ISSN: 2472-3703 DOI: 10.6911/WSRJ.202006_6(6).0024

Carbon Emission Prediction Under Regression Model

Yashi Nie[1, a, *]

[1]Baoding campus of North China Electric Power University, Hebei, China.

[a]Corresponding author e-mail: 1332439799@qq.com

Abstract

Since the first industrial revolution, the rapid increase of greenhouse gas emissions, such as carbon dioxide, has been considered as the main cause of global climate deterioration. CVS survey shows that the global average concentration of greenhouse gases is still growing, China has surpassed the United States to become the largest carbon emission country, the emission reduction warning has been sounded, low-carbon emission reduction is urgent. Based on the actual situation of our country, this paper analyzes the main factors affecting carbon emissions, and makes suggestions for the low-carbon emission reduction in China. First of all, combined with China's economic development, we analyze the changing trend and reasons of China's energy consumption structure from the perspective of industrial structure. Then, we will establish two main models, namely growth rate regression model and prediction model. In the regression model, we will focus on the analysis and prediction of carbon emissions. In order to simplify the calculation, we use the principal component analysis method to calculate the growth rate of each component according to the existing data and the data of the first 15 years (2000-2014) in the China Statistical Yearbook. Here, we use the growth rate as the dependent variable, by calculating the contribution rate of each influencing factor, we better select three factors that contribute more than 85% of the cumulative annual growth rate of carbon emissions from the eight influencing factors, namely, international trade, industrial structure and energy consumption structure, and use these three main components to replace the original eight influencing factors for subsequent analysis. Then, we use multiple regression method to calculate the partial regression coefficient relative to each factor, and finally get a better growth rate regression model. In the prediction model, we first establish the prediction model for the three principal components. Using the data from 2000 to 2014, we get three principal component trend equations with high fitting degree by Fourier function fitting, and combine these three equations with the regression model in the previous paper to get our final carbon emission prediction model. In order to verify the reliability of the prediction model, we use the prediction model to simulate the carbon emissions in 2015 and 2016, and compare with the real value. Then, we analyze the error, advantages and disadvantages of the model. Finally, we analyze and summarize the results of the model. Combined with China's current economic development and energy consumption structure, we put forward a resource allocation method based on quantitative analysis and give our low-carbon emission reduction suggestions by discussing the change trend of carbon emissions.

Keywords

Principal component analysis; multiple regression; Fourier function.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202006_6(6).0024

1. INTRODUCTION

With the development of economy and the increase of energy consumption, its products, such as greenhouse effect, acid rain, industrial smoke and dust, have a very bad impact on our surrounding environment, which is undoubtedly contrary to our "sustainable development" strategy.

The CVS survey shows that the global average concentration of greenhouse gases continues to grow. Since its first measurement in 1976, Cape grim has monitored a carbon dioxide increase of at least 15%, but about 40% compared to pre industrial levels.

At present, China has surpassed the United States to become the largest carbon emitter. Facing the increasingly serious carbon emission problems, China has actively carried out emission reduction work, carried out research on low-carbon emission reduction strategies, and joined the Paris climate change agreement on September 3, 2016, and supervised and worked together with all parties to achieve low-carbon transformation.

In the following content, we will start from the energy consumption structure of our country, the influencing factors of carbon emissions, and combine with the viewpoint of sustainable development, state the current situation of carbon emissions in China, and establish a prediction model, so as to further put forward suggestions for energy conservation and emission reduction in China.

2. PROPERTIES

2.1. Problem Analysis

The reason why China has such a large amount of carbon emissions is not only because China is in the stage of rapid development, but also closely related to the proportion of China's industrial structure. The survey shows that China's long-term coal based energy consumption structure is exactly the reason why China's pollutant emissions rank second in the world and carbon emissions rank first in the world. We are committed to contributing to China's emission reduction work. First of all, we need to understand the sources of China's carbon emissions.

For problem one, we think about the impact of industrial structure change on carbon emissions caused by economic growth and national strength enhancement, and analyze the energy consumption structure in different economic development situations in China. (Note: the economic development is not only the economic growth, for example, the growth rate is fast from 2000 to 2005, but slower from 2018 to 2020)

For the second problem, we use PCA to analyze the influencing factors of carbon emissions, calculate the contribution rate of each factor to carbon emissions, and then establish a prediction model to predict. In order to test the reliability of our model, we will also carry out model test and error analysis.

Finally, the paper concludes the development direction of reducing carbon emission to the maximum extent under the premise of economic development in China.

2.2. Symbols

1. C1 :change rate of population factor (total population)
2. C2 :change rate of urbanization rate (urban population / total population)
3. C3 :change rate of economic development level (per capita GDP)
4. C4 :change rate of carbon emission per capita (total carbon emission / total population)
5. C5 :change rate of energy consumption intensity (energy consumption / GDP)
6. C6 :change rate of energy consumption structure (coal proportion)

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202006_6(6).0024

7. C7 :change rate of industrial structure (proportion of secondary industry)
8. C8 :rate of change in international trade (exports / GDP)
9. x_1 :international trade
10. x_2 :industrial structure
11. x_3 :energy consumption structure

2.3. Hypothesis

1. The eight influencing factors are independent of each other.

2. Since science and technology are relatively developed and the protection measures for disasters are relatively sound, we do not consider the large-scale population reduction caused by sudden large-scale disasters.

3. Since the state will carry out macro-control over the economy, we will not consider the sudden economic change.

4. There is no carbon emission in the production and consumption of primary power and other energy.

2.4. Energy Structure Analysis

From the data given in Annex I, we can see that: GDP grows approximately linearly with the growth of the year, and GDP can directly reflect the economic development of a country. Therefore, here, we can use GDP to measure the change of China's economic development with the year (see Figure 1).

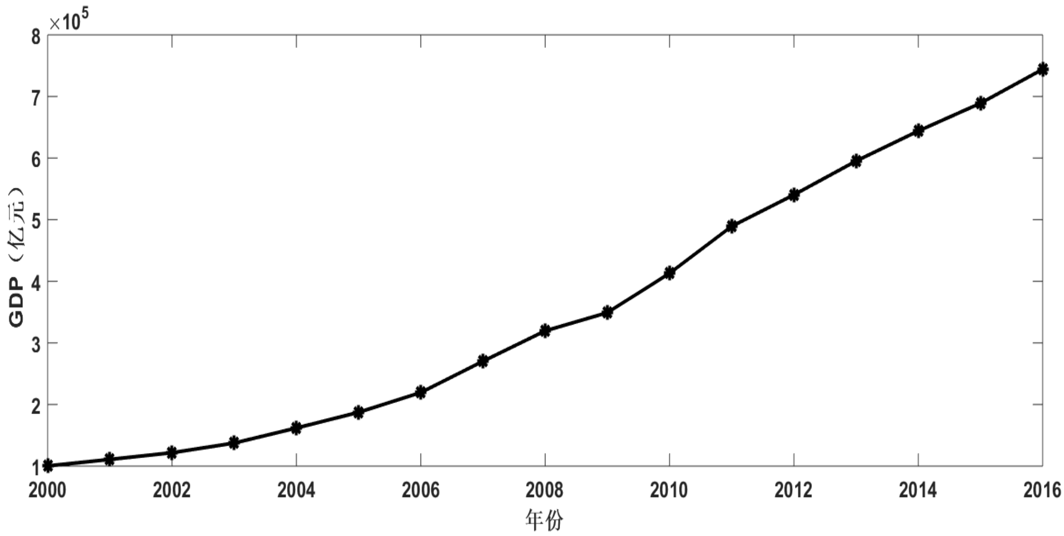

Figure 1. Average annual growth of China's GDP

We look up the data to find the classification of the three major industries and analyze the functions of each industry.

It can be seen from Figure 2 that the tertiary industry shows a trend of rapid growth with the growth of GDP, while the growth of the primary industry is slow, and the secondary industry shows a trend of rapid growth to moderate. We read a lot of information and learned that coal, oil and natural gas are mainly consumed by the secondary industry. If the secondary industry continues to develop, with the continuous expansion of economic scale, it will inevitably bring the increase of resource and environmental pressure and carbon emissions. The carbon emission of the tertiary industry, which is dominated by service industry, is far less than that of the secondary industry, but it can continuously and steadily promote economic growth. The growth rate of the second industry is lower than that of the third industry, which is conducive

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202006_6(6).0024

to the transformation of China's industry-oriented economy to service-oriented economy and reduces the burden of China's emission reduction.

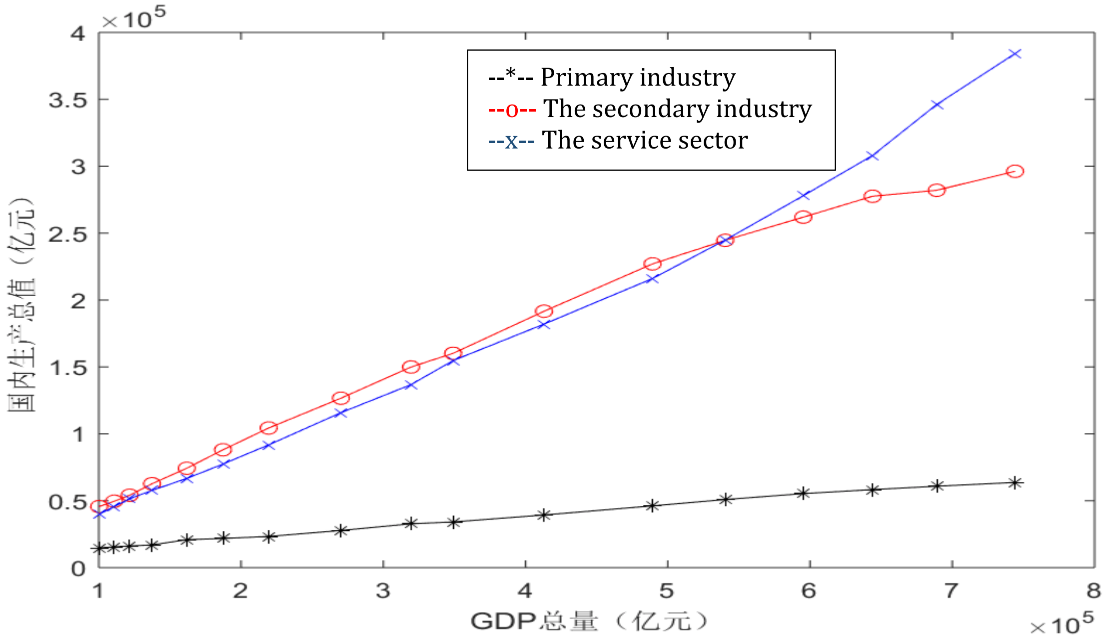

Figure 2. The relationship between GDP of three industries and total GDP

After analyzing the table data, it is not difficult to find out the consumption of coal, oil and natural gas with the economic growth. As can be seen from Figure 3, in the initial stage, China's economy is growing rapidly, and all energy consumption is increasing, especially coal, which has the fastest growth rate, almost three times that of oil, which has the second growth rate. Moreover, the consumption of coal is far greater than that of other energy sources.

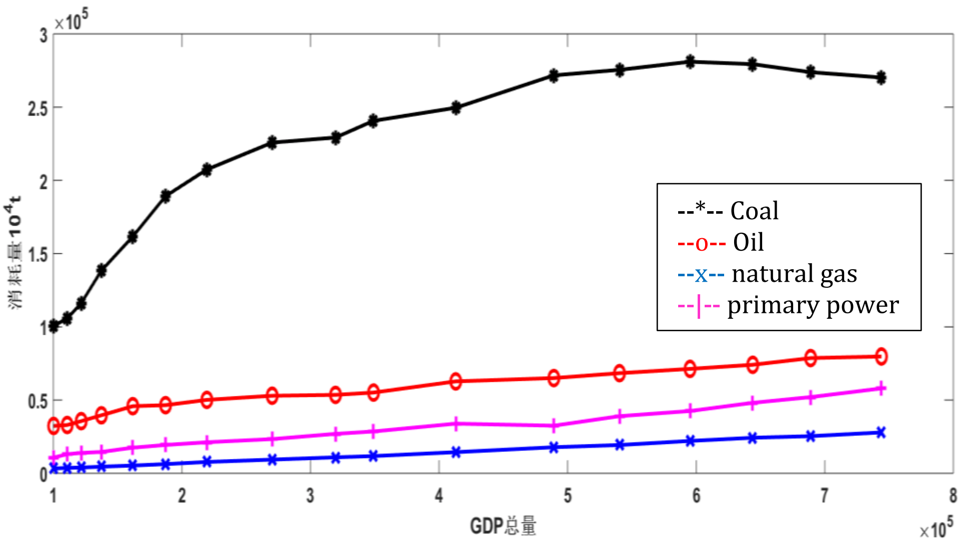

Figure 3. Relationship between consumption of coal, oil, natural gas and primary power and GDP

As can be seen from Figure 2, the consumption of coal is far greater than that of other energy sources, accounting for more than 50% of the total, almost three times more than that of the second largest energy source. The high carbon emission coefficient and necessary high consumption of coal are the most difficult problems for China to carry out emission reduction work on the premise of meeting the demand of economic development. With China's economic growth, the consumption of coal has declined slightly under the general trend, while the proportion of clean energy such as natural gas and primary power in China's energy consumption has increased year by year. We analyze that this is related to the improvement of China's environmental awareness and industrial transformation.

Although the consumption of coal has shown a downward trend, and has continued to decline under the slogan of carbon reduction, it is far from the ideal goal of carbon reduction in China. So, how to maintain the continuous decline and accelerated decline of coal consumption, the most fundamental solution, we need to increase the research input of clean energy, to produce a large number of clean energy at a low cost, shake the position of coal energy output.

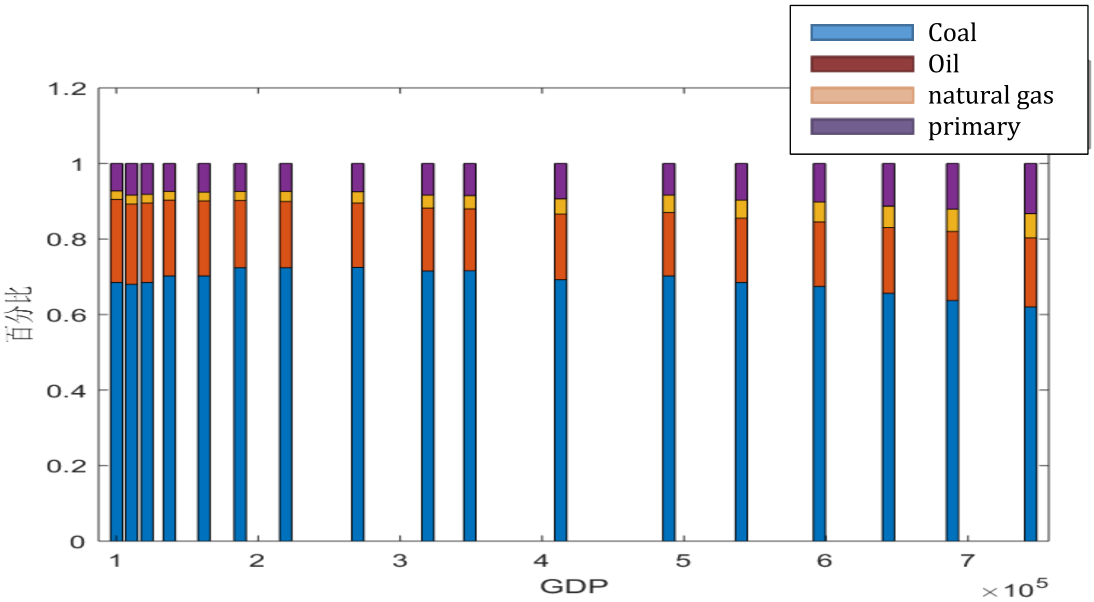

Figure 4. Relation between consumption percentage of coal, oil, natural gas and primary power and GDP

At present, high carbon emission energy is still the main energy consumption in China, but we have considered the impact of carbon emission on the environment and are committed to change this situation. At the end of 2014, the strategic action plan for energy development issued by China pointed out the optimized structural path of China: reducing the proportion of coal consumption and vigorously developing renewable energy.

2.5. Model Establishment

2.5.1 Regression model of growth rate

2.5.1.1 Annual carbon emissions

We calculate the energy consumption and carbon emission coefficient from the table, and use the IPCC method to calculate the carbon emission.

$$C = \sum E_i * F_i$$

Where C is the carbon emission, For all types of energy consumption (coal, oil, natural gas), Is the carbon emission coefficient of each energy source.

Bring in the specific data:

$$C = \text{Annual consumption of coal} * 0.7476 + \text{Annual oil consumption} * 0.5825 + \text{Annual consumption of natural gas} * 0.4435$$

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202006_6(6).0024

2.5.1.2 Growth rate calculation

Since these eight factors are changeable, for the convenience of calculation, we calculate the annual growth rate of these eight factors according to the data in Annex I to carry out the subsequent carbon emission analysis. (see attached table 1 for details)

2.5.1.3 Principal component analysis

We selected the data from 2000 to 2014 for analysis, standardized the growth rate of each influencing factor, and obtained an eight dimensional random vector $X = (X_1, X_2, \ldots, X_8)^T$ 15 samples in years$x_i = (x_{i1}, x_{i2}, \ldots, x_{i8})^T, i = 1,2, \ldots 15$.To construct a sample array, standardize the sample array as follows:

$$Z_{ij} = \frac{x_{ij} - \overline{x_j}}{s_j}, i = 1,2, \ldots, 15; j = 1,2, \ldots, 8$$

Among them, $\overline{x_j} = \frac{\sum_{i=1}^{n} x_{ij}}{n}, {s_j}^2 = \frac{\sum_{i=1}^{n}(x_{ij} - \overline{x_l})^2}{n-1}$, we get the normalized matrix Z.

Next, we find the correlation coefficient matrix R for the standardized matrix Z, and the formula is as follows:

$$R = [r_{ij}]_8 x_8 = \frac{Z^T Z}{n-1}$$

Among them, $r_{ij} = \frac{\sum z_{ij} . z_{ij}}{n-1}, i = 1,2, \ldots, 15; j = 1,2, \ldots 8$.

Then, we use matlab to solve the characteristic equation of the sample correlation matrix R, that is

$$\left|R - \lambda I_p\right| = 0 \text{ (p = 8)}$$

The characteristic roots and contribution rates of 8 influencing factors are obtained as follows:

Table 1. The characteristic roots and contribution rates of 8 influencing factors

influence factor	Characteristic root	Contribution rate
C8	3.679953043	0.459994130
C7	1.975128144	0.246891018
C6	1.242128815	0.155266102
C5	0.629546145	0.078693268
C4	0.255803925	0.031975491
C3	0.147600000	0.018400000
C2	0.069038448	0.008629806
C1	8.26E-04	1.03E-04

We follow $\frac{\sum_1^m \lambda_j}{\sum_1^8 \lambda_j} \geq 0.85$ The number of principal components M = 3 is determined, and the information utilization rate is more than 85%. In this way, most of the information of the original data can be extracted basically. In order to simplify the calculation, we use these three main components, namely, international trade, industrial structure and energy consumption structure, instead of the original eight factors, to obtain the general form of multiple linear regression model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$$

Where y is the dependent variable, β_0 As a constant. $\beta_i (i = 1,2,3)$is the partial regression coefficient with respect to each variable, ϵ For residuals.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202006_6(6).0024

According to the idea of theoretical regression equation.

$$\frac{\partial y}{\partial x_i} = \beta_i (i = 1,2,3)$$

The partial regression coefficient of each component is determined, and then the final regression model is obtained by using the least square method

$$y = 0.078217955 + 0.19853978 * x_1 + 0.246784686 * x_2 + 1.873965673 * x_3$$

At the same time, we get the correlation coefficient R ^ 2 is 0.8418, the significance probability index p value is 5.9713e-04, which is far less than 0.05, which shows that our model has a better overall fitting effect.

2.5.2 Prediction model

We have completed principal component analysis and successfully extracted the principal components we need. Next, we will fit the annual change trend of these three principal components. In order to find the fitting function suitable for our model, we made many attempts, and finally chose the function with the best fitting effect, the general form of which is: Fourier

$$y = a_0 + a_1 * \cos(wx) + b_1 * \sin(wx) + a_2 * \cos(2wx) + b_2 * \sin(2wx) + \cdots + a_n * \cos(nwx) + b_n * \sin(nwx)$$

Among them, $a_i, b_i (i = 1,2, \dots, n)$Constant coefficient, w is constant.

Next, according to the data characteristics of each principal component, we analyze and fit the prediction model suitable for each principal component.

2.5.2.1 International trade forecast model

$$x_1 = 0.2547 - 0.23 * \cos(t * 0.4145) + 2.3 * \sin(t * 0.4145) - 0.02418 * \cos(2 * t * 0.4145) + 0.04 * \sin(2 * t * 0.4145)$$

Correlation coefficient: 1

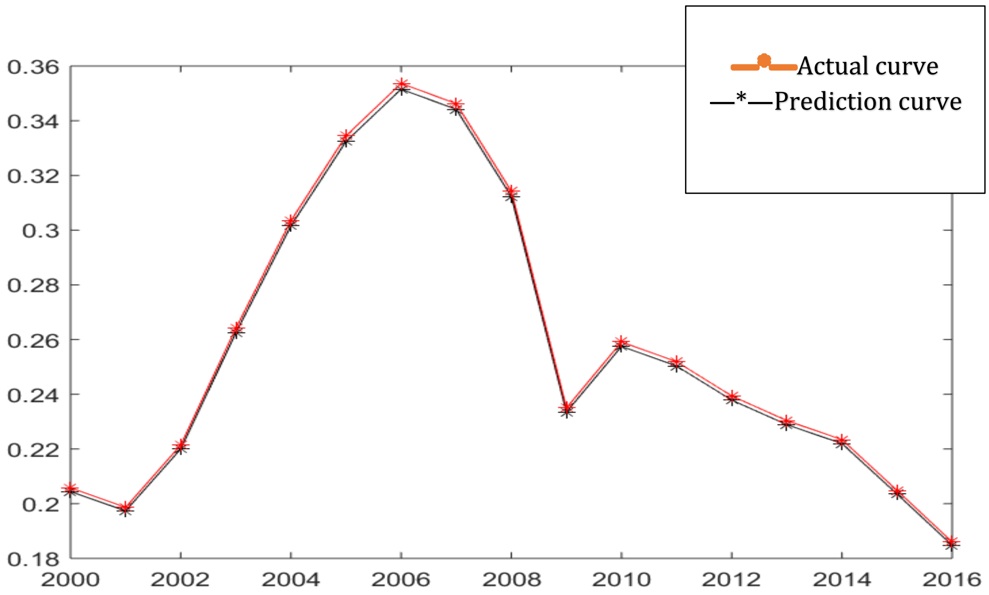

Figure 5. International trade forecast model and actual comparison

2.5.2.2 Prediction model of industrial structure

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202006_6(6).0024

$$x_2 = 0.4409 - 0.027 * \cos(t * 0.2519) + 0.022 * \sin(t * 0.2519) + 0.004 * \cos(2 * t * 0.2519) + 0.016 * \sin(2 * t * 0.2519) + 0.003 * \cos(3 * t * 0.2519) + 0.009 * \sin(3 * t * 0.2519) - 0.0035 * \cos(4 * t * 0.2519) + 0.0007 * \sin(4 * t * 0.2519)$$

Correlation coefficient: 0.9895

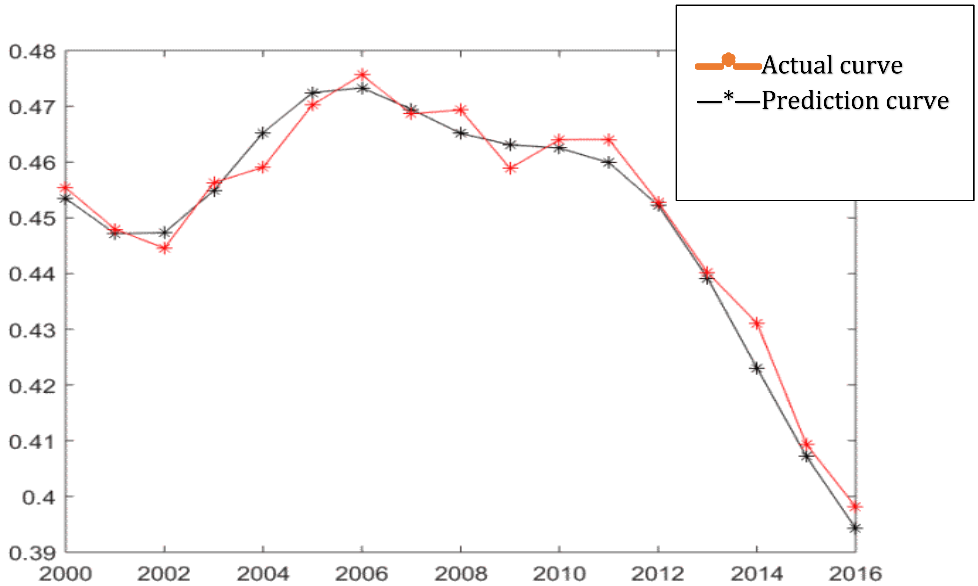

Figure 6. Industrial structure prediction model and actual comparison

2.5.2.3 Energy consumption structure model

$$x_3 = 0.68 - 0.029 * \cos(t * 0.2519) + 0.03 * \sin(t * 0.2519) + 0.0109 * \cos(2 * t * 0.2519) + 0.017 * \sin(2 * t * 0.2519) + 0.006 * \cos(3 * t * 0.2519) + 0.016 * \sin(3 * t * 0.2519)$$

Correlation coefficient: 0.9820

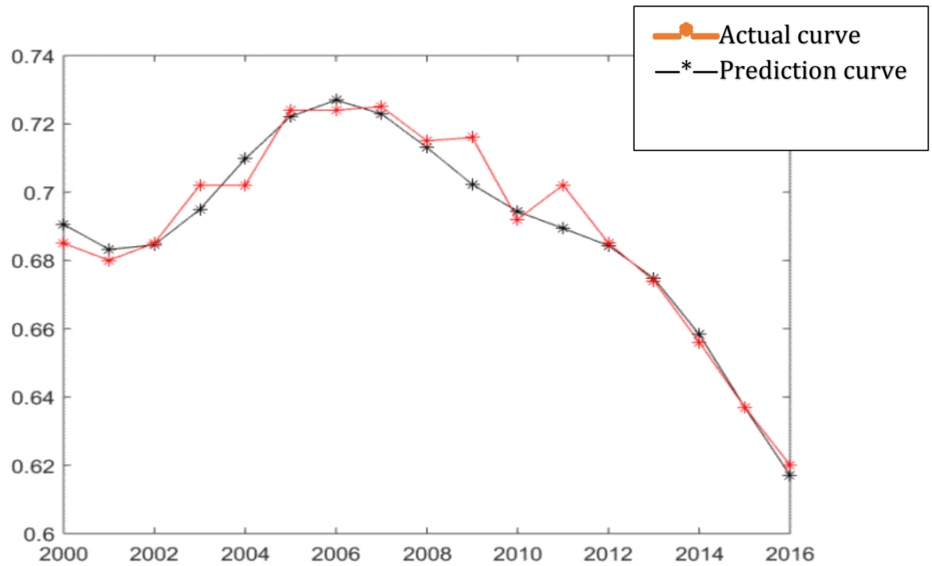

Figure 7. Comparison between prediction model of energy consumption structure and Practice

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202006_6(6).0024

3. RESULT AND DISCUSSION

3.1. Carbon Emission Prediction Model

By combining the above three principal component prediction models and the regression model calculated above, we can deduce the annual growth rate of carbon emissions in year t:

$$\begin{cases} y(t) = 0.078217955 + 0.19853978 * x_1 + 0.246784686 * x_2 + 1.873965673 * x_3 \\ x_1(t) = 0.2547 - 0.23 * \cos(t * 0.4145) + 2.3 * \sin(t * 0.4145) - 0.02418 * \cos(2 * t * 0.4145) + 0.04 * \sin(2 * t * 0.4145) \\ x_2(t) = 0.4409 - 0.027 * \cos(t * 0.2519) + 0.022 * \sin(t * 0.2519) + 0.004 * \cos(2 * t * 0.2519) + 0.016 * \sin(2 * t * 0.2519) + 0.003 * \cos(3 * t * 0.2519) + 0.009 * \sin(3 * t * 0.2519) - 0.0035 * \cos(4 * t * 0.2519) + 0.0007 * \sin(4 * t * 0.2519) \\ x_3(t) = 0.68 - 0.029 * \cos(t * 0.2519) + 0.03 * \sin(t * 0.2519) + 0.0109 * \cos(2 * t * 0.2519) + 0.017 * \sin(2 * t * 0.2519) + 0.006 * \cos(3 * t * 0.2519) + 0.016 * \sin(3 * t * 0.2519) \\ x_i = \dfrac{f_i(t) - f_i(t-1)}{f_i(t-1)}, i = 1,2,3 \end{cases}$$

Similarly, the growth rate y(i) of carbon emissions in any I-year can also be obtained.

From the general formula of growth rate:

$$\mathrm{y(n)} = \frac{f(n) - f(n-1)}{f(n-1)} * 100\%, \mathrm{n} = 1,2, \ldots$$

We can calculate the total carbon emission in year t:

$$\mathrm{C(t)} = \mathrm{C(0)} * \prod_{i=1}^{t} (y(i) + 1)$$

When the total amount of carbon emission in the T-1 year is known, the above formula can be simplified as:

$$\mathrm{C(t)} = (\mathrm{y(t)} + 1) * \mathrm{C(t-1)}$$

Among them, C (2000) is the total carbon emission in 2000 and y(i) is the growth rate of carbon emission in i year.

Based on the calculation formula of total carbon emission and its annual growth rate, we get the prediction model of final carbon emission, and draw the prediction curve and the real value in the same coordinate system, and get the following curve chart:

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202006_6(6).0024

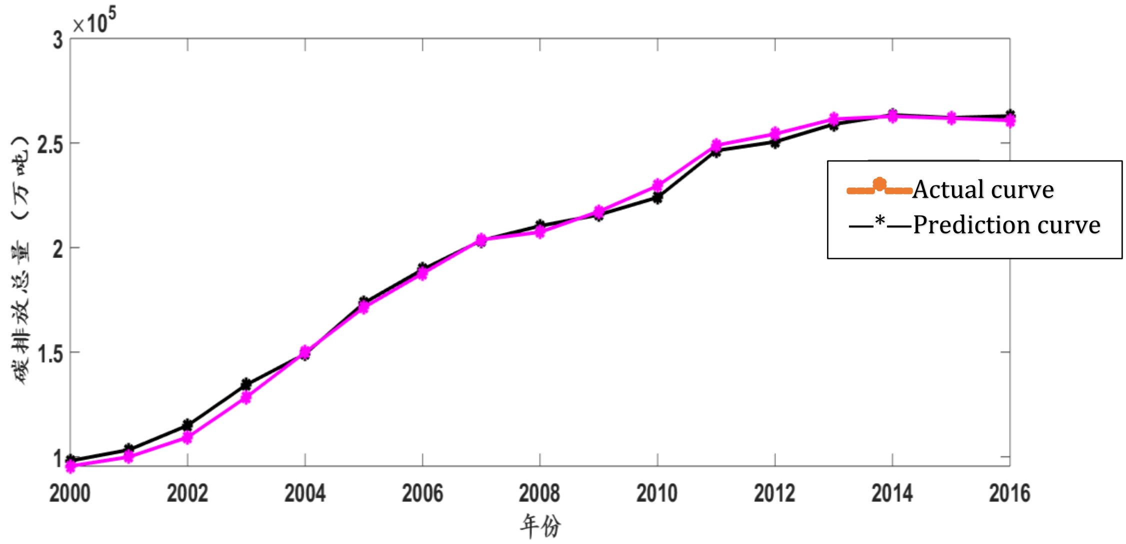

Figure 8. Carbon emission prediction model and actual comparison

3.2. Error Analysis

Our prediction model is based on the data from 2000 to 2014. Next, we will use our prediction model to calculate the total carbon emissions in 2015 and 2016, and compare with the real value to verify the accuracy of the prediction model.

According to the prediction model, C (2015) = 262084.4178, error:

$$\Delta = \frac{|261805.7824 - 262084.4178|}{261805.7824} = 0.1064\%$$

C (2016) = 261812.5617, error:

$$\Delta = \frac{|261812.5617 - 260834.8387|}{260834.8387} = 0.3748\%$$

The error is less than 0.5%, all of which are within the allowable range, which proves that our carbon emission prediction model has high reliability.

3.3. Model Evaluation

Advantage:

1. In the growth rate regression model, we use the method of principal component analysis to find the contribution rate of each influencing factor to carbon emissions, and extract the principal component whose cumulative contribution rate to carbon emissions is more than 85%, and the significance probability index p value is far less than 0.05, so the fitting effect is good. This idea of "dimension reduction" enables us to determine the change trend of China's carbon emissions in fewer factors, which is not only conducive to our group's analysis of carbon emissions, but also ensures the accuracy of the model.

2. In the prediction model, we integrate the characteristics of our own model, choose the fitting function suitable for each prediction model, and the correlation coefficient of each model can reach more than 0.95.

Disadvantages:

1. When we choose the principal component, we assume that these eight factors are independent of each other, but in fact, they are not completely independent; and when we analyze the contribution rate of the principal component, we assume that it is a linear

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202006_6(6).0024

relationship with carbon emissions, and in the actual fitting process, we find that the linear fitting effect is not perfect.

2. We assume that there is no carbon emission in the process of generation and use of primary power and other energy, but there will be carbon emission in its actual life cycle. Its proportion is small but the total amount is huge, so it cannot be ignored.

4. CONCLUSION

We found that among the factors affecting carbon emissions, the proportion of energy consumption structure, industrial structure and international trade ranked the top three. It can be seen that adjusting energy consumption structure, accelerating industrial transformation and reasonably controlling GDP growth play an important role in the realization of emission reduction.

For the regression results and prediction analysis of the model, we put forward the following suggestions:

1. Control the total amount of carbon emissions and reduce the proportion of coal in China's energy consumption structure. Due to the high carbon emission coefficient and high consumption of coal, this is the biggest challenge for our emission reduction target. Increase the development of natural gas, electric power and other energy, call for the use of low-carbon energy, and strive to achieve coal substitutability, which will shake China's position of coal energy output, effectively change the energy consumption structure, and achieve low-carbon.

2. Accelerate industrial transformation, promote economic growth and reduce carbon emissions at the same time. When our country is in the industry-oriented economy, with the economic growth, it will inevitably be accompanied by a large number of energy consumption and carbon emissions. According to the survey, the tertiary industry can promote consumption, promote GDP growth, and carbon emissions are far less than that of the secondary industry. If China can successfully transform into a service-oriented economy, it will certainly contribute a lot to China's carbon reduction.

3. We will increase research on clean energy and increase its penetration rate. In the table, we can see that the proportion of primary electric power and other energy in energy consumption increases year by year, but the growth rate is extremely slow, which is closely related to the cost of clean energy and technology. Increase the investment in clean energy research, make clean energy achieve "low cost, high yield", realize the replacement of coal energy output, and fundamentally reduce carbon emissions.

4. Increase low-carbon propaganda, popularize the concept of low-carbon life, and encourage low-carbon behavior. Due to the huge population in China, the implementation of the propaganda of "low-carbon life" and "green travel" can make the concept of low-carbon deeply rooted in the hearts of the people, achieve low-carbon for the whole people, and effectively reduce the total amount of carbon emissions in China.

REFERENCES

[1] Wang Bing, Zhang Xian, carbon emission prediction research based on improved principal component regression model, global sustainable development report, gsdr brief No. 37cn, 2015.

[2] Jiang Jinhe, carbon emission measurement and analysis of influencing factors in China, resource science, 1007-7588 (2011) 04-0597-082011.

[3] Fu Zhihong, analysis of calculation methods of implicit problems based on input-output table, research on quantitative economy, technology and economy, f233; f2052018.