THE American Journal of Humanities and Social Sciences Research (THE AJHSSR) 2024
E-ISSN: 2581-8868
Volume-07, Issue-05, pp-08-22
www.theajhssr.com Crossref DOI: https://doi.org/10.56805/ajhssr
Research Paper Open Access

Carbon emission combination prediction based on IOWGA operator

[1]ZhenMu Hong, [2]Junyan Qi
[1,2]*Anhui University of Finance and Economics.*

ABSTRACT

In this paper, the H-W seasonless model, Arima model and multiple linear regression model are constructed to fit the carbon emission of our country from 1990 to 2022. Then the combination forecasting model based on Shapley value and the combination forecasting model based on IOWGA operator are constructed, by comparing and analyzing the single forecasting model and the combined forecasting model, it is found that the forecasting precision and error of the two combined forecasting models are better than that of the single forecasting model, especially, the combination forecasting model based on IOWGA operator is more effective. Finally, we find that our carbon emissions from 2023 to 2027 are still on an upward trend in the next five 2027, but the growth rate is decreasing. At present, China should continue to actively and steadily push forward the goal of carbon peak and carbon neutrality, and push forward the process of building a beautiful China in an all-round way, so as to achieve sustainable economic and social development.

KEYWORDS: degree Carbon emission, Shapley value, IOWGA operator, combined prediction.

1. INTRODUCTION

As more and more countries around the world industrialize, human behavior goes far beyond natural cycles and has a huge impact on climate change. According to the World Meteorological Organization report, the annual temperature of the 2023 increased by about 1.4℃ from its pre-industrial baseline, carbon dioxide levels are also 50 percent higher than pre-industrial levels, indicating a severe trend in global warming. According to the 2023 China Climate Bulletin, the 2023 national average temperature reached 10.71℃, 0.82℃ higher than usual and the highest since 1951. Against the background of global climate change, China has realized the urgency of integrating global resources and jointly coping with climate crisis. At the same time, based on its own ecological environment status quo, committed to promoting the construction of ecological civilization of the new development to achieve the vision of a beautiful China. So in September 2020, China set out a clear vision of reaching "Peak carbon" by 2030 and "Carbon neutrality" by 2060.

For carbon emission prediction, different scholars adopt different prediction methods according to different regions. Zhao Aiwen and Li Dong [1] made a short-term prediction of China's carbon emissions through the grey GM(1,1) model. Song Jiekun [2] used data from 1980 to 2009 to forecast China's carbon emissions by using partial least squares regression method. Qu Shenning and Guo Chaoxian[3] used STIRPAT model to predict the future carbon emission peak in China. Zhao Xi [4] et al used discrete second-order difference equations to predict China's carbon emissions in 2020. With the development of prediction methods, the prediction of carbon emission has evolved from simple linear model to more complex nonlinear combination model. Zhao Chengbai and Mao

Chunmei [5] used the combination model of Arima model and BP neural network to predict the carbon emission intensity of our country. Gao Jie et al. [6] established Richards-BP neural network combination model. Zhu Wanbing and Lv Xuebin [7] predicted the ARMI-SSA-LSTM combined model. Hu Linlin and Jia Junsong [8] used combined ESARIMA model to predict. All the above studies show that the combined model has higher prediction accuracy and more accurate results

This paper selects China's carbon emissions data from 1990 to 2022. Firstly, three single prediction models, Holt-Winters non-seasonal model, ARIMA model and multiple linear regression model, are used to forecast China's carbon emissions from 2023 to 2027. Then, the combined prediction model based on Shapley value and the combined prediction model based on IOWGA operator are constructed to predict China's carbon emissions. Finally, the development trend of China's carbon emissions from 2023 to 2027 is predicted according to the optimal results.

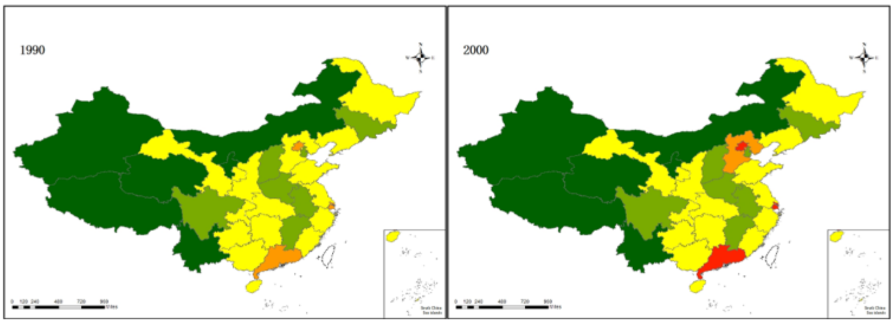
2. CARBON EMISSION STATUS AND SINGLE PREDICTION MODEL

2.1 Original data analysis

In this paper, carbon emission data from 30 provinces (autonomous regions and municipalities directly under the Central Government) from 1990 to 2022 are aggregated, excluding Hong Kong, Macao, Taiwan and Tibet, as the total carbon emission data of China from 1990 to 2022.

2.1.1 Spatial distribution of provinces

In 1990, the average carbon emissions of all provinces were 89.52 million tons, in 2000, 175.86 million tons, in 2010, 267.01 million tons, and in 2022, 367.77 million tons, with an approximate increase of 10,000 tons of carbon emissions per decade. As can be seen from the figure below, the carbon emissions of all provinces in the country are growing rapidly, and present a distribution feature of "high in the east and low in the west". The eastern region has a booming economy and a high level of industrialization, and the demand for energy is particularly strong, especially the consumption of fossil fuels. This high consumption has led to a continuous rise in carbon emissions and exacerbated the pressure of climate change in our country. At the same time, the central region, as an important producing area of coal and other fossil resources, with the acceleration of industrialization, energy consumption is also increasing, especially in 2022, the carbon emissions in the central region jumped to the highest level in the country. In recent years, the state has increased economic support for the western region, but due to the impact of geographical location and ecological environment, its carbon emissions are still increasing, but the growth rate is slower than that of the eastern and central regions.



(Note: The approval number is GS Jing (2022) 1873. Data for Tibet Autonomous Region and Hong Kong, Macao and Taiwan are currently lacking. The same below)

Figure 1. Spatial distribution of carbon emissions in China in 1990 and 2000

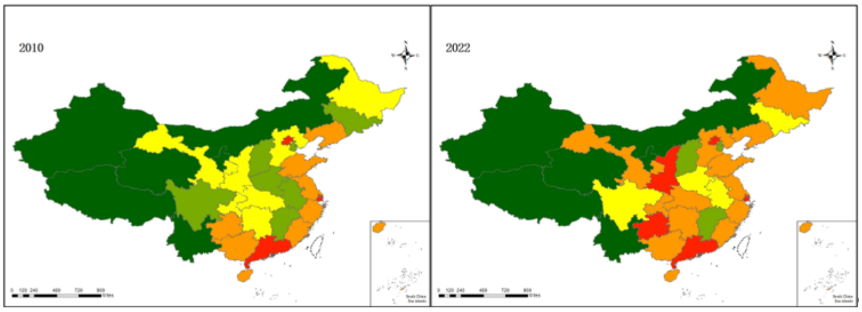


Figure 2. Spatial distribution of carbon emissions in China in 2010 and 2022

2.1.2 Horizontal distribution

As can be seen from Figure 3, China's carbon emissions have basically shown a steady upward trend over the past three decades, rising from 2672.64 million tons of carbon dioxide in 1990 to 1140098 million tons of carbon dioxide in 2022, an increase of about 3.3 times. This is mainly influenced by the development of industrialization in our country. With the continuous advancement of Chinese-style modernization, China's industrialization process has gradually accelerated, and the large consumption of fossil energy has caused the emission of carbon dioxide and other greenhouse gases, resulting in the continuous growth of carbon emissions. However, the sequential growth rate showed a downward trend, which is due to the further development of China's economy and society, the country began to change the existing economic development model, no longer rely on unreasonable energy consumption and industrial development model, the whole society's environmental awareness has been improved, people began to pay attention to ecological environmental issues. This makes China's carbon emissions are still growing, but the growth rate of carbon emissions began to decline year by year.

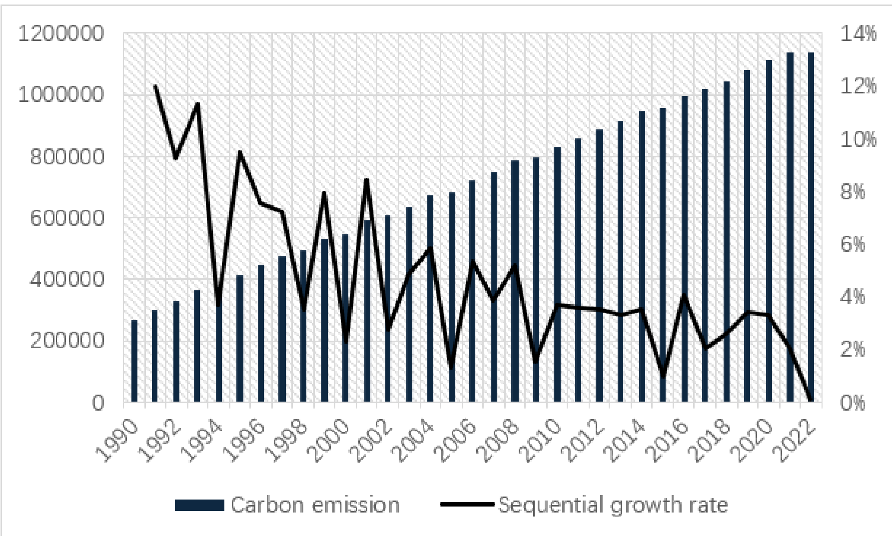


Figure 3. Horizontal distribution of carbon emissions in China from 1990 to 2022

2.2 Single forecasting model

2.2.1 Exponential smoothing method

Exponential smoothing is a time series analysis and prediction technique based on the principle of moving average method. The core of this method is to calculate the exponential smoothing value and combine the specific time series prediction model to achieve the accurate prediction of the future trend. At present, exponential smoothing method has been widely used in many fields, especially in production forecasting and short and medium term economic development trend forecasting, showing its unique advantages and accuracy.

The Holt-Winters model is based on exponential smoothing, which uses a weighted average of historical data to predict future values. This paper selects the Holt-Winter seasonless model, namely the quadratic exponential smoothing method, and constructs the model as follows:

$$\hat{y}_{t+k} = a_t + b_t k \tag{1}$$

$$a_t = \alpha y_t + (1-\alpha)(a_{t-1} + b_{t-1}) \tag{2}$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \tag{3}$$

Where, α and β are damping factors. $\alpha \in [0,1]$, $\beta \in [0,1]$.

The predicted values are calculated as follows:

$$\hat{y}_{T+k} = a_T + b_T k \tag{4}$$

Where a_T is the intercept term, b_T is the slope, T is the end value of the sample to be estimated, and k is the interval between T period and prediction period($k > 0$).

After the data of our country's carbon emission are processed by logarithm, the $\alpha = 0.64$, $\beta = 0.58$ are calculated by Eviews 13 software, and the H-W seasonal model is shown as (5).

$$in\hat{y}_{T+k} = 6.062 + 0.007k \tag{5}$$

2.2.2 ARIMA model

Because most of the economic time series are not stable, it often needs to be differential to mine a large number of data information, showing the characteristics of a stable time series [11]. The time series of the model is called the differential integrated moving average autoregressive model, which is denoted as ARIMA(p, d ,q):

The AR (p) model can be expressed as:

$$\begin{cases} x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p} + \varepsilon_t \\ \varphi_p \neq 0 \\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_S), s \neq t \\ E(x_s \varepsilon_t) = 0, \forall s < t \end{cases} \tag{6}$$

The MA (q) model can be expressed as:

$$\begin{cases} x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \\ \theta_q \neq 0 \\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_S), s \neq t \end{cases} \tag{7}$$

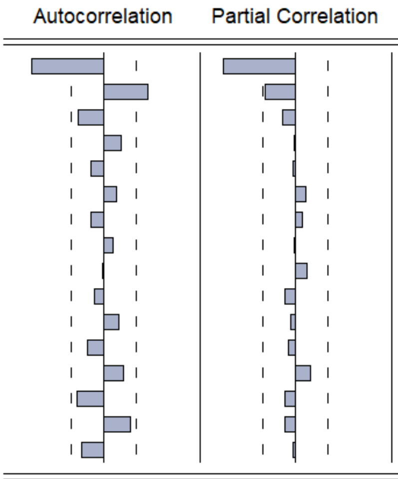
The ARIMA model is structured as follows:

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_S), s \neq t \\ E(x_s \varepsilon_t) = 0, \forall s < t \end{cases} \tag{8}$$

Where, $\nabla^d = (1-B)^d$; $\Phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$ is the autoregressive coefficient polynomial of the stationary reversible ARMA (p, q) model;

$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_p B^p$ is the moving average coefficient polynomial of the stationary reversible ARMA (p, q) model.

Firstly, the original data are processed by logarithm, the logarithm sequence is non-stationary sequence, which needs to be differential processed, the autocorrelation function and partial autocorrelation function are shown in Figure 4.

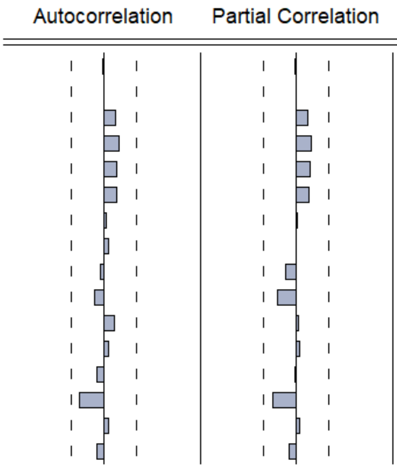


Lag	AC	PAC	Q-Stat	Prob
1	-0.773	-0.773	20.374	0.000
2	0.468	-0.322	28.099	0.000
3	-0.280	-0.139	30.954	0.000
4	0.185	-0.009	32.257	0.000
5	-0.138	-0.023	33.006	0.000
6	0.146	0.111	33.880	0.000
7	-0.140	0.069	34.711	0.000
8	0.098	-0.014	35.139	0.000
9	-0.011	0.125	35.144	0.000
10	-0.105	-0.121	35.677	0.000
11	0.163	-0.050	37.031	0.000
12	-0.172	-0.076	38.632	0.000
13	0.215	0.159	41.263	0.000
14	-0.289	-0.117	46.286	0.000
15	0.282	-0.107	51.354	0.000
16	-0.231	-0.029	55.008	0.000

Figure. 4 Autocorrelation and partial autocorrelation function diagram of second order difference sequence

Based on the optimal AIC criterion, $p=1,\ q=1$ are obtained through the analysis of Eviews13 software, and the residual error is also stable. Then the optimal ARIMA model is ARIMA(1,2,1), and the model is shown in Equation (9).

$$\hat{y}_t^{(2)} = -0.003 - 0.601\hat{y}_{t-1}^{(2)} + \varepsilon_t - 0.622\varepsilon_{t-1} \tag{9}$$



Lag	AC	PAC	Q-Stat	Prob
1	-0.009	-0.009	0.0029	0.957
2	-0.002	-0.002	0.0030	0.998
3	0.123	0.123	0.5556	0.907
4	0.160	0.165	1.5229	0.823
5	0.139	0.152	2.2797	0.809
6	0.132	0.138	2.9904	0.810
7	0.030	0.009	3.0277	0.882
8	0.048	-0.005	3.1316	0.926
9	-0.037	-0.122	3.1954	0.956
10	-0.103	-0.203	3.7142	0.959
11	0.107	0.027	4.3023	0.960
12	0.044	0.036	4.4080	0.975
13	-0.077	-0.014	4.7407	0.980
14	-0.268	-0.248	9.0545	0.828
15	0.043	0.036	9.1749	0.868
16	-0.076	-0.076	9.5705	0.888

Figure. 5 Residual autocorrelation function and partial autocorrelation function

2.2.3 Multiple linear regression model

Considering that multiple variables will have an impact on carbon emissions, a multiple linear regression model is selected, and its general model is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon \tag{10}$$

Where, $x_1, x_2, \ldots x_p$ is p independent variables $(p \geq 2)$, y is the dependent variable, β_0 is the regression constant, $\beta_1, \beta_2, \ldots \beta_p$ is the regression coefficient, and $\varepsilon \sim N(0, \sigma^2)$ is the random error.

Based on the existing studies on the main factors affecting carbon emissions [9], the following factors are selected: (1) economic development level (gdp). At present, there are a lot of researches on the effect of economic development on carbon emission in China. In this paper, GDP per capita is selected to represent the level of economic development. (2) Industrial structure (ISE/ISO). The change of industrial structure will have a direct or indirect impact on carbon emissions. The secondary industry occupies a major position in China's economic development, fossil energy consumption is huge, and the increase of the proportion of the tertiary industry will reduce carbon emissions, so choose the proportion of the secondary industry added value to GDP (ISE) and the proportion of tertiary industry added value to GDP (ISO). (3) State intervention (gov). The country may have an impact on carbon emissions by introducing incentive policies and increasing investment in environmental protection, which is represented by public financial expenditure in this paper. (4) Urbanization level (urban). The acceleration of urbanization will produce a large amount of energy consumption, which may have an impact on carbon emissions to a certain extent. This paper selects the proportion of permanent urban population in total population to represent the level of urbanization. (5) Total energy consumption (en) The process of energy consumption is accompanied by carbon dioxide emissions, and the proportion of total energy consumption in GDP is chosen to measure. (6) Openness to the outside world. Foreign trade can promote economic growth. With the expansion of China's foreign economic and trade scale, it may have a certain impact on carbon emissions. This paper selects the proportion of total import and export trade in GDP to express it. (7) Population size (peo). The increase of population size will lead to the increase of demand and become one of the main factors affecting carbon emissions. In this paper, the number of permanent residents is selected to represent it. All the above 8 variables and carbon emission intensity were logarithmically processed, and the following model was obtained after removing insignificant variables by step-by-step regression method:

$$\begin{aligned} \ln y_t &= -65.582 + 6.727 \ln peo_t + 0.294 \ln ISE_t + \varepsilon_t \\ &\qquad\qquad (133.902) \qquad\quad (6.577) \end{aligned} \tag{11}$$

In the formula, the coefficients of each independent variable are significant at the level of 1%, $\bar{R}^2 = 0.999$, and the model has a good degree of fitting.

For the above three models, the unit root test is performed on their residual sequences. If the residual sequence is listed as a stationary sequence, it indicates that the co-integration test is passed; if the residual sequence is listed as a non-stationary sequence, it indicates that the model cannot perform the subsequent combinatorial prediction.

Table 1 Residual sequence unit root test results table

Parameter	Multiple linear regression model	ARIMA model	Holt-Winter seasonless model
ADF statistics	−3.559	−5.261	−5.801
5% critical value	−2.957	−2.964	−2.960
P	0.013	0.000	0.000
Unit root	No	No	No
Result	Pass	Pass	Pass

3. CONSTRUCTION OF COMBINATION PREDICTION MODEL

3.1 Combined prediction model based on Shapley value

The core concept of Shapley value method is to fairly distribute the benefits or costs of cooperation according to the actual contributions made by each participant in the cooperation. This method adheres to the principle that the benefits or costs borne by each participant should match their contributions in the cooperation [10]. Suppose that there are n prediction methods for combined prediction, and the total error of combined prediction is E.

$$E = \frac{1}{m}\sum_{i=1}^{m} E_i \qquad (12)$$

Where, E_i is the ith prediction model.

The distribution model formula based on Shapley value is as follows:

$$E_i = \sum_{i \in s} \frac{(|s|-1)!(m-|s|)!}{m!} \times [E(s) - E(s-\{i\})] \qquad (13)$$

Where, E_i is the error of the ith model, namely the Shapley value; s is the set of models; $|s|$ is the number of prediction models in the combination; m is the total number of prediction models involved in combined prediction; $s-\{i\}$ means to remove member i from the combination.

The weight calculation formula of a single method in the combined model is as follows:

$$\omega_i = \frac{1}{m-1} \square \frac{E - E_i}{E} \qquad (14)$$

Firstly, the Shapley value of the corresponding member 1 in the combination is calculated according to equation (13) :

$$\begin{aligned} E_1 &= \left(\frac{0!2!}{3!}\right)[E\{1\} - E\{1\} - \{1\}] + \frac{1!1!}{3!}[E\{1,2\} - E(\{1,2\} - \{1\})] \\ &+ \frac{1!1!}{3!}[E\{1,3\} - E(\{1,3\} - \{1\})] + \frac{2!0!}{3!}[E\{1,2,3\} - E(\{1,2,3\} - \{1\})] \qquad (15) \\ &= 5456.877 \end{aligned}$$

In the same way, it can be calculated that the error amount that should be shared by prediction method 2 and 3 in combination prediction is $E_2 = 2083.758$ and $E_3 = 2839.572$ respectively. The sum of the three apportionment results is exactly the total error deferred value, and the calculation is correct. According to the above results, the weights of each method in the combined prediction are calculated, and $\omega = (0.237, 0.400, 0.363)$ is obtained by substituting the data.

$$Y_t = 0.237\hat{y}_{1t} + 0.400\hat{y}_{2t} + 0.363\hat{y}_{3t} \tag{16}$$

Where, $\hat{y}_{1t}$ is the predicted value of the H-W non-seasonal model, $\hat{y}_{2t}$ is the predicted value of the ARIMA model, and $\hat{y}_{3t}$ is the predicted value of the multiple linear regression model.

3.2 Combination prediction model based on IOWGA operator

3.2.1 OWGA and IOWGA operators

Let $G_W : R^m \to R$ be an m -ary function, $W = (w_1, w_2, \dots, w_m)^{\mathrm{T}}$ is the weighting vector associated with G_W ,satisfying $\sum_{i=1}^{m} w_i = 1, w_i \geq 0$, $i = 1, 2, \dots, m$, if $G_W(a_1, a_2, \dots, a_m) = \prod_{i=1}^{m} b_i^{w_i}$, where b_i is the i largest number in $a_1, a_2, \dots, a_m$ in order from largest to smallest, then the function G_W is said to be an m -dimensional ordered weighted geometric average operator, abbreviated as the OWGA operator.

Let $\langle u_1, a_1 \rangle, \langle u_2, a_2 \rangle, \cdots, \langle u_m, a_m \rangle$ is m two-dimensional arrays, so:

$$G_W\left(\langle u_1, a_1 \rangle, \langle u_2, a_2 \rangle, \dots \langle u_m, a_m \rangle\right) = \prod_{i=1}^{m} a_{u-\text{index(i)}}^{w_i} \tag{17}$$

It is said that the function G_W is the m -dimensional induced ordered weighted geometric average operator generated by $u_1, u_2, \dots, u_m$, briefly denoted as the IOWGA operator, u_i is the induced value of a_i, where $u - \text{index}(i)$ is the subscript of the i t

h largest number in $u_1, u_2, \dots, u_m$ in order from largest to smallest, $W = (w_1, w_2, \dots, w_m)^{\mathrm{T}}$ is the weighted vector of OWGA, satisfying $\sum_{i=1}^{m} w_i = 1, w_i \geq 0, i = 1, 2, \dots, m$. Formula (17) shows that the IOWGA operator is to carry out an ordered weighted geometric average of the number in $a_i, a_1, a_2, \dots, a_m$ corresponding to the induced value $u_1, u_2, \dots, u_m$ after it is arranged in the order from largest to smallest, and w_i has nothing to do with the size and position of a_i, but with the position where the induced value is located.

3.2.2 Optimal combination prediction model based on IOWGA operator

Based on the concept of induced ordered weighted geometric mean (IOWGA) operator, the following combinatorial prediction model is considered. Let:

$$p_{it}=\begin{cases}1-|(x_t-x_{it})/x_t| & \text{when}\,|(x_t-x_{it})/x_t|<1\\ 0 & \text{when}\,|(x_t-x_{it})/x_t|\geq 1\end{cases}\qquad (18)$$
$$i=1,2,\ldots,m,\quad t=1,2,\ldots,N$$

Where p_{it} represents the prediction accuracy at time t of the i prediction method, $p_{it}\in[0,1]$, and x_t represents the actual value; x_{it} represents the predicted value. In this way, the prediction accuracy of the t time of m single prediction method and its corresponding predicted value in the sample interval constitute m two-dimensional arrays. $\langle p_{1t},x_{1t}\rangle,\langle p_{2t},x_{2t}\rangle,\cdots,\langle p_{mt},x_{mt}\rangle$. $W=(w_1,w_2,\ldots,w_m)^{\mathrm{T}}$ is the weighting vector of OWGA, and the prediction accuracy sequence $p_{1t},p_{2t},\cdots p_{mt}$ at the t moment of m single prediction methods is arranged in order from largest to smallest, Let $p-\text{index}(it)$ be the subscript of the i largest prediction accuracy at time t, according to equation (18), let:

$$\text{IOWGA}\left(\langle p_{1t},x_{1t}\rangle,\langle p_{2t},x_{2t}\rangle,\ldots,\langle p_{mt},x_{mt}\rangle\right)=\prod_{i=1}^{m}x_{p-\text{index(it)}}^{w_i}\qquad (19)$$

Then formula(19) is the predicted value of IOWGA combination at time t generated by prediction precision sequence $p_{1t},p_{2t},\cdots p_{mt}$.

Let: $e_{p-\text{index}(it)}=\ln x_t-\ln x_{p-\text{index}(it)}$. Therefore, the sum of the logarithm squared error of the total combinatorial prediction for the N period is:

$$\begin{aligned}S&=\sum_{t=1}^{N}(\ln x_t-\ln\prod_{i=1}^{m}x_{p-\text{index}(it)}^{w_i})^2\\&=\sum_{t=1}^{N}(\ln x_t-\sum_{i=1}^{m}w_i\ln x_{p-\text{index}(it)})^2\\&=\sum_{i=1}^{m}\sum_{j=1}^{m}w_iw_j(\sum_{t=1}^{N}e_{p-\text{index}(it)}e_{p-\text{index}(jt)})\end{aligned}\qquad (20)$$

Therefore, the combinatorial prediction model based on IOWGA can be expressed as:

$$\min S(W)=\sum_{i=1}^{m}\sum_{j=1}^{m}w_iw_j(\sum_{t=1}^{N}e_{p-\text{index}(it)}e_{p-\text{index}(jt)})$$
$$s.t\begin{cases}\sum_{i=1}^{m}w_i=1\\ w_i\geq 0,i=1,2,\ldots,m\end{cases}\qquad (21)$$

Let $E_{ij}=\sum_{t=1}^{N}e_{p-\text{index}(it)}e_{p-\text{index}(jt)}$, $i,j=1,2,\ldots,m$, then $E=(E_{ij})_{m\times m}$ is a square matrix of combinational prediction logarithmic error information of m -order IOWGA, and equation (20) can be expressed as equation (21) :

$$\min S(W) = W^{\mathrm{T}}EW$$
$$s.t\begin{cases} R^{\mathrm{T}}W = 1 \\ W \geq 0 \end{cases} \tag{22}$$

Where, $R = (1,1,\ldots,1)^{\mathrm{T}}$.

3.2.3 Error index evaluation system

(1) Error sum of squares SSE: $\mathrm{SSE} = \sum_{t=1}^{N}(x_t - \hat{x}_t)^2$;

(2) mean square error RMSE: $\mathrm{RMSE} = \frac{1}{N}\sqrt{\sum_{t=1}^{N}(x_t - \hat{x}_t)^2}$;

(3) Average absolute error MAE : $\mathrm{MAE} = \frac{1}{N}\sum_{t=1}^{N}| x_t - \hat{x}_t |$;

(4) Average absolute percentage error MAPE : $\mathrm{MAPE} = \frac{1}{N}\sum_{t=1}^{N}| (x_t - \hat{x}_t)/x_t |$;

(5) Mean square percentage error MSPE: $\mathrm{MSPE} = \frac{1}{N}\sqrt{\sum_{t=1}^{N}[(x_t - \hat{x}_t)/x_t]^2}$.

4. EMPIRICAL ANALYSIS

4.1 Comparative analysis of the results of each model

By replacing the predicted values and prediction accuracy of each model into the above formula, the IOWGA combined prediction model with the smallest sum of the square of the predicted logarithm error is obtained as follows:

$$\min S = 0.003w_1^2 + 0.004w_1w_2 + 0.014w_1w_3 + 0.008w_2^2 + 0.016\mathrm{w}_2w_3 + 0.033w_3^2$$

$$s.t\begin{cases} w_1 + w_2 + w_3 = 1 \\ w_1、w_2、w_3 \geq 0 \end{cases}$$

The optimization model is solved by LINGO18.0, and the corresponding optimal weight coefficients $w_1^* = 0.857$, $w_2^* = 0.143$, $w_3^* = 0$ are obtained. According to the optimal weight coefficient and the combination prediction model of IOWGA operator, the combined prediction value of each year is obtained.

Table 2 Carbon emission projections for China from 1990 to 2022

year	original data	Holt-Winter seasonless model	ARIMA model	Multiple linear regression model	Combined prediction model based on Shapley value	Combinatorial prediction model based on IOWGA operator
1990	267264	267264	267264	270604	268477	267264
1991	299328	284258	299328	296188	286413	294614

1992	327098	302436	326035	323805	305811	319628
1993	364117	344888	362777	356783	347446	356357
1994	377580	388911	398494	384631	390281	391186
1995	413385	441606	419733	414455	438478	423003
1996	444685	453916	440301	445519	451969	445245
1997	476765	483353	478558	476687	482668	479015
1998	493571	516600	511288	502759	515841	509450
1999	532855	551060	531222	529712	548223	535378
2000	545176	561204	557594	558057	560688	558619
2001	591322	597009	581842	581953	594840	585479
2002	607746	605409	610562	606395	606146	607826
2003	637164	652558	644412	636250	651393	643379
2004	674290	673208	662026	663064	671609	665055
2005	683090	698434	698670	694781	698468	697201
2006	719581	739385	717517	722277	736258	724432
2007	747344	743847	736092	744715	742738	741063
2008	786251	774028	772327	771028	773785	772259
2009	798434	806020	808428	791604	806365	801746
2010	827736	852597	831274	820351	849548	832363
2011	857441	863715	846311	855059	861226	853616
2012	887827	884301	877859	892584	883379	884735
2013	916984	913693	909386	921140	913077	914677
2014	949158	947572	939800	956611	946460	947749
2015	958741	980060	971189	973442	978791	974111
2016	997870	1014927	986643	1007812	1010883	1001040
2017	1018136	1017390	1008787	1048449	1016160	1025233
2018	1044438	1050709	1041056	1074132	1049328	1055359
2019	1080261	1073429	1061462	1089346	1071717	1074428
2020	1116192	1098211	1095265	1093715	1097790	1095401
2021	1138923	1138303	1135880	1108375	1137956	1126464
2022	1140098	1183628	1164228	1104200	1180854	1147025

The prediction accuracy of each year's combination is shown in Table 3:

Table 3 Prediction accuracy of China's carbon emissions by model from 1990 to 2022

year	Holt-Winter seasonless model	ARIMA model	Multiple linear regression model	Combined prediction model based on Shapley value	Combinatorial prediction model based on IOWGA operator
1990	1.000	1.000	0.988	0.995	1.000
1991	0.950	1.000	0.990	0.957	0.984

1992	0.925	0.997	0.990	0.935	0.977
1993	0.947	0.996	0.980	0.954	0.979
1994	0.970	0.945	0.981	0.966	0.964
1995	0.932	0.985	0.997	0.939	0.977
1996	0.979	0.990	0.998	0.984	0.998
1997	0.986	0.996	1.000	0.988	0.995
1998	0.953	0.964	0.981	0.955	0.968
1999	0.966	0.997	0.994	0.971	0.995
2000	0.971	0.977	0.976	0.972	0.975
2001	0.990	0.984	0.984	0.994	0.990
2002	0.996	0.995	0.998	0.997	1.000
2003	0.976	0.989	0.999	0.978	0.990
2004	0.998	0.982	0.983	0.996	0.986
2005	0.978	0.977	0.983	0.977	0.979
2006	0.972	0.997	0.996	0.977	0.993
2007	0.995	0.985	0.996	0.994	0.992
2008	0.984	0.982	0.981	0.984	0.982
2009	0.990	0.987	0.991	0.990	0.996
2010	0.970	0.996	0.991	0.974	0.994
2011	0.993	0.987	0.997	0.996	0.996
2012	0.996	0.989	0.995	0.995	0.997
2013	0.996	0.992	0.995	0.996	0.997
2014	0.998	0.990	0.992	0.997	0.999
2015	0.978	0.987	0.985	0.979	0.984
2016	0.983	0.989	0.990	0.987	0.997
2017	0.999	0.991	0.970	0.998	0.993
2018	0.994	0.997	0.972	0.995	0.990
2019	0.994	0.983	0.992	0.992	0.995
2020	0.984	0.981	0.980	0.984	0.981
2021	0.999	0.997	0.973	0.999	0.989
2022	0.962	0.979	0.969	0.964	0.994

It can be seen from Table (4) that the error index values of the combined prediction model based on IOWGA operator are significantly lower than the error index values of each single prediction and Shapley combined prediction, indicating that the combined prediction method based on IOWGA operator is significantly better than each single prediction method and effectively overcomes the limitations of a single prediction model. The prediction effect is significantly improved.

Table 4 Evaluation of error indexes of each forecasting model

	SSE	RMSE	MAE	MAPE	MSPE
Holt-Winter seasonless model	8233689352	2749.687	12377.015	0.021	0.005

ARIMA model	4085648556	1936.942	9003.897	0.013	0.003
Multiple linear regression model	6123956259	2371.385	9759.710	0.013	0.003
Combined prediction model based on Shapley value	6933338533	2523.232	11494.986	0.019	0.005
Combinatorial prediction model based on IOWGA operator	2627403641	1553.279	7301.955	0.011	0.003

4.2 Carbon emission forecast for the next five years

When constructing IOWGA combinatorial forecasting model, the traditional approach is to assign weight to individual forecasting methods according to prediction accuracy. However, for the prediction of future carbon emissions, because the actual value is unknown, it is impossible to directly calculate the prediction accuracy and corresponding weights of future time points. Therefore, according to the weight coefficients of the three single forecasting methods in each year, the average weight of each single forecasting method is obtained as: $w_1^* = 0.316$, $w_2^* = 0.333$, $w_3^* = 0.351$.

Table 5 Weight coefficients of individual forecasting methods from 1990 to 2022

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Holt-Winter seasonless model	0.857	0	0	0	0.143	0	0	0	0	0	0
ARIMA model	0.143	0.857	0.857	0.857	0	0.143	0.143	0.857	0.143	0.857	0.857
Multiple linear regression model	0	0.143	0.143	0.143	0.857	0.857	0.857	0.143	0.857	0.143	0.143
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Holt-Winter seasonless model	0.857	0.143	0	0.857	0.143	0	0.143	0.857	0.143	0	0.143
ARIMA model	0.143	0	0.143	0	0	0.857	0	0.143	0	0.857	0
Multiple linear regression model	0	0.857	0.857	0.143	0.857	0.143	0.857	0	0.857	0.143	0.857
	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Holt-Winter seasonless model	0.857	0.857	0.857	0	0	0.857	0.143	0.857	0.857	0.857	0
ARIMA model	0	0	0	0.857	0.143	0.143	0.857	0	0.143	0.143	0.857
Multiple linear regression model	0.143	0.143	0.143	0.143	0.857	0	0	0.143	0	0	0.143

Combined with each individual prediction model, the individual carbon emission forecast value of 2023-2027 was calculated, and then the combined forecast value of 2023-2027 was obtained by multiplying and summing the weight coefficients.

Table 6 Predicted values of China's carbon emissions from 2023 to 2027

year	2023	2024	2025	2026	2027
Holt-Winter seasonless model	1208434	1242844	1278234	1314632	1352066
ARIMA model	1164853	1170822	1183631	1187588	1192109
Multiple linear regression model	1096628	1085838	1071970	1055161	1035562
Combined prediction model based on Shapley value	1150408	1157034	1165508	1169616	1173181
Combinatorial prediction model based on IOWGA operator	1154678	1163752	1174333	1181252	1187707

According to the individual forecast model and the combined forecast model, China's carbon emissions will still show a growing trend in the next five years, but the annual growth rate will decrease significantly. The average annual growth rate of each model is 2.70%, 0.48%, －1.27%, 0.45% and 0.12%, respectively.

5. CONCLUSION

Based on China's carbon emissions data from 1990 to 2022, this paper first uses three single prediction models, namely Holt-Winters seasonless model, ARIMA (1,2,1) model and multiple linear regression model, to forecast China's carbon emissions from 2023 to 2027. Then the combined prediction model based on Shapley value and the combined prediction model based on IOWGA operator are constructed to predict China's carbon emissions. Finally, the results of comparison and analysis with three single forecasting models show that:

First, for a single forecasting model, the accuracy of different forecasting methods at the same forecasting point is significantly different. At the same time, for the same forecasting model, its accuracy at different forecasting points also presents certain fluctuations.

Second, by comparing the three single prediction models and two combined prediction models through the error index system, it is found that the average prediction accuracy of the combined prediction model is higher than that of the single prediction model, and the error term is lower than that of the single prediction model. In particular, the combined prediction model based on the IOWGA operator performs better, indicating that the combined prediction model has better prediction effect. The disadvantages of each single forecasting model are overcome.

Third, through the prediction of China's carbon emissions from 2023 to 2027 by each individual forecasting model and combined forecasting model, it can be seen that China's carbon emissions still show an upward trend in the next five years, but the growth rate is lower than before.

As a major carbon emitter in the world, China is seeking a balance between high-quality economic development and carbon reduction. In order to achieve the "double carbon" goal, China should optimize the energy structure and promote industrial upgrading, accelerate the development of renewable energy, and encourage and support the development of green and low-carbon industries. The government needs to expand financial support to promote green technology innovation and reduce energy consumption in the production process. At the same time, increase publicity and education, advocate low-carbon life, and improve public awareness and participation in green and low-carbon development.

REFERENCES

1. Zhao Aiwen, Li Dong. Grey prediction of carbon emission in China [J]. Mathematics in Practice and Understanding,2012,42(04):61-69.
2. Song Jiekun. Carbon emission prediction model based on STIRPAT and partial least squares regression [J]. Statistics and Decision,2011(24):19-22.
3. Qu Shenning, Guo Zhaoxian. Prediction of China's carbon emission peak based on STIRPAT model[J]. China Population, Resources and Environment,2010,20(12):10-15.
4. Zhao Xi, Qi Jianmin, Liu Guangwei. Prediction of carbon emissions in China based on discrete second-order difference algorithm [J]. Journal of Arid Land Resources and Environment,2013,27(01):63-69.
5. Zhao Chengbai, Mao Chunmei. Prediction of carbon emission intensity in our country based on Arima and BP neural network combined model [J]. Resources and environment of the Yangtze River basin, 2012,21(06) : 665-671.
6. Gao Jie, Wang Jianwei, Li Linna. A combined prediction model of transportation carbon emissions based on Richards-BP neural network [J]. Journal of Chang 'an University (Natural Science Edition),2013,33(04):99-104+109.
7. Chui Wan-Bing, Lv Xue-Bin. Carbon emission trading price prediction using ARMI-SSA-LSTM model [J]. Journal of Xi 'an University of Science and Technology, 2019,43(05):1025-1034.
8. Hu Linlin, Jia Junsong. Carbon emission prediction of Jiangxi tourism industry based on combined ESARIMA model [J]. Journal of Beijing International Studies University,2014,36(01):34-38.
9. Wang Shaojian, Xie Zihan, Wang Zehong. Spatial and temporal evolution and influencing factors of county carbon emissions in China [J]. Acta Geographica Sinica, 2019,76(12):3103-3118.
10. Chen Du, Zhang Ting. Research on financial income distribution model of PPP project based on Shapely value method: A case study of Beijing-Shanghai high-speed Railway Project [J]. Journal of Finance and Accounting,2022(18):154-160.
11. Wu Qingqing, Yang Guiyuan. Portfolio forecasting of tertiary industry GDP based on IOWGA operator [J]. Statistics and Decision,2015(19):68-71.