The study about the water strategy of America which is based on the data analysis of the MATLAB software

Yilin Li

Department of Power Engineering, North China Electric Power University, Baoding 071000, China

email:18932696516@163.com

Keywords: MATLAB software, water strategy, data analysis, statistic stimulation model

Abstract. This article explains and demonstrates the study about the water strategy of American which is based on the data analysis of the MATLAB software. This society is an era which is consisted of data. The formulation of some policies and the reform of some systems need the help of data simulation. In this article, the problem of the distribution and allocation of the water resources in America is discussed in a professional manner. In the procedure of solving problems of America, the correspondence analysis is used to classify the whole American into five main parts.The standard of the classification is the amount of water resources that they have. After that, the neural network model is used to predict water consumption. That's the new way to set the water resources policy by using the MATLAB other than ill-considered discuss details or implement test.

Introduction

This society is an era which is consisted of data, the formulation of some policies and the reform of some system need the help of data simulation. Strategy of the water resources is a hot problem that need solving in the scope of the world. For example, this century is faced with the situation of the distribution of water resources and water amount supply in the United States.

So the main task of this paper is to comprehensively consider several factors, including the transportation, the storage, the desalination of water and the water protection and set an optimal model about the best strategy of the United States in different areas and different time. In the procedure of this work, the MATLAB software is used to simulate and predict the data. The method which is linked to the mathematics model and the information science is convenient and not ill-considered. So it worth a deeper and wider study.

Research contents

General assumptions. (1)The elements that we take into consideration play a vital role in the evaluation.(2)Elements that we ignore do not influence the ranking.(3)The data that we have collected is sufficient and accurate.

Collect data. The data is founded in the official website of the American, which include the original research data of several years.

Table1 Sample of data for the 56 states[1]

	Alabama	Alaska	Arizona	……
Total area	135767	1723337	295234	……
Water total area	4597	245383	1026	……
Inland water area	2740	49997	1026	……
Coastal water area	1340	67647	0	……
Great lakes water area	0	0	0	……
Territorial water area	516	127739	0	……

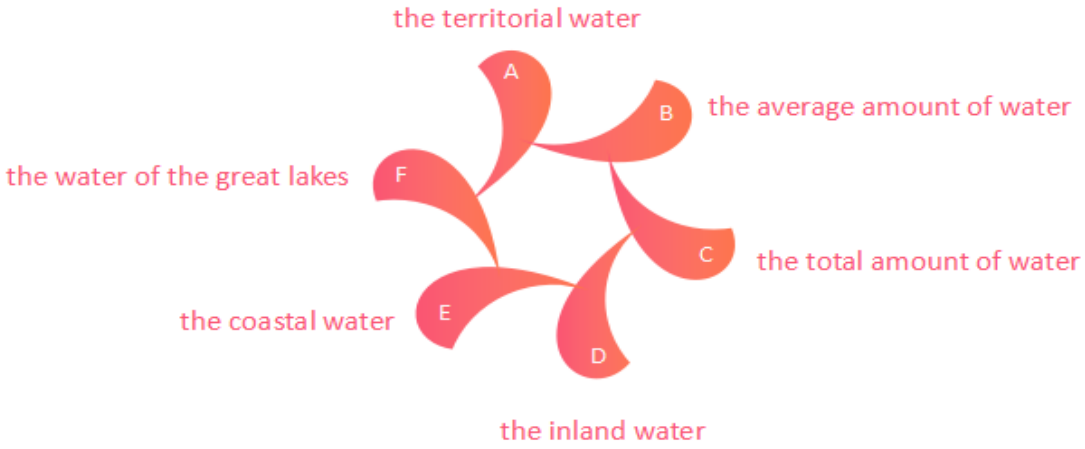

Fig.1. The map of the six evaluation norms

Specify Evaluation Norms. As for the evaluation standard for the water resources of the United States, there are mainly six aspects count: the average amount of water, the total amount of water, the inland water, the coastal water, the water of the great lakes and the territorial water. The 52 states and 4 islands of the United States are numbered from 1to 56 in the procedure of solving this problem.

Correspondent analysis for classification. Adopt the method of correspondent analysis to classify the 56 different places into five categories.This model is chosen for classification of the 56 states for several reasons. When try to obtain the weight of the six aspects of the evaluation norms, the subjective judgment is ill-considered. So we choose the correspondent analysis as the way to combine the weighting coefficients of the six aspects to evaluate the tremendous amount of water in different areas.

The correspondent analysis is developed on the basis of the R type and Q-type factor statistical analysis[2], which is also known as R-Q type factor analysis. The correspondent analysis decreases the number of the factors which influence the evaluation of the degree of moisture content and grasp the relationship between these things.

In this article, the method of correspondent analysis is used to divide the 56 states into the main 5 parts. The richness of water resources is different in each part. We explain the concrete steps as follow. There have 56 samples, each sample has six indicators. We perform data transformation at first.

$$B = \left(b_{ij}\right)_{56\times 6} \tag{1}$$

$$b_{ij} = \frac{p_{ij} - p_{i.}p_{.j}}{\sqrt{p_{i.}p_{.j}}} = \frac{a_{ij} - a_{i.}a_{.j}/T}{\sqrt{a_{i.}a_{.j}}}, i = 1,2,\cdots,5,6;\ j = 1,2,\cdots,6 \tag{2}$$

$$a_{i.} = \sum_{j=1}^{6} a_{ij}, a_{.j} = \sum_{i=1}^{56} a_{ij} \tag{3}$$

$$S_R = B^T B, S_Q = BB^T \tag{4}$$

Where the S_R stands for the total moment of inertia and the χ^2 stands for statistic. The weighted square distance between the i and j is expressed as follow:

$$D^2(k,l) = \sum_{j=1}^{6}\left(\frac{p_{kj}}{p_k} - \frac{p_{lj}}{p_l}\right)/\, p_{.j} = \left(R_k - R_l\right)^T D_c\left(R_k - R_l\right) \tag{5}$$

The characteristic value is set as

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \geq \lambda_m > 0 \tag{6}$$

It can be designed that the standardization of corresponding feature vectors is

$$\eta_1, \eta_2 \cdots \eta_m \tag{7}$$

It can be selected that the cumulative contribution of which is higher than 98% of the items[3]:

$$\frac{\lambda_1 + \lambda_2 + \lambda_3 + \cdots \lambda_l}{\lambda_1 + \lambda_2 + \cdots \lambda_l + \cdots \lambda_m} \geq 0.98 \tag{8}$$

Finally, we calculate the coordinate G of the line contour and the coordinate F of the column outline.

$$F = \begin{bmatrix} -0.214 & -1.01 & -0.0982 & -2.61e-6 & -5.76e-6 & 5.64e-17 \\ -0.214 & -1.01 & -0.0982 & -2.5e-6 & 5.04e-7 & 6.46e-17 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -0.493 & 0.66 & -0.672 & -9.48e-6 & -2.01e-6 & 8.65e-17 \\ -0.485 & 0.611 & -0.607 & -8.48e-6 & -2.58e-6 & -2.79e-17 \end{bmatrix} \tag{9}$$

$$G = \begin{bmatrix} -1.07e-6 & 4.77e-7 & 4.02e-6 & -3.48e-5 & 3.19e-8 & 9.38e-17 \\ -0.276 & -0.78 & -0.0479 & 3.48e-5 & -4.6e-8 & 9.38e-17 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -0.637 & 0.515 & -0.332 & 3.48e-5 & -4.6e-8 & 9.38e-17 \\ -3.97e-4 & 1.81e-4 & -4.04e-5 & 1.92e-4 & 7.06e-4 & 9.38e-17 \end{bmatrix} \tag{10}$$

Table 2. The cumulative contribution rate of the most important three factors

eigenvalue	Chi-square statistic	contribution rate rate	cumulative contribution
0.417	582704.16	0.666	0.666
0.15	208987.82	0.24	0.90
0.06	83279.82	0.10	1.00

The scatter diagram of the ranks of the coordinates can show the results of the debate as follow:

Table 3. The category results of the 56 states in American

Category	Serial number
E	52、53、55、54、56、12
C	7、21、31、5、40、8、28、2、22
B	20、19、41、25、1、11、30、34、44
AF	33、14、39、15、24
D	50、36、23

Neural network model to predict water consumption. Artificial neural networks has a strong learning and adaptive capacity[4]. We use this advantage to build a nonlinear mapping relationship of the input and output data. That's our water allocation decision model.The flow charting of this BP model is as follow:

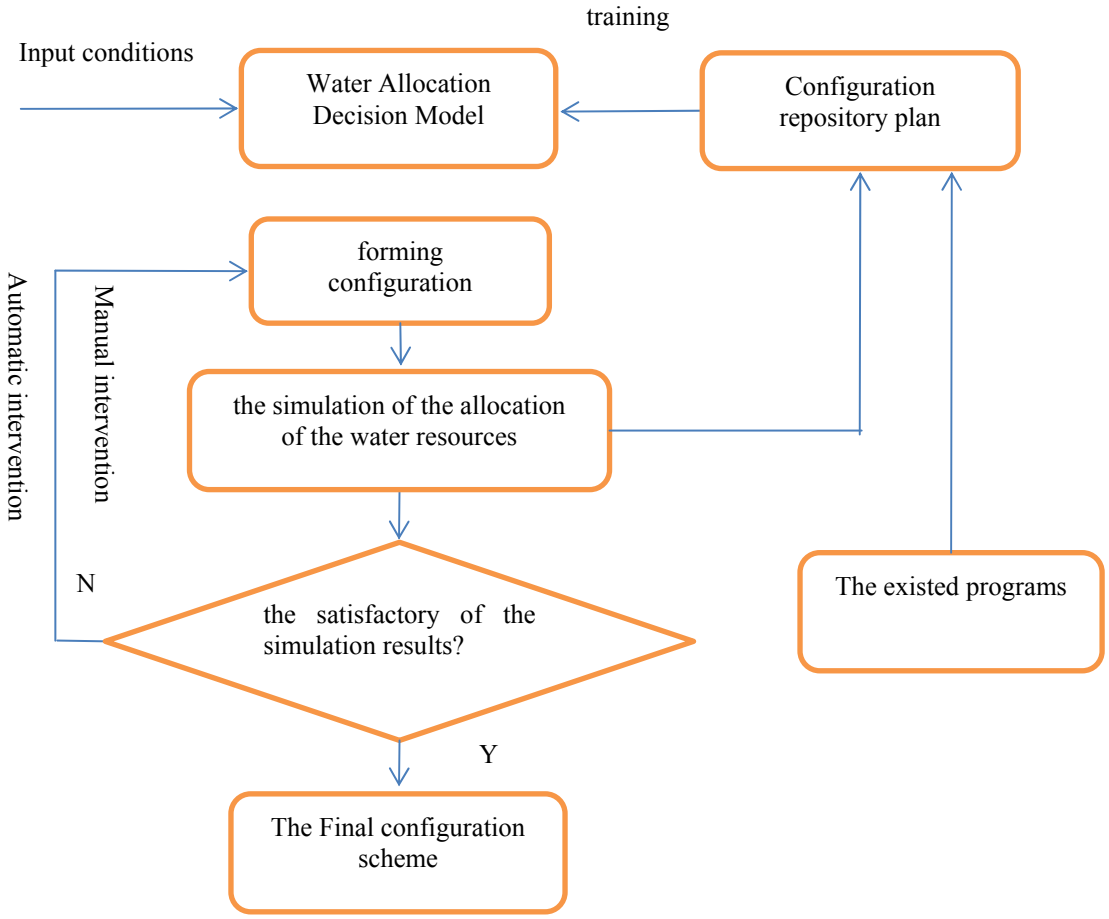

Fig.2. The flow charming of the neural network model

Conclusion and analysis of the results.

After the prediction and screening of the neural network model, it is not hard to evaluate the amount of water that are needed in ten years of the five parts in American. The water resources allocation simulation model, with the establishment of physical mechanism, on one hand can simulate the changing process within a region, including changes in water quality, water loss and so on. On the other hand, several key parts of the data can be used to set a evaluation model to evaluate the water quality and other water resources utilization. The model can be established to consider the problems such as social, economic,ecological and water use cases and so on.

To extend the model of this article, the establish of a set of evaluation index systems is essential. Using the application of AHP (Analytic Hierarchy Process, AHP) [5]to calculate the weight of each index, and then apply fuzzy comprehensive evaluation of program ratings to review the results of the simulation of water allocation.

Acknowledgement

This research has been supported by “A research on a new energy-efficient indoor radiator” (201510054124), which belongs to the innovation and Entrepreneurship Program for College Students in North China Electric Power University.

References

[1] Information on http://www.census.gov/

[2] H.J. Ji, Y.Z. Zhu, X.S. Wu, Correspondence cluster analysis and its application in exploration geochemistry, J. Geoehem. Explore. 55 (1995) 137.

[3] K.R. Joreskogkg. Method in geomathematics, Geological factor analysis[M].(1976)

[4] Information on http://www.mcm.edu.cn/

[5] G.F. Peng, S.C. Li, M.K. Sheng, Study on government performance evaluation index heavy weight determined by analytic hierarchy process, China soft science, 2004-06-28.