Environment, Resource and Ecology Journal (2023)
Clausius Scientific Press, Canada
DOI: 10.23977/erej.2023.070502
ISSN 2616-3756 Vol. 7 Num. 5

Research on Carbon Emission Levels of 30 Provinces in China Based on Factor Analysis

Ziyao Dong

School of Economics and Management, Hebei University of Technology, Tianjin, 300131, China

Keywords: Carbon emission level, factor analysis, principal component analysis

Abstract: Based on the research background of the "dual carbon" goals, this study explores the carbon emission levels of provinces across the country. This article uses principal component analysis and factor analysis to study the carbon emission levels of 30 provinces in China from 2016 to 2020. The research results indicate that: (1) The level of economic development is the main factor affecting the carbon emission levels of 30 provinces in China based on factor analysis. (2) Overall, carbon emissions in the eastern and central regions have increased, while carbon emissions in the northwest and northeast regions have decreased. Therefore, the following measures should be taken: strengthen regional communication and optimize resource allocation.

1. Introduction

On September 22, 2020, China proposed for the first time the "dual carbon" target, an important commitment made by China to the international community, which not only concerns the sustainable development of the Chinese nation, but also the common destiny of human beings all over the world. Achieving peak carbon and carbon neutrality is an extensive and profound economic and social systemic change, facing unprecedented difficulties and challenges.

At present, scholars at home and abroad mainly study carbon emissions in the following three aspects: first, to study the influencing factors of carbon emissions; second, to study the spatial and temporal characteristics of carbon emissions; and third, to design as well as predict the peak path. Yang[1]et al. based on the geographically weighted regression method to study the influencing factors of carbon emissions in China's provinces, among which total electricity consumption and total fossil energy consumption have the greatest influence on carbon emissions. Liu[2]et al. analyzed the panel data of 30 provinces and cities nationwide from 2000-2018 using the panel quantile STIRPAT model, and the results showed that per capita disposable income and industrial structure play a promoting effect, and urbanization level, average household size, and technological innovation level play a suppressing effect on carbon emissions. Jiang[3]et al. used the Probit model to find that the carbon peaking time varies significantly among the provincial regions in China, with a north-south strip aggregation in the spatial pattern. Wang[4]et al. analyzed the spatial effects of carbon emission intensity in 30 Chinese provinces from 2015 to 2017 based on the EKC model and STIRPAT model using exploratory spatial data analysis and the spatial Durbin model. The results showed that from 2005 to 2017, China's carbon emission intensity gradually decreased from east to west and from south to north, China's inter-provincial carbon emission intensity showed a clustering effect in space, and the clustering effect gradually weakened over time. Based on the LMDI model and the predefined targets of economic and social development announced by relevant policies, Chen[5]et al. predicted CO2 emissions to peak in 2027 at about 110.87 Mt. Li[6]et al. used the stochastic effects of population, affluence and technology regression to predict China's 2015 to 2035 s peak carbon emission model, and the results showed that China's peak carbon emissions would be reached in 2024, 2027, and 2030 under three different scenarios of low, medium, and high scenarios, respectively.

In the previous studies, it was found that scholars mostly focused their research on carbon emissions on the above three aspects, and less analysis was done on the comprehensive evaluation of carbon emission levels. At the same time, there are problems such as excessive disparity and uneven distribution of carbon emission levels among various provinces and regions in China, so it is especially important to evaluate the provincial carbon emission levels objectively. Based on this, this article selects the data related to carbon emissions of 30 provinces in China from 2016 to 2020 and calculates the comprehensive score of carbon emission levels of 30 provinces in China by using principal component analysis and factor analysis methods. At the same time, to visually represent the characteristics of carbon emission level changes, ArcGIS software is used to draw carbon emission zoning maps with dynamic changes in 2016, 2020, and 2016-2020, respectively, to analyze the changes in carbon emission levels in each province during the five years, and to analyze the reasons for spatial distribution in combination with the characteristics of the spatial distribution of carbon emission levels nationwide, and then put forward relevant suggestions.

2. Model introduction

2.1 Principal component analysis

The principal component analysis is a commonly used statistical method for data that transforms multiple indicators into several representative principal components, which are linear combinations of the original variables, and the new principal components are not only uncorrelated with each other but also reflect the vast majority of information from the original data[7].

2.2 Factor analysis

Factor analysis refers to the principle of dimensionality reduction based on the premise of ensuring the maximum retention of the original data, replacing the intricate and high-dimensional variables with public factors, i.e., grouping variables according to their correlation magnitude, increasing the correlation of variables in the same group, and transforming variables in different groups into uncorrelated or lower correlation, each group then representing a public factor, and finally decomposing the original variables into two parts and forms. One is a linear function composed of a few common factors, and the other is a special factor. The advantage of this approach is the ability to identify a few representative principal factors from many variables with overlapping information and complex relationships, thus simplifying the complex problem[8].

The model for the factor analysis is as follows.

$$
\begin{cases}
x_1 = a_{11} F_1 + a_{12} F_2 + \ldots + a_{1m} F_m + \xi_1 \\
x_2 = a_{21} F_1 + a_{22} F_2 + \ldots + a_{2m} F_m + \xi_2 \\
\bullet \\
\bullet \\
\bullet \\
x_n = a_{n1} F_1 + a_{n2} F_2 + \ldots + a_{nm} F_m + \xi_n
\end{cases}
\tag{1}
$$

In Eq. (1), X_1, X_2, X_3,..., X_n are called the original variables, F_1, F_2, …, F_m are called the common factors, which are the factors appearing in the expressions of the original variables; a_{n1}, a_{n2}, …, a_{nm} are called the factor loadings, and the larger the absolute value of aij ($|a_{ij}|\leq 1$) means the greater the loading of the common factor F_j on X_i. ξ_1, ξ_2, …, ξ_n are special factors, and this part cannot be explained by the original variables.

The steps of the factor analysis study are shown in Figure 1.

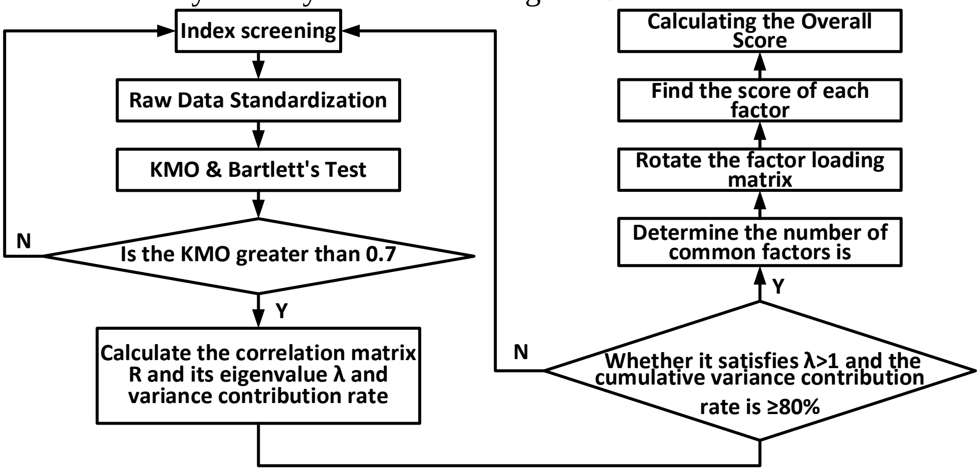

Figure 1: Flow chart of factor analysis steps.

3. Example analysis

3.1 Indicator selection and data sources

In this paper, concerning previous research literature, a total of nine indicators were selected and considered to construct the indicator system as shown in Table 1. And 30 provincial regions in China (except Tibet, Taiwan, Hong Kong, and Macao) are taken as the research objects, and the factors influencing the level of provincial carbon emissions in China in 2016-2020 are studied. Among them, socioeconomic development data were obtained from the 2016-2020 China City Statistical Yearbook[9] and Statistical Communique of the People’s Republic of China on the 2016-2020 National Economic and Social Development of the 30 regions, and energy consumption data were obtained from the 2016-2020 China Energy Statistical Yearbook[10]. Due to the different indexes with different measurement levels, the raw data were standardized to ensure the validity of the data. The standardization formula is as follows:

$$
Z_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \tag{2}
$$

In Eq. (2), X_i is the original data and Z_i is the normalized data.

Table 1: Content of indicators.

Influencing Factors	Indicator Name	Unit	Reference
Economic Development Level	Real GDP percapita(X_1)	yuan	[11]
Per capita income level	Disposable income per capita for all residents(X_2)	yuan	[12]
Urbanization level	Urban population as a share of the resident population(X_3)	%	[13]
Industry Structure	Share of secondary industry in GDP(X_4)	%	[14]
Technology Innovation Level	Internal expenditure of R&D funds as a percentage of GDP(X_5)	%	[12]
Foreign trade level	Total imports and exports of the operating unit location as a percentage of GDP(X_6)	%	[15]
Traffic level	Road passenger traffic(X_7)	10000 persons	[16]
Energy intensity	Total energy consumption as a share of GDP(X_8)	%	[17]
Level of environmental regulation	Industrial pollution control completed investment as a proportion of GDP(X_9)	%	[18]

3.2 Bartlett's sphericity test

To ensure the validity of the factor analysis results, it is necessary to conduct Bartlett's sphericity test on the variables to judge whether the Sig value and KMO value meet the requirements, and the test results are shown in Table 2. As can be seen from Table 2, the KMO value is 0.774, which is greater than the critical value of 0.6 and suitable for factor analysis; Bartlett's sphericity test results show that the significance level is much less than 0.05, indicating that the above data meet the requirements of factor analysis.

Table 2: KMO and Bartlett's Test.

KMO sampling suitability quantity		0.774
Bartlett Test	Approx. Chi-Square	1186.193
	Degrees of freedom	36
	Significant	0.000

3.3 Example measurement

Principal component analysis was performed on all indicator data (X_1 to X_9) after standardization to extract principal components. The calculated eigenvalues and variance contribution rates of each factor are shown in Table 3.

From Table 3, we can see that there are three eigenvalues with eigenvalues greater than 1, λ_1, λ_2, and λ_3, whose values are 3.794, 1.881, and 1.661, respectively, and the variance contribution rates corresponding to the three eigenvalues are 42.153%, 20.895%, and 18.457%, respectively, and the cumulative variance contribution rate is 81.505%, which indicates that these three principal components explain 81.505% of the total variance, indicating that the loss of data information was less, and the factor extraction was more successful, so the three principal components were extracted as Y_1, Y_2, and Y_3. Among them, the first principal component contains more information and has a greater influence on the provincial carbon emission level.

The rotated component matrix is shown in Table 4, from which it can be seen that X_1, X_2, X_3 , and X_6 have larger loadings on the first principal component Y_1, which mainly reflects the basic economic situation and urbanization rate, and can be named the economic development level factor (F_1); X_7, X_8 , and X_9 have larger loadings on the principal component Y_2, which mainly reflects the transportation and energy consumption situation and environmental pollution control level, which can be named as environment and energy utilization efficiency factor (F_2); X_4 and X_5 have larger loadings on the third principal component Y_3, which mainly reflects the industrial structure and R&D input level, and can be named as industrial structure and innovation level factor (F_3). Among them, F_1 is the main influencing factor of carbon emission level, and the rapid economic development is generally accompanied by the increase in carbon emission level, and the different economic development levels among different provinces and regions will inevitably bring about different carbon emission levels.

Finally, to accurately analyze the carbon emission levels of the 30 provincial areas, a weighting operation is required to obtain a composite score. Through the Score process, the expression of the combined carbon emission level score I is calculated as:

$$I = \frac{0.42153}{0.81505} * F_1 + \frac{0.20895}{0.81505} * F_2 + \frac{0.18457}{0.81505} * F_3 \tag{3}$$

Table 3: Total Variance Explained

factor	Eigen			% of Variance(Unrotated)			% of Variance(Rotated)		
	Eigen Value	% of Variance	Cumulative % of Variance	Eigen Value	% of Variance	Cumulative % of Variance	Eigen Value	% of Variance	Cumulative % of Variance
1	4.440	49.333	49.333	4.440	49.333	49.333	3.794	42.153	42.153
2	1.728	19.197	68.530	1.728	19.197	68.530	1.881	20.895	63.048
3	1.168	12.974	81.505	1.168	12.974	81.505	1.661	18.457	81.505
4	0.674	7.489	88.993						
5	0.446	4.961	93.954						
6	0.280	3.110	97.064						
7	0.150	1.666	98.729						
8	0.085	0.944	99.673						
9	0.029	0.327	100.000						

Table 4: Component Score Coefficient Matrix

	Component		
	1	2	3
Real GDP percapita (X_1)	0.934	-0.144	-0.196
Disposable income per capita for all residents (X_2)	0.936		-0.273
Urban population as a share of the resident population (X_3)	0.949		-0.109
Share of secondary industry in GDP (X_4)	0.918	-0.205	-0.104
Internal expenditure of R&D funds as a percentage of GDP (X_5)		-0.821	0.280
Total imports and exports of the operating unit location as a percentage of GDP (X_6)	-0.313	0.832	0.196
Road passenger traffic (X_7)		0.664	0.523
Total energy consumption as a share of GDP (X_8)	-0.127		0.890
Industrial pollution control completed investment as a proportion of GDP (X_9)	0.423		-0.585

3.4 Results and Analysis

Through the above calculations, a total of 150 composite scores of carbon emission levels for 30 provincial domains were obtained. First, the descriptive statistics of the composite score I are shown in Table 5, and to ensure the statistical distribution of the data samples, the I values are normalized and the expression is:

$$\text{Normalized } I = \frac{I-(-0.97)}{1.62-(-0.97)} \tag{4}$$

Using ArcGIS, the natural intermittent point grading method (Jenks) is selected to classify the normalized I values, and four categories are selected in this paper. According to the classification results, the normalized I values are classified into low, lower, high, and higher according to different levels of carbon emission levels, and the statistical results are shown in Fig.2. According to Fig.2, it can be seen that the number of provinces with high carbon emission levels is the least in 2016, accounting for only 10%. In 2020, the number of provinces with high carbon emission levels increases significantly, increasing by 200% compared to 2016.

Table 5: Composite Score I Descriptive Statistics.

	N	Range	Min	Max
Score	150	2.59	-0.97	1.62
Number of effective cases (in columns)	150			

The evaluation levels of carbon emission levels of 30 provincial areas in 2016 and 2020 are shown in Fig.3. and Fig.4. As shown in Fig.3 and Fig.4, the carbon emission levels of about 27% of the national provinces in 2020 compared with 2016 have increased and 20% of the provinces have decreased. Among them, the carbon emission level of the southeast coastal provinces improves overall, such as Jiangsu, Shanghai, Zhejiang, Fujian, and Guangdong, all improve from higher to high levels. A few inland provinces, such as Beijing and Chongqing, also saw an increase in carbon emissions. The provinces with unchanged carbon emission levels are concentrated in the western and northeastern regions of China.

From the change of carbon emission level composite score I, the provinces with a two-year difference within ±5% are defined as the provinces with flat carbon emission levels, greater than 5% is the level increase, and less than -5% is the level decrease. The change in the carbon emission level of 30 provinces from 2016 to 2020 is shown in Fig.5, from Fig.5, it can be seen that in 2020, compared with 2016, about 43% of the provinces have an increase in carbon emission level. About 27% of the provincial areas have the same carbon emission level, and 30% of the provincial areas have a decrease in carbon emission level. Most of the provinces with decreasing carbon emission levels are in the western region, such as Inner Mongolia, Qinghai, and Guangxi. The Northeast region also shows a decreasing trend in carbon emission levels.

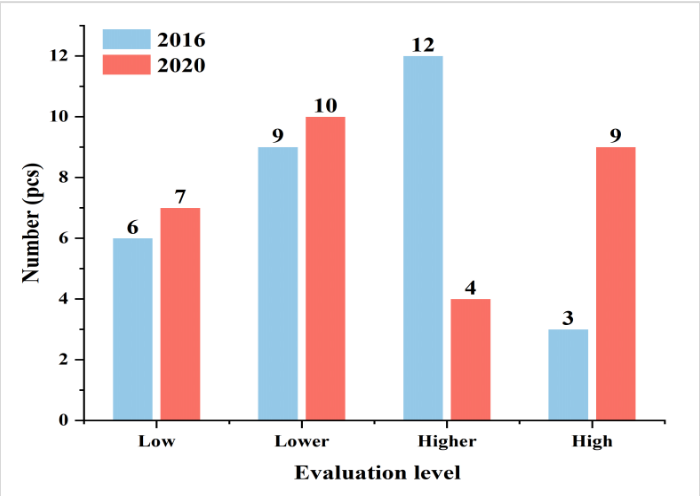

Figure 2: Carbon emission levels for 2016 and 2020

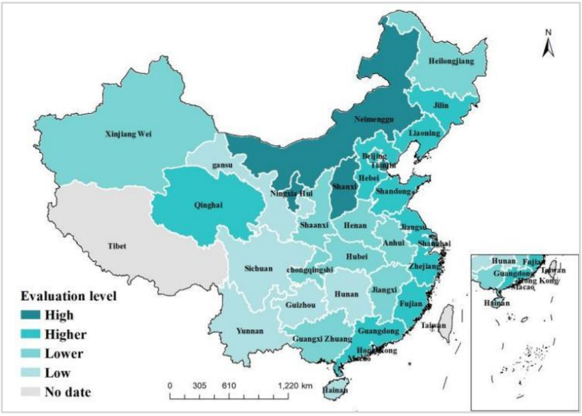

Figure 3: 2016 Carbon emission level evaluation level

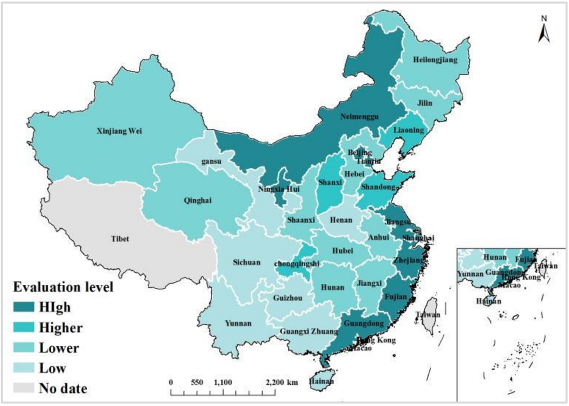

Figure 4: 2020 Carbon emission level evaluation level

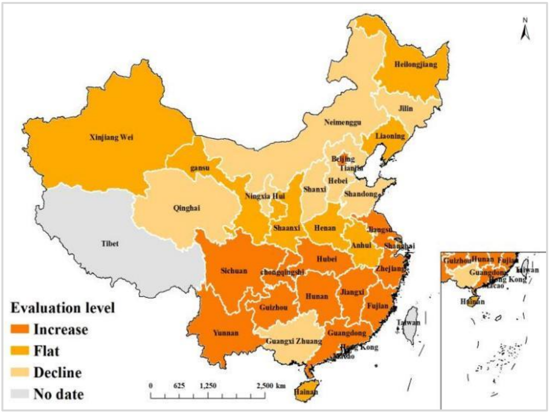

Figure 5: Change in carbon emission level from 2016 to 2020.

4. Conclusions and Recommendations

China's provincial carbon emissions show an overall increase in the level of the eastern and central regions and a decrease in the level of the northwest and northeast regions. The level of economic development is the main reason for the above spatial distribution of carbon emission levels. In recent years, with the gradual expansion of foreign trade and the improvement of science and technology innovation in the eastern coastal region, the level of economic development has been improving. The economic development level of southwest and central regions has also been improving due to their location advantages and steady development. However, the northwest and northeast regions have slowed down their economic development in recent years due to their natural geographical conditions and lack of development momentum, respectively. Based on the above conclusions, the following recommendations are obtained:

(1) Strengthening regional communication and optimizing resource allocation. The Beijing-Tianjin-Hebei region has abundant resources and talents and should build a reasonable industrial chain and supply chain by optimizing the allocation of resources, which is conducive to narrowing the development gap between Beijing, Tianjin, and Hebei and improving the overall efficiency of resource utilization. To focus on the high energy consumption and high pollution industries in Hebei, to fundamentally improve the coal-based energy industrial system, we should vigorously develop clean energy and renewable energy; at the same time, we should vigorously develop a low-carbon economy, introduce new technologies and develop new energy sources to reduce high-carbon energy consumption. Beijing, Tianjin, and Hebei should strengthen regional communication to achieve energy-saving and emission reduction technology sharing, and environmental pollution management, and promote the collaborative high-quality development of the three regions, which will help to enhance the carbon emission regulation ability of Hebei.

(2) Since the economic development of the northwest region cannot be separated from the consumption of fossil energy, the resource economy accounts for a large proportion. We should accelerate the transformation of economic development mode, and realize the transition from resource dependence to technology dependence through technology introduction and talent introduction so that economic development is no longer dependent on fossil energy consumption, which will help to improve the level of economic development in Northwest China. Meanwhile, in Northwest China, heavy industries are the main source of carbon emissions, and the government should strongly support the transformation and upgrading of traditional industries to low-carbon industries to promote the development of a green economy in Northwest China, which helps to achieve sustainable development.

(3) The Northeast should take advantage of its location, take the initiative to optimize and improve the business environment and enhance the momentum of regional economic development. Seize the important opportunities of "One Belt, One Road" and RCEP, expand the opening to the outside world, actively expand the import and export trade and modern trade industry, and obtain new economic growth points in foreign trade. At the same time, as one of the important old industrial bases in China, the northeast region has the problem of a single industrial structure. It should take into account the development of the times and strive to build high-tech industries such as information technology and biotechnology to break the single industrial layout and achieve new economic development in the northeast region.

References

[1] Yang Q, Peng R, Liu X, et al. Study on influencing factors of provincial carbon emission based on geographically weighted regression [J]. Journal of Environmental Engineering Technology, 2023, 13(01):54-62.

[2] Yunlong L, Xianlin C, Chengfeng H. Research and Analysis on the Influencing Factors of China's Carbon Emissions Based on a Panel Quantile Model[J]. Sustainability, 2022, 14(13).

[3] Jiang Y, Zhong S, Wang Y, et al. Spatio-temporal characteristics and influencing factors of carbon emission peak by province of China [J]. Journal of Natural Resources, 2022, 37(05):1289-1302.

[4] Yingdong W, Yueming Z. Spatial effects of carbon emission intensity and regional development in China. [J]. Environmental science and pollution research international, 2020, 28(11).

[5] Chen T, Li X, Chen B. Decomposition of influencing factors and peak prediction of carbon emissions in China[J]. Journal of Safety and Environment, 2022:1-12.

[6] Li L, Lei Y, He C, et al. Prediction on the Peak of the CO2 Emissions in China Using the STIRPAT Model[J]. Advances in Meteorology, 2016, 2016.

[7] Chen Y, Cheng Q, Fang X, et al. Principal component analysis and long short-term memoryneural network for predicting dissolved oxygen in water for aquaculture[[J]. Transactions of the Chinese Society of Agricultural

Engineering, 2018, 34(17):183-191.
[8] Xiaowei B I, Manman L, Dongsheng F, et al. A Target Threat Sequencing Method Based on Factorial Analysis [J]. Electronic Information Warfare Technology, 2023, 38(2):1-7.
[9] China City Statistical Yearbook-2016-2020 Editorial Board and Editorial Board China City Statistical Yearbook-2016-2020 Editorial Board and Editorial Board [Z]. China City Statistical Yearbook, 2016-2020: 4-5. [C].
[10] China Energy Statistical Yearbook-2016-2020 Editorial and publishing staff China Energy Statistical Yearbook-2016-2020 Editorial and publishing staff [Z]. China Energy Statistical Yearbook, 2016-2020: 4-5. [S].
[11] Yin N, Wang C, Yan P. Provincial carbon emission efficiency and driving factors along the Belt and Road in China under carbon peak and carbon neutrality [J]. Coal Engineering, 2022, 54(09):182-186.
[12] Wang Q, Wang S, Dong N. Research on the Influencing Factors and SpatialSpillover of Carbon Emission Intensity Under the Background of Carbon Peak—Based on the Empirical Analysis of Chongqing[J]. Soft Science, 2022, 36(07):97-103.
[13] Liu Z, Zhang Y. Temporal and Spatial Transition Characteristics, Influencing Factors and Peak Path Design of Carbon Emissions in China [J]. Journal of Southwest University (Social Sciences Edition), 2022, 48(06):99-112.
[14] Wang Z, Li Z, Wu W. Spatio-Temporal Evolution and Influencing Factors of Carbon Emissions in Different Grade Cities in the Yangtze River Economic Belt [J]. Research of Environmental Sciences, 2022, 35(10):2273-2281.
[15] Hu J, Wang K. Spatial and Temporal Differences and Spatial Convergence of Carbon Emission Efficiency in Chinese Provinces [J]. Journal of Management, 2022, 35(04):36-52.
[16] Zhao Q, Zhou Y, Fang Q, et al. Spatial-temporal evolution of carbon emissions and its influencing factors inChina central region[J]. Acta Scientiae Circumstantiae, 2023, 43(02):354-364.
[17] Wang S, Ye L, Zhang W. Dynamic Effect Analysis and Scenario Prediction of Factors Influencing Carbon Emission Intensity in Beijing, Tianjin and Hebei[J]. Urban Problems, 2021(05):42-51.
[18] Liu Z, Xu J. Equity and influence factors of China's provincial carbon emissions under the "Dual Carbon" goal [J]. Scientia Geographica Sinica, 2023, 43(01):92-100.