

Prediction of Carbon Emission Right Price Based on XGBoost Algorithm

Fuyun Zhu¹, Peiyuan Liu², Ping Hu³

¹ Professor, School of Economics & Management, Shaanxi University of Science & Technology, Xi'an, China

² Master, School of Economics & Management, Shaanxi University of Science & Technology, Xi'an, China

³ Bachelor, School of Economics & Management, Shaanxi University of Science & Technology, Xi'an, China

Abstract: Reasonable carbon market price prediction can facilitate the carbon market participants, such as physical producers, to achieve the goal of efficient emission reduction through the market mechanism. In this paper, we use XGBoost, an integrated algorithm in machine learning, to forecast the domestic carbon price from 2013 to 2021. Pearson coefficient is utilized to calculate the correlation of data features, and perform PCA dimensionality reduction on the features with high correlation coefficients. Before using PCA to reduce the dimensions, in order to make the feature data more suitable for later model training, the feature data is standardized. After the standardization, PCA selects super parameters with maximum likelihood estimation and outputs the dimension reduction results of features. Finally, the integrated algorithm XGBoost is used to form a prediction model for the carbon price. RMSE output from cross validation are used to evaluate the accuracy and error of the prediction results. The results show that this paper confirms that the integrated algorithm XGBoost model has a good prediction ability for carbon price, and provides a novel idea for the field of carbon emission price prediction. It is expected to provide some rational basis for the carbon market participants to make investment decisions, so as to avoid the carbon market risk caused by the change of carbon price.

Keywords: Prediction of carbon price, Machine learning, Data processing and analysis, XGBoost, Bayes optimization.

1. Introduction

The introduction should provide background information (iCarbon emission trading market (carbon market) is an important policy tool for the international community to address climate change. Excessive emission of greenhouse gas, mainly carbon dioxide, has led to a gradual rise in global temperature, and an over-focus on reducing carbon dioxide emissions will certainly bring high treatment costs. Quantifying carbon emissions means that the government provides carbon emission quotas to enterprises, and enterprises with surplus carbon emissions can transfer the surplus part to enterprises with surplus carbon emissions through trading to achieve optimal allocation of resources and achieve emission reduction target at a lower cost.

In June 1992, the United Nations Framework Convention on Climate Change encouraged developed countries to reduce carbon dioxide emissions and set the ultimate goal of "stabilizing the concentration of atmospheric greenhouse gases at a level that would prevent dangerous anthropogenic interference with the climate system. This level should be achieved within a time frame sufficient to make the ecosystem sustainable." In December 1997, the Third Conference of the Parties was held in Kyoto, Japan. The signing of the Kyoto Protocol forced major industrialized countries to reduce their carbon dioxide emissions from 2008 to 2012 by 5.2% from 1990. With the Kyoto Protocol and the provision of three flexible performance mechanisms, namely Joint Implementation Mechanism (JI), Clean Development Mechanism (CDM) and International Emissions Trading (IET), these three flexible mechanisms have given birth to the international carbon emission trading market. The EU established the EU Emission Trading System (EU-ETS) in April 2005, and has made remarkable achievements in environmental and economic benefits,

becoming the most influential global carbon emission trading market in the world.

In October 2011, the Notice of the General Office of the National Development and Reform Commission on the Pilot Project of Carbon Emission Trading identified that Beijing, Tianjin, Shanghai, Chongqing, Hubei, Guangdong and Shenzhen should carry out the pilot work of carbon emission trading. In 2019, China's carbon dioxide emissions reached 9.826 billion tons, of which 80% will come from coal combustion, and the carbon dioxide emissions of power, industry and transportation industries will account for 90% of the total emissions. At the 75th United Nations General Assembly in 2020, China pledged to peak its carbon dioxide emissions by 2030 and strive to achieve carbon neutrality by 2060. At the National People's Congress and the Chinese People's Political Consultative Conference in 2021, Premier Li Keqiang pointed out that efforts should be made to achieve the vision of "carbon neutrality". On July 16, 2021, the national carbon market officially started online trading. On March 15, 2022, the Ministry of Ecology and Environment issued the Notice on Key Work Related to the Management of Enterprises' Greenhouse Gas Emission Reports in 2022, which established the list of key emission units in the power generation industry in the second compliance cycle of the national carbon market. According to the verification results, the annual carbon emissions in 2020 or 2021 will reach 26,000 tons of carbon dioxide equivalent, It also has key emission units in the power generation industry that meet the quota management standards for units, and is included in the list of key emission units under quota management in the national carbon emission trading market in 2022.

The research on carbon trading market price prediction has focused on two main points: one is to use only the historical data of carbon trading market price to predict

carbon price [1-5], and the other is to predict carbon price by combining the influencing factors of carbon trading market price [6-10]. However, since there are few articles that apply machine learning to carbon price prediction in China. Based on this, this paper will use XGBoost model to predict the carbon price, which will provide new ideas for the carbon price prediction field, and provide some practical basis for the carbon market participants to make investment decisions, so as to avoid the carbon market risk caused by the carbon price change.

2. Research Method

2.1. XGBoost algorithm principle

Extreme Gradient Boosting (XGboost) is an end-to-end gradient lifting tree system proposed by Dr. Tianqi Chen of Washington University. XGBoost implements a generalized Tree Boosting algorithm. One representative of this algorithm is the Gradient Boosting Decision Tree (GBDT). XGBoost adopts the idea of integration, adds regular terms on the basis of GBDT, and extends the loss function from square loss to Taylor's second-order loss.

The following article will briefly describe the principle of XGBoost algorithm:

2.1.1. Boosting additive model representation

$$\hat{y}_i^{(m)} = \sum_{j=1}^m f_j(x_i) \quad (1)$$

Where $\hat{y}_i^{(m)}$ refers to the predicted value of the i th sample equal to the sum of m regression trees, and x_i refers to the characteristic value of the i th sample. Forward distribution algorithm:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i) \quad (2)$$

Where t represents the t -th tree to be optimized, $t-1$ represents the $t-1$ tree that has been optimized, and $f_t(x_i)$ represents the t -th tree to be optimized.

objective function:

$$obj^{(t)} = \sum_{i=1}^N L(y_i, \hat{y}_i^{(t)}) + \sum_{j=1}^t \Omega(f_j) \quad (3)$$

Where $\sum_{i=1}^N L(y_i, \hat{y}_i^{(t)})$ represents the total loss of N samples, and $\sum_{j=1}^t \Omega(f_j)$ is a regular term (adding the complexity of t regression trees). The purpose of adding a regular term is to prevent over fitting.

Optimization objective:

$$(w_1^*, w_2^*, \dots, w_r^*) = \arg \min obj^{(t)} \quad (4)$$

Set T leaf nodes, determine their node values, and minimize the loss function.

2.1.2. Processing of regular terms

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (5)$$

Where, γ , λ is a super parameter and a penalty item for variables. T is equal to leaf nodes of the regression tree.

$\sum_{j=1}^T w_j^2$ represents the sum of the squares of each node value.

When the number of leaf nodes T of the regression tree is larger, the depth of the tree is deeper, and the over fitting phenomenon is more likely to occur. Therefore, T should be punished, and the punishment intensity should be controlled by γ ; When $\sum_{j=1}^T w_j^2$ is larger, that is, when the node value is larger, the proportion of the node value of this regression tree in all regression trees is too large. At this time, the phenomenon of over fitting also occurs. Therefore, $\sum_{j=1}^T w_j^2$ is punished, and the punishment is determined by λ .

2.1.3. Treatment of $\sum_{j=1}^t \Omega(f_j)$ in objective function

$obj^{(t)}$

$$\sum_{j=1}^t \Omega(f_j) = \sum_{j=1}^{t-1} \Omega(f_j) + \Omega(f_t) \quad (6)$$

The regular term in the objective function is decomposed into the optimized first $t-1$ term (Constant) and the term to be optimized. In the optimization process of the objective function, that is, the minimization process, constant is not required to be considered:

$$\sum_{j=1}^t \Omega(f_j) = \Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (7)$$

Assume that the sample loss is the square error loss:

$$L(y_i, \hat{y}_i) = (y_i - \hat{y}_i)^2 \quad (8)$$

Take the minimum loss calculation of node 1 as an example:

$$\begin{aligned} L_{\text{node 1}} &= L_3 + L_5 = (y_3 - w_1)^2 + (y_5 - w_1)^2 \\ &= (y_3^2 + y_5^2) - 2(y_3 + y_5)w_1 + w_1^2 \end{aligned} \quad (9)$$

Where y is a real value and a constant. Minimum W can be obtained:

$$\min w_1 = -\frac{b}{2a} = y_3 + y_5 \quad (10)$$

That is, when $W_1 = y_3 + y_5$, the loss of L -node 1 is the minimum. (The same with W_2 , W_3 and W_4)

Review of objective function and regular term formula:

$$obj^t = \sum_{i=1}^N L(y_i, \hat{y}_i^{(t)}) + \sum_{j=1}^t \Omega(f_j) \quad (11)$$

$$\sum_{j=1}^t \Omega(f_j) = \Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (12)$$

The regular term is calculated by leaf node j. At this time, in order to ensure the consistency of the calculation method before and after the formula, the following transformation is made:

Convert the loss $\sum_{i=1}^N L(y_i, \hat{y}_i^{(b)})$ calculated by sample to the loss calculated by leaf node:

$$\sum_{i=1}^N L(y_i, \hat{y}_i^{(t)}) = \sum_{j=1}^T \sum_{i \in I_j} L(y_i, \hat{y}_i^{(t)}) \quad (13)$$

Substituted into the objective function:

$$obj^{(t)} = \gamma T + \sum_{j=1}^T \sum_{i \in I_j} L(y_i, \hat{y}_i^{(t)}) + \frac{1}{2} \lambda w_j^2 \quad (14)$$

where $\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + w_j$, in summary:

$$obj^{(t)} = \gamma T + \sum_{j=1}^T [\sum_{i \in I_j} L(y_i, \hat{y}_i^{(t-1)} + w_j)] + \frac{1}{2} \lambda w_j^2 \quad (15)$$

In order to facilitate the calculation of node value W, the objective function is divided into unified leaf node calculation.

The above assumption is that the sample loss is a square error loss, but the loss functions corresponding to different loss scenarios are different, and it is likely that the loss function cannot be changed into a polynomial containing W^2 for solution, so Taylor second-order processing is performed on the loss function $L(y_i, \hat{y}_i^{(t-1)} + w_j)$:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2$$

2.1.4. Taylor processing of $L(y_i, \hat{y}_i^{(t-1)} + w_j)$

$$L(y_i, \hat{y}_i^{(t-1)} + w_j) = L(y_i, \hat{y}_i^{(t-1)}) + L'(y_i, \hat{y}_i^{(t-1)}) w_j + \frac{1}{2} L''(y_i, \hat{y}_i^{(t-1)}) w_j^2$$

, Where, W_j is the unique variable, and other items are constants. These constants need not be considered in the optimization of loss function, and the first step and the second step are g_i and h_i respectively:

$$L(y_i, \hat{y}_i^{(t-1)} + w_j) \approx g_i w_j + \frac{1}{2} h_i w_j^2 \quad (16)$$

Substituted into the objective function:

$$obj^{(t)} = \gamma T + \sum_{j=1}^T [\sum_{i \in I_j} (g_j w_j + \frac{1}{2} h_j w_j^2)] + \frac{1}{2} \lambda w_j^2 \quad (17)$$

$$\sum_{i \in I_j} (g_i w_j + \frac{1}{2} h_i w_j^2) = w_j \sum_{i \in I_j} g_i + \frac{1}{2} w_j^2 \sum_{i \in I_j} h_i \quad (18)$$

in summary:

$$obj^{(t)} = \gamma T + \sum_{j=1}^T [w_j \sum_{i \in I_j} g_i + \frac{1}{2} w_j^2 (\lambda + \sum_{i \in I_j} h_i)] \quad (19)$$

In each leaf node, the sum of a step degree g_i of all samples is called G_j ; In each leaf node, the sum of two steps of all samples is called H_j .

To sum up, the objective function is obtained:

$$obj^{(t)} = \gamma T + \sum_{j=1}^T [w_j G_j + \frac{1}{2} w_j^2 (\lambda + H_j)] \quad (20)$$

2.1.5. Minimize the node value of the objective function

$$\begin{aligned} (w_1^*, w_2^*, \dots, w_j^*) &= \arg \min obj^{(t)} \\ w_1^* &= \arg \min [w_1 G_1 + \frac{1}{2} w_1^2 (\lambda + H_1)] \\ w_2^* &= \arg \min [w_2 G_2 + \frac{1}{2} w_2^2 (\lambda + H_2)] \\ &\vdots \\ w_j^* &= \arg \min [w_j G_j + \frac{1}{2} w_j^2 (\lambda + H_j)] \end{aligned} \quad (21)$$

At this time, the optimal solution can be obtained:

$$w_j^* = -\frac{b}{2a} = -\frac{G_j}{\lambda + H_j} \quad (22)$$

Substitute the optimal solution into the objective function to further obtain the optimal objective function:

$$obj^{(t)*} = \min obj^{(t)} = \gamma T - \frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{\lambda + H_j} \quad (23)$$

2.2. Variable Design and Descriptive Statistics of Samples

The data in this paper comes from the database of China Research Data Service Platform (CNRDS), which covers the daily data of carbon emission trading information in several provinces and cities (Beijing, Fujian, Guangdong, Hubei, Shanghai, Shenzhen, Tianjin, Chongqing) from August 5, 2013 to October 8, 2021. The data indicators include: trading date, city name, trade variety abbreviation, opening price, highest price, lowest price, average transaction price (yuan/ton) The closing price, trading volume (tons), and turnover were 10615 data samples. (Table 1)

Table 1. Characteristic descriptive statistics

	O price	H price	m price	AT price (yuan/ton)	C price	TV (ton)	turnover
count	10615	10615	10615	10615	10615	10615	10615
mean	30.7335968	31.46141875	30.29406029	30.73066886	31.07137164	23412.55315	612664.6003
std	17.44338215	17.62580027	17.34321116	17.43692862	18.35309228	126957.9379	4539548.708
min	1	1	1	1	1	1	2
25%	18.5	19	18.255	18.405	18.52	50	949.89
50%	27.7	28.49	27.02	27.73	27.9	1648	46380
75%	39.5	40.2	39	39.5	39.995	13002.5	377307.5
max	130.9	143.99	119	122.97	555.43	8474383	354157396.3

2.3. Research Method Design

In order to improve the prediction accuracy of carbon emission right price and the robustness of the model, the modeling in this paper is divided into the following three steps: ① divide the data set, take the first 10508 groups of data as the training set, and the last 107 groups of data as the test set, and predict the closing price of the last 107 groups of data. ② The Pearson correlation coefficient is utilized to analyze the correlation degree between features. In order to ensure that the data fit the model better, the data is standardized, and the features with high correlation coefficient are reduced by the MLE of PCA. ③ Building

XGBoost model to make an initial prediction on data. Use Cross Validate and Bayes optimization to optimize XGBoost model and get the final prediction results. Finally, MAPE indicators are used to evaluate the robustness of the prediction model, and RMSE and MAE indicators are used to evaluate the accuracy of the prediction results.

3. Empirical Analysis

3.1. EDA Analysis

In order to better fit the model training of features, this paper uses Pearson correlation coefficient index to analyze the correlation degree of feature variables (Table 2) (Figure 1).

Table 2. Pearson correlation coefficient between features

	O price	H price	m price	AT price (yuan/ton)	C price	TV (ton)	turnover
O price	1	0.994613748	0.996937241	0.997143961	0.958351867	-0.0402325	0.034012596
H price	0.994613748	1	0.992208845	0.995677459	0.961935042	-0.037980419	0.036354499
m price	0.996937241	0.992208845	1	0.997399303	0.958658076	-0.04180245	0.033424528
AT price (yuan/ton)	0.997143961	0.995677459	0.997399303	1	0.960931526	-0.040839603	0.034046921
C price	0.958351867	0.961935042	0.958658076	0.960931526	1	-0.038659397	0.032932777
TV (ton)	-0.0402325	-0.037980419	-0.04180245	-0.040839603	-0.038659397	1	0.897452236
turnover	0.034012596	0.036354499	0.033424528	0.034046921	0.032932777	0.897452236	1

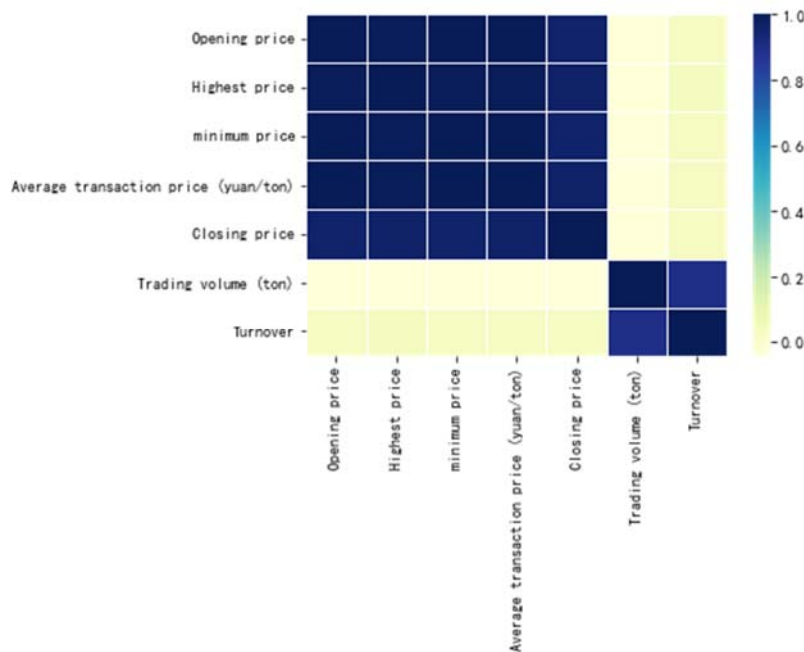


Figure 1. Heat-map of Pearson correlation coefficient among features

In this paper, the P-value between features is positively correlated with the correlation coefficient, so the P-value matrix will not be described separately. It can be seen more clearly from the thermodynamic chart that the opening price,

the highest price, the lowest price, and the average transaction price (yuan/ton) are highly linear ($0.7 \leq |r| < 1$), and the transaction volume is highly linear ($0.7 \leq |r| < 1$) with the transaction amount.

Secondly, establish the relationship between the

characteristics and the target closing price (Figure 2).

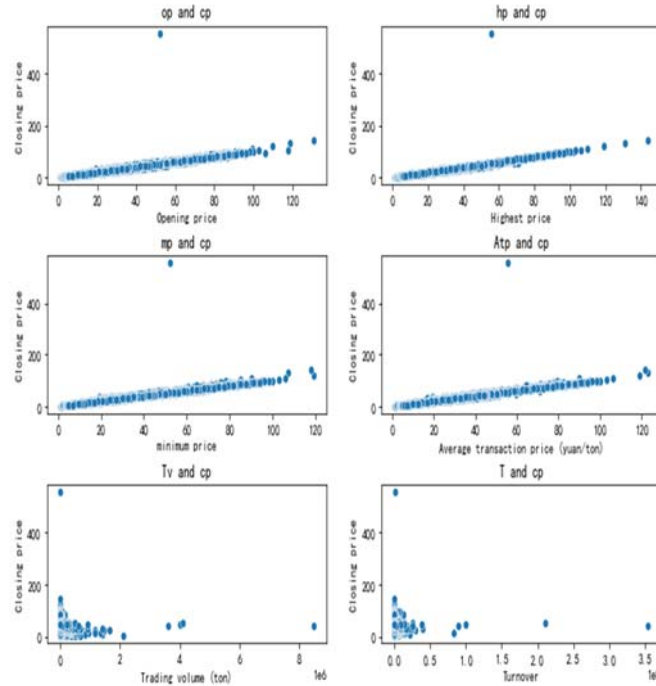


Figure2. Relationship between characteristics and closing price

As shown in Figure 2, the opening price, the highest price, the lowest price, and the average transaction price under the similar slope have a positive correlation with the closing price. There is no obvious linear relationship between the trading volume, the transaction volume, and the closing price.

Based on EDA analysis, the correlation between features

is relatively high. Although there are no obvious outliers, the possibility cannot be excluded. It is an indispensable application link of machine learning to solve data anomaly and data preprocessing behavior with extreme feature correlation. Therefore, feature engineering is also needed for data before building models.

Table 3. Parameter Space Tuning of XGBoost Model

Parameters	Parameter meaning	search space	Tuning results
learning_rate	Update the contraction step in the learning process	[0.01,3,0.01]	0.05
max_depth	The maximum depth of the tree to avoid over fitting	[2,30,1]	8
min_child_weight	Minimum sample weight sum of leaf nodes to avoid over fitting	[1,10,1]	1
n_estimators	Control the number of weak learners	[20,200,1]	171
subsample	Control the proportion of random sampling for each tree	[0.01,0.8,0.01]	0.39

3.2. Feature Engineering and XGBoost Model Building

This experiment is performed on a computer with Intel i5-7300HQ 2.5GHz CPU, 8 GB RAM and Windows10 operating system. Under the Anaconda environment, Python language is employed for programming. The IDE is Jupyter Notebook and GPU acceleration is not used.

The optimization process of the model takes the form of seeking to minimize the loss function, and so is Bayes optimization. Generally, there are several ideas to find the minimum value of a function $f(x)$ within the value range of x : let the first derivative of the function be 0, and determine the minimum value of $f(x)$ by solving the value of x at this time, but this method requires that $f(x)$ is differentiable and the differential equation has solutions; The minimum value of $f(x)$ can also be obtained by gradient descent and other optimization methods, but this method requires that $f(x)$ is differentiable and convex; Another method is to substitute x into $f(x)$ to obtain all possible values of the function and then determine the minimum value. Obviously, the workload of this method is extremely large. The main idea of Bayes

optimization is to randomly select points of loss function $f(x)$, and continuously reduce the confidence interval and improve the confidence level between these points to obtain a new loss function $b(x)$ that is infinitely close to the real loss function $f(x)$. At this time, the minimum value of $b(x)$ is regarded as the minimum value of $f(x)$. Bayes optimization idea solves the problem of finding the minimum value when the loss function is smooth and uniform and complex and nondifferentiable.

Before modeling, divide the data set, take the first 10,508 groups of data as the training set, and predict the closing price of the last 107 groups of data. In order to solve the problem of feature outliers and better fit the algorithm model for feature data, standardize the divided data set, that is, transform the original data into a mean value of 0, and the standard deviation is within the range of 0-1. For features with high correlation, the maximum likelihood estimation of principal component analysis (PCA) is used to reduce the feature dimension. Instantiate XGBoost Regressor, use Cross Validate to optimize the model, and KFold uses 50% cross validation to avoid disrupting the sample sequence, and use TPE (Tree Parzer Estimator) based Bayes optimization to

optimize the model parameters, and finally output the optimized prediction results.

3.3. Evaluating Indicator

In this paper, the mean absolute error (MAE) and root mean square error (RMSE) are used to evaluate the prediction results, and the mean absolute percentage error (MAPE) is used to evaluate the stability of the prediction model.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (24)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (25)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (26)$$

Among them, the more novel the MAE and RMSE are, the better the prediction results are. The more novel the MAPE is, the stronger the interpretation ability of the characteristics to the target value is. The better the fitting degree of the model to the data is.

4. Experimental Results and Analysis

From August 5, 2013 to October 8, 2021, there are 10615 daily time series data of the opening price, the highest price, the lowest price, the average transaction price, the closing price, the trading volume and the transaction amount of each domestic carbon emission trading market. The rise and fall of some of the closing prices are shown in Figure 4.

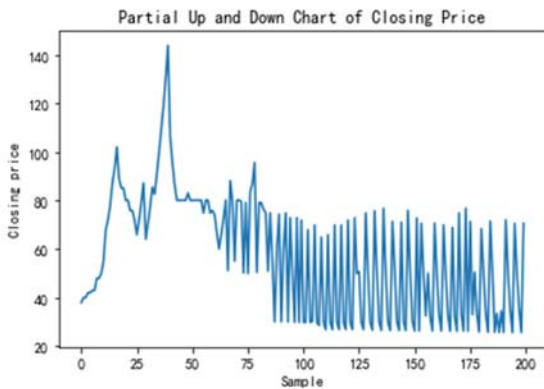


Figure 4. Partial Up and Down Chart of Closing Price

The initialized XGBoost model is adjusted by Bayes optimization based on TPE to make the model prediction results more accurate. ①Define the objective function, that is, instantiate the regression and train the model. Add a 50% cross validation that does not disrupt the order of samples to the objective function, and the output Scoring is the negative root mean square error. ② Define the parameter space. The super parameters in XGBoost algorithm are divided into three types: general parameters (for macro function control), Booster parameters (for controlling Booster details), and learning objective parameters (for determining learning strategies). This paper selects the learning that has the greatest impact on the algorithm performance_rate、max_depth、

min_child_weight、n_estimators and samples are used as the parameter space.③The optimization function is defined.The optimization agent model adopts TPE algorithm. When the loss stops for 100 consecutive times, the training is stopped. The loss is set to root mean square error (RMSE) and printed when the training is stopped. The results of optimization are shown in Table 3.

Learning_rate、max_depth、min_child_weight、n_ The final tuning results of estimators and subsamples are 0.05, 8, 1, 171 and 0.39 respectively. Both the stability and prediction effect of the Bayes optimized model are significantly improved compared with the traditional XGBoost model (Table 4) and the data fitting effect is better (Figure 5).

Table 4. Comparison of Algorithm Evaluation Indicators

Algorithms	RMSE	MAE	MAPE
Original-XGBoost	48.42	5.15	0.02
Bayes-XGBoost	1.27	0.55	0.02

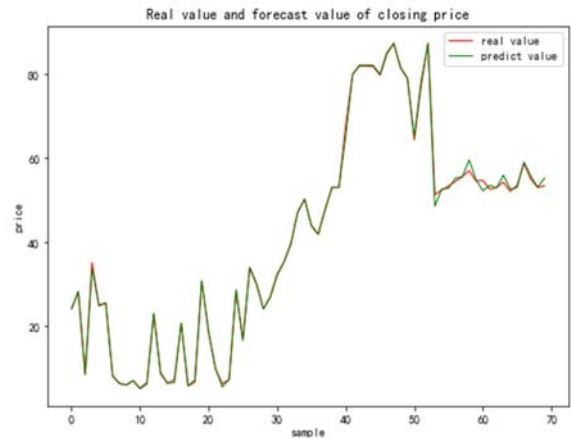


Figure 5. Final Forecast Results and Real Results

5. Conclusion

a. The XGBoost model based on Bayes optimization proposed in this paper can effectively predict the carbon price, and is more accurate than the traditional XGBoost prediction method.

b. The regular terms incorporated in XGBoost can reduce the complexity of the model, effectively prevent over fitting, reduce the amount of calculation, and significantly improve the prediction efficiency of the model. Therefore, it also reflects the advantages of using XGBoost combined model for prediction.

To sum up, the Bayes-XGBoost model proposed in this paper provides a new idea in the field of carbon emission price prediction, which can provide some practical basis for carbon market participants to make investment decisions, so as to avoid carbon market risks caused by carbon price changes.

References

- [1] Chevallier J. Nonparametric modeling of carbon prices [J]. Energy Economics, 2011, 33(6):1267-1282.
- [2] Chevallier J, Sévi B. On the realized volatility of the ECX CO2 emissions 2008 futures contract: distribution, dynamics and forecasting [J]. Ann Finance, 2011, 7:1-29.
- [3] Koop G, Tole L. Forecasting the European carbon market [J]. Journal of the Royal Statistical Society: Series A (Statistics in Society), 2013, 176(3):723-741.
- [4] Atsalakis G S. Using computational intelligence to forecast carbon prices [J]. Applied Soft Computing, 2016, 43:107-116.

- [5] Sun W, Zhang C C. Analysis and forecasting of the carbon price using multi—resolution singular value decomposition and extreme learning machine optimized by adaptive whale optimization algorithm [J]. *Applied Energy*, 2018, 231:1354-1371
- [6] Wei Yu, Zhang Jiahao Research on the Prediction Method of China's Carbon Emission Trading Price Based on DMS and DMA -- Empirical Evidence from Hubei Carbon Market [J]. *System Engineering*,2022, 40(04):3-15
- [7] Zhuang Yingdong. Analysis on the influencing factors of carbon emission trading price -- taking five representative open pilot projects in China as an example [J]. *Journal of Shenyang Institute of Engineering (Social Science Edition)*,2018, 14(03):349-357.
- [8] Abdi A, Taghipour S. Forecasting carbon price in the Western Climate Initiative market using Bayesian networks [J]. *Carbon Management*, 2019, 10(3):255-268.
- [9] Han M, Ding L L, Zhao X, Kang W L. Forecasting carbon prices in the Shenzhen market, China: The role of mixed-frequency factors [J]. *Energy*, 2019, 171:69-76.
- [10] Zhang Yun. Research on Price Drivers of China's Carbon Trading -- A Dual Perspective Based on Market Fundamentals and Policy Information [J]. *Social Science Series*,2018. 01:111-120.