DLINE JOURNALS

Journal of Digital Information Management

ISSN: 0972-7272

JDIM 2025: 23 (1)

Regional Carbon Emission Prediction and Low-carbon Path Analysis Based on BP Neural Network Model

Xuanke Zhang[1], Lijun Wu[1], Jun Yang[1], Runbin Xue[1], Yong Tong[2], Yali Lei[2], Wenya Fang[2], Ronghua He[2]

[1]Xinjiang Oil and Gas Storage and
Transportation Company, CNPC, Karamay
Xinjiang, 834000, China

[2]Southwest Petroleum University School of Economics and Management
Chengdu, Sichuan, 610500, China
419323813@qq.com

ABSTRACT

To attain carbon emission control and sustainable economic development, tailored low-carbon policies must be adapted to distinct regional contexts. Given disparities in key industries, economic growth, and resource availability, variations in carbon emissions across China's eastern, central, and western regions necessitate divergent low-carbon strategies. To comprehensively grasp regional carbon emissions and provide precise recommendations, a Lasso regression method identified five influential factors from seven, including population, per capita GDP, total energy consumption, energy mix, industrial makeup, urbanization rate, and forest coverage. Analyzing the link between carbon emissions and total output, a predictive model employed a GA-optimized BP neural network to forecast emissions. Findings indicate higher carbon emissions in the developed eastern region due to rapid economic growth, industrial production, and energy use. While the east region maintains emission leadership, emission growth rates converge, reflecting nationwide progress in reduction efforts. Future strategies should focus on regional development, exploring low-carbon paths through energy restructuring, urban optimization, and synergy of energy strengths, thereby achieving shared low-carbon objectives.

Subject Categories and Descriptors: [I.5 PATTERN RECOGNITION] Neural nets; [B.5 REGISTER-TRANSFER-LEVEL IMPLEMENTATION] Data-path design

General Terms: Neural Networks, Carbon Path Analysis

Received: 7 September 2024, Revised 25 November 2024, Accepted 4 December 2024

Keywords: Carbon Emissions, Regional Differences, Lasso Regression, GA-BP Neural Network, Low-carbon Path

Review Metrics: 0/6; Review Score: 4.92; Inter-reviewer Consistency: 74.6%

DOI: https://doi.org/10.6025/jdim/2025/23/1/47-66

1. Introduction

China has proposed the dual carbon goal for 2020 and 2060, aiming to achieve a win-win situation of carbon emission control and sustainable economic development through measures such as controlling carbon emissions and promoting clean energy. The dual carbon goals are a powerful tool for China to fulfil its social responsibility and control global climate change. Coordinated improvement in pollution reduction and carbon reduction is an important approach for China to promote its overall green-oriented transition to economic development, and it is an inevitable choice to achieve sustainable development [1]. Due to the vast land area of China, there are differences in the rapid growth of industrialization, urbanization, and transportation in different regions due to geographical location. As a result, the carbon emissions issues in the eastern, central, and western regions also have various causes.

In this study, we examined the carbon emissions in the eastern, central, and western regions of China and assessed the differences in the total amount of carbon emissions and the intensity of carbon emissions among these regions. Furthermore, we have explored the reasons behind these differences and aimed to find low-carbon emission paths for each region. We utilized Lasso regression and GA algorithm to optimize the BP neural network model and simulate the variable values from 2023 to 2030. This helped us gain insights into the current situation and characteristics of carbon emissions in the eastern, central, and western regions. We hope to provide scientific evidence for setting carbon reduction goals and plans in these regions and promote the overall reduction of carbon emissions in China.

2. Literature Review

Once the "dual carbon" goal was proposed, it became a hot topic in the academic community. The literature review shows that the discussions among researchers on carbon emission prediction can be divided into two categories: research methods and research content. Firstly, in terms of research methods, different approaches are used to improve the accuracy of predictions. Wang [2] considers utilising a combination of neural algorithms to improve prediction accuracy from the perspective of model accuracy. Fang [3] used the IPAT model to address the current issues of complexity and poor accuracy in the model, and they made predictions on carbon emissions in green buildings. Han [4] uses skip connections in the internal residual blocks of RESNET to alleviate the gradient disappearance caused by the increasing depth of deep neural networks, thereby improving the accuracy of the prediction model. Regarding research content, many scholars focus on different factors that affect carbon emissions based on their selected research objects. They use models to determine the most important key factors when identifying multiple potential influencing factors. For example, Guo et al. [5] used grey relational analysis to fit different influencing factors and identify more precise factors related to carbon emissions in the construction industry. Kong et al. [6] employed a two-stage feature selection method composed of a partial autocorrelation function and *ReliefF* to select appropriate input variables for the following prediction process. Zhao J et al. [7] used quadratic programming regression analysis to analyze the influencing factors of carbon emissions from the perspective of regional differences.

This study adopts the dual carbon goal to ensure the coherence of regional selection and achieve more accurate carbon emission prediction. It introduces the GA optimization algorithm into the BP network to establish a carbon emission prediction model. It predicts the total carbon emissions in three regions. Additionally, the Lasso regression is used to identify the five factors that significantly impact the total carbon emissions in each region.

The relationship between these factors and carbon emissions is analyzed to reflect the future differences and trends in carbon emissions in different regions of China. It supports the orderly implementation of China's dual carbon goal and developing region-specific dual carbon policies.

3. Analysis of Factors Influencing Carbon Emission Prediction Models

3.1 Data Source and Factors Influencing Selection

Considering the historical accuracy of the data, the data in this article is based on the China Energy Statistical Yearbook and the China Statistical Yearbook from 2002 to 2021. The data is aggregated by dividing the provinces (municipalities directly under the central government) in eastern, central, and western China, excluding the Hong Kong Special Administrative Region, Macau Special Administrative Region, and Taiwan Province. The east region of China includes Beijing, Tianjin, Hebei, Liaoning, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, and Hainan. The central region includes Heilongjiang, Jilin, Shanxi, Henan, Anhui, Hubei, Jiangxi, and Hunan. The western region includes Inner Mongolia, Guangxi, Chongqing, Sichuan, Yunnan, Guizhou, Shaanxi, Gansu, Ningxia, Qinghai, and Xinjiang. Tibet is not included in the Western region due to incomplete statistical data. It consists of the total carbon emissions, Gross Domestic Product (GDP), population size, and total energy consumption from the East, Central, and West regions. The calculation formula and parameters are based on the "2006 IPCC Guidelines for National Greenhouse Gas Inventories" (2019 revised edition).

After summarizing the researchers' selection of factors affecting carbon emissions, the scope of content is very extensive. Cui et al. [8] utilized previous experiences to select GDP, per capita consumption level, total import and export volume, urbanization rate, and total energy consumption as the dataset for the experiment. Shi et al. [9] utilized LASSO regression to select important influencing factors such as coal and oil consumption, flat glass, pig iron, and crude steel production. Based on previous research, Zhu al. [10] have constructed an initial indicator system, using per capita GDP, national income, total energy consumption, urban population, and industrial added value as factors influencing carbon emissions. Based on China's carbon emission intensity targets and macro development goals in 2020 and 2030, we will construct a carbon emission prediction model using a BP neural network improved by genetic algorithms. In this model, we will consider the factors influencing carbon emissions in China, such as population size, per capita GDP, total energy consumption, energy consumption structure, industrial structure, urbanization rate, and forest coverage.

Carbon emissions are dynamic and time-dependent, [11] and shifting to low carbon emissions is the only way to maintain eco-balance. [12,13,14]. Many research studies have analysed neural network models for predicting residential carbon emissions. [15,16,17,18] The ANN models are proposed, validated, and reliable for predicting the growth of Carbon emission intensity for countries including China. [19] The improved residual neural network is an incremental model proposed by Han et al. [20]

3.2 Determination of influencing factors of the prediction model

Because many factors determine carbon emission, and the influence of these factors is not equal, there is a very complex interaction between the factors. By finely analyzing influencing factors, the wrong assessment of their impact on carbon emissions can be reduced, and the formation mechanism of carbon emissions can be strengthened.

Lasso regression is a linear regression model that can solve the problem of variable selection in high-dimensional data sets. It compresses the model's coefficients through the penalty function, compresses some coefficients to zero, and screens out significant variables [21]. By using the L1 regularization method, the minimized objective function is divided into two parts: one is a mean square error, and the other is the L1 regularization term. Different coefficient estimates can be obtained by adjusting regularization coefficients. The objective function of Lasso regression can be expressed as:

$$\min_{\beta}\frac{1}{2n}\|y-X\beta\|_2^2+\lambda\|\beta\|_1$$

Including, y is and n- dimensional vector, representing the dependent variable; X is an $n*p$-dimensional matrix, representing the independent variable, which is listed as the feature of the independent variable; β is a p dimensional vector representing the regression coefficient to be found. $\lambda \geq 0$ is the hyperparameter that adjusts the weight of the regularization term, $||\beta||_1$ indicates the $L1$ norm. The first term of the objective function is the loss function of least squares, and the second term is the regularization penalty term. The norm penalty term will make part of the feature coefficients 0 to achieve the purpose of feature selection or model compression. The coordinate rotation algorithm can be used to solve for the minimum value of the above objective function.

4. Construction of Total Carbon Emission Prediction Model

4.1 BP Neural Network

BP neural network is a kind of directed graph, and its structure is composed of three parts: input layer, hidden layer, and output layer. BP neural network is a standard feedforward artificial neural network for pattern recognition, classification and prediction. The basic idea is to use the method of gradient descent to constantly adjust the weight and bias in the network so that the output result of the network is as close as possible to the actual result of the training data by propagating the error from the output layer forward, calculating the error signal of neurons in each layer to update the network weight and bias.

During operation, the connection weights ω_{ji} 、ω_{kj} between the input layer, hidden layer and output layer are initialized, and the thresholds a_j and b_k of hidden layer and output layer are initialized. Then, calculate the output H_j of the hidden layer, calculate the output switch Angle θ_k of the output layer and calculate the error between the output switch Angle θ_k of the network and the expected output switch Angle $\theta_{hope.k}$. If e_k does not meet the requirements, The error is reversed-transmitted and updated to correct the weight ω_{ji} between the input layer and the hidden layer, the weight ω_{kj} between the hidden layer and the output layer, as well as the threshold of the hidden layer and the threshold b_k of the output layer until the preset error is satisfied:

$$\omega'_{ji} = \omega_{ji} + \eta H_j(1 - H_j)x_i \textstyle\sum_{k=1}^{2M} \omega_{kj} e_k \quad (1)$$

$$\omega'_{kj} = \omega_{ji} + \eta H_j e_k \quad (2)$$

$$a'_j = a_j + \eta H_j(1 - H_j) \textstyle\sum_{k=1}^{2M} \omega_{kj} e_k \quad (3)$$

$$b'_k = b_k + e_k \quad (4)$$

Among the formulas: η is the learning rate; ω'_{ji}、ω'_{kj}、a'_j、b'_k are the updated values, respectively. The weights and thresholds are adjusted in the direction of a negative gradient.

4.2 Genetic Algorithm

A genetic algorithm represents the problem to be solved as a specific optimization problem, then designs a gene representation scheme and maps the scheme to a fitness function. This function defines the fitness of each individual in the population. The fitness function is gradually optimized through continuous crossover, mutation, and selection, and the optimal solution is obtained.

Genetic algorithms usually involve three main processes: selection, crossover and mutation. The GA algorithm iteratively generates new populations through population selection, crossover, variation and fitness evaluation, etc., until the stop criterion is met. Through selection operation, the GA algorithm selects the retained individuals according to the fitness function value. Common selection methods include roulette selection, tournament selection, etc. Roulette selection formula G is expressed as:

$$P_i = \frac{f_i}{\sum_{j=1}^{n} f_j}$$

Including, P_i represents the probability of the i *th* individual being selected, and f_i is the fitness function value of the *ith* individual. Then, the GA algorithm swaps the chromosome information of the two parent individuals and generates new offspring through the crossover operation. Common crossing modes include single-point crossing, multi-point crossing and so on. Single point crossover formula:

$$c_1 = a_1 + b_2 - b_1$$
$$c_2 = b_1 + a_2 - a_1$$

Including, a_1、a_2 are the two genes of the first parent chromosome, b_1、b_2 are the two genes of the second parent chromosome, and c_1、c_2 are the two genes of the new offspring. Then, the mutation operation is carried out, and the GA algorithm can randomly change individual chromosomes through the mutation operation to increase the diversity of the population. Common variation modes include single-point variation, multi-point variation and so on. Single point variation formula:

$$c_i = !\, a_i$$

Including, a_i is the i[th] gene on the chromosome of the original individual, and c_i is the gene after mutation.

Finally, the fitness evaluation is carried out. The GA algorithm evaluates the adaptability of each individual in the environment through the fitness function to determine which individuals can be selected as the parent individuals of the next generation. Fitness function formula:

$$f(x) = \sum_{i=1}^{n} \omega_i x_i$$

Including, x_i is the individual to be evaluated, n is the dimension of the individual and ω_i is the weight coefficient of the dimension i. Combined Figure 1 shows that the genetic algorithm takes the individuals (chromosomes) in an initial population as the initial solution to the problem and calculates their fitness. Then, through the selection operation, select a part of the individuals with the highest fitness to carry out crossover and mutation operations to generate new individuals. Finally, the optimal solution is obtained through the iterative crossover, mutation and selection process until the predetermined termination conditions are met [22]. The genetic algorithm obtains the optimal solution and assigns the weight threshold to the neural network.

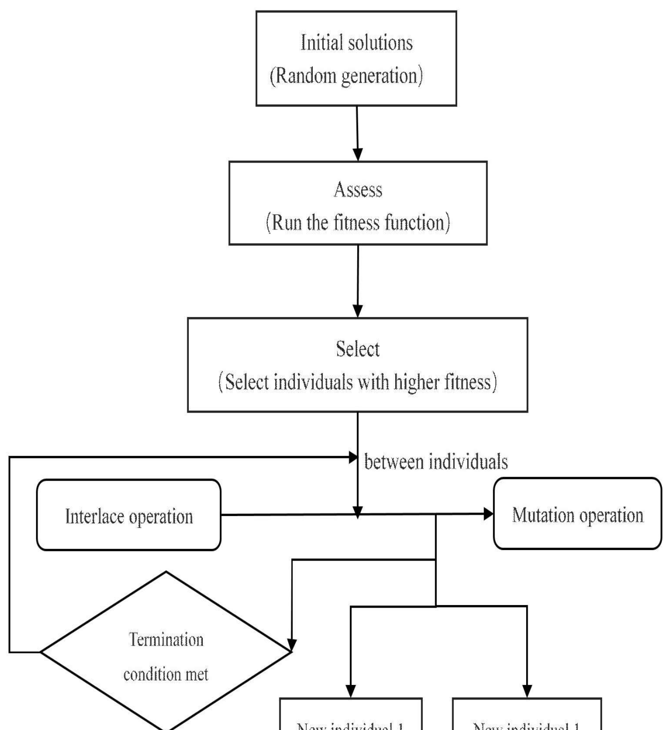

Figure 1. Process structure of genetic algorithm

5. Design of Scenario Value of Carbon Emission Prediction and Determination of Influencing Factors

5.1 Scenario Value Design Of Carbon Emission Prediction

The setting of carbon emission forecast scenario values is based on China's overall development planning objectives and regional development plans for East, Central and Western regions to design reasonable growth for each variable and strive to help the government, enterprises and society to understand the trend and change of carbon emissions in the future, achieve the goals of reducing carbon emissions, energy saving and emission reduction, protect the environment and promote sustainable development [23]. Part of the latest data required in this paper will be collected until 2020, and the index data values for 2022 will be set according to the provincial planning of the Eastern and Western Ministries.

(1) Population Size

According to the 2019 World Population Prospects of the United Nations Department of Economic and Social Affairs (UNDESA, 2019), China's population in 2030 is forecast to be 1.46 billion. The population developmentplan (2018-2030) and the country's total population are divided into the proportion of population development in the east and the west.

(2) Per Capita GDP

According to the research on the policy documents of the national future development plan, the GDP growth rate proposed by Zhang [24] should be a five-year development cycle. Combined with the 14th Five-Year plan and the 2035 vision goal plan of each province, the per capita GDP development of the eastern, central, and western regions should fluctuate according to each province's development status.

(3) Total Energy Consumption

Concerning China Economic Weekly and China's Energy Development in the New Era, this paper sets the total amount of energy consumption in the East, Central, and Western regions under the low-carbon policy. The index is energy consumption that generates carbon emissions and is finally converted into 10,000 tons of standard coal.

(4) Energy Consumption Structure

This paper uses the proportion of coal consumption in total energy to represent the energy structure. Based on the calculation of Wei Y et al. [25], it is estimated that by 2025, the proportion of coal in China will be reduced to 50%- 52%. The energy consumption structure of the eastern, central, and western regions is based on China's coal consumption demand. It is set according to the regional provinces' development and energy plans.

(5) Industrial Structure

In this paper, the data on the proportion of the secondary industry in GDP represents the industrial structure, and the initial data source is the panel data of the China Energy Statistical Yearbook. The future development of industrial structure refers to Zhang Xiliang's forecast of China's economic structure change from 2020 to 2060, and the industrial structure of the eastern and western regions is estimated according to the industrial development planning of each province.

(6) Urbanization Rate

By the end of 2020, the urbanization rates in the eastern and central regions will be 70.3 per cent, 57.1 per cent, and 49.2 per cent. According to China's 2014 Strategy, the country is set to reach a balanced state by 2050, 68 percent by 2025, and 71 percent by 2030. The proportion of growth in the eastern, central, and western regions shall be set based on strategic objectives and in combination with the development goals of regional provinces.

(7) Forest Coverage Rate

According to the China Environmental Statistical Yearbook data, the forest coverage rate in the eastern and western regions was 37.07%, 25.36%, and 19.3%, respectively. According to the Outline of the Plan for Ecological Civilization Construction (2016-2030), China's forest coverage rate is expected to reach 24.1% by 2025 and 25% by 2030. The proportion of forest coverage growth in the eastern, central, and western regions will be set according to the goals and plans of the regional provinces.

5.2 Determination of Influencing Factors

R analysis software was used to select the optimal lambda value into the Lasso model and fit the coefficients of the Lasso regression model. According to the model cycle, the variable coefficients decreased until unreasonable influencing factors were eliminated. The calculation results are shown in Table 1. Finally, population size (X_1), total energy consumption (X_3), energy consumption structure(X_4), industrial structure (X_5) and urbanization rate (X_6) are selected as the main influencing factors of the model. The data on influencing factors in the East, West and Middle regions are shown in Tables 2-4.

variable	x_1	x_2	x_3	x_4	x_5
Eastern coefficient	0.14	0.48	-0.13	-0.08	0.104
Middle coefficient	0.11	0.40	-0.1	-0.02	0.07
Western coefficient	0.06	0.42	-0.09	-0.06	0.07

Table 1. Lasso regression coefficient

A given year	x_1(per thousand people)	x_3(ten thousand metrictons)	x_4	x_5	x_6	Total Carbon emissions (ten thousand metric tons)
2002	49,466	132,648.96	0.757	0.454	52.50%	224,842.28
2003	49,928	146,900.51	0.765	0.474	53.61%	254,826.20
2004	50,464	166,482.83	0.752	0.488	53.98%	296,425.34
2005	50,949	193,262.16	0.740	0.478	54.26%	356,002.15
2006	51,709	239,439.58	0.728	0.485	54.86%	388,764.22
2007	52,458	227,973.76	0.712	0.485	55.72%	425,296.75
2008	53,169	232,279.96	0.690	0.489	56.68%	443,090.76
2009	53,889	250,122.43	0.679	0.461	57.59%	466,521.54

A given year	x_1(per thousand people)	x_3(ten thousand metrictons)	x_4	x_5	x_6	Total Carbon emissions (ten thousand metric tons)
2010	55,040	272,479.15	0.675	0.466	60.02%	512,612.01
2011	56,020	287,004.05	0.664	0.464	61.01%	553,680.61
2012	56,797	299,905.52	0.652	0.454	62.16%	566,738.69
2013	57,433	305,129.42	0.641	0.446	63.09%	566,156.63
2014	58,097	310,066.73	0.629	0.430	63.90%	570,744.10
2015	58,541	320,589.83	0.618	0.408	64.95%	588,158.51
2016	59,114	329,177.89	0.616	0.389	66.05%	597,594.37
2016	59,114	329,177.89	0.616	0.389	66.05%	597,594.37
2017	59,595	338,260.86	0.605	0.384	66.99%	608,610.72
2018	59,997	346,637.32	0.597	0.378	67.79%	618,648.67
2019	60,351	356,897.71	0.589	0.356	68.50%	629,885.54
2020	60,690	358,924.71	0.576	0.365	69.20%	647,024.34
2021	60,834	364,758.14	0.568	0.365	69.80%	657,176.94
2022	61,118	369,213.07	0.561	0.360	70.50%	664,079.45

Table 2. Influencing factors of carbon emission in Eastern China

A given year	x_1(per thousand people)	x_3(ten thousand metrictons)	x_4	x_5	x_6	Total Carbon emissions (ten thousand metric tons)
2002	42,086	66,828.03	0.558	0.475	32.81%	145,178.97
2003	42,265	73,600.95	0.574	0.489	33.45%	163,681.61
2004	42,486	83,990.40	0.582	0.503	34.85%	178,193.18
2005	41,738	97,312.35	0.591	0.474	36.12%	204,642.47
2006	41,797	107,191.69	0.590	0.488	37.56%	222,033.29

A given year	x_1(per thousand people)	x_3(ten thousand metrictons)	x_4	x_5	x_6	Total Carbon emissions (ten thousand metric tons)
2007	41,847	118,872.81	0.592	0.492	39.04%	238,599.67
2008	42,025	123,737.43	0.586	0.502	40.53%	245,617.08
2009	42,169	126,795.50	0.589	0.485	41.90%	257,634.65
2010	42,276	143,044.71	0.575	0.513	43.31%	280,506.10
2011	42,317	158,631.00	0.585	0.524	44.85%	310,309.41
2012	42,326	162,735.42	0.562	0.504	46.24%	317,166.83
2013	42,304	163,749.05	0.549	0.492	47.56%	310,510.66
2014	42,327	166,671.46	0.541	0.468	48.95%	317,547.01
2015	42,323	167,145.91	0.533	0.430	50.27%	316,294.45
2016	42,361	168,484.13	0.536	0.408	51.52%	314,308.43
2017	42,369	172,387.87	0.522	0.403	52.76%	323,019.21
2018	42,318	179,955.95	0.515	0.387	54.11%	338,595.89
2019	42,276	184,183.11	0.507	0.371	55.43%	341,858.07
2020	42,015	186,181.75	0.523	0.375	55.94%	349,957.91
2021	41,945	189,903.71	0.520	0.371	56.47%	348,504.53
2022	42,088	191,837.32	0.512	0.365	56.88%	349,330.15

Table 3. Influencing factors of carbon emission in Central China

A given year	X_1(per thousand people)	X_3(ten thousand metrictons)	X_4	X_5	X_6	Total Carbon emissions (ten thousand metric tons)
2002	35,499	48,461.23	0.526	0.424	31.00%	97,250.81
2003	35,725	58,210.35	0.585	0.442	31.24%	122,125.36
2004	35,796	68,996.81	0.577	0.461	32.35%	144,762.35
2005	35,637	82,071.28	0.572	0.438	33.45%	163,686.41

A given year	X_1(per thousand people)	X_3(ten thousand metrictons)	X_4	X_5	X_6	Total Carbon emissions (ten thousand metric tons)
2006	35,732	91,611.56	0.575	0.462	34.61%	186,325.87
2007	35,799	101,640.34	0.574	0.473	35.81%	206,356.00
2008	35,948	109,544.72	0.563	0.488	37.02%	225,862.76
2009	36,089	119,174.66	0.569	0.473	38.24%	246,990.68
2010	35,769	133,340.06	0.562	0.495	39.46%	271,691.26
2011	36,041	149,311.25	0.564	0.503	40.61%	312,841.26
2012	36,256	161,602.54	0.559	0.495	41.83%	335,986.24
2013	36,442	171,058.89	0.525	0.490	43.17%	344,402.26
2014	36,668	177,081.41	0.513	0.470	44.35%	352,370.38
2015	36,903	173,416.86	0.496	0.443	45.42%	341,332.78
2016	37,188	177,223.88	0.496	0.430	46.63%	347,206.63
2017	37,466	186,487.17	0.488	0.414	47.73%	363,446.89
2018	37,641	196,075.06	0.479	0.407	48.92%	381,186.95
2019	37,817	208,515.23	0.478	0.379	50.17%	400,604.58
2020	37,942	212,598.65	0.479	0.383	51.23%	412,399.25
2021	37,915	215,969.31	0.474	0.3823	52.84%	419,813.91
2022	38,021	219,710.60	0.468	0.378	53 .25%	423,707.22

Table 4. Influencing factors of carbon emission in western China

6. Prediction of Carbon Emissions in Eastern, Central, and Western China

The influencing factor sequences determined separately for the eastern, central, and western regions were used to predict the carbon emissions for these three areas using the GA-BP neural network. In this study, data from 2002 to 2022 were randomly shuffled to create 21 data sets, with the first 16 sets used as the testing set and the remaining five as the training set. The initial maximum iteration count of the neural network was set to 10000, with an error threshold of $1.0\times10^{-6}$and a learning rate of 0.01. The initial population of the genetic algorithm was 50, with a population size of 5, a selection function parameter of 0.09, and a crossover function parameter of 2. After performing the calculations with the model, the prediction errors of the carbon emissions model for the eastern region were shown in Figures 2 and 3 for the testing and training sets, respectively; for the central region, the errors were shown in Figures 4 and 5; and for the western region, the errors were shown in Figures 6 and 7."

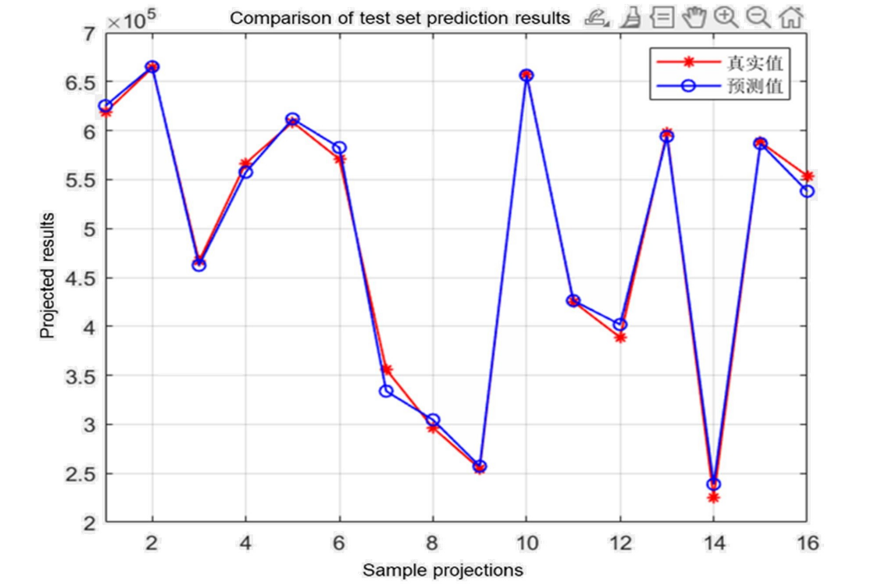

Figure 2. Training Set Error of Carbon Emission Model for Eastern Region

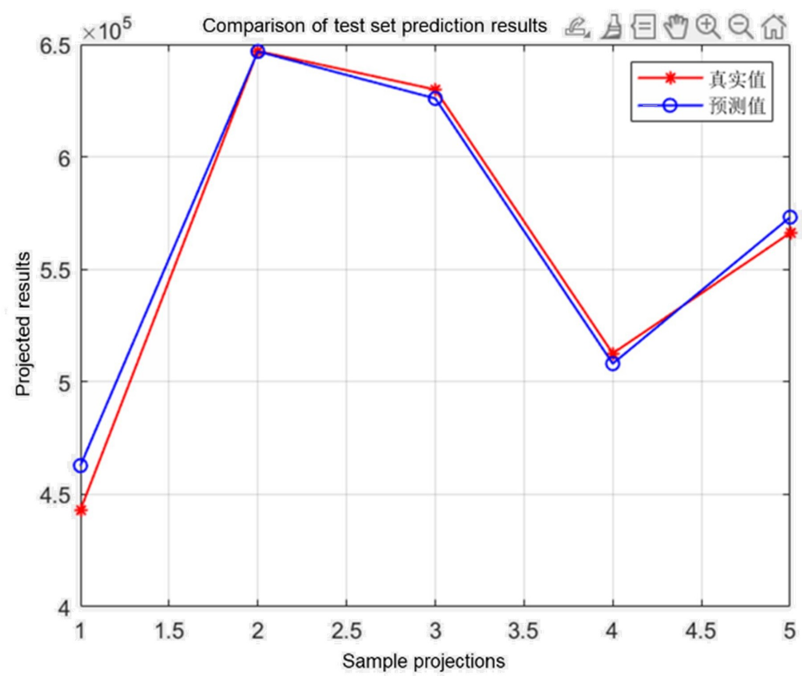

Figure 3. Testing Set Error of Carbon Emission Model for Eastern Region

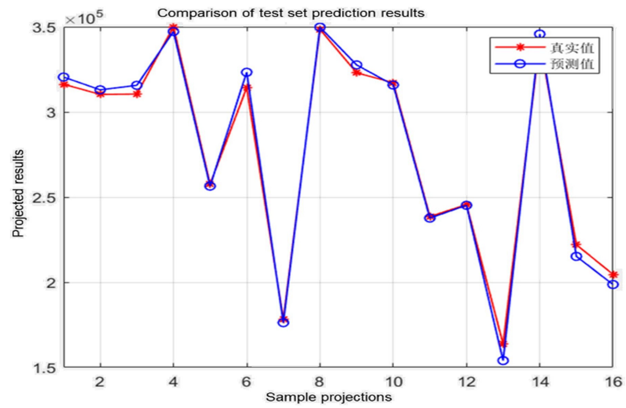

Figure 4. Training Set Error of Carbon Emission Model for Central Region

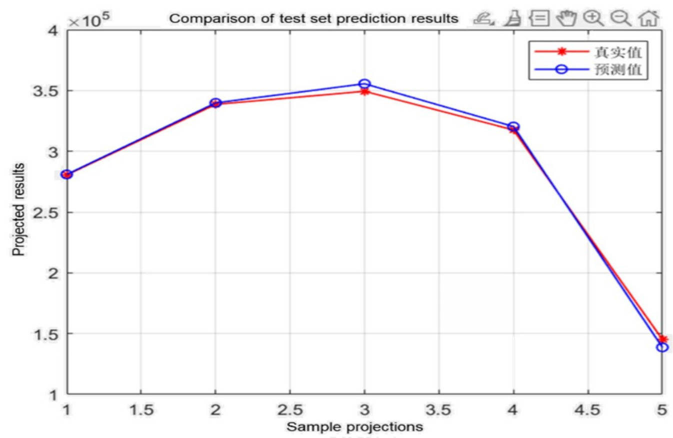

Figure 5. Testing Set Error of Carbon Emission Model for Central Region

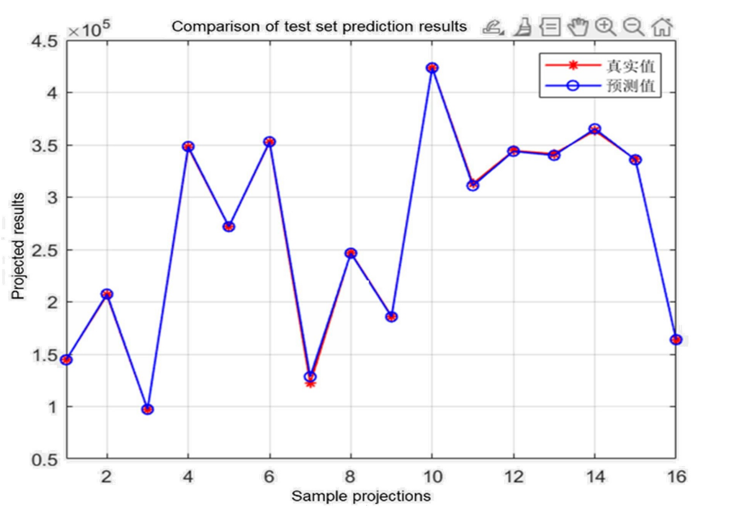

Figure 6. Training Set Error of Carbon Emission Model for Western Region

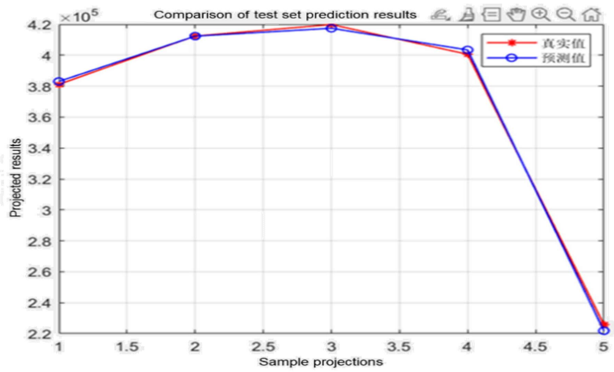

Figure 7. Testing Set Error of Carbon Emission Model for Western Region

Based on the training results of the GA-BP neural network prediction model, the training set accuracies for the eastern, central, and western regions were 0.99326, 0.99744, and 0.99965, respectively. Using these models, the testing set accuracies for the eastern, central, and western regions were 0.99497, 0.99212, and 0.99873, respectively. The optimal fitness values were 43.2101 for the east region, 57.8212 for the central region, and 55.8456 for the western region. The best validation performance was 1.04×10^{-4} for the east region, 4.71×10^{-5} for the central region, and 2.07×10^{-4} for the region of the west. The mean square error results are presented in Figure 8 for the eastern region, Figure 9 for the central region, and Figure 10 for the western region. The specific results of the model's predictions are provided in Table 5.

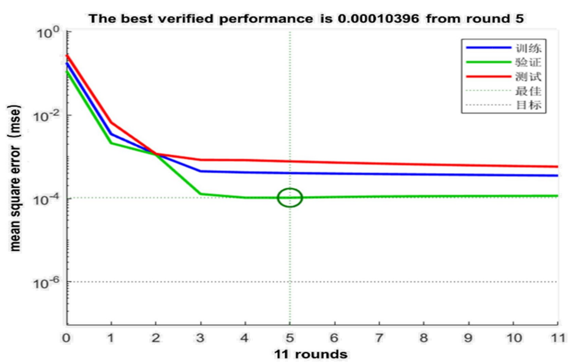

Figure 8. Mean Square Error of Carbon Emission Data Prediction for Eastern Region

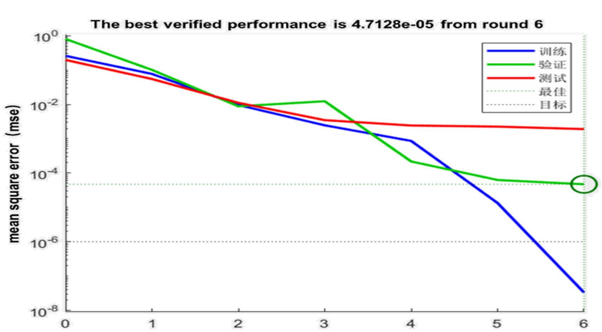

Figure 9. Mean Square Error of Carbon Emission Data Prediction for Central Region

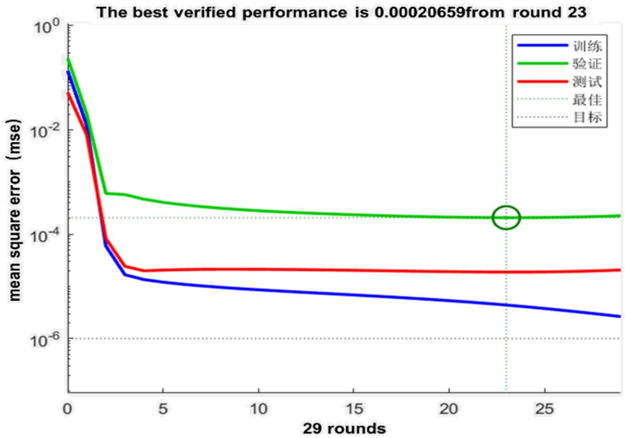

Figure 10. Mean Square Error of Carbon Emission Data Prediction for Western Region

Year	Eastern Region			Central Region			Western Region		
	actual value/ ten thou-sand metric tons	projected value/ ten thousand metric tons	MAPE	actual value/ ten thou-sand metric tons	projected value/ ten thousand metric tons	MAPE	actual value/ ten thou-sand metric tons	projected value/ ten thousand metric tons	MAPE
2002	224,842.28	229,133.22	1.91%	145,178.97	143,625.44	1.07%	97,250.81	97,277.47	0.03%
2003	254,826.20	251,757.01	1.20%	163,681.61	157,242.11	3.93%	122,125.36	128,500.64	5.22%
2004	296,425.34	308,092.56	3.94%	178,193.18	178,237.25	0.02%	144,762.35	144,597.14	0.11%
2005	356,002.15	334,716.68	5.98%	204,642.47	212,645.25	3.91%	163,686.41	163,871.04	0.11%
2006	388,764.22	399,428.31	2.74%	222,033.29	221,988.47	0.02%	186,325.87	185,695.53	0.34%
2007	425,296.75	421,962.12	0.78%	238,599.67	245,452.17	2.87%	206,356.00	207,415.52	0.51%
2008	443,090.77	461,501.53	4.16%	245,617.08	245,595.32	0.01%	225,862.76	221,851.70	1.78%
2009	466,521.54	466,296.94	0.05%	257,634.65	252,778.05	1.89%	246,990.68	246,450.41	0.22%
2010	512,612.02	508,593.50	0.78%	280,506.10	275,814.14	1.67%	271,691.26	271,816.47	0.05%
2011	553,680.61	536,421.19	3.12%	310,309.41	310,311.99	0.00%	312,841.26	310,934.95	0.61%

2012	566,738.70	556,676.19	1.78%	317,166.83	307,376.54	3.09%	335,986.24	335,682.20	0.09%
2013	566,156.64	570,055.74	0.69%	310,510.66	310,520.78	0.00%	344,402.26	343,801.37	0.17%
2014	570,744.11	579,489.62	1.53%	317,547.01	317,568.15	0.01%	352,370.38	352,907.56	0.15%
2015	588,158.52	588,603.31	0.08%	316,294.45	316,326.82	0.01%	341,332.78	339,928.03	0.41%
2016	597,594.37	592,209.72	0.90%	314,308.43	314,346.08	0.01%	347,206.63	348,360.72	0.33%
2017	608,610.73	609,270.83	0.11%	323,019.21	323,062.05	0.01%	363,446.89	365,287.84	0.51%
2018	618,648.67	621,877.40	0.52%	338,595.89	338,634.55	0.01%	381,186.95	383,045.00	0.49%
2019	629,885.55	625,392.53	0.71%	341,858.07	340,617.48	0.36%	400,604.58	403,403.93	0.70%
2020	647,024.34	643,848.74	0.49%	349,957.91	350,011.51	0.02%	412,399.25	412,333.43	0.02%
2021	657,176.95	653,898.38	0.50%	348,504.53	351,251.98	0.79%	419,813.91	417,426.52	0.57%
2022	664,079.45	661,221.21	0.43%	349,330.15	349,385.35	0.02%	423,707.22	423,550.30	0.04%
Average			**1.54%**			**0.94%**			**0.59%**

Table 5. Results of Carbon Emission Total Prediction and Error for Eastern, Central, and Western Regions

According to Table 5, the carbon emission prediction model for the eastern region can explain 98.46% of the variability in the explained variable. In comparison, the model for the central region can explain 99.06%, and for the western region, it can explain 99.41%. The overall Mean Absolute Percentage Error (MAPE) is less than 10.0%, indicating that the model's predictive accuracy is 'excellent,' and the overall predictions are entirely satisfactory [20]. The training results of the model are effective. Thus, the established model holds predictive significance. The carbon emission total predictions for the eastern, central, and western regions from 2023 to 2030, forecasted using the GA-BP neural network model, are presented in Table 6.

With the strict implementation of our country's strategy for clean, low-carbon, high-quality energy development and environmental protection policies, carbon emissions in China's eastern, central, and western regions will gradually increase, achieving a peak in the future and attaining carbon neutrality by 2060. According to Table 6, it is evident that from 2023 to 2030, carbon emissions in the eastern, central, and western regions are expected to show an upward trend. Still, certain variations exist due to differences in the primary industries, economic development, and resource conditions among these regions. The carbon emissions in the eastern region will remain considerably higher than those in the central and western regions. This is associated with the rapid economic development in the advanced eastern region, characterized by significant industrial output and higher per capita energy consumption. The central region's carbon emissions will experience relatively modest growth mainly due to its focus on manufacturing and agriculture. These industries tend to have lower carbon emissions and carbon dioxide is absorbed by plants, soil, and livestock within the agricultural sector, leading to lower emissions than other industries. In the upcoming years, while carbon emissions in the eastern region will continue to rise, the growth rate will slightly decrease. Emissions in the western region will maintain a smaller increase primarily due to its relatively weaker development foundation. Factors such as limited availability of relevant technologies and funding, along with a delayed start in developing environmental protection industries, constrain the pace of emission reduction in the western region. Overall, the total carbon emissions in the eastern region will remain ahead of those in the west and central regions. Still, all regions' carbon emission growth rates will gradually approach each other. This indicates that emission reduction efforts across various regions of China progressively yield results.

Total carbon emissions (ten thousand metric tons)			
Year	**Eastern Region**	**Central Region**	**Western Region**
2023	665,197.99	344,398.60	429,484.54
2024	668,461.43	351,373.41	435,675.75
2025	674,136.91	359,504.76	441,725.22
2026	678,838.81	359,896.86	447,809.33
2027	682,155.87	364,533.65	454,057.87
2028	686,166.37	370,189.33	460,313.58
2029	690,224.25	374,315.58	466,674.18
2030	695,495.55	376,467.18	473,059.26

Table 6. Forecasted Results of Carbon Emissions for the Eastern, Central, and Western Regions from 2023 to 2030

7. Conclusion

Fully understand that there are certain differences in the total amount of carbon emissions in regional development, and on this basis, use GA-BP neural network and scenario design to forecast the total amount of carbon emissions in eastern, central and western China, which can not only put forward regionally targeted policies for the realization of China's dual carbon goal but also build a modern energy system for energy industries in different regions and relevant governments. Explore a low-carbon path to provide data reference. The analysis is described below.

(1) The eastern region is one of China's most economically developed regions. However, total carbon emissions have increased because most enterprises still use traditional energy consumption methods, such as coal-fired

power generation and fossil fuels. To effectively control carbon emissions, the country can promote the low-carbon transformation of the manufacturing industry, increase the proportion of clean energy, such as solar, wind, hydropower, etc., and reduce reliance on coal. The government can guide enterprises to promote clean energy, formulate preferential policies, adjust industrial structure, develop service and technology industries, and establish carbon markets to regulate emission reduction behavior. In addition, energy enterprises should promote clean production, adopt efficient energy technologies, encourage using clean fuels such as natural gas, and promote carbon emission reduction.

(2) The central region is a substantial agricultural, scientific, and educational base in China, but there is still a problem of carbon emission growth caused by traditional energy dependence. The state can promote industrial transformation, increase the proportion of medium and high-end manufacturing and modern service industries, reduce conventional heavy and energy-intensive industries, and reduce carbon emissions through technological innovation and efficiency improvement. The government can also promote the green transformation of cities, encourage the construction of green buildings, improve public transport, and reduce the carbon emissions caused by urbanization. Encourage enterprises to develop green and low-carbon industries, such as environmental protection and energy conservation, and provide incentives and support for technological innovation and energy efficiency improvement for traditional industries. Promote resource conservation and environmental protection, improve energy efficiency, and encourage enterprises to reduce emissions. Energy enterprises should choose low-carbon products and clean energy consumption, promote intelligent energy-saving and emission-reduction technologies, and reduce carbon emissions. Establish a carbon footprint standard and monitoring system to achieve transparency and reduce carbon emissions. Enterprises promote emission reduction measures in coordination, achieve complementary carbon emission reduction services, and reduce regional carbon emissions.

(3) The primary industries in the western region are Resource-based, agriculture and animal husbandry, aerospace, clean energy, and new energy industries. To reduce carbon emissions in the region of the West, the national government can encourage and support clean energy transformation, increase investment and development of renewable energy, increase its proportion in the energy consumption structure, simplify the industrial structure, adopt efficient production processes and equipment, and reduce energy waste. The government can implement policies such as tax incentives, technical support and carbon emission trading markets to promote the development of low-carbon industries. For the traditional energy industry, it can encourage efficient conventional fuels such as clean renewable coal and liquefied natural gas, the application of energy-saving technologies, such as energy-saving boilers and heat pumps, and carbon capture technology to reduce carbon emissions.

References

[1] Xu, Y., Zhao, H. (2022). Study on the environmental costs of petrochemical enterprises based on double carbon targets. *Environment, Resource and Ecology Journal, 6*(5), 90–97.

[2] Wang, B. (2022). Application of carbon emission prediction based on a combined neural algorithm in the control of coastal environmental pollution in China. *Frontiers in Ecology and Evolution, 10*, 1043976.

[3] Fang, P. (2023). Short-term carbon emission prediction method of green building based on IPAT model. *International Journal of Global Energy Issues, 45*(1), 1–13.

[4] Han, Y., Cao, L., Geng, Z., et al. (2023). Novel economy and carbon emissions prediction model of different countries or regions in the world for energy optimization using improved residual neural network. *Science of The Total Environment, 860*, 160410.

[5] Guo, Z., Yin, J. (2022). Carbon emission forecast of construction industry based on grey theory. *Academic Journal of Computing & Information Science, 5*(10), 66–69.

[6] Kong, F., Song, J., Yang, Z. (2022). A daily carbon emission prediction model combining two-stage feature selection and optimized extreme learning machine. *Environmental Science and Pollution Research, 29*(58), 87983–87997.

[7] Zhao, J., Kou, L., Wang, H., et al. (2022). Carbon emission prediction model and analysis in the Yellow River Basin based on a machine learning method. *Sustainability, 14*(10), 6153.

[8] Cui, X., E, S., Niu, D., et al. (2021). Forecasting of carbon emission in China based on gradient boosting decision tree optimized by modified whale optimization algorithm. *Sustainability, 13*(21), 12302.

[9] Shi, M. (2022). Forecast of China's carbon emissions under carbon neutrality. *Environmental Science and Pollution Research, 29*(28), 43019–43033.

[10] Zhu, H., Wen, C., Xu, W., et al. (2022). A study on carbon emission forecasting in China based on PSO-BP neural network. *Academic Journal of Environmental Earth Science, 4*(2).

[11] Feng, W., Chen, T., Li, L., Zhang, L., Deng, B., Liu, W., Li, J., Cai, D. (2024). Application of neural networks on carbon emission prediction: A systematic review and comparison. *Energies, 17*, 1628.

[12] Guo, M. X. (2023). Quantitative assessment of the contribution of fossil energy reduction to pollution reduction and carbon reduction. *Ecological Economics, 39*, 184–190+207.

[13] Chen, J. J. (2023). Discussion on energy conservation and emission reduction ideas in the refining and chemical industry under "double carbon." *Modern Chemistry, 43*, 7–12.

[14] Xu, F., Pan, Q., Wang, Y. N. (2022). Research on the impact of green and low-carbon transformation on corporate profitability under the "dual carbon" goal. *Macroeconomic Research, 1*, 161–175.

[15] Jin, H. (2021). Prediction of direct carbon emissions of Chinese provinces using artificial neural networks. *PLOS ONE, 16*(5), e0236685. https://doi.org/10.1371/journal.pone.0236685

[16] Ye, H., Ren, Q., Hu, X., Lin, T., Shi, L., Zhang, G., et al. (2018). Modeling energy-related CO2 emissions from office buildings using general regression neural network. *Resources, Conservation and Recycling, 129*, 168–174.

[17] Balki, M. K., Çavuþ, V., Duran, Ý. U., Tuna, R., Sayin, C. (2018). Experimental study and prediction of performance and emission in an SI engine using alternative fuel with artificial neural network. *International Journal of Automotive Engineering and Technology, 7*(1), 58–64.

[18] Liu, Y., Tian, Y., Chen, M. (2017). Research on the prediction of carbon emission based on the chaos theory and neural network. *International Journal of Bioautomation, 21*(4), 339–348.

[19] Acheampong, A. O., Boateng, E. B. (2019). Modelling carbon emission intensity: Application of artificial neural network. *Journal of Cleaner Production, 225*, 833–856.

[20] Han, Y., Cao, L., Geng, Z., Ping, W., Zuo, X., Fan, J., Wan, J., Lu, G. (2023). Novel economy and carbon emissions prediction model of different countries or regions in the world for energy optimization using improved residual neural network. *Science of The Total Environment, 860*, 160410.

[21] Nalbantov, G., Bonvalot, S. (2021). Reply to "When to consider logistic LASSO regression in multivariate analysis?". *European Journal of Surgical Oncology, 47*(8), 2207.

[22] Chen, H., Liu, S., Magomedov, R. M., et al. (2021). Optimization of inflow performance relationship curves for an oil reservoir by genetic algorithm coupled with artificial neural intelligence networks. *Energy Reports, 7*, 3116–3124.

[23] Liu, X., Wang, X., Meng, X. (2023). Carbon emission scenario prediction and peak path selection in China. *Energies, 16*(5), 2276.

[24] Zhang, X., Huang, X., Zhang, D. (2022). Energy economy transition path and policy research under carbon neutral target. *Management World, 38*, 35–66.

[25] Wei, Y., Wang, Z., Wang, H., et al. (2021). Compositional data techniques for forecasting dynamic change in China's energy consumption structure by 2020 and 2030. *Journal of Cleaner Production, 284*, 124702.