THE INPUT AND OUTPUT RELATIONSHIP OF WATER RESOURCE IN JILIN FROM 2004 TO 2017

by

Bo-Yang SUN and Xiao-Hua YANG *

State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing, China

Original scientific paper
https://doi.org/10.2298/TSCI2004337S

Jilin is an important grain producing area and heavy industry base in China. With the economic development and improvement of people's livelihood, water resource issues have been becoming increasingly prominent, Jilin Province comprehensive utilization of water resources is not efficient. The traditional method of water consumption per 10000 Yuan of output value cannot objectively reflect the water-use efficiency of different local industrial structures. Therefore, this paper builds the data envelopment analysis model based on the input-output relationship. In order to meet the requirements of data envelopment analysis model for data analysis, this paper introduces the principal component analysis method, and reduces the 16 water resources input indicators and 14 water resources output indicators of Jilin Province from 2004 to 2017 to 3 input principal components and 2 output principal components. The multi-model data envelopment analysis method is used to analyze the water-use efficiency of Jilin Province, and the results show that with the rapid economic growth, the water resources efficiency in 2010-2013 was relatively poor, there was a waste of water resources, and the management technology was backward, but with the deepening of the industrial transformation and upgrading, farming modernization and revitalizing the strategy of the old industrial base in Northeast China, the water-use efficiency and the water resources carrying capacity of Jilin Province has been improved.

Key words: *water use efficiency, principal component analysis, data envelopment analysis, input-output analysis*

Introduction

Jilin is an important grain producing area in China and a heavy industrial base. With the advancement of the agricultural modernization and the revitalization of the old industrial bases in Northeast China, the pressure on water resources in the region has been continuing to increase, and the contradiction between supply and demand of water resources will become increasingly prominent. Water resources are essential for daily life of residents and important production materials in economic production activities. Therefore, it is extremely important to study the input-output relationship between water resources and social economy, and to improve water use efficiency for policy advice. These measures have important strategic significance for Jilin's future economic development, industrial transformation and upgrading, and improvement of people livelihood [1, 2].

* Corresponding author, e-mail: xiaohuayang@bnu.edu.cn

Data envelopment analysis (DEA) is a non-parametric technical efficiency analysis method and a new field of cross research in operations research, management science, and mathematical economics. This method was first proposed by Charnes, *et al.* [3] in the United States in 1978. Because the DEA method has a wide range of applications and is suitable for the special advantages of multiple inputs and multiple outputs, it is widely used in agriculture, environment, macroeconomics and other fields [4, 5]. Therefore, this paper selects the DEA method to compare and analyze the water resources input and output efficiency of Jilin Province from 2004 to 2017. The data source of this paper contains 16 input indicators and 14 output indicators, as shown in tab. 1. Excessively high dimensionality of raw data can lead to complex calculations in the DEA model and distortion of results [6]. By introducing the principal component analysis (PCA), not only can the data dimension reduction meet the algorithm requirements, but also the influence of the dimension and original data correlation on the results can be eliminated, and the analysis result is more accurate.

Table 1. Water resources input and output indicators

Symbol	Input indicator	Symbol	Output indicator
X1	Agricultural water	Y1	Total GDP
X2	Industrial water	Y2	Agricultural value added
X3	Water for live	Y3	Industrial added value
X4	Ecosystem	Y4	Financial added value
X5	Total water consumption	Y5	Real estate added value
X6	Per capita water	Y6	Construction industry added value
X7	Total water supply	Y7	Wholesale and retail value added
X8	Surface water supply	Y8	Warehousing logistics added value
X9	Groundwater supply	Y9	Accommodation and food added value
X10	Other water supply	Y10	Per capita GDP
X11	Total water resources	Y11	Total employment population
X12	Surface water resources	Y12	Total wastewater discharge
X13	Groundwater resources	Y13	Chemical oxygen demand
X14	Repeated calculation	Y14	Ammonia nitrogen
X15	Per capita water resources		
X16	Water supply pipe length		

Data and methods introduced

Jilin Province water resources input and output data

This paper selects the water supply and demand data of Jilin Province from 2004 to 2017 as the input index of water resources, and selects the gross domestic product (GDP) of the industry, the employment population and the amount of pollutants entering the river as output indicators. The data source is the National Bureau of Statistics of China (China Statistical Yearbook). The water consumption by industry is shown in fig. 1.

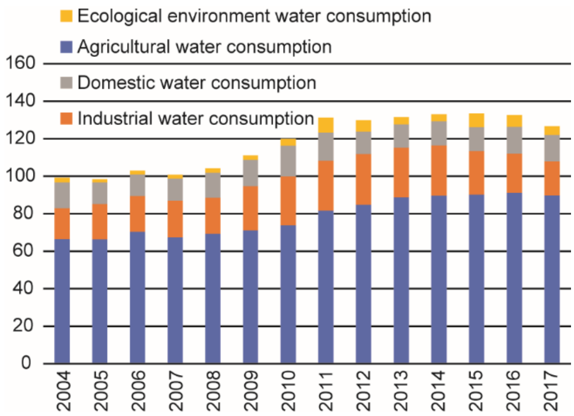

Figure 1. Water consumption data

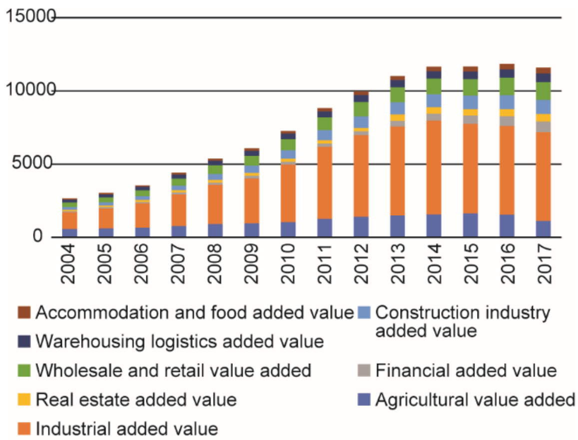

Figure 2. The GDP data

As can be seen from fig. 1, Jilin Province, as a major agricultural province and an important industrial base, has a large proportion of agricultural water and industrial water, of which agricultural water consumption accounts for between 60% and 70%. Industrial water consumption accounts for between 16% and 21%. In the last four years, the proportion of agricultural water increased slightly, the proportion of industrial water decreased, and the total water consumption has decreased year by year. The GDP statistics of Jilin Province by industry are shown in fig. 2. As can be seen from fig. 2, the top three sources of GDP contribution in Jilin Province are industrial, agricultural and construction industries. The added value of industry shows the growth trend before 2014 and slight decrease trend since 2014. The added value of agriculture has a slow increase before 2015 and has a downward trend after 2015. The added value of the construction industry increased before 2015, and declined after 2015. According to the above analysis, the input-output indicators of the water resources in Jilin Province has undergone a significant trend change between 2014 and 2015. This phenomenon is related to the transformation and upgrading strategy of local industries. In order to further analyze the input-output relationship of water resources in Jilin Province, it is necessary to analyze the water use efficiency of Jilin Province year by year. The water pollutants in Jilin Province were relatively stable before 2011, and increased significantly between 2011 and 2015, and decreased rapidly after 2015. This is related to the effective local pollution control methods, and is also closely related to industrial transformation and upgrading.

The DEA method

The principle of this method is to treat each object of evaluation as a decision-making unit (DMU) by keeping the output or input of DMU unchanged, and determining the relatively effective production frontier, and then projecting each DMU onto the production frontier, their relative efficiency is, therefore, evaluated by comparing the deviation from the DEA production frontier [7, 8].

The DEA has a variety of measurement models, including CCR, BCC, ST, FG, *etc.* [9]. Suppose there are n DMU in a system, each of them has m input indicators (x_{m1}, x_{m2}, …, x_{mj}) and s output indicators (y_{s1}, y_{s2}, … y_{sj}), and the DEA model is:

$$\min[\theta-\varepsilon(e_1^T S^+ + e_2^T S^-)]$$
$$s.t\begin{cases}\sum_{j=1}^{n}\lambda_j X_j + S^- - \theta X_{j0} = 0\\ \sum_{j=1}^{n}\lambda_j Y_j - S^+ - Y_{j0} = 0\\ \lambda_j \geq 0, \left(j=1,2,\ldots,n\right)\\ S^+ \geq 0\\ S^- \geq 0\\ 0\leq\theta\leq 1\end{cases} \tag{1}$$

where θ is the effective value, ε – the non-Archimedean infinitesimal, S^+ – the slack variables of m input, S^- – the slack variables of s output, λ_j – the weight vector of input and output, n – the number of DMU, and $e_1^T = (1, 1,…,1)1 \times$ m, $e_2^T = (1, 1,…,1)1 \times$ s.

The economic meaning is:

– if $\theta = 1$, while $S^+ = S^- = 0$, then DEA is effective, and
– if $\theta < 1$, then DEA is invalid; when $\Sigma_{j=1}^{n}\lambda = 0$, technical efficiency, otherwise technical inefficiency.

Here $K = 1/(\theta\Sigma_{j=1}^{n}\lambda_j)$, when $K = 1$, scale efficiency, when $K < 1$, increasing returns to scale, when $K > 1$, decreasing return to scale [10]. Wu and He [11] established a variational principle for economic activity by the semi-inverse method [12-14]. The optimization given in eq. (1) is difficult to be solved by the variational method or the least square method, so the PCA method is adopted in this paper.

The PCA method

The PCA is a multivariate statistical method that reduces multiple variables into a few principal components (*i. e.*, integrated variables) by dimensionality reduction techniques. Those principal components can reflect most of the information of the original variables, and represent the linearity of the original variables. In order to make the information contained in these principal components do not overlap each other, the principal components are required not related to each other [15]. By normalizing the raw data and calculating the covariance matrix, can get the eigenvalues of the correlation matrix $\lambda_1 \geq \lambda_2 \geq \lambda_3 \ldots \geq \lambda_p$, and find the corresponding regular unit eigenvectors $u_a = u_{a1}, u_{a2}, \ldots, u_{ap} (a = 1, 2, \ldots, p)$ according to each eigenvalue, and then converting the normalized index variable into a main component: $F_a = u_{a1}z_1 + u_{a2}z_2 + \ldots + u_{ap}z_p (a = 1, 2, \ldots, p)$, where F_1 is the first principal component, F_2 is the second principal component,..., F_p is the p^{th} principal component. According to the theory of principal component analysis, if the eigenvalue is greater than or equal to 1, it is considered to be a principal component, which can reflect a large amount of information of the data. In general, if the contribution rate of the first k principal components reaches 85%, it means that the first k principal components basically contain the information of all the measurement indicators, which reduces the number of variables and facilitates the analysis of actual problems.

Data calculation and result analysis

Input indicator principal component extraction

Firstly, the original 16 input index data (corresponding to the relationship shown in tab. 1) are standardized by MATLAB software. Then, based on the standardized input index data, the eigenvectors of the correlation coefficient matrix and the correlation coefficient matrix are calculated. The eigenvectors of the correlation coefficient matrix represent the importance of each principal component, and the ordering is as shown in fig. 3.

As can be seen from fig. 4, the feature roots of the four principal components are greater than 1, so at most only four principal component factors need to be extracted. The variance contribution rates of the first four principal components are shown in 58.80%, 21.38%, 8.18%, and 7.55%. It can be seen that the variance contribution rate of the first three principal components has reached 88%, which is greater than 85%. Therefore, in order to simplify the calculation, only the first three principal components need to be selected as input indicators for analysis. The coefficient matrix of the first three principal components is shown in tab. 2.

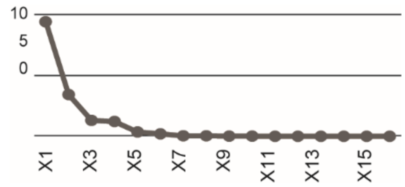

Figure 3. Input indicator gravel distribution

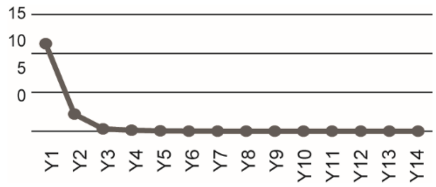

Figure 4. Distribution index gravel distribution

Table 2. Input and output principal component coefficient matrix

	F1	F2	F3		G1	G2
X1	0.305306	–0.12689	–0.1754	Y1	0.29748	0.020919
X2	0.219429	0.016732	0.537844	Y2	0.279251	–0.15164
X3	0.104539	0.045292	0.332162	Y3	0.295257	–0.06318
X4	0.260435	–0.16262	0.095541	Y4	0.270297	0.250224
X5	0.31868	–0.10332	0.065229	Y5	0.283151	0.154036
X6	0.319017	–0.10266	0.053332	Y6	0.297132	0.011631
X7	0.318659	–0.10339	0.065563	Y7	0.295729	0.049142
X8	0.316806	–0.10369	0.06126	Y8	0.295559	0.067256
X9	0.297296	–0.08777	0.111289	Y9	0.29041	0.142543
X10	0.274646	–0.13812	–0.25975	Y10	0.297421	0.025366
X11	0.106304	0.509658	–0.0142	Y11	0.288212	0.130689
X12	0.097524	0.514397	–0.00426	Y12	0.232724	–0.24225
X13	0.274854	0.252502	–0.13975	Y13	–0.09243	0.63555
X14	0.296648	0.173617	–0.12848	Y14	–0.10699	0.615233
X15	0.101458	0.512236	–0.02151			
X16	–0.1272	0.071032	0.654807			

According to the principal component coefficient matrix, the expressions of the three input principal components are:

$$F1 = 0.3053X1 + 0.2194X2 + 0.1045X3 + 0.2604X4 + 0.3187X5 + 0.319X6 + \\ + 0.3187X7 + 0.3168X8 + 0.2973X9 + 0.2746X10 + 0.1063X11 + 0.0975X12 + \\ + 0.2749X13 + 0.2966X14 + 0.1015X15 - 0.1272X16$$

$$F2 = -0.1269X1 + 0.0167X2 + 0.0453X3 - 0.1626X4 - 0.1033X5 - 0.1027X6 - \\ - 0.1034X7 - 0.1037X8 - 0.0878X9 - 0.1381X10 + 0.5097X11 + 0.5144X12 + \\ + 0.2525X13 + 0.1736X14 + 0.5122\,X15 + 0.071X16$$

$$F3 = 0.1754X1 + 0.5378X2 + 0.3322X3 + 0.0955X4 + 0.0652X5 + 0.0533X6 + \\ + 0.0656X7 + 0.0613X8 + 0.1113X9 - 0.2598X10 - 0.0142X11 - 0.0043X12 - \\ - 0.1398X13 - 0.1285X14 - 0.0215X15 + 0.6548X16$$

The input principal component index values calculated according to the previous expression are shown in tab. 3.

Table 3. Input and output principal component indicator values

	F1	F2	F3	G1	G2
2004	0.114835	0.315508	0.329118	0.1	0.513684
2005	0.181153	0.735609	0.176532	0.150993	0.434626
2006	0.137362	0.372387	0.566867	0.17433	0.462737
2007	0.1	0.326376	0.424146	0.240607	0.500132
2008	0.185399	0.299829	0.463296	0.327189	0.497569
2009	0.313738	0.226765	0.638008	0.382817	0.527139
2010	0.676441	0.9	0.781154	0.460864	0.521978
2011	0.69895	0.131046	0.9	0.601457	0.1
2012	0.795494	0.401819	0.558353	0.684675	0.137759
2013	0.889166	0.68341	0.444221	0.789115	0.240811
2014	0.756609	0.102255	0.459337	0.839875	0.264854
2015	0.827117	0.1	0.274761	0.889269	0.305711
2016	0.9	0.445053	0.190668	0.849007	0.9
2017	0.732674	0.298703	0.1	0.9	0.881795

Output indicator principal component extraction

Similar to the principal component analysis of input indicators, MATLAB is also used to standardize the raw data of 14 output indicators, see tab. 1, and then calculate the corresponding correlation coefficient matrix and correlation coefficient based on the standardized output indicator data. The eigenvector of the matrix. The eigenvectors of the correlation coefficient matrix represent the importance of each principal component. The eigenvectors of the correlation coefficient matrix of the output indicators obtained after such processing represent the importance of the principal components of each output indicator, and the ranking is as shown in fig. 4.

As can be seen from fig. 4, the characteristic roots of the two principal components are greater than 2, and the remaining principal components are all less than 1, and the principal component extraction result is very significant. The variance contribution rates of the first two principal components are 80.35% and 15.76%, respectively, and the total variance contribution rate of the two is greater than 95%, which satisfies the requirements. Therefore, in order to simplify the calculation, only the first two principal components need to be selected as the output indicators for analysis. The coefficient matrix of the first two output principal components is shown in tab. 2.

According to the principal component coefficient matrix, the expressions of the two input principal components are:

$$\begin{gathered} G1 = 0.2975Y1 + 0.2793Y2 + 0.2953Y3 + 0.2703Y4 + 0.2832Y5 + 0.2971Y6 + \\ + 0.2957Y7 + 0.2956Y8 + 0.2904Y9 + 0.2974Y10 + 0.2882Y11 + \\ + 0.2327Y12 - 0.0924Y13 - 0.107Y14 \end{gathered}$$

$$\begin{gathered} G2 = 0.0209Y1 - 0.1516Y2\text{-}0.0632Y3 + 0.2502Y4 + 0.1541Y5 + \\ + 0.0116Y6 + 0.0491Y7 + 0.0673Y8 + 0.1425Y9 + 0.0254Y10 + \\ 0.1307Y11 - 0.2422Y12 + 0.6355Y13 + 0.6152Y14 \end{gathered}$$

The output principal component indicator values calculated according to previous expression are shown in tab. 3.

Water input and output based on DEA method

According to the three input principal components F1, F2, F3, and the two output principal components G1 and G2 extracted previously, combined with the input-oriented BCC model, the water resources input-output efficiency analysis model of Jilin Province is constructed, and Jilin Province is calculated. The water use efficiency of each year is shown in tab. 4.

Table 4. Water use efficiency of Jilin Province

DMU	Technical efficiency score (CRS)	Pure technical efficiency score (VRS)	Scale efficiency score	RTS
2004	1	1	1	Constant
2005	1	1	1	Constant
2006	0.776726	0.850006	0.913789	Increasing
2007	1	1	1	Constant
2008	1	1	1	Constant
2009	1	1	1	Constant
2010	0.448073	0.46025	0.973543	Decreasing
2011	0.759384	0.96159	0.789717	Increasing
2012	0.676359	0.676601	0.999642	Decreasing
2013	0.666	0.704348	0.945555	Decreasing
2014	1	1	1	Constant
2015	1	1	1	Constant
2016	0.773376	1	0.773376	Decreasing
2017	1	1	1	Constant

It can be seen from tab. 4 that the water use efficiency of Jilin Province was in the state of optimal production scale and optimal technical contribution rate in 2004 and 2005. In 2006, water use efficiency decreased and scale returns increased. This shows that the scale of production in Jilin Province was not in place and the production factors such as water resources and capital technology were not fully utilized. In 2007-2009, the water use efficiency of Jilin Province was in an optimal state. From 2010 to 2013, the water use efficiency of Jilin Province was significantly reduced. As shown in figs. 1 and 2, these four years are also the fastest economic growth in Jilin Province and the fastest increase in water consumption, so the water efficiency is reduced. Mainly due to the extensive economic growth, the corresponding technology management is not in place, of which only the scale income in 2011 is increasing, and the rest of the years are diminishing returns to scale. This shows that the economic growth of Jilin Province is too fast during this period, and the corresponding industrial technology Transformation and upgrading are not matched. In 2014-2017, the water use efficiency level returned to the optimal level, and the pure technical efficiency reached the best for four consecutive years. Only in 2016, the scale efficiency diminished, but it was quickly adjusted to the optimal state.

Conclusion

According to the aforementioned analysis, the water use efficiency of Jilin Province has been maintained at a good level before 2010, in synchronization with the level of economic growth. However, with the rapid expansion of the economic scale after 2010, there will inevitably be inadequate management. The problem of technological transformation and upgrading is not timely. Therefore, it can be seen that the utilization efficiency of water resources in Jilin Province during this period is extremely low, and there are serious problems of water resource waste and water pollution. However, the above analysis also shows that Jilin Province has started the industrial transformation and upgrading in a timely manner, implemented the national agricultural modernization and revitalized the old industrial base in Northeast China, so the water use efficiency after 2014 has gradually improved to a good level. If the current policy can be successfully implemented, the water resources problem in Jilin Province will be alleviated and the water carrying capacity will be improved.

Acknowledgment

This work was supported by the National Key Research Program of China (No. 2016YFC0401305, 2017YFC0506603), the State Key Program of National Natural Science of China (No. 41530635), and the Project of National Natural Foundation of China (No. 51679007, 51379013).

References

[1] Wang, X. F., Study of Water Resources Efficiency by Considering the Regional Water Carrying Capacity (in Chinese), *Urban and Environmental Studies*, *1* (2018), 2, pp. 97-110
[2] Yang, G. L., *et al.,* Review of Data Envelopment Analysis (in Chinese), *Journal of Systems Engineering, 28* (2013), 4, pp. 840-860
[3] Charnes, A., *et al*., Measuring the Efficiency of Decision Making Units, European Journal of Operational Research, *2* (1978), 6, pp. 429-444
[4] Yang, X. H., *et al.*, Chaos Gray-Coded Genetic Algorithm and Its Application for Pollution Source Identifications in Convection-Diffusion Equation, *Communications in Nonlinear Science and Numerical Simulation*, *13* (2008), 8, pp. 1676-1688
[5] Li, Z. M., Liao, H. C., Input and Output Analysis of Water Resources Across China in 2010 (in Chinese), *Resources Science*, *34* (2012), 12, pp. 2274-2281

[6] Yang, X. H., *et al.*, Improved Gray-Encoded Evolution Algorithm Based on Chaos Cluster for Parameter Optimization of Moisture Movement, *Thermal Science*, *21* (2017), 4, pp. 1581-1585
[7] Sun, F. H., *et al*., Evaluation of Utilization Efficiency of Regional Agricultural Water Resources Based on Three-Stage DEA-Malmquist Model (in Chinese), *Journal of Economics of Water Resources*, *37* (2019), 2, pp. 53-58
[8] Cao, F, L., Analysis of Industrial Water Resources Utilization Efficiency Based on DEA Method (in Chinese), *Energy Conservation & Environmental Protection*, *7* (2018), 10, pp. 64-67
[9] Banker, R. D., *et al*., Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, *Management science, 30* (1984), 9, pp. 1078-1092
[10] Xu, J. H., Prediction of Urban Water Demand Based on Improved Principal Component Analysis Method (in Chinese), *Water Resources Development and Management*, *1* (2019), 3, pp. 23-25
[11] Wu, Y., He, J.-H. A Remark on Samuelson's Variational Principle in Economics, *Applied Mathematics Letters, 84* (2018), Oct., pp. 143-147
[12] He, J. H. A Modified Li-He's Variational Principle for Plasma, *International Journal of Numerical Methods for Heat and Fluid Flow*, On-line first, https://doi.org/10.1108/HFF-06-2019-0523, 2019
[13] He, J. H. Variational Principle for the Generalized KdV-Burgers Equation with Fractal Derivatives for Shallow Water Waves, *J. Appl. Comput. Mech.*, *6* (2020), 4, pp. 735-740
[14] He, J. H., Sun, C., A Variational Principle for a Thin Film Equation, *Journal of Mathematical Chemistry, 57* (2019), 9, pp. 2075-2081
[15] Jolliffe, I. T., *Principal Component Analysis*, 2nd ed., New York: Springer-Verlag New York Inc., USA, 2002

Paper submitted: April 28, 2019
Paper revised: June 19, 2019
Paper accepted: August 8, 2019