**PAPER • OPEN ACCESS**

# Evaluation and Prediction of Water Resources Utilization Structure Based on Projection Pursuit Method

To cite this article: Jie Wu 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **300** 022164

# Evaluation and Prediction of Water Resources Utilization Structure Based on Projection Pursuit Method

**Jie Wu**

School of Automation, Wuhan University of Technology, Wuhan 430070, China

qyzhang @whut.edu.cn

**Abstract**. All Based on the statistics of water use in the past five years, this paper uses the projection pursuit method to evaluate the rational use of water resources in various provinces and cities in China, and establishes corresponding prediction models for the completion of red line indicators. Perform preliminary statistical analysis on the data, select relevant indicators to establish a projection pursuit evaluation model, and use the genetic algorithm to solve the problem, display the projection values of each province and city on the map, and point out the existing problems. Four of the red line indicators were exponentially fitted and predicted. The final water quality compliance rate did not reach the red line target value of 2030, and the deviation from the target was 13.68%.

## 1. Introduction

At this stage, China's water shortage, serious water pollution and deterioration of water ecology have become important factors restricting China's economic development. Strengthening the rational planning of domestic water resources use and improving the efficiency of water resources development and utilization are our urgent development goals. How to evaluate the use of water resources in various provinces and cities has always been a difficult and hot topic of current research. In recent years, mathematical models of mathematical statistics have been widely used to analyze the impact of water resources. The projection pursuit [1, 2] model is an evaluation model that is simple in calculation and can project multiple influencing factors into a single integrated element. It not only effectively avoids the subjective influence of the previous evaluation method by the expert scoring method, but also selects different correlation coefficient solving methods according to the number of sample points, which effectively improves the objectivity of the final result. This study combines the State Council's “Three Red Lines”[3] standards to analyze the status quo and existing problems of China's water resources development and utilization, with a view to analyzing and contrasting the indicators with the national 2020 and 2030 targets, and rationally developing water resources. Provide a reference basis.

## 2. Model establishment

### *2.1. Analysis of the problem*

According to national special planning and research, China Water Resources Bulletin, national water census and national related industry statistics, the status quo of China's water resources development

IOP Conf. Series: Earth and Environmental Science **300** (2019) 022164 doi:10.1088/1755-1315/300/2/022164

and utilization and the main problems are analyzed, and the deviations from the 2020 and 2030 goals proposed by the state are compared and compared. First, sort out the relevant indicators of the "three red lines" proposed by the state, as shown in Table 1.

**Table 1.** Three red line indicator requirements

| | **Total national Water consumption** | **Water consumption of ten thousand yuan of industrial added value** | **Effective utilization coefficient of farmland irrigation water** | **Water quality compliance rate** |
|---|---|---|---|---|
| 2020 | ≤6700 | 65 | 0.55 | 0.8 |
| 2030 | ≤7000 | 40 | 0.6 | 0.95 |

Firstly, statistical analysis is carried out on some indicators in the 2012-2016 Water Resources Bulletin[4] to observe the trends in five years and conduct a preliminary analysis of the utilization of water resources in China. After that, the indicators are sorted and established using the mind map method. The projection pursuit evaluation model evaluates the water use situation of 31 provinces and cities in China and marks it on the map. Finally, according to the data of 2012-2016, the four indicators of the above table are predicted and completed for 2020 and 2030. Judgment of the annual compliance status.

### *2.2. Model Hypothesis*

**Hypothesis 1:** Unify national industrial and agricultural water prices, ignoring the difference in pricing between different provinces and setting the price gradient.
**Hypothesis 2:** China did not implement major water conservancy projects or make large allocations of water resources before 2030.
**Hypothesis 3:** All data sources used in the text are true and reliable.
**Hypothesis 4:** The total amount of precipitation in China before 2030 will not be disturbed by some sudden factors, and thus there will be large fluctuations.

### *2.3. Establishment of evaluation model based on projection pursuit*

The projection pursuit model is used to analyze and evaluate the water resources utilization of 31 provinces and cities in China, and the use efficiency of water resources is converted into the value of the objective function value in the projection. The larger the projection value, the better the development and utilization of water resources. The establishment of a specific model is as follows.

*2.3.1. Selection of evaluation indicators.* Through the integration of data in the Water Resources Bulletin, it is proposed to evaluate the use of water resources in various provinces and municipalities in combination with the rationality of ecological environment, the rationality of water use structure and the rationality of water use. This study will use 10 indicators [5, 6] to evaluate the status quo of water resources development and utilization in various provinces and cities, as shown in Figure 1.

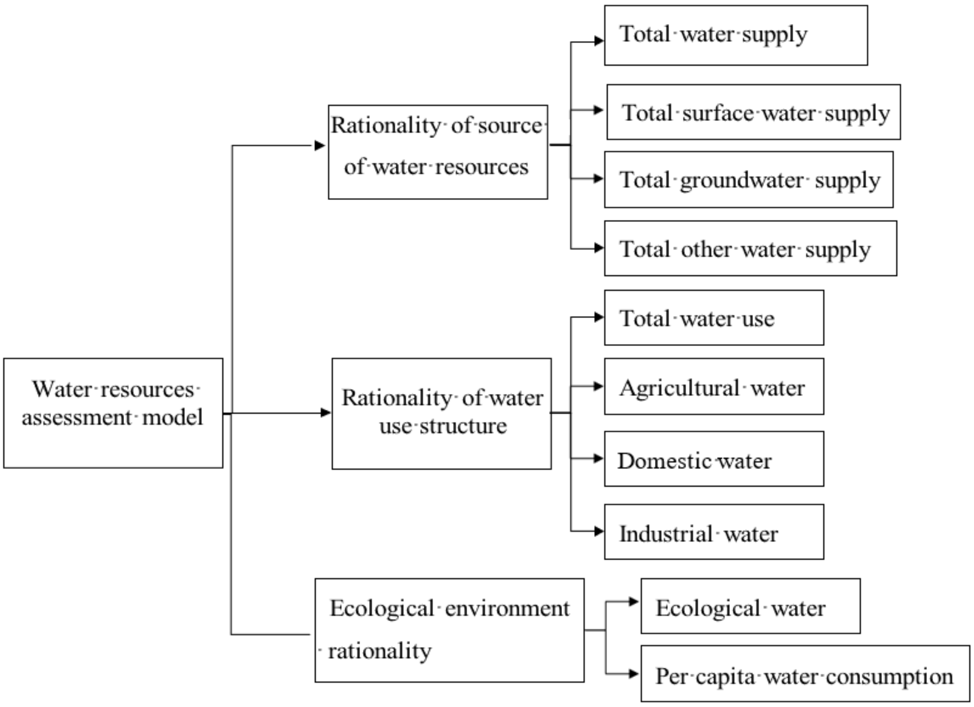


**Figure 1.** Structure diagram of water resources evaluation model

*2.3.2. Normalization of evaluation index values.*

P data is extracted from the bulletin for overall evaluation and analysis. The hypothetical solution research set is:

$$\{x^*(i, j) \mid i = 1, 2, ..., n; j = 1, 2, ..., p\}$$

Where x* (i, j) represents the jth evaluation indicator for the year i. In order to eliminate the dimension of each evaluation index and unify the range of variation of each evaluation index, the maximal and superior evaluation index is subjected to extreme value normalization.

$$x(i, j) = \frac{x^*(i, j) - x_{\min}(i, j)}{x_{\max}(j) - x_{\min}(j)} \tag{1}$$

The evaluation index for the smaller and better type can be subjected to the extreme value normalization as follows.

$$x(i, j) = \frac{x_{\max}(j) - x^*(i, j)}{x_{\max}(j) - x_{\min}(j)} \tag{2}$$

Where $x_{\max}(j), x_{\min}(j)$ are the minimum and maximum values of the jth indicator in the data of different years. After normalization, x* (i, j) can be converted into an evaluation index in the interval [0,1].

*2.3.3. Construction projection index function.*

The selected p-dimensional data $\{x*(i, j) i = 1, 2, ..., 5; j = 1, 2, ..., 11\}$ combined into a unit length vector $a = (a(1), a(2), a(3), ..., a(p))$ is the one-dimensional projection value of the direction.

$$z(i) = \sum_{j=1}^{p} a(j)x(i,j) \tag{3}$$

According to the one-dimensional scatter diagram of z(i), the local projection points are as dense as possible, and it is better to condense into a plurality of clusters; and the projection points on the whole are scattered as much as possible, and the construction index function is as follows:

$Q(a) = S_z D_z$

Where $S_z$ is the standard deviation of the projection value z(i) and $D_z$ is the local density of the projection value z(i):

$$S_z = \sqrt{\sum_{i=1}^{5} \frac{(z(i)-E)^2}{n-1}} \tag{4}$$

$$D_z = \sum_{i=1}^{5}\sum_{j=1}^{p}(R-r_{ij})u(R-r_{ij}) \tag{5}$$

$E_z$ is the mean of the system z(i); R is the window radius of the local density, which is 0.1SZ; *u* (R r*ij)* is a unit step function, when $u(R-r_{ij}) \geq 0$, its function value is 1. When $u(R-r_{ij}) \leq 0$, its function value is 0. $r_{ij}$ is the distance and the expression looks like this:

$$r_{ij} = |z(i) - z(j)|$$

*2.3.4. Determining the projection objective function.* The optimal projection direction is selected to maximize the projection direction of a certain type of feature structure of high-dimensional data. The optimal projection direction can be estimated by solving the problem of maximizing the projection index function.

Maximize the objective function:

$$\begin{aligned} &max: Q(a) = S_z D_z \\ &s.t.\sum_{j=1}^{p} a^2(j) = 1(2-6) \end{aligned} \tag{6}$$

*2.4. Establishment of red line indicator prediction model*

*2.4.1. Determination of the fitting function.* In view of the small amount of data, the gray prediction or the time series prediction effect is poor. When the four indicators defined by the red line are predicted in the 2012-2016 data, and the results are compared with the future goals, the exponential function is selected. Fit the four indicators. The prediction function of two of the water consumptions is as follows:

$$y_i = a_i e^{b_i x_i}$$

In the calculation of the effective utilization coefficient of two farmland irrigation waters and the compliance rate of the water functional zones, in order to make the result not exceed 1, the prediction function is as follows:

$$y_i = 1 - a_i e^{b_i x_i}$$

$y_i$ represents the i-th indicator, $a_i$ and $b_i$ represents the coefficient corresponding to the i-th indicator, $x_i$ indicates the year corresponding to the indicator. The prediction function finally fits the change rule of the index value with the increase of the year.

*2.4.2. Predictive bias analysis.* Bring 2020 and 2030 into the fitting function and calculate the deviation from the national red line:

$$D_{ij} = \frac{y_{ij} - A_{ij}}{A_{ij}} \; j = \{2012, 2013\}$$

Indicates the deviation from the target of the indicator for the first indicator, indicating the red line target for the indicator year.

## 3. model solving

### *3.1. Solving the Projection Pursuit Model Based on Genetic Algorithm*

*3.1.1. Algorithm Implementation.* Based on the literature [7-10], the projection pursuit model is substituted into the genetic algorithm for solving. The solution to the problem is encoded into a chromosome and then iteratively optimized. The final algorithm converges to the best chromosome, which is the optimal solution to the problem. specific
Proceed as follows:
*3.1.2. Statistics of data results.* Perform multiple calculations on the projection value, and select the value with the highest fitness for the next analysis.
The local optimum situation occurs. The convergence result of the algorithm is shown in Figure 2.

After many calculations, the highest fitness value is selected, that is, the projection value calculated by the right image is used as the final solution result. It can be seen from the right graph that the number of iterations reaches 1600 times and gradually converges.

Combined with the 2016 water resources statistics, 31 provincial and municipal evaluation projection values were calculated and ranked. The optimal and worst five projection values for the provinces and cities are shown in Table 2.

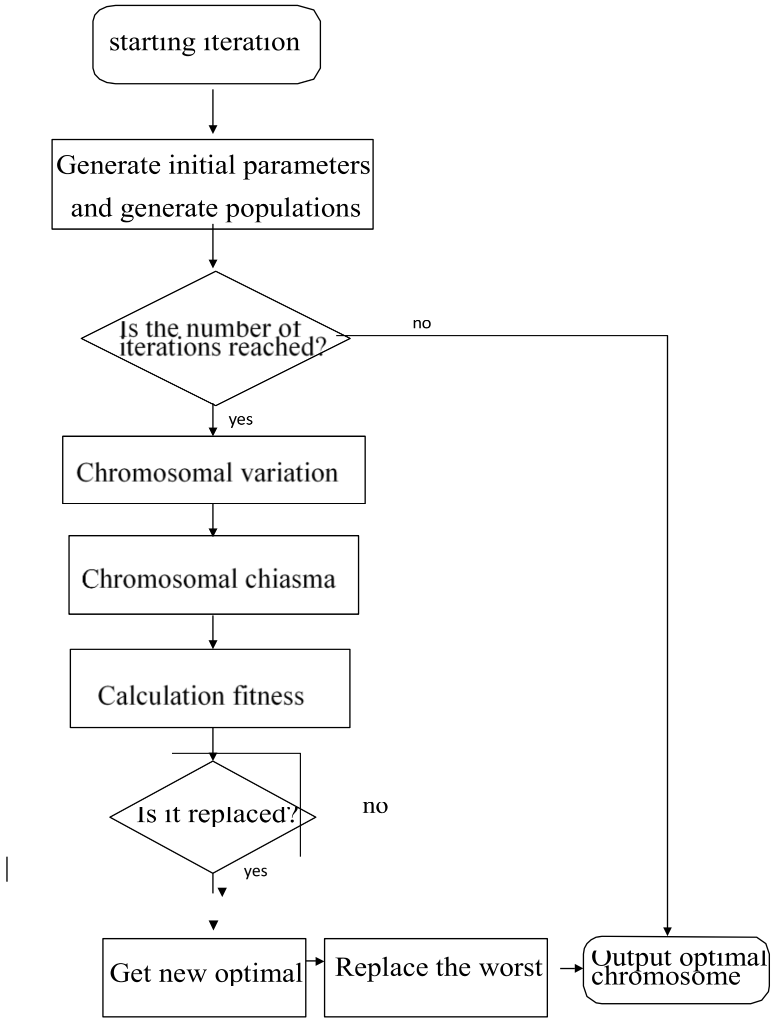


**Figure 2.** genetic algorithm solution flow diagram

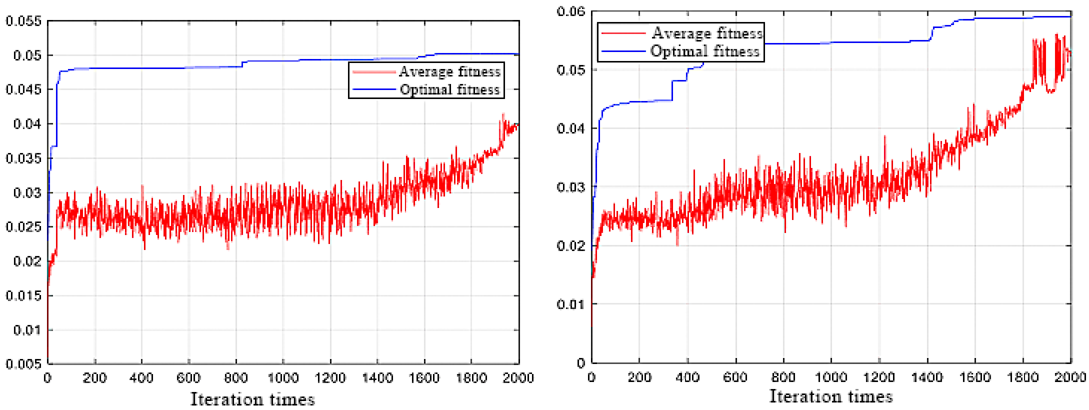


**Figure 3.** Convergence results of different batch calculations

**Table 2.** Projection of some provinces and cities

| Serial number | Province | Projection value | Serial number | Province | Projection value |
|---|---|---|---|---|---|
| 1 | Guangdong | 313 055.0737 | 27 | Gansu | 22 664.52 |
| 2 | Jiangsu | 206 733.3468 | 28 | Chongqing | 15 018.66 |
| 3 | Shandong | 169 316.1766 | 29 | Ningxia | 12 195.85 |
| 4 | Zhejiang | 143 824.9805 | 30 | Qinghai | 9 477.207 |
| 5 | Henan | 1342 80.0313 | 31 | Tibet | 2 955.379 |

### *3.2. Solution of the prediction model*

#### *3.2.1. Forecast of national water consumption and water consumption per 10,000 yuan of industrial added value.*

Considering that the national water consumption is composed of industrial water, agricultural water, domestic water and artificial ecological water consumption in the country, and the national water consumption data within the five-year span is relatively discrete and has strong volatility, two prediction methods [11] are adopted. Those with better predictions are the final predictors of national water use. The two prediction methods are respectively the method using the sum of the four indicators and the individual prediction methods, see Figure 4 and Figure 5.

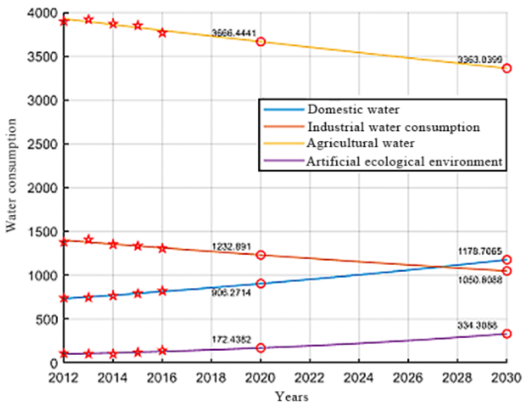


**Figure 4.** Four indicator prediction chart

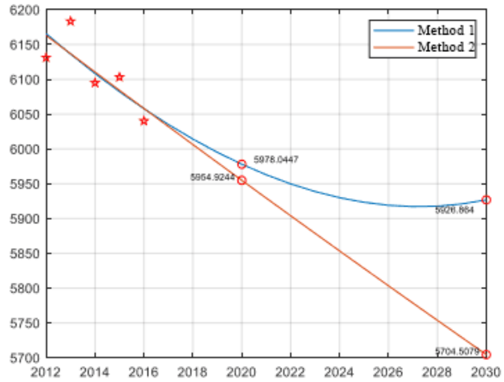


**Figure 5.** Comparison of two methods

The results of Method 1 were found to be more reasonable by $R^2$ test. The fitting function of the four indicators is:

$$y_1 = 7.93\times10^{-21}e^{0.03x_1}$$
$$y_2 = 1.29\times10^{17}e^{-0.02x_2}$$
$$y_3 = 1.38\times10^{11}e^{-0.01x_i}$$
$$y_4 = 1.44\times10^{-56}e^{0.07x_i}$$

The R2 test value is 0.956 and the fit curve formula is as follows:

$$y = 5.79\times10^{60}e^{-0.667x}$$

#### *3.2.2. Forecast of industrial added value, water quality compliance rate and agricultural irrigation coefficient of 10,000 yuan.*

In view of the prediction of farmland irrigation coefficient, the scatter trend is close to linear, so linear fit and the index fits two options. The linear fit R2 test value is 0.990 and the exponential fit R2test value is 0.998. Finally, the exponential type is used for fitting, as shown in Figure 6.

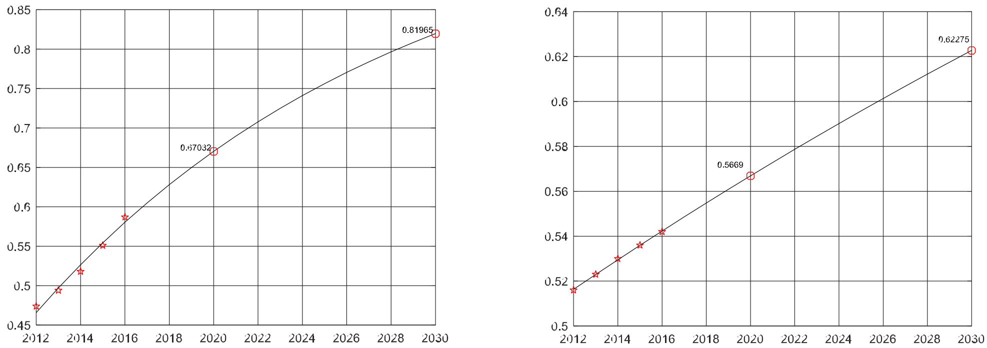


**Figure 6.** Water quality compliance rate (left) farm irrigation coefficient fitting curve

The three fitting curve coefficients and the $R^2$ test values are shown in Table 3.

**Table 3.** Fitting curve coefficient table

| Index | $R^2$ | Coefficient a | Coefficient b |
|---|---|---|---|
| water qualification rate | 0.976 | $2.737x10^{52}$ | -0.06 |
| Farmland irrigation coefficient | 0.999 | $5.6\ x10^{11}$ | -0.014 |

Therefore, the water quality compliance rate formula is:

$$y = 1 - 2.737 \times 10^{52} e^{-0.06x}$$

The farmland irrigation coefficient formula can be expressed as:

$$y = 1 - 5.6 \times 10^{11} e^{-0.06x}$$

## 4. Result analysis

### *4.1. Analysis of national water resources utilization*

The projection evaluation values of 31 provinces and municipalities are marked on the map. According to the projection calculation, the larger the projection index and the darker the color, the higher the utilization efficiency of water resources as shown in Figure 7.


**Figure 7.** evaluation chart of the development and utilization of water resources in mainland China

Among them, the projection value calculated by the projection is larger, and the darker the color, the higher the utilization efficiency of water resources.

At present, the most reasonable province for water resources utilization in China is Guangdong Province. The water resources utilization in the southeast coastal areas is relatively reasonable, while the utilization of water resources in North China, Northwest China, Southwest China (except Sichuan) and Northeast China needs to be strengthened. It can be seen from the projection values of different indicators that the water use structure has the greatest impact on the overall reasonable degree of evaluation. Industrial water, agricultural water and per capita water consumption are the three core indicators for assessing the rationality of water resources.

### *4.2. Deviation and comparison analysis of forecasting targets*

The forecast results of various indicators are compared with the target values of the red line in 2020 and 2030. The results are shown in Table 4.

**Table 4.** Forecast results and target statistics

| Index | 2020 forecast | 2020 goal | 2030 forecast | 2030 goal |
|---|---|---|---|---|
| Total national water consumption | 5978.04 | <=6700 | 5926.86 | <=7000 |
| Water consumption of ten thousand yuan of industrial added value | 40.72 | 65 | 20.75 | 40 |
| Farmland irrigation coefficient | 0.57 | 0.55 | 0.62 | 0.6 |
| Water qualification rate | 0.67 | 0.8 | 0.82 | 0.95 |

Through statistical analysis, it is found that the total water consumption in the country, the water consumption per 10,000yuan of industrial added value, and the effective coefficient of farmland irrigation can all meet the requirements of the national red line. The water quality compliance rate prediction results can be successfully achieved in 2020, but there is a gap between the 2030 results and the forecast, which is 13.68% deviation from the target value.

## 5. Summary

Using the projection pursuit method to evaluate the rationality of water resources utilization in 31 provinces across the country, the water quality compliance rate prediction results can be successfully achieved in 2020, but there is a gap between the results and predictions in 2030, and the deviation from the target value is 13.68%, and the level Compared with the evaluation methods such as analysis method, the subjective factor evaluation brought by the expert scoring is avoided, and the overall evaluation result has better scientific and objective. For the adjustment of water resources utilization, the state can refer to the index weights of the indicators in this model, and prioritize the indicators with higher weights. The solution to the problem can be used to evaluate and guide the reasonable amount of water used in each province and city, and provide a certain reference value for the national water resources target planning.

## Acknowledgments

This paper is Supported by "the Fundamental Research Funds for the Central Universities" (ID: 2018-186811012-A1-01).

## References

[1] Xu Qiushi and Ren Futian, "Interchange-based interchange selection method based on projection pursuit model," Journal of Beijing Polytechnic University, pp. 193-196, 2007.
[2] Yin Faqiang, Li Dan, Song Xuyan, and He Jiawen, "Projection Tracking Regression (PPR) Method for Establishing Mathematical Model of Cigarette Tar Prediction," Tobacco Science and Technology, pp. 15-18+25, 2009.

[3] Qi Dianbin, "Water resources management must strictly observe the "three red lines" - Interview with Hu Siyi, Vice Minister of Water Resources," Decision and Information, pp. 40-42, 2012.
[4] "Appendix 2012 China Water Resources Bulletin (Excerpt),".
[5] W. Guoying, "Research on Multi-objective Optimization Linear Programming Problem of Power Generation under Carbon Emission Trading," in Research on Multi-objective Optimization Linear Programming Problem of Power Generation under Carbon Emission Trading, Guangzhou, Guangdong, China, 2017, p. 5.
[6] M. Zhu, "Application of Interval Fuzzy Multi-objective Strategy Model in Water Resources Optimal Allocation," in Application of Interval Fuzzy Multi-objective Strategy Model in Water Resources Optimal Allocation, Shanghai. China, 2017, p. 5.
[7] Ge Jike, Qiu Yuhui, Wu Chunming, and Pu Guolin, "Review of Genetic Algorithm Research," Computer Application Research, pp. 2911-2916, 2008.
[8] Ji Genlin, "Overview of Genetic Algorithms," Computer Applications and Software, pp. 69-73, 2004.
[9] Ma Yongjie and Yun Wenxia, "Progress in Genetic Algorithms," Computer Application Research, vol. 29, pp. 1201-1206+1210, 2012.
[10] Qi Weimin, Xi Yugeng, Chai Tianyou, "Overview of Genetic Algorithms," Control Theory and Applications, pp. 697-708, 1996.
[11] Liu Yanyan, "Recycling Water Life Cycle Comprehensive Value Model and Quantification,": Xi'an University of Architecture and Technology, 2011.
[12] Gao Xiaodong, Yan Leihua, Liu Heng, Jiang Yulei, and Tong Kun, "Analysis of the accessibility of water consumption in the "Twelfth Five-Year Plan" national industrial added value," Water Conservancy Economy, vol. 31, pp. 1-5+ 75, 2013