Contents lists available at YXpublications

International Journal of Applied Mathematics in Control Engineering

Journal homepage: http://www.ijamce.com

The Prediction of Carbon Emissions in Hebei Based on the Grey Prediction Model and Ridge Regression Model

Tianyi Chen [a], Yuehua Fang[a*], Xinru Qiao[a], Jingyi Zhang[a]

[a] College of Mathematics and Computer Science, Hengshui University, Hengshui City, Hebei Province, 053000, China,)

ARTICLE INFO

Article history:
Received 15 July 2024
Accepted 21 September 2024
Available online 21 September 2024

Keywords:
Grey Prediction Model
Multicollinearity
Ridge Regression
Carbon Emissions

ABSTRACT

With the aggravation of the negative impact of carbon emissions, the research of carbon emissions has gradually become a hot topic. Accurate prediction of carbon emissions serves as a crucial foundation for attaining carbon neutrality and carbon peaking, and the prognosis of carbon emissions is of paramount significance. By collecting the data of nine major energy sources in Hebei Province from 2005 to 2019, this study adopts the Ridge Regression model to study the influencing factors of carbon emissions. The results show that coal consumption, electricity and GDP have a significant positive effect on carbon emissions, while natural gas has a significant negative effect on carbon emissions. Based on the heat conversion, the nine major energies were converted into carbon emission data. The grey prediction model was employed to prognosticate the carbon emissions in the subsequent five years, and conclusions and recommendations were presented in accordance with the research results.

1. Introduction

In recent years, the adverse effects of carbon emissions on the global level have been escalating progressively. Phenomena such as global warming, glacier melting, sea level rise, and El Niño occur frequently. The augmentation of carbon emissions not only triggers environmental issues, but also bears a significant impact on national societal development and people's welfare. Consequently, research on carbon emissions has progressively emerged as a focal topic in academic discourse[1]. According to the data of the United Nations Environment Programme, the global carbon emissions are on a continuous upward trend. In China, the annual carbon dioxide emissions released into the atmosphere exceed 600 million tons, and the level of carbon emissions remains persistently high. In the report released by the International Energy Agency in March 2024, it is proposed that the global energy - related carbon dioxide emissions in 2023 reached a record - high of 37.4 billion tons, increasing by 410 million tons compared with the previous year, with an increase rate of 1.1%. Over the past several years, the total amount of global carbon emissions has generally exhibited an ascending trend[2]. The increase in carbon emissions directly impacts on human bodies and daily life, and brings about tremendous challenges to the environment. It is indicated in the "Emissions Gap Report 2023: Breaking Records" that the temperature has reached a new peak, and the rapid global warming has resulted in increasingly frequent weather issues. In the face of the colossal challenges posed by the escalating carbon emissions, carbon and emission reduction constitutes a primary matter to be contemplated in economic and social development, and greater efforts should be exerted to implement emission reduction measures.

China has always been a practitioner of ecological civilization. As early as 2022, China proposed new nationally determined contributions of "carbon peaking" and "carbon neutrality": achieving carbon peaking by 2030 and carbon neutrality by 2060. During the Two Sessions in March 2024, the targets of carbon peak and carbon neutrality were mentioned repeatedly, and Premier Li Qiang emphasized the importance of strengthening ecological civilization construction and promoting green and low-carbon development to support the goals of carbon peak and carbon neutrality proposed by China in 2030 and 2060, respectively[3]. Accurate prognostication of carbon emissions constitutes a crucial basis for attaining carbon neutrality and carbon peaking. Hence, the prediction of carbon emissions holds significant importance.

Numerous scholars have carried out research on the prediction of carbon emissions. SU et al. probed into the main

* Corresponding author.
E-mail addresses: fangyuehua717@163.com(Y. Fang)
Digital Object Identifiers: https://doi.org/10.62953/IJAMCE.873797

driving factors behind carbon emissions in 60 countries along the Belt and Road Initiative by applying the extended STIRPAT model, determined the main driving factors for controlling carbon emissions, and the main factors influencing carbon emissions were per capita GDP and the energy consumption structure. The population factor exerted a dual influence on carbon emissions[4]. Peng et al. established a prediction and assessment framework for energy security risks by employing methods such as the grey prediction model and predicted that the consumption of coal in China exhibits a downward tendency[5]. Luo et al. utilized the least squares support vector machine model to prognosticate the carbon emissions of the new energy power system[6]. Samuel and Vladimir employed methods such as the ordinary least squares method and SIMPLS analysis to analyze the influence of Australia's energy mix on environmental pollution and carbon dioxide emissions [7]. Anis Omri utilized the simultaneous equation model to verify the presence of a one-way causal relationship between energy consumption and carbon dioxide, and there was no feedback between the two[8].

Research on carbon emissions is often focused on national or major regional carbon emissions, and there is less research on carbon emissions at the provincial level. As a major energy consumer and carbon emitter, China's carbon emissions are concentrated in certain industries and regions, thus it is essential to pay attention to the higher-emitting industries and regions. According to the China Carbon Accounting Database (2019), the carbon emissions of Shandong, Hebei, Inner Mongolia, Guangdong, and Shanxi are ranked in the top five. Faced with such a severe situation, Hebei Province should take the initiative to shoulder the burden of emission reduction and promote high-quality ecological development.

Hebei Province is one of the crucial regions in the integration of Beijing-Tianjin-Hebei. As a major province in energy consumption and carbon emissions, Hebei Province is an important industrial base in China. Industries such as steel, machinery, textiles, and chemicals occupy significant positions across the country. Nevertheless, its overweight industrial structure, coal-dominated energy structure, and the relatively low proportion of non-fossil energy have brought about a series of difficulties and challenges for the carbon emission reduction target, leading to consistently high carbon emissions. Hence, this research utilizes ridge regression and the grey prediction model to prognosticate the carbon emissions of Hebei Province and conducts an analysis based on the research outcomes.

2. Analysis of the Factors Influencing Carbon Emissions

To facilitate the study, the nine major carbon energy sources in Hebei Province were converted, and carbon energy data in Hebei Province from 2005 to 2019 were collected from the China Energy Database for analysis. It can be observed from Figure 1 that the overall energy consumption in Hebei Province exhibits an upward tendency. Coal energy remains the foremost energy consumed in Hebei Province, followed by electric energy, and then successively by crude oil, coke, diesel, gasoline, natural gas, fuel oil, kerosene, and coal energy. The table also shows that coal energy has increased steadily from 2005 to 2013, but it began to decrease from 2013 to 2019. In contrast, the consumption of coke, crude oil, gasoline, fuel oil, natural gas, and electricity has generally increased. The aforementioned nine primary energy sources constitute the major contributors to energy consumption in China. It is evident that non-renewable energy sources continue to dominate China's energy consumption, while renewable energy sources such as natural gas and electricity still account for a minority share. This highlights the suboptimal energy consumption structure in China, which requires significant improvement. The urgent need for the development and utilization of renewable energy sources underscores the imperative to reduce carbon emissions, indicating that the path towards decarbonization remains a formidable challenge.

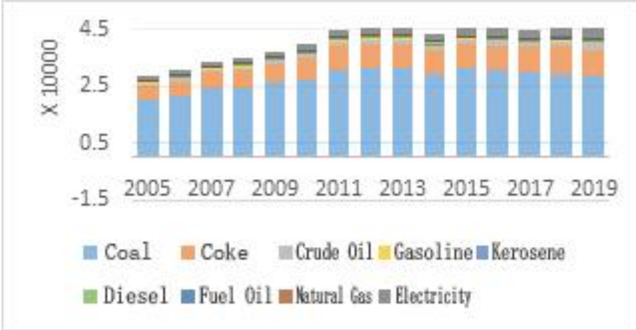

Fig. 1 Stacked Graph of the Consumption of Nine Major Energy Sources

In order to elucidate the primary determinants of carbon emissions, the present study utilizes ridge regression analysis to identify the key factors influencing carbon emissions. Linear regression is an effective method for predicting the impact of independent variables on dependent variables, but when there is multicollinearity among independent variables, consider using ridge regression to eliminate the problem of collinearity. Ridge regression is an improvement upon the traditional ordinary least squares estimation method, incorporating a regularization term into the loss function to prevent overfitting. It is a linear regression method designed to deal with the issue of multicollinearity.

Coal and crude oil, as the principal consumed energy sources of non-renewable energy, and natural gas and electricity, as the representative energy sources of green energy, accordingly, these four variables are included in the equation that might affect carbon emissions. GDP, representing the level of economic development, is also chosen to be included in the equation. Secondly, other factors that might have an impact on carbon emissions are also incorporated into the equation. Firstly, multiple linear regression is conducted on the data, and the results are as follows.

Tab. 1 The Coefficient Table of Multiple Linear Regression

Model	B	Standard Error	Standardized Coefficient	Significance	Tolerance	VIF
constant	-299.4	99198.775		0.778		
x1	2.475	0.47	0.551	0.1	0.33	3.433
x2	3.75	2.772	0.68	0.218	0.17	9.339
x3	-57.89	6.632	-0.163	0.371	0.9	17.551
x4	7.52	4.65	0.354	0.17	0.7	136.599
x5	-0.2	0.2	-0.54	0.332	0.11	9.874
x6	5.856	12.619	0.72	0.657	0.11	89.881

x7	0.632	0.522	0.316	0.266	0.4	253.645

x1: Coal Consumption；x2: Crude Oil Consumption;x3: Natural Gas Consumption;x4: Electricity Consumption;x5: Forest Area;x6: Population;x7: GDP

From the observation of Table 1, it is evident that the Variance Inflation Factor (VIF) values are excessively high, predominantly exceeding 10, signifying a significant level of multicollinearity among the dataset. Consequently, to deal with this issue of multicollinearity, the method of ridge regression is employed. The process of ridge regression involves the following steps:

Step 1: Generate the corresponding ridge trace plot and utilize it to determine the optimal value of the regularization parameter K.

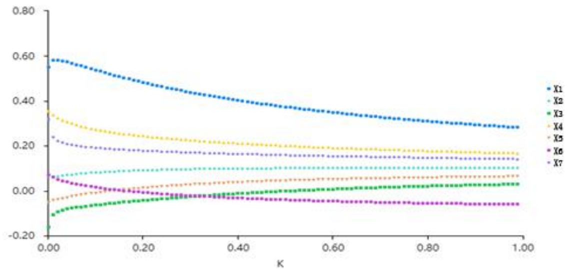

Fig. 2 Ridge trace plot

The selection of the K value is the minimum K value when the standardized coefficients of each independent variable tend to stabilize, and the smaller the K value, the better. It can be observed from the above that eight variables, such as coal consumption, electricity consumption, natural gas consumption... GDP, are regarded as independent variables, and carbon emissions are considered as the dependent variable. It can be directly perceived that each coefficient tends to stabilize when K = 0.20, thereby determining the K value to be 0.20.

Step 2: Substitute the determined K = 0.20 into the equation and conduct ridge regression analysis to yield the following results.

Tab. 2 Model Summary Table

Model Summary			
Sample size	R 2	Improve R 2	model error
15	.998	.995	769.876

From the ridge regression analysis results presented in Table 3, it can be observed that when k=0.20, the model's adjusted R-squared value is 0.998, indicating a strong fit. This suggests that the ridge regression model effectively explains the influencing factors of carbon emissions. The adjusted R-squared value of 0.998 indicates that coal, electricity, natural gas, crude oil, forest area, population, and GDP can explain 99.8% of the variation in carbon emissions.

Tab. 3 Ridge Regression Table

Table					
	Sum of Squares	*df*	Mean square	F	p value
regression	384741315.934	7	549628759.419	432.748	0.000
residual	889635.556	7	1279.794		
total	3856291951.49	14			

From Table 3, it can be observed that the value is 432.748, and the significance level p is less than 0.05, indicating that the model is statistically significant and further analysis can be conducted. This also suggests that variables such as coal, electricity, natural gas, crude oil, forest area, and population have an impact on carbon emissions.

As evidenced by Table 4, the Variance Inflation Factor (VIF) values are observed to be less than 10, indicating that the established ridge regression model effectively addresses the issue of collinearity within the equation. When the p-value is less than 0.1, it can be concluded that the factor has a significant impact on carbon emissions and should be included in the equation. It is evident from the table that the p-values for coal, electricity, and GDP are all less than 0.05, while the p-value for natural gas is 0.064, which is less than 0.1. At a significance level of 0.1, it can be inferred that coal consumption, electricity consumption, GDP, and natural gas consumption have a significant impact on carbon emissions. However, the p-values for crude oil, forest area, and population are all greater than 0.05, indicating that they do not have a significant impact on carbon emissions. Therefore, the following standard equation is established:

$$y = .581x_1 - .293x_2 + .325x_4 + .221x_7$$

Tab.4. The Result Table of Ridge Regression Analysis

Regression Analysis Results						
	Non standardized coefficient		Standardized Coefficient	T value	P value	VIF 值
		Standard Error				
constant	-188.898	35638.75	-	-.528	.614	-
x1	2.613	.144	.581	18.163	.000**	3.11
x2	3.56	2.67	.64	1.722	.129	4.247
x3	-1.76	15.65	-.293	-2.196	.064	5.43
x4	6.88	.955	.325	7.23	.000**	6.169
x5	-.2	.2	-.39	-.997	.352	4.634
x6	4.154	4.158	.51	.999	.351	7.979
x7	3.442	.83	.221	5.354	.001**	5.19
F(7,7)=432.748						
dependent variable：y						
* p<0.05 ** p<0.01						

It can be discerned from Table 4 that the element exerting the most significant influence on carbon emissions is coal, followed by the electricity factor. It is imperative to facilitate the transformation of energy patterns for green ecological construction, curtail the utilization of non-renewable energy, optimize the energy structure, and employ more green and clean energy. The development of GDP also has a profoundly marked impact on carbon emissions. It is essential to advance high-quality economic development, encourage the establishment of green ecological enterprises, and promote green and sustainable development with the economy driving technology. Natural gas has a negative effect on carbon emissions, that is, augmenting the usage of clean energy such as natural gas will reduce carbon emissions. Moreover, in conjunction with Figure 1, it is observable that among the major nine energy sources in China, the proportion of natural gas is relatively low, and there remains a considerable potential for growth.

3. The prediction of carbon emissions in Hebei Province

Based on the non-equivalence of coal consumption, electricity consumption, and natural gas consumption with carbon emissions, the study employs a thermal energy conversion coefficient to convert the nine primary energy sources into carbon emission data for analysis. The grey prediction model GM(1,1) is utilized to forecast the carbon emissions in Hebei Province.

Step 1: Establish the time series of the average energy consumption.

"X" ^（"0" ）" = [" "x" ^(("0")) ("1")"," "x" ^(("0")) ("2")",⋯⋯," "x" ^(("0")) "(15)]"

= (75963.4335，81496.1292，959.35，92773.494，99626.4779，16693.553，1226.6959，1227.657，122981.8862，11692.9875，121489.8529，12786.3965，12482.5411，12242.3825，121167.8625).

Step 2: Derive the grade ratio:

$$\sigma(k)=x^{(0)}(k-1)/x^{(0)}(k)$$

According to the data, the level ratios are as follows

$\sigma=[\sigma(2),\sigma(3),\ldots\ldots\sigma(15)]$

=(.932，.9，.976，.931，.934，.887，.985，.993， 1.52，.962，1.6，1.3，.984），

the results are presented as shown in Table 5.

Table .5. GM**Tab.5.** GM(1,1) Model Grade Ratio Table

Serial Number	original value	Grade ratioλ
1	75963.433	-
2	81496.129	0.932
3	90509.03	0.9
4	92773.494	0.976
5	99626.478	0.931
6	106693.553	0.934
7	120260.696	0.887
8	122070.657	0.985
9	122981.886	0.993
10	116902.988	1.052
11	121489.853	0.962
12	120786.397	1.006
13	120482.541	1.003
14	122402.383	0.984
15	121167.863	1.01

Step 3: Conduct a grade ratio test,

$$\sigma(k)\epsilon(e^{\frac{-2}{n+1}},e^{\frac{2}{n+1}})=(0.882,1.133)$$

The grade ratio test values fall within the standard range interval of (.882, 1.133), indicating that the research data is suitable for the GM(1,1) model. Consequently, a GM(1,1) model can be developed for predictive purposes.

Consequently, a GM(1,1) model can be developed for predictive purpose:

$$X^{(1)}(K)=\sum_{M=1}^{K}X^{(0)}(M)(K=1\ldots\ldots15) \quad (2)$$

The cumulative sum sequence is calculated:

$X^{(1)}=[X^{(1)}(1),\ldots\ldots X^{(1)}(15)]$

=(75963.4335，157459.5627，247968.5932，34742.872，44368.5651，54762.1181，667322.814，789393.471，912375.3572，129278.345，115768.198，1271554.594，139237.135，1514439.518，163567.38)

Step 5: Construct the data matrix and the data vector:

$$z^{(1)}(2)=1/2[X^{(1)}(1)+X^{(1)}(2)]=116711.5$$

$$z^{(1)}(3)=1/2[X^{(1)}(2)+X^{(1)}(3)]=22714.8$$

……

$$z^{(1)}(15)=1/2[X^{(1)}(14)+X^{(1)}(15)]=157523.45$$

The ensuing vector and matrix are derived as follows:

$$Y=\begin{bmatrix}81496.1292\\90509.0305\\92773.494\\99626.4779\\106693.553\\120260.6959\\122070.657\\122981.8862\\116902.9875\\121489.8529\\120786.3965\\120482.5411\\122402.3825\\121167.8625\end{bmatrix} \qquad B=\begin{bmatrix}-116711.4981 & 1\\-202714.078 & 1\\-294355.3402 & 1\\-390555.3262 & 1\\-493715.3416 & 1\\-607192.4661 & 1\\-728358.1425 & 1\\-850884.4141 & 1\\-970826.851 & 1\\-1090023.271 & 1\\-1211161.396 & 1\\-1331795.865 & 1\\-1453238.326 & 1\\-1575023.449 & 1\end{bmatrix}$$

We use the least squares methodology to estimate the parameters:

$$(a,b)^T=(BB^T)^{-1}B^TY=(-.024862277,91306.32861) \quad (3)$$

So we get a =-.024862277, b =91306.32861

Step 6: Establish the GM(1,1) model:

$$X^{(0)}(K)-0.024862277Z^{(1)}(K)=91306.32861 \quad (4)$$

Hence we btain the time response sequence as follows:

$$X^{(1)}(K+1)=[X^{(0)}(1)-\frac{b}{a}]e^{-aK}+\frac{b}{a} \quad (5)$$

$$=75963.433e^{0.024862277K}+3672484.542$$

So we generate a sequence of numbers "X" ^(("1")) ("K+1") and model reduction values ,The K Value is substituted into the time response to obtain the result "x" ^(("1")) ("K"), where x^((1)) (1)=75963.433.

In order to obtain the fitted value of the true value, the sequence is generated according to the cumulative subtraction method as follows:

$$X^{(1)}(K)=X^{(1)}(K)-X^{(1)}(K-1) \quad (6)$$

Therefore, we get the reduced value

$$X^{(0)}(K)=（X^{(0)}(1),X^{(0)}2\ldots\ldots X^{(0)}15）.$$

By transformation, we calculate the values of the original sequence $X^{(0)}(K)$ according to the formula.

$X^{(1)}(K)$=(75963.433，94363.133，96738.623，99173.914，1167.51，14229.955，16853.832，19543.762，11231.48，115128.475，11826.71，12997.95，12443.896，127166.568，13367.849)

In order to test the prediction effect of the grey prediction model, we calculate the residual, relative error, and level ratio deviation of the grey prediction model. One criterion is the relative error, when it is less than 0.2, it indicates that the model is basically valid. If it is less than 0.1, it meets higher requirements, the smaller the relative error , the better the model prediction. The other criterion is level ratio deviation. When the level ratio deviation is less than 0.2, it meets the requirement, and if it is less than 0.1, it higher requirements. the smaller the level ratio deviation, the better the model prediction. According to the data of this study, we calculate the residual value, relative error value, and level ratio deviation value, as shown in Table

6.

Table.6. GM(1,1) model test

Number	Residual	Relative Error	Deviation Grade Ratio
1	0	0.000%	-
2	-12867.001	15.788%	0.044
3	-6229.59	6.883%	0.077
4	-6400.417	6.899%	0
5	-2044.03	2.052%	0.045
6	2463.58	2.309%	0.043
7	13406.866	11.148%	0.09
8	12526.897	10.262%	-0.01
9	10680.48	8.685%	-0.018
10	1774.514	1.518%	-0.078
11	3463.144	2.851%	0.014
12	-211.508	0.175%	-0.031
13	-3561.355	2.956%	-0.028
14	-4764.185	3.892%	-0.009
15	-9199.986	7.593%	-0.036

As shown in Table 6, the values of relative error are all less than 20%, meeting the error requirements, and the model fitting effect meets the expectations. Therefore, the GM(1,1) model can be adopted for prediction.

We used the gray prediction model to predict the carbon emissions in Hebei Province over the next 10 years, and the results are as follows. The predicted value of the GM（1,1) model is (133649.719，13714.26 ， 14463.391 ， 143999.45 ， 147624.435 ， 15134.721,155150.563,159056.311,163060.383,167165.252).

The GM(1,1) model is adopted for prediction. The fitting values and predictions of the carbon emission model are shown in Figure 3. It can be seen from Figure 3 that the carbon emissions in Hebei Province will continue to increase in the next five years, and the task of reducing carbon and emissions remains arduous.

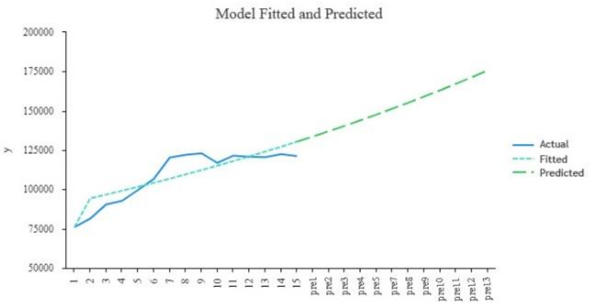

Fig.3. Carbon emission model fitting and prediction

4. Result Analysis

The data of the nine major energy sources in Hebei Province collected in this research suggest that the primary energy sources in Hebei Province remain concentrated in non-renewable energy sources like coal and coke, while the utilization ratios of natural gas and electricity, as representatives of renewable energy sources, remain comparatively low. And among the nine major energy sources, renewable energy still accounts for a minority, and the proportion of non-renewable energy remains rather considerable. The results of ridge regression analysis indicate that the consumption of coal, the consumption of electricity, GDP, and the consumption of natural gas all have significant impacts on carbon emissions. The coal factor is the predominant factor influencing carbon emissions. The influence of natural gas on carbon emissions is negative, suggesting that augmenting the utilization of natural gas can curtail a part of carbon emissions, followed by electricity and GDP. It also indicates that we can promote economic development, transform the energy structure, and thereby build a green ecology. The GM(1,1) grey prediction model is adopted to predict the carbon emissions of Hebei Province. The research findings reveal that in the upcoming five years, the carbon emissions of Hebei Province will persist in escalating, and the task of carbon and emission reduction remains formidable.

Based on the above conclusions, this study puts forward the following suggestions: (1) Innovation-driven, introduce new low-carbon technologies. Strengthen the intensity of low-carbon technological innovation, enhance technological research and development in the field of low-carbon environmental protection, increase the conversion rate of scientific and technological achievements. The government should provide financial and policy support to the aforementioned cities and offer better conditions for scientific research and innovation. Learn advanced energy conservation and emission reduction technologies, strengthen domestic and international exchanges and cooperation, promote technology sharing and collaborative carbon reduction, take advanced science and technology as an efficient way to achieve carbon reduction governance, improve the utilization rate of energy, and reduce carbon emissions.(2) Promote the popularization of clean and renewable energy and strengthen government subsidies and preferential systems. The national government should increase the promotion and use of clean and renewable energy such as natural gas and solar energy. In comparison with the traditional price of coal, natural gas is more cost-effective. Hence, the government should offer corresponding price subsidies and preferential policies for it, remit or exempt taxes on the production and use of natural gas, and introduce corresponding policies for regulation and incentives. (3) Intensify the publicity and education efforts on low-carbon and environmental protection. Strengthen the publicity in society, communities, schools, etc., about the main sources of carbon emissions, the various harms they bring, and what should be done. Enhance the public's awareness of carbon emissions. People should establish an environmental protection awareness, transform the cognition of low-carbon into actions, and advocate a low-carbon life.

References

[1] Li, G., Wu, H. & Yang, H. A multi-factor combination prediction model of carbon emissions based on improved CEEMDAN. *Environ Sci Pollut Res* 31, 20898–20924 (2024). https://doi.org/10.1007/s11356-024-32333-x

[2] Raj P .China's emissions ease but coal-based capacity addition to hinder climate goals: report[J]，https://usa.chinadai-ly.com.cn/world/201411/24/content_18963296.htm

[3] Hu, A. (2024). China’s Carbon Emission Peak Goals, Strategies and Policies. In: China International United Petroleum & Chemicals Co., Ltd., Chinese Academy of Social Sciences, Peking University (eds) Annual Report on China’s Petroleum, Gas and New Energy Industry (2022–2023). Current Chinese Economic Report Series. Springer, Singapore. https://doi.org/10.1007/978-981-99-7289-0_2

[4] Sun L, Cui H, Ge Q. Driving Factors and Future Prediction of Carbon Emissions in the 'Belt and Road Initiative' Countries[J]. Energies, 2021,14,5455. https://doi.org/10.3390/en14175455

[5] Peng, C., Chen, H., Lin, C., Guo, S., Yang, Z., & Chen, K. (2021). A framework for evaluating energy security in China: Empirical analysis of forecasting and assessment based on energy consumption. *Energy*, *234*, 121314. https://doi.org/10.1016/j.energy.2021.121314.

[6] Luo, Y., Yang, S., Niu, C. et al. An adaptive prediction method for carbon emissions of power systems containing new energy based on least-squares support vector machine. Sustainable Energy res. 11, 36 (2024). https://doi.org/10.1186/s40807-024-00131-1

[7] Sarkodie, Samuel Asumadu; Strezov, Vladimir . (2018). Assessment of contribution of Australia's energy production to CO_2 emissions and environmental degradation using statistical dynamic approach. Science of The Total Environment, 639, 888–899. doi: 10.1016/j.scitotenv.2018.05.204

[8] Omri A. CO2 emissions, energy consumption and economic growth nexus in MENA countries: Evidence from simultaneous equations models[J]. Energy Economics, 2013, 40:657-664. https://doi.org/10.1016/j.eneco.2013.09.003.

Tianyi Chen , female, studying in Hengshui University majoring in Applied statistics. Participated in the China International College Student Innovation Competition and National college students Market research and analysis Competition. Good at data collection and application of a variety of statistical software for analysis.

Yuehua Fang, female, a teacher of Hengshui University. She obtained her Ph.D. in February 2024. In recent years, she has been devoted to teaching and research work. She has taught the courses of applied multivariate statistical analysis, applied stochastic processes, and applied time series analysis. She has won several teaching awards and has been focusing on research in optimization algorithms statistical modeling. She has published 3 core papers and more than ten other papers.

Xinru Qiao, female, studying in Hengshui University majoring in Applied statistics. Participated in the Market Research and Analysis Competition and the Mathematical Modeling Competition organized by the school. Mainly responsible for the data collection and analysis part, using statistical knowledge and statistical analysis software such as SPSS for analysis, and programming languages like MATLAB and R for data processing and analysis, being able to efficiently handle large-scale data.

Jingyi Zhang, female, studying in Hengshui University majoring in Applied statistics. Participated in the Market Research Competition and Statistical Modeling Contest, harboring a passionate and keen interest in scientific research, and genuinely applying the acquired knowledge to practical scenarios. Boasting basic competencies for scientific research, encompassing the ability to search for and peruse literature, the capacity to compose papers, and the comprehension of relevant software and languages, such as Python, MATLAB, etc.