcity has become very serious all over the world, accurate assessment of water resources has an important realistic demand for social development. The water poverty index (WPI) is an interdisciplinary approach to assess water stress. At present, there are some scholars in this area. In the work by Luo Jun-gang et al. [1], they put in the entropy theory is introduced to combine with the traditional fuzzy comprehensive assessment method to establish an entropy fuzzy comprehensive assessment model. Based on the principle of nonlinear multi-objective programming a model for optimal allocation of regional water resources with multi-consumers was developed in Sun Zhi-lin et al [2]. work and more the validity and adaptability of the model are verified by experiments. based on the maximum information entropy principle, Han Yu-ping et al. [3] aiming at the lower resolutions of the evaluation results what existed in the risk analysis methods of water resources system, the risk rate, vulnerability, possibility of recovery, recurrence interval and risk level of regional water resources shortage were chose as the evaluation indexes to establish the risk evaluation model of regional water resources shortage. Based on the fuzzy entropy weight

model comprehensive evaluation method, Han Yun-hong et al [4]. evaluated the water resources carrying capacity of Fuyang City, Anhui Province in 2009-2013, and more obtained the conclusion that the water resources carrying capacity is close to saturation. Based on WPI model, Zhang Zhe et al. [5] built an assessment index system of water resources security, they calculated the comprehensive security index of water resources and the five subsystems scores during 1997-2013 in Chongqing city using weighting function method, and conducted comparative analysis by using the method of factor analysis. Furthermore, they established factors rating model and classified the five subsystems into superior, medium and inferior. In this paper, based on the above related research, with the aid of the principal component analysis method, the optimization of the classical water resources model makes the data processing more close to the reality. The second step is choose Shan-gan-ning district in China as our study object. In order to give a comprehensive analysis of cause and seriousness of the water scarcity in this region, we study both the urban and rural areas in the region. The analysis shows Shan-gan-ning district belongs to the "not safe" class. The third step, we build a prediction model incorporating feature weight adjustment based on feedback information, which proves to be more holistic and useful in dealing with the impact on citizens' lives. At last, we build up a model based on effect index to express intervention priority in different region. In our chosen region, capacity scores -2.5 and ranks the first. Then the paper introduces an intervention plan based on univariate intervention analysis model aimed at increase the WPI value to 60 by 2030.

## II. ANALYSIS ON THE BASIC FACTORS OF WPI EVALUATION SYSTEM

In general, The water poverty index (WPI) is composed of 5 key issues, Resource (R), Access (A), Capacity (C), Use (U), Environment (E), which are essential to the assessment of water provision.

TABLE I. WPI VALUE, CORRESPONDING CLASS AND DESCRIPTION

WPI value	Class	Description
$WPI > 62$	Very safe	Water resources and water environment system is in harmony with the social and economy system and operate efficiently. The degree of satisfaction is the highest.
$56 < WPI < 62$	Safe	Water resources and water environment is cooperating with the social and economy system well. The degree of satisfaction is relatively higher.
$48 < WPI < 56$	Generally safe	Water resources, water environment system and society, economy can develop coordinately. The degree of satisfaction is normal.
$35 < WPI < 48$	Not safe	Water resources, water environment system cannot be coordinated with the social and economic development and it has been a threat to the community's sustainable economic development.
$WPI < 35$	Extremely unsafe	Water resources, water environment system is under deterioration, and become serious impediment to social and economic sustainable development.

The WPI is primarily designed to provide a tool by which water managers can evaluate the water situation in different locations in a holistic way.

According to the authority data, taking the current social and economic development level into account, water safety can be divided into five classes by the WPI value [6]:

The WPI calculation consists of 2 levels. The relationship between each level is shown as the Figure 1 below:

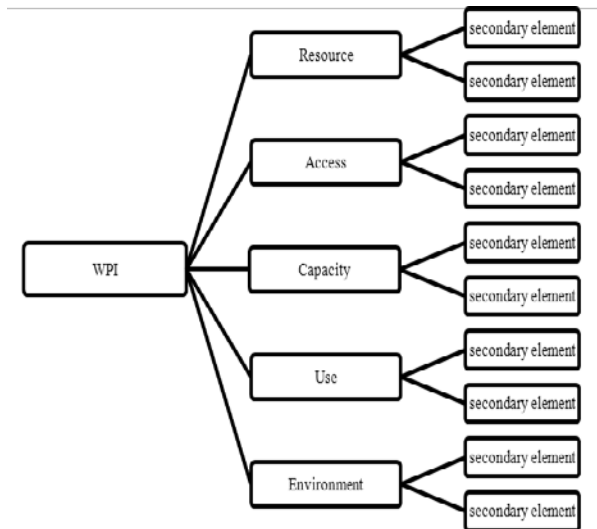


Figure 1. Two levels structure of WPI

The mathematical structure on which the composite index version of the WPI is based, is expressed as follows[7]:

$$WPI = \frac{\sum_{i=1}^n \omega_i X_i}{\sum_{i=1}^n \omega_i} \quad (1)$$

Where WPI standing for the Water Poverty Index value for a particular region, is the weighted sum of five components Resource, (R) Access (A) Use (U) Capacity (C)

and Environment (E), each having a value ranging between 0 and 100. The weight  $\omega_i$  is applied to each component (X) of the WPI structure, for that region, with  $X_i$  referring to the value of each component.

To standardize the result and produce a WPI value of between 0 and 100, the sum needs to be divided by the sum of weights as shown:

$$WPI = \frac{\omega_r R + \omega_a A + \omega_c C + \omega_u U + \omega_e E}{\omega_r + \omega_a + \omega_c + \omega_u + \omega_e} \quad (2)$$

In order to eliminate the influence of the dimension of each evaluation index, we also have to carry out standardized processing to the sample data set. We should keep the information as more as possible, so we are going to divide the index into two different catalogs:

For some indexes, bigger value means better water situation result (such as the sum of water resources within the region etc.). These indexes adopt formula as follows:

$$v(i, j) = \frac{x\{i, j\}}{x_{\max}\{j\} + x_{\min}\{j\}} \quad (3)$$

For the other indexes, smaller value means the better water situation (such as the water pollution etc.). These indexes adopt formula as follows:

$$v(i, j) = \frac{x_{\max}\{j\} + x_{\min}\{j\} - x\{i, j\}}{x_{\max}\{j\} + x_{\min}\{j\}} \quad (4)$$

Where  $v(i, j)$  stands for the  $j$  th index value in the  $i$  th sample data set after the standardized process,  $x(i, j)$  stands for the initial  $j$  th index value in the  $i$  th sample data set. Since the WPI value is between 0 and 100, so the data should be multiplied by 100 to participate in the formula (3) to get the final WPI value.

### III. THE DESIGN AND OPTIMIZATION OF WPI ASSESSMENT MODEL

#### A. Model based on entropy weight and G1 method to calculate WPI value

In this section we propose a method of calculating water poverty index based on entropy weight and G1 method [8]. There are many methods to determine the weight. G1 method is a subjective weighting method which requires the expert to sort the importance of each index, giving a relatively good reflection of the expert's views. But the experts always have a hard time to judge between the indexes which have adjacent importance. To solve this problem, we adopt a model based on the entropy weight and G1 method to calculate WPI value. By calculating the size of the entropy to determine the importance of adjacent index, the ratio reflects the amount of information which the data contains, making the combination weights reflect not only the opinions of the experts and also the changes in the index data, thus more objective and rational. The steps of the new model based on the entropy weight and G1 method are as follows:

- (1) The ranking the importance of indicators was done by experts. For evaluation index set  $\{x_1, x_2, \dots, x_m\}$ , and the ranking each indicator to form a new set with order  $\{X_1, X_2, \dots, X_m\}$  was done by experts, where  $X_1$  has the greatest importance on the water situation.
- (2) We then calculate the entropy of each indicator after the ranking.  $e_j$  is the  $j$ th indicator's entropy, then it can be calculated by formulas as follows:

$$f_{ij} = \frac{v_{ij}}{\sum_{i=1}^n v_{ij}} \quad (i=1, 2, \dots, n; j=1, 2, \dots, m) \quad (5)$$

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n f_{ij} \ln(f_{ij}) \quad (6)$$

Where  $f_{ij}$  stands for the  $j$ th indicator's feature ratio among the  $i$ th object,  $v_{ij}$  stands for the  $j$ th indicator's observation data among the  $i$ th object. For a given  $j$ , the greater the difference of  $v_{ij}$  is, the more information this indicator contains.

- (3) Then we should to determine the importance ratio between  $x_{k-1}$  and  $x_k$ :

$$r_k = \begin{cases} \frac{e_{k-1}}{e_k}, & e_{k-1} \geq e_k \\ 1, & e_{k-1} < e_k \end{cases} \quad (7)$$

- (4) And more we should to calculate the modified weight.

According to the  $r_k$  from the formula above, we can get the modified weight by formulas as follows:

$$r_k = \begin{cases} \frac{e_{k-1}}{e_k}, & e_{k-1} \geq e_k \\ 1, & e_{k-1} < e_k \end{cases} \quad (8)$$

$$\omega_{k-1} = r_k \omega_k \quad (k = m, m-1, m-2, \dots, 3, 2) \quad (9)$$

- (5) By introducing these modified weight into the formula 1, we get the WPI value under the situation where weight differs between each indicator, The model one:

$$WPI = \frac{\omega_r R + \omega_a A + \omega_c C + \omega_u U + \omega_e E}{\omega_r + \omega_a + \omega_c + \omega_u + \omega_e} \quad (10)$$

#### B. Model improvement based on principal component analysis

Though the model based on the entropy and G1 method does solve the problem that weight should differ according to the importance of each indicator. But the importance actually comes from the expert's opinion, so inevitably it may become subjective. This is the shortcoming of the opinion-based model.

To make the evaluation more objective, we adopt the Principal Component Analysis in the five components. Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.

Actually, the five components of WPI, Resource, Access, Use, Capacity and Environment, do have correlation with each other. So what we have done in this model improvement is to carry out Principal Component Analysis to eliminate the correlation between the components and then establish a new method to calculate the WPI value.

The advantage of this method is that it makes full use of the previous data. When we have a large amount of data, we can carry out the improvement as the following steps [9]:

- (1) To standardize the raw data of all the years using the formula as follow:

$$x_i = (X_i - \bar{x}_i) / S_i \quad (11)$$

- (2) To consider the five elements of WPI, calculate correlation coefficient matrix for the five components:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{21} & r_{22} & r_{23} & r_{24} & r_{25} \\ r_{31} & r_{32} & r_{33} & r_{34} & r_{35} \\ r_{41} & r_{42} & r_{43} & r_{44} & r_{45} \\ r_{51} & r_{52} & r_{53} & r_{54} & r_{55} \end{bmatrix} \quad (12)$$

Where  $r_{ij}$  ( $i, j = 1, 2, 3, 4, 5$ ) stands for the correlation value between each component.

$$r_{ij} = \frac{\sum_{k=1}^p (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^p (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^p (x_{kj} - \bar{x}_j)^2}} \quad (13)$$

- (3) To compute Eigenvalues and Eigenvectors:  
Solve the equation to get the eigenvalues with Jacobi method:  $|\lambda E - R| = 0$

And then sort them by size:  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_5 \geq 0$

Figure out the corresponding eigenvectors  $e_i$  ( $i = 1, 2, 3, 4, 5$ ), which satisfy the equation:

$$\|e_i\| = 1, \text{ that is } \sum_{j=1}^p e_{ij}^2 = 1.$$

- (4) To compute contribution rate and cumulative contribution rate for each components, the contribution rate can be calculated as follow:

$$\lambda_i / \sum_{k=1}^p \lambda_k, (i = 1, 2, \dots, p) \quad (14)$$

The cumulative contribution rate:

$$\sum_{k=1}^i \lambda_k / \sum_{k=1}^p \lambda_k, (i = 1, 2, \dots, p) \quad (15)$$

- (5) We compute principal component load and score  
 $z_{ij} = p(z_i, x_j) = \sqrt{\lambda_i} e_{ij}$  ( $i, j = 1, 2, \dots, p$ )

$$Z = \begin{bmatrix} z_{11} & z_{12} & z_{13} & z_{14} & z_{15} \\ z_{21} & z_{22} & z_{23} & z_{24} & z_{25} \\ z_{31} & z_{32} & z_{33} & z_{34} & z_{35} \\ z_{41} & z_{42} & z_{43} & z_{44} & z_{45} \\ z_{51} & z_{52} & z_{53} & z_{54} & z_{55} \end{bmatrix} \quad (16)$$

Then we can introduce this Z matrix into the calculation of WPI value.

#### IV. CASE ANALYSIS AND MODEL COMPARISON

##### A. Analysis of WPI status in Beijing

We take Peking's water situation data from 2005 to 2015 as study object to verify our model and difference between each method.

TABLE II. PEKING'S WATER SITUATION DATA FROM 2005 TO 2015[10]

Year	Resources	Access	Capacity	Use	Environment
2005	43.5	38.5	39.7	30.3	38.4
2006	38.4	35.8	37.3	33.6	39.4
2007	41.7	39.6	42.4	37.3	41.6
2008	39.9	42.2	45.8	39.8	37.9
2009	46.3	41.1	34.7	31.7	38.9
2010	36.1	37.7	43.4	36.5	43.4
2011	39.7	44.6	47.6	39.3	38.1
2012	38.9	43.9	45.1	42.1	37.4
2013	42.4	40.2	43.5	45.8	39.6
2014	38.1	39.7	46.5	43.5	42.7
2015	37.8	41.7	43.9	46.7	41.8

Now we are going to use Peking's WPI value in 2015 to evaluate Peking's ability to provide clean water to meet its demand.

By applying the entropy-weight and G1 method, Peking's WPI value in 2015 is 48.7. By analyzing the data from 2005 to 2014, we applied principal component analysis on each components and get a new indicator. By using the new indicators, we calculate Peking's WPI value in 2015 is 46.9.

TABLE III. THREE METHODS TO CALCULATE WPI VALUE

area	Entropy-weight and G1		method		principal component analysis	
	WPI		class		WPI	class
Peking	48.7	Generally safe			46.9	Not safe

So we have table above to compare these two methods, we can see that the Peking's WPI value calculated from the former model still belong to the "generally safe" class, while result coming from the principal component analysis shows that Peking's water situation has dropped to "not safe" class. According to the UN water scarcity map [11], Peking is one of the heavily water-overloaded area, so we can see that principal component analysis provides a better result.

##### B. The WPI differences between Urban and Rural Areas based on the PCA

In the case of country or region level assessment using the composite index approach, there are significant differences between urban and rural areas. Where possible, the same variables have been used in each case, but due to differences between the sites, this has not always been

possible. For example, some of the indicators used for the environmental attribute are not representative of the state found in urban areas, whereas they are appropriate for the rural areas. The differences between urban and rural sites mean that although the same WPI framework can be used, some of the sub-components are different, making direct comparisons of the values less reliable between urban and rural sites. So, when we pick the country or region, we are going to discuss the different sites of the area.

Meanwhile, in order to give a comprehensive analysis about the seriousness and cause of water scarcity, we must take into the climate situation into consideration. In the certain area, it may have the dry and wet season, which will bring great impact to the WPI value. Though WPI value gives an overall evaluation of the water situation. But when we want to take a close look at the causes of the scarcity, it is worthwhile to study the WPI value in different season.

Case study and WPI value results. We pick Shan-gan-ning district in China to carry out our case study. Based on the analysis above, we will divide our analysis into 4 parts [10]:

- 1) WPI value for urban locations in dry seasons;
- 2) WPI value for urban locations in wet seasons;
- 3) WPI value for rural locations in dry seasons;
- 4) WPI value for rural locations in wet seasons.

From the information showing these seasonal WPI values for urban areas, we can see that during the wet season, much better access rates brings the WPI scores up. This is because in some areas, water supply points become unreliable in the dry season, generating lower WPI values.

These figures give some indication of how the WPI scores may vary between communities. The communities presented here are fairly representative of thousands of communities in Africa and South Asia, in urban and rural areas. From this information we can see that on this measure, the WPI scores range from 32 to 35.6 for urban areas in the dry season, while in rural areas, WPI scores range from 28.8 to 32.1 for the same period. As would be expected, wet season scores tend to be higher (better), tending to provide evidence supporting the robustness of the WPI methodology.

TABLE IV. WPI COMPONENTS OF URBAN AREAS IN DRY SEASON

Village	Component values(Dry season)					
(Urban)	Resources	Access	Capacity	Use	Environment	WPI
Lanzhou	28.0	46.7	45.9	23.2	39.4	35.6
Longnan	20.2	26.5	48.7	22.9	42.2	32
Zhangye	23.2	51.8	61.9	22.9	28.7	32.7

TABLE V. WPI COMPONENTS OF URBAN AREAS IN WET SEASON

Village	Component values(Wet season)					
(Urban)	Resources	Access	Capacity	Use	Environment	WPI
Lanzhou	34.7	48.8	46.1	19.7	42.2	40.2
Longnan	25.3	25.9	50.6	18.9	43.5	35.6
Zhangye	28.2	53.2	65.3	21.6	31.3	36.7

TABLE VI. WPI COMPONENTS OF RURAL AREAS IN DRY SEASON

Village	Component values(Dry season)					
(Rural)	Resources	Access	Capacity	Use	Environment	WPI
BanqiaoXiang	20.8	39.5	42.1	21.2	40.1	32.1
Baiyang Xiang	19.8	25.6	40.4	25.9	39.9	30.3
Shangwan Xiang	20.1	42.3	42.3	25.9	38.4	28.8



TABLE VII. WPI COMPONENTS OF RURAL AREAS IN WET SEASON

Village	Component values(Wet season)					
(Rural)	Resources	Access	Capacity	Use	Environment	WPI
Banqiao Xiang	25.6	40.5	41.9	20.3	39.1	35.6
Baiyang Xiang	22.3	29.8	38.7	36.5	36.8	33.1
Shangwan Xiang	23.3	24.7	41.1	28.9	35.9	30.9

## V. THE PREDICTION MODEL OF THE WATER SITUATION

We will still use the WPI value to evaluate the water situation, but we can not predict the WPI value directly. Suppose we have the WPI data for Shan-gan-ning district from 2005 to 2015. If we place the WPI value in a figure and then adopt curve fitting to get the following 15 years' WPI value, that will be quite meaningless because on the one hand, the way to calculate WPI value differs from year to year, on the other hand, we can't give a holistic analysis of the water situation for the resident within the region with only a WPI value.

So we will start from the basic elements of the WPI value. We collect data for each component of WPI value in Shan-gan-ning district from 2006 to 2015. Use these data to predict the development of each elements in the following 15 years. Then use these values to calculate the WPI value for each year.

### A The prediction of basic elements of WPI

Different elements take on different model to predict its future value. Basically we can divide these 5 elements into two catalogs:

- (1) As time goes by, the value obeys its previous tendency, but the speed of changing becomes slower. These elements include access, capacity and use. For this catalog, we will take the logarithmic model [12], because the logarithmic model has a decreasing derivative value.
- (2) As time goes by, the value will not change too much unless great change take place in other elements. These elements are relatively stable, including resources and environment. For this catalog, we will take the Autoregressive Integrated Moving Average Model (ARIMA) [13] since it has stable tendency.
- (3) In the first catalog, we take the element "capacity" as an example. We have the data for capacity in Shan-gan-ning district from 2006 to 2015 as follows:

TABLE VIII. CAPACITY VALUE OF SHAN-GAN-NING DISTRICT FROM 2006 TO 2015[10]

2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
35.7	37.2	39.6	41.7	43.3	44.6	45.9	47.1	48.0	48.6

We assume the tendency obeys the model like this:

$$C(t) = a \ln x + b \quad (17)$$

Use the data above to carry out curve fitting, and we can get result as follows:

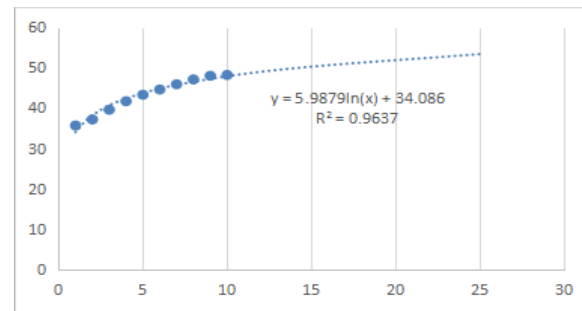


Figure 2. Predicted curve of component "capacity"

So we will get our prediction data as the following chart:

TABLE IX. PREDICTED VALUE OF COMPONENT "CAPACITY"

2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
48.4	48.9	49.4	49.8	50.3	50.8	51.0	51.3	51.7	52.0	52.3	52.5	52.8	53.1	53.3

In the second catalog, we take the element "resource" as an example. We have the data for environment in Shan-gan-ning district from 2006 to 2015 as follows:

TABLE X. RESOURCE VALUE OF SHAN-GAN-NING DISTRICT FROM 2006 TO 2015[10]

2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
26.7	25.8	26.0	27.8	25.6	25.9	26.3	26.5	27.0	25.9

Before we carry out the prediction, we will first give a brief introduction of the ARIMA. This model is fitted to time series data either to better understand the data or to predict future points in the series (forecasting). They are applied in some cases where data show evidence of non-stationarity, where an initial differencing step (corresponding to the "integrated" part of the model) can be applied to reduce the non-stationarity.

The element environment shows non-stationarity in the data we get, but the real condition is that it should be

relatively stable because it's not that to change the environmental factors in a short period.

We apply the Autoregressive Integrated Moving Average Model (ARIMA)[13] and get the prediction value of environment as follows:

TABLE XI. PREDICTED VALUE OF COMPONENT 'RESOURCE'

2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
26.4	25.9	27.4	26.8	25.3	24.8	25.0	25.3	25.7	26.0	24.3	25.5	24.8	25.1	25.3

So we apply the logarithmic model to element access, capacity and use, while element environment and resource can be predicted with ARIMA. We can get all the prediction value of the basic elements in the next 15 years. Due to the length limitation, here we will not give out the prediction value of the other three elements.

#### B The prediction of WPI without considering the weight change during the following 15 Years

After we have got the basic elements of WPI value, we can carry out calculating the WPI value for the next 15 years. However, here the weight becomes another problem to be solved. We can assume that the weight stays as the same with the weight we use before. This method is simple and easy, but quite rough.

The WPI values calculated without considering the weight change during the next 15 years are as follows:

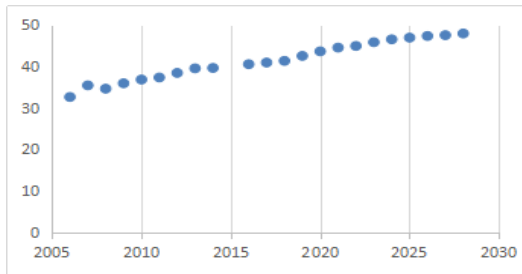


Figure 3. Predicted value of WPI without considering the weight change

#### C The prediction of WPI incorporating the weight change during the following years

The previous prediction does not take the weight change into consideration. It is not reasonable since the weight which forms the WPI value is highly changeable. For example, if element use is reaching a rather high level, then its contribution to WPI value may fall gradually. So what we are trying to address here is the model to carry out weight adjustment.

The model we use is the feature weight adjustment method based on feedback Information [14]. It is a method which is used in text categorization, but here it can also be applied to adjust the weight for each element's contribution to the WPI value. This model's advantage is that it is closely related to the change of the elements, which will help to give more information than the traditional model. The

detailed model explanation can be obtained in literature[14]. Here due to the length limitation, we won't give steps of the model.

We have our two prediction models' result shown in the table below:

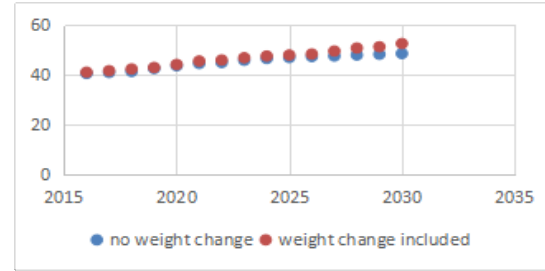


Figure 4. Comparison between two kinds of predicting method

In the figure above, we can see that the prediction value of the second model is a little bit higher than the first model, which means better water situation. The second model does contain more information than the first model, and is highly flexible to the time series. We believe the second model is thus better than the first one. The following contents of this paper are going to adopt the second model in value prediction.

#### D Model to describe the impact on citizens' lives in the region

Impact on citizens' lives can be described with a function related with the access and environment elements, among which access is the most important index to describe the citizens' lives.

So we build a function to describe the citizens' satisfaction index [15] to the water situation as follows:

$$G = \frac{\omega_1 * access + \omega_2 * environment}{\omega_1 + \omega_2} (\omega_1 > \omega_2) \quad (18)$$

Where  $G$  represent the satisfaction index to the water situation of the local citizens.

Using the formula above and the data we predicted before, we get the index from 2016 to 2025 as follows:

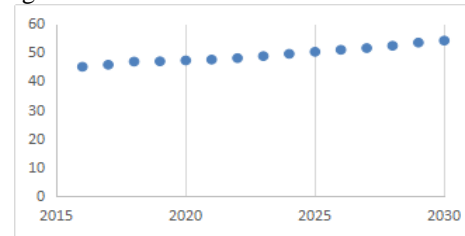


Figure 5. Predicted value of satisfaction index

We can see from the figure above that: during the following 15 years, the situation has a good impact on citizens' lives, which brings higher satisfaction index towards water availability.

## VI. INTERVENTION PLAN DESIGN AND MODEL BASED ON CORRELATION ANALYSIS TO EVALUATE THE IMPACT ON THE SURROUNDING AREAS

For different region, we should have different intervention priority since the weight and influence of each element is different in the WPI calculating process.

First, we are going to define effect index of each element to the WPI value to determine the intervention priority, then we will give our intervention model based on univariate intervention analysis model.

As it has been mentioned above, in the five elements of WPI, access, capacity and use are highly changeable and can be controlled with intervention plan, while the other two, resources and environment reflect the effect of intervention plan on the surrounding areas and the entire water ecosystem. So our intervention plan will focus on the intervention with access, capacity and use.

A *Model of intervention priority and intervention plan*  
we introduce the effect index as follows:

$$N_i = (PWI - X_i) * \omega_i (i=1,2,3) \quad (19)$$

Where  $N_i$  represents the effect index of access, capacity and use accordingly.

If  $N_i \geq 0$ , we believe that this element has negative effect on the WPI value, and we need intervention on the element. The greatest  $|N_i|$  represents that the  $i$ th elements rank the first in the intervention priority.

After figuring out the intervention priority, we are going to explain our intervention model. Our intervention model belongs to the continuous intervention variables. It means when intervention takes place at time T, after T the impact continue to exist, which can be described as follows[16]:

$$S_i^T = \begin{cases} 0, & \text{before event } (t < T) \\ 1, & \text{after event } (t \geq T) \end{cases} \quad (20)$$

And the influence of the intervention can be described as:

$$Y_t = \omega S_i^T \quad (21)$$

Where  $\omega$  stands for the unknown parameter of the intervention intension. When  $Y_t$  is not stable enough, we can apply the difference model to generate a Stationary series, then the model can be adjusted to:

$$(1 - B) \cdot Y_t = \omega S_i^T \quad (22)$$

where  $B$  stands for backward shift operator.

Suppose the Univariate series after stationary process satisfies the following model:

$$y_t = \frac{\theta(B)}{\phi(B)} a_t \quad (23)$$

And the influence of intervention case is:

$$Z_t = \frac{\omega(B)}{\delta(B)} I_t^T \quad (24)$$

Then, the Univariate intervention analysis model can be describe as:

$$\begin{aligned} y_t &= \frac{\omega(B)}{\delta(B)} I_t^T + \frac{\theta(B)}{\phi(B)} a_t \\ &= \psi(B) I_t^T + \varepsilon_t \end{aligned} \quad (25)$$

Where:  $\psi(B) = \frac{\omega(B)}{\delta(B)}$ ,  $\varepsilon_t = \frac{\theta(B)}{\phi(B)} a_t$

The steps are as follows

- (1) By using the data before the intervention, the time series model of the single variable is established. Then this model is used to predict the extrapolation, and the predicted value is obtained as a numerical value that is not affected by the intervention.
- (2) Subtract the expected value after intervention from the predicted value without intervention to get the parameter in the intervention model. Here in Shan-gan-ning district, our expected value of WPI after intervention is 60.
- (3) Identify and estimate the time series model of a single variable, by excluding the whole data after the intervention.
- (4) Establish the whole intervention model.

In Shan-gan-ning district, the effect index of each element and the intervention priority is shown in the table below:

TABLE XII. INTERVENTION PRIORITY FOR ACCESS, CAPACITY AND USE

element	access	capacity	use
Effect index	-1.3	-2.5	0.8
Intervention priority	2	1	3

Capacity ranks the first in the intervention priority. So we set the capacity as the first element to intervene.

### B *Correlation analysis of the key components*

We want to figure out the impact on surrounding areas, so we study the correlation relationship between element resource and environment with the element we are going to intervene. Specially, the correlation matrix in Shan-gan-ning district, which has been acquired in task 1, is shown below:



TABLE XIII. CORRELATION MATRIX

	Resource	Access	Capacity	Use	environment
Resource	1	-0.56	0.62	0.84	0.95
access	-0.56	1	0.23	-0.15	-0.76
capacity	0.62	0.23	1	0.04	0.85
use	0.84	-0.15	0.04	1	0.82
environment	0.95	-0.76	0.85	0.82	1

From the correlation matrix above we can see, both the resource and environment are negatively related to access. That is because, when the extent to water for human use becomes larger, the resource will inevitably fall. Also, in order to expand the way to get water, man will build transport line, which will do harm to the environment nearby.

But both the resource and environment are positively related to capacity and use. As society and technology advance, the efficiency, utilization ratio and recycling ration increase, bringing positive effect to both the resource and environment.

#### C The effect on mitigating the water scarcity

Since we directly control the single elements by the intervention plan, the effect of the intervention on mitigating the water scarcity is not that clear. So in this part we introduce an index called remission rate for each elements.

The remission rate can be expressed as follows:

$$\rho_i = \frac{|\Delta X_i|}{|\Delta WPI|} (i=1,2,3) \quad (26)$$

In Shan-gan-ning district, we have the remission rate shown in the following chart:

TABLE XIV. REMISSION RATE FOR EACH COMPONENT

element	access	capacity	use
Remission rate	0.21	0.54	0.13

So we say intervening the capacity is the most effective way to mitigate the water scarcity. Our intervention plan mainly focuses on controlling the element capacity to bring a higher WPI value and in this way, the water situation can be improved.

#### VII. THE CONCLUSION OF THE ARTICLE

The assessment of the poverty situation of water resources has the urgent practical needs, and the accurate data analysis is of great significance to the development of the society. The optimization WPI model based on principal component analysis is presented in this paper, which has a holistic and comprehensive analysis of the data, and is more

complete in the local water resources. In this paper, the optimization of the WPI evaluation model is tested and analyzed. The experimental results verify the effectiveness of the proposed optimization model. And more, we use satisfaction index to describe the impact on citizen's life. But the model is a little bit rough and that's where we are going to work further.

#### REFERENCES

- [1] Luo Jun-gang, Xie Jian-cang, Ruan Ben-qing, Fuzzy comprehensive assessment model for water shortage risk based on entropy weight[J], Journal of Hydraulic Engineering, 2008, 39(9):1092-1097
- [2] Sun Zhi-lin, Xia Shan-shan, Xu Dan, Optimal allocation model of regional water resources[J], Journal of Zhejiang University(Engineering Science) 2009, (43)2:344-349
- [3] Han Yu-ping, Comprehensive Evaluation of Regional Water Resources Shortage Risk Based on the Maximum Entropy Principle[J], Journal of Anhui Agricultural Sciences 2011 39(1):397-399
- [4] Han Yun-hong, Tang De-shan, Li Ao-dian, Meng Zhen-zhu, Application of fuzzy comprehensive evaluation model for evaluating water resources carrying capacity in Fuyang City [J], Hydropower and Energy Science, 2015(5):26-29
- [5] Zhang Zhe, He Tai-rong, Shu Rui-qin, Liu Huan, Analysis of Water Resources Security in Chongqing City Based on WPI Model[J], Journal of Yangtze River Scientific Research Institute, 2016, 33(4):1-5
- [6] JIANG Ning, FU Qiang, SUN Yingna, Application of water poverty index in water security evaluation in north east region of China[J], Journal of Northeast Agricultural University, June 2013, 44(2), 88-91
- [7] Cao Jian-ting, Water poverty index and its application in the development and utilization of water resources[J], Water and electricity energy science, 2011(9):23-26
- [8] Liu Ren-tao, Fu Qiang, Li Han-yang, Bian Xi-long, Yu Jing-yang, Qi Shi-hua, A new method of calculating water poverty index based on entropy weight G1 method and its using research[J], Journal of Engineering of Heilongjiang University, 2014(6):232-237
- [9] Fu Xiang, Ji Chang-ming, The comprehensive evaluation of regional water resources carrying capacity and the application of principal component analysis[J], Resources and Environment in the Yangtze Basin, 1999(5):323-328
- [10] China Water Resources, <http://www.chinawater.com.cn/>
- [11] UN water scarcity map, <http://www.unep.org/dewa/vitalwater/jpg/0222-waterstress-overuse-EN.jpg>
- [12] Bao Chao, Fang Chuang-lin, Quantitative study on the relationship between urbanization and water resources utilization in the Hexi Corridor[J], Journal of Natural Resources, 2006(3):32-38
- [13] Zhang Bo, Liu Xiu-li, Dynamic ecological footprint simulation and prediction based on ARIMA Model: a case study of Gansu Province[J], acta ecologica sinica, 2011(10):83-89
- [14] Li Yan-ling, Dai Guan-zhong, Yu Mei, feature weight adjustment method based on feedback information[J], Computer Engineering, 2009(1):23-29
- [15] Liu Wu, research on the public services customer satisfaction index models[J], Northeastern University, 2008(12):54-60
- [16] Fen Wen-quan, Cai Ji-dong, Intervention analysis model and its application, Journal of Wuhan University, 1994(5):61-70