Energy Demand Forecast of Hubei Logistics Industry Based on RBF Neural Network

[1st]Rong Zhou[a], JiangXue Di[*b]

[a]906480609@qq.com, [*b]Corresponding author: 791424893@qq.com

School of management Hubei University of Chinese Medicine Wuhan, China

Abstract—With the rapid economic development in Hubei Province, the logistics industry demand has grown fast and its scale has been expanded continuously, which leads to an increase in energy consumption. It is conducive to the development of energy-saving work in the logistics industry and alleviating energy pressure by studying the energy consumption level and energy demand of the logistics industry in Hubei Province.11 main factors that affect the energy demand of the logistics industry are selected in this paper. According to the RBF neural network method, the related data of the energy demand of the logistics industry in Hubei Province from 2009 to 2017 is simulated and emulated. Based on it, the energy demand of the logistics industry in Hubei Province are predicted in 2021 and 2022. The results show that: (1) From 2009 to 2017, the total energy consumption of the logistics industry in Hubei Province continued to increase.With the further development of the logistics industry, by 2021 and 2022, the total energy consumption of the logistics industry will reach 23.098 million tons of standard coal and 23.438 million tons of standard coal.(2) The RBF neural network has higher prediction accuracy than the multiple linear regression model prediction model In solving the energy demand forecasting problem of the logistics industry.(3) It is found that the factor ofthe total retail sales of social consumer goods hasa greatimpact on energy consumption of the logistic industry through Grey Relation Analysis.

Keywords-Logistics Industry;Energy Demand Predictio; RBF Neural Networ;Multiple Linear Regression

1. INTRODUCTION

As an emerging productive service industry, the logistics industry has connected various economic and social industrial sectors, which is not only a basic industry supporting regional economic development, but also a new growth point for the national economy. Meanwhile,the logistics industry is also one of the most important industries ofenergy consumption in China,which has driven its energy consumption to continue to increase especially because of the rapid development of the logistics industry in recent years[1]. As a large economic province with a regional GDP accounting for nearly 5% of the country, Hubei Province is developing at a faster rate than the national level. According to the calculations of the China Energy Statistical Yearbook and the Hubei Statistical Yearbook, the proportion of the energy consumption of the logistics industry in the total energy consumption of Hubei Province increased from 8.79% in 2009 to 10.85% in 2017[2].. Therefore, under the background of issues including energy security and energy pricesand the development of a low-carbon economy, It is of great significance tosave energy in Hubei logistics industry, improve energy efficiency and

MSEA 2023, May 26-28, Nanjing, People's Republic of China

DOI 10.4108/eai.26-5-2023.2334201

reduce carbon emissionsthrough forecasting the energy demand of the logistics industry in Hubei Province and estimating the energy efficiency of the logistics industry, which can also provide an effective basis for rational energy planning and policy formulation in Hubei Province[3].

At present, researches on energy consumption and demand forecasting,energy efficiencyand energy conservation in the logistics industry at home and abroad are relatively rare[4].The existing literature mainly includes the following three aspects: (1)Research on the energy demand forecasting of the logistics industry applies the main methods such as two-stage genetic algorithm, neural network algorithm, gray prediction model, time series prediction model, etc. However, the influencing factors of energy consumption and demand in logistics industryhave not been systematically summarized and selected and there are few studies taking the energy consumption of logistics industry in Hubei Province as a research object. (2)The main entry points to evaluate the energy efficiency of the logistics industryare the supply chain performance evaluation system and the energy consumption index evaluation system of the logistics storage system. (3)From the perspective of sustainable development,the energy conservation of the logistics industry is explored. From the point of view of economic and environmental benefits in existing research studies, energy conservation in supply chain management is analyzed.

It can be seen that even though there have been some research achievements in the field of energy consumption and demand forecasting in the logistics industry, shortcomings still exist.Therefore, regarding the energy demand system of the logistics industry in Hubei Provinceaffected and restricted by various factors,energy consumption and demand of the logistics industry in Hubei Province are simulated and compared in this paper with the radial basis (RBF) neural network model of Matlab software and multiple linear regression in SPSS software[5]. The energy efficiency of the logistics industry is measured to increase the research results on energy consumption and demand and energy efficiency of the logistics industry in Hubei Province and to provide scientific decision support for energy managers.

2. RESEARCH METHODS AND DATA SOURCES

2.1. Research Methods

1) Grey Relation Analysis. Rey correlation analysis is a method to analyze and determine the degree of influence among system factors or the contribution of factors to the main behavior of the system through the gray correlation degree. The calculation steps are:

Let the total energy consumption of the logistics industry for several years be the original data. After dimensionless processing, a reference data column is generated and recorded as formula (1):

$$X_0=[X_0(1),X_0(2),\cdots,X_0(n)] \tag{1}$$

The index data of the influencing factors of the energy demand in the logistics industry is generated through a dimensionless process and is compared as a formula (2):

$$X_i=[X_i(1),X_i(2),\cdots,X_i(n)](i=1,2,\cdots,m) \tag{2}$$

Then the real number:

$$\xi_i(\mathrm{K}) = \frac{\min_i \min_K |X_0(\mathrm{K}) - \mathrm{X}_i(\mathrm{K})| + \rho \max_i \max_K |X_0(\mathrm{K}) - \mathrm{X}_i(\mathrm{K})|}{|X_0(\mathrm{K}) - \mathrm{X}_i(\mathrm{K})| + \rho \max_i \max_K |X_0(\mathrm{K}) - \mathrm{X}_i(\mathrm{K})|} \tag{3}$$

It is called the correlation coefficient of Xi for X0 at the K point. In formula (3), $\rho \in [0,1]$, ρ is usually 0.5. Then the correlation between the influencing factors of the energy demand in the logistics industry and the total energy consumption is formula (4):

$$\gamma_i = \frac{1}{n} \sum_{K=1}^{n} \xi_i(\mathrm{K}) \tag{4}$$

The larger the value of γi in the formula, the greater the degree of correlation between the comparison sequence and the reference sequence. In general, when $0<\gamma i\le 0.4$, the correlation is weak; $0.4<\gamma i\le 0.6$ is the acceptable correlation; when $0.65<\gamma i\le 0.85$, the correlation is strong; When $0.85 <\gamma i\le 1$, the correlation is extremely strong.

2) RBF neural network. Radial Basis Function (RBF) neural network is a kind of forward network with good performance with characteristics such as simple network structure, fast training speed, strong local approximation ability, high approximation accuracy, adaptability and generalization ability, which is able to find the non-linear mapping relationship between the influencing factors of energy consumption and energy demand in the logistics industry by studying historical data[6].

The RBF network is a three-layer forward network: the first layer is the input layer and consists of signal source nodes. Second floor is the hidden layer, and the transformation function of the hidden unit is a locally distributed non-negative and non-linear function, which is radially symmetrical and attenuated to the center point[7]. The number of units of the hidden layer is determined by the needs of the problem described. The third layer is the output layer, and the output of the network is a linear weighting of the hidden unit output. The transformation from the input space of the RBF network to the hidden layer space is nonlinear, while the transformation from the hidden layer space to the output layer space is linear. Without loss of generality, it is assumed that the output layer has only one hidden unit, and the training sample pair of the network is {Xn, dn} (n=1,2, ...,N), where Xn=[Xn1, Xn2,...,Xnm] T, (n=1,2,...,N)is the input of the training sample, dn (1,2,...,N) is the expected output of the training sample, and the corresponding actual output Yn (n=1,2,...,N); the basis functionφ(X, ti) is the output of the i-th hidden unit, ti= [ti1, ti2,...,tim,...,tiM] (i =1,2,I) is the center of the basis function; wi (i=1,2,...,I) is the weight between the i hidden unit and the output unit. The topology of a single-output RBF network is shown in Figure 1:

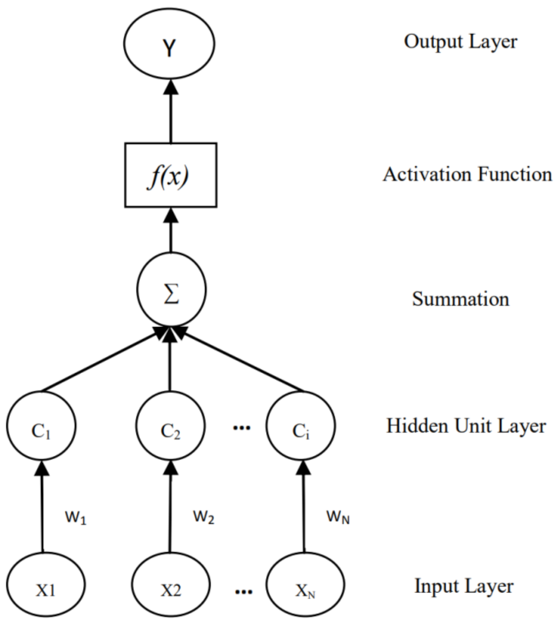

Figure1 The topology of a single-output RBF network

When the network inputs training samples Xn, the actual output of the network is:

$$Y(\mathrm{X}_n) = \sum_{i=1}^{I} w_i \varphi(\mathrm{X}_n, \mathrm{t}_i) \tag{5}$$

As for the RBF neural networks, a Gaussian function is usually selected as the radial basis activation function, and its expression is:

$$f(\mathrm{x}) = \exp\left(-\frac{\left\| x_n - c_i \right\|^2}{2\sigma^2}\right) \tag{6}$$

It can be obtained that the output of the RBF neural network, that is, the energy demand prediction model of the logistics industry is:

$$Y_{j=}\sum_{i=1}^{m} \omega_i \exp\left(-\frac{\left\| x_n - c_i \right\|^2}{2\sigma^2}\right) \tag{7}$$

In the formula,Yj is the energy consumption in the j-th year; ωi is the network weight; ||.|| is the Euclidean norm; σ is the variance of the Gaussian function; xn is the data of the n-th training year; ci is the center of the hidden layer node in the network.

2.2. Establishment of a Forecast Model for Energy Demand in the Logistics Industry in Hubei Province

The energy demand system of the logistics industry is a complex system[8]. In order to analyze the influencing factors of the energy demand of the logistics industry in Hubei Province,the impact of various factors including economic, social, and technical factors relating to its consumption behavior must be comprehensively considered and the existence of these factors or interrelated or mutually restrictive influence must be fully took into account.

Combining the research results in the field of energy consumption and demand of the logistics industry in the paper, based on the analysis of the actual data from 2009 to 2017 with the principles of comprehensiveness, comparability and availability, the main factors affecting the energy demand of the logistics industry in Hubei Province are summarized and selected. Takingthe data of the 11 main influencing factors as inputs including the X1 value-added industry(Billion Yuan), X2 fixed asset investment(Billion Yuan), X3 energy efficiency(Ten thousand Yuan/t), X4 transportation equipment ownership(car), X5 freight volume(Ten thousand t), X6 cargo turnover(Billion t km), X7 gasoline and diesel kerosene power consumption ratio, X8 number of employees(Ten thousand person), X9 total retail sales of consumer goods(Billion Yuan), X10 R&D funding(Ten thousand Yuan), X11 total business volume of postal industry(Billion Yuan) and the total energy consumption(million tons of standard coal) is taken as the output and represented by Y.Hence, a multi-input and single-output RBF neural network of energy demand forecast model of the logistics industry is constructed. The details are shown in Table 1.

Table1. Logistics Industry Energy Influencing Factors And Energy Consumption Observations In 2009-2017 Of Hubei Province

Year	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	Y
2009	642.72	957.23	0.533	466160	82714	2808.46	0.7715	360873	5929.07	231760	41.29	1205.794858
2010	753.61	1147.77	0.589	538402	97007	3370.37	0.6426	316548	7014.45	337088.8	53.45	1279.138397
2011	869.48	1152.61	0.572	602705	110168	4044.45	0.5879	316928	8363.34	356179.2	42.02	1520.320785
2012	934.96	1266.58	0.617	631795	125392	4693.61	0.6305	308375	9682.35	420838.8	55.46	1513.589666
2013	1078.11	1733.95	0.74	691138	139740	4883.01	0.7426	408357	11035.94	577699.6	74.15	1456.444957
2014	1181.58	2157.93	0.801	732044	154736	5798.12	0.7867	425599	12449.27	707785	98.66	1474.319273
2015	1242.34	2794.23	0.822	705957	156357	5908.4	0.7752	408142	14003.24	840294	137.41	1511.209606
2016	1297.48	2892.65	0.712	701110	165125	6159.9	0.8225	441775	15649.22	865057.8	192.1	1821.905594
2017	1420.01	2925.05	0.763	746795	190845	6589.58	0.8182	457725	17394.1	994801.6	265.74	1859.87785

2.3. Sources of Data

The added value of transportation, warehousing and postal industry comes from the data selected form the screening channel directly provided by the official website of the National Bureau of Statistics of the People's Republic of China. R&D funding is the sum of R&D funding of Railway, Ship, Aerospace, and Other Transportation Equipment Systemsand Auto Manufacturingin the Hubei Statistical Yearbook. There are differences in the statistical units of gasoline, diesel, keroseneand electricity, therefore, the statistical result is data obtained by converting the four types of energy according to the reference coefficients of various energy conversion standards in the China Energy Statistics Yearbook. Energy efficiency is the ratio of industrial added value to total energy consumption. All other data comes from the direct data of the Hubei Statistical Yearbook.

3. RESULTS AND ANALYSIS

3.1. Data Preprocessing

Because of the differences in quality of the dimensions of the identified influencing factors and their index data, in order to obtain more accurate prediction results,the original data of each influencing factor needs to be dimensionless processed, that is, all index data is transformed into [0,1]. Dimensionless processing of each indicator data:

$$x=[xij-\min(xj)]/[\max(xj)-\min(xj)] \quad (8)$$

In the formula, x is the result of dimensionless processing of a certain index data (xij); xij is the actual value of the i-thinfluencing factor index in the j-th year; max (xj) is the i-thinfluencing factor with the largest value being the actual value of the j-th year; min (xj) is the i-thinfluencing factor with the smallest value being the actual value of the j-th year.

3.2. Analysis of Correlation between Influencing Factors and the Total EnergyConsumption

After performing the dimensionless processing on the data in Table 1, the total energy consumption of the logistics industry and the 11 influencing factors were analyzed with Matlab2017a for gray correlation analysis. The results are shown in Table 2 and they demonstrate that the order of correlation between these 11 influencing factors and the total energy consumption of the logistics industryis X9>X6>X5>X1>X2>X10>X4>X11>X8>X3>X7. Because these 11 correlation values are greater than 0.6, the 11 factors summarized and selected in this study have a strong correlation with the energy demand of the logistics industry, which is suitable for forecast.

TABLE2. Grey Correlation Degree Of Each Influencing Factor And Total Energy

Indexes	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
Correlation value	0.7818	0.7797	0.6270	0.7419	0.7877	0.7975	0.6163	0.6635	0.8095	0.7713	0.7391

3.3. Multiple Linear Regression Analysis

Multiple regression analysisrefers to the method of establishing a linear prediction model for the correlation analysis of two or more than two independent variables and one dependent variable. The 11 influencing factors selected above was applied as the independent variable with SPSS19.0 and the total energy consumption as the dependent variable in this paper. A multiple linear regression model is constructed with the stepwise independent variable entry method.

As can be seen from the results, the model has completed convergence after 2 iterations. From the adjusted the square of R, the fit of the model is reaching 0.943,indicating that the 94.3% changecan be determined by the 2 independent variables entering the model. DW = 1.717, close to 2, indicates that the residuals are independent of each other. From the results of the method analysis, the observed value of model F is 67.745 and the probability P is 0.In the case of a significant level of 0.05, it can be considered that there is significant linear relationship among the total retail sales of consumer goods, the total R & D expenditure and total energy consumption. Meanwhile, the multiple linear regression equation is established according to the model:

$$Y=578.230+.194*X9-0.002*X10 \tag{9}$$

The constants in the equation and the probability P values of the two partial regression coefficients are 0.001,0.001and 0.004, respectively. After the T test, according to the given significance level of 0.05, it has significant significance.The total energy consumption of Hubei Province's logistics industry in 2016 and 2017 predicted using this model is shown in Table 3 and the degree of fitting is shown in Figure 2.

3.4. RBF Neural NetworkAnalysis Results and Accuracy Comparison of the Two Methods

Based on the working principle of the RBF neural network model, 11 energy demand influencing factors of the logistics industry with Matlab2017a is taken as input sample and the total energy consumption as the output sample. Based on the availability of data,the energy consumption data of the logistics industry from 2009 to 2017 is taken as training samples for simulation and emulation in this paper. After repeated experiments, the number of nodes in the hidden layer is finally determined. That's to say, when Ci in formula (7) is5, the best training error value of the model is 8.56299e-4 and the error is almost 0, which shows that the output of the network can approximate the non-linear function Yj well, namely, the RBF neural network model has an excellent fitting ability.

In order to verify its effectiveness, the relevant data of the logistics industry's energy consumption in 2016 and 2017 are selected for prediction. The prediction results and simulation results are shown in Table 3 and Figure3. The closer the simulated value and actual value obtained by the model is and the higher the coincidence between the simulated value curve and the actual value curve is. As can be seen from Figure 3, the simulation and prediction effects of the RBF neural network are the best.

TABLE 3. Energy Demand Forecast Of Logistics Industry In 2016 And 2017

	Observation value	RBF neural network prediction value	Multiple linear regression prediction value
2016	1821.9055938	1818.0384835153	1787.4691824967845
2017	1859.8778502	1861.32648070066	1851.7876180123412

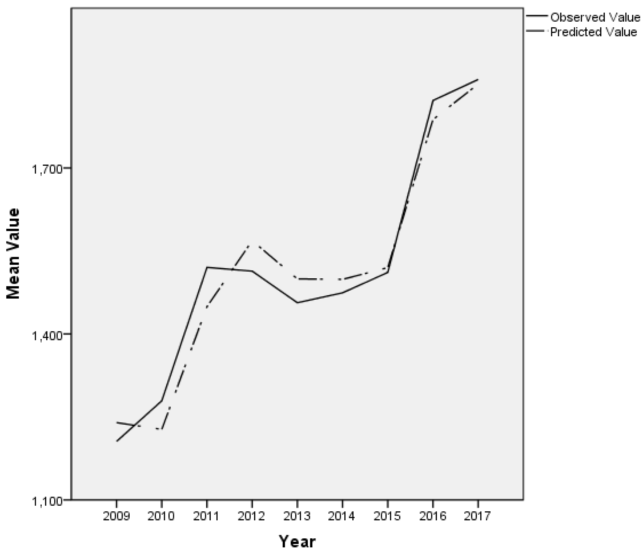

Figure 2 Simulation results of multiple linear regression

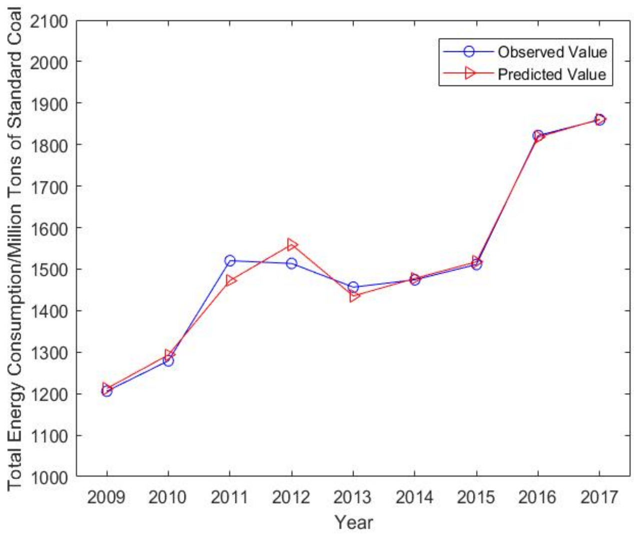

Figure 3 Simulation results of RBF neural network

The three absolute error estimation methods of average absolute error (MAE), root mean square error (RMSE)and average absolute percent error (MAPE) are adopted in this paper to evaluate the prediction accuracy of the three prediction methods. Their expressions are:

$$MAE = \frac{1}{t}\sum_{j=1}^{t}\left|Y_j - \hat{Y}_j\right| \tag{10}$$

$$RMSE = \sqrt{\frac{1}{t}\sum_{j=1}^{t}(Y_j - \hat{Y}_j)^2} \tag{11}$$

$$MAPE = \frac{1}{t}\sum_{j=1}^{t}\left|\frac{Y_j - \hat{Y}_j}{Y_j}\right| \times 100\% \tag{12}$$

In the formulaŶj is the actual energy consumption value of the logistics industry; Yj is the simulated value; t is the total number of training samples and testing samples. After calculation, the prediction accuracy evaluation results of the three prediction methods are shown in Table 4.Among the two prediction methods, the average absolute error, root mean square error and average absolute percentage error value of the RBF neural network method are the smallest, which indicates that the established RBF neural network model has higher prediction accuracy than the multiple linear regression prediction method with a stronger promotion ability value.

TABLE4. Statistical Errors Of Different Prediction Methods

Prediction method	MAE	RMSE	MAPE
RBFneural network	16.68	23.82	1.13%
multiple linear regression	36.71	110.12	2.53%

Therefore, the RBF model is adopted for the prediction of the energy demand of the logistics industry in 2021 and 2022and results are 23.098 million tons of standard coal and24.385 million tons of standard coal. The annual growth rate is 5.57%.

4. CONCLUSIONS AND RECOMMENDATIONS

4.1. Conclusion

Taking the energy demand of the logistics industry in Hubei Province as the research object and grey relation analysis, RBF neural network, and multiple linear regression analysis as the research methods in this paper, the following conclusions are obtained:

*1)*Conducting a correlation analysis between the 11 main factors affecting the energy consumption of the logistics industry and the energy consumption of the logistics industry, the results show that the correlation degree of these 11 factors with the energy consumption of the logistics industry is above 0.6 and is applicablethe forecast of energy demand in the logistics industry, among which the correlation of retail sales of consumer goods, cargo turnoverand freight volume is the strongest, and the correlations are 0.8095、0.7975and 0.7877 respectively.

2) The logistics industry energy demand forecasting model in Hubei Province is established with the RBF neural network method and multivariate linear regression method. The relevant data of Hubei Province logistics industry energy consumption from 2009 to 2017 is selected for analysis showing that the RBF neural network method with high prediction accuracy is conductive to providing valuable reference information and enhancing the rationality of future energy demand planning for decision-makers and to promoting the sustainable and healthy economic development of HubeiProvince. Based on this, the established RBF neural network model is applied to predict that the energy demand of China's logistics industry in 2021 and

2022 will be 23.098 million tons of standard coal and24.385 million tons of standard coal, respectively, with an average annual growth rate of 5.57%.

4.2. Recommendations

*1)*The investment structure of fixed assets in the logistics industry should be optimized. On the one hand, there is a close relationship between the energy consumption of vehicles and the road conditions. Good road conditions can not only save energy consumption of vehiclesbut also extend the mileage of vehicles.Therefore, the improvement of transportation road conditions should be attached great importance to actively invest in building high-quality Road; On the other hand, logistics machinery and equipmentas an important tool in the logistics industryprovides support for the smooth implementation of logistics activities. Funds should be invested to upgrade logistics machinery and equipment to reduce their energy consumption.

2) The energy consumption structure of the logistics industry should be optimized by promoting clean energy. It is necessary to actively promote clean energy in the logistics industry including fully developing renewable energy like wind energy and solar energy,gradually adopting new energy transportation tools,green storage facilities and other logistics equipment, maximizing the clean energy consumption ratio in the logistics industry and reducing the consumer demand for non-renewable resources such as refined oil and coal.

3) Transportation structure and transportation means should be optimized to reduce energy consumption. The transportation mode determines the logistics efficiency and energy consumption. Therefore, it is necessary to reasonablyequip with various transportation modes, actively carry out various transportation, increase the proportion of railway and water transportation and reduce the proportion of road transportation in logistics transportation to reduce vehicle fuel consumption.

4) Technological innovation and the improvement of logistics operating should be relied on to continuously increase energy efficiency. Taking scientific and technological progress as a tool, new energy-saving logistics technologies should be actively developed and promoted, unreasonable links and processes in logistics and supply chain management should be transformed to improve the management level of the logisticsindustry and to provide necessary technical support for improving its energy efficiency.

REFERENCES

[1] X. B. Li and J. M. Xu , "Traffic condition discrimination of Expressway Based on RBF neural network,"*Computer simulation*, vol. 28, no. 2, pp. 350-353, 2011.

[2] J. J. Li and W. P. Xu, "Study on product performance evaluation model based on improved grey correlation analysis," YunnanJournal of University (NATURAL SCIENCE EDITION) , vol. 28, no. 21, pp. 115-120, 2006.

[3] Y. G. Ban , "Prediction of urban water demand based on RBF neural network," Groundwater, vol. 34,no. 1,pp. 114-116, 2012.

[4] D. Ma andB. G. Shan, "Comparative study of 2030 world energy outlook based on global energy outlook report," China Energy, vol. 39, no. 2, pp. 21-24, 2017.

[5] "National Bureau of Statistics",Simple query of annual data by province, 2019. [Online]. Available: http://data.stats.gov.cn/easyquery.htm?cn=E0103. [Accessed: 10- Mar- 2020].
[6] "Statistical yearbook of Hubei Provincial Bureau of Statistics",Hubei Provincial Bureau of Statistics, 2019. [Online]. Available: http://tjj.hubei.gov.cn/tjsj/sjkscx/tjnj/gsztj/whs/. [Accessed: 12- Mar- 2020].
[7] Halldorsson and M. Svanberg, "Energy resources:Trajectories for supply chain management," Supply Management, vol. 18, no. 1, pp.66-73, 2013.
[8] N. Apergis, G. C. Aye and C. P, "Barros Energy efficiency of selected OECD countries :A slacks based model with undesirable outputs," Energy Economics, vol. 51, no. 9, pp.45-53, 2015.