Original Research Article

Coal Price Forecast Based on ARIMA Model

Xiaofan Zhang*, **Chao Liu, Yuhang Qian**

School of Finance, Anhui University of Finance and Economics, Bengbu 233030, Anhui, China. E-mail: 305871261@qq.com

Abstract: This paper analyzes and determines the decision variables and constraints, establishes the EECM-ARAMA model to analyze and research coal price forecasts. Firstly, we first confirm the influencing factors. Then, we conduct correlation coefficient tests on price and various factors, and get the strength of the correlation between each factor and price. The second is to establish a coal price prediction model. Firstly, we use the EEMD method to transform the original price series into a stable time series, and then formulate three ARIMA models by comparing the size of the influencing factors and the parameter estimation results. After testing, we finally choose the ARIMA 03 model to predict the next 31 days, 35 Weekly and 36-month coal prices. Finally, we combine the models and ideas of the above issues to obtain factors that affect coal price changes and related price prediction models, and combine experience to give some feasible policy recommendations.

Keywords: Coal Price Forecast; ARAMA Model; EECM Algorithm

1. Background

1.1 Background introduction

As the second largest fuel in the world, coal accounts for 70% of my country's primary energy consumption and is the most important energy source in China. China is a major coal producer and consumer in the world. In 2018, my country's coal production accounted for 45.72% of global production, and its consumption accounted for 50.58% of the world. In the next ten years, coal will still account for more than half of my country's energy consumption. Coal will play a pivotal role in my country's economy. Therefore, changes in coal prices have also received extensive attention from the country. The price of coal is not only subject to the supervision of relevant state departments, but also to the influence of the domestic coal market.

1.2 Research significance

(1) Reduce the production cost of coal-using enterprises

General coal market participants' forecasts of coal prices depend on their experience in the industry, or their tracking of futures, and their judgment on the peak and off-season of thermal coal. But, due to the one-sided nature of the information obtained, market participants cannot accurately judge prices.

(2) Help the transportation sector adjust capacity

Coal transportation is an important part of coal trading. If coal prices can be accurately predicted, coal prices are predicted to rise, more transportation capacity can be deployed for coal transportation to ease the shortage of coal

supply.

(3) Contribute to the formulation of coal-related policies

Accurate forecasting of coal prices helps the formulation of policies of relevant government agencies. In order to ensure the supply of thermal coal, the National Development and Reform Commission and other relevant agencies must also predict the coal market in advance to avoid large fluctuations in thermal coal prices.

2. Literature review

Sun Jihu and Peng Jianping are based on econometrics and statistics, using time series analysis and forecasting methods to apply stochastic theory, probability theory and linear difference equations to coal market price forecasting[1].

Ning Yuncai, Zhang Dongri, and Li Xiangyi established a composite wavelet neural network model to study the price forecast of the international coal market and discuss the application of this model[2].

Zhang Huan and Jiang Zuobin used the ARIMA model to predict my country's coal prices from 1977 to 2005[3].

3. Model hypothesis

Firstly, the influence of climate on the price of coal is finally reflected in temperature;

Secondly, the demand for coal prices in Qinhuangdao is mainly affected by China's coal consumption;

Thirdly, The authenticity and reliability of official data collection;

Fourthly, the unconsidered factors have little effect on coal prices.

4. Noun explanation and symbolic account

4.1 Noun explanation

(1) CPI: Consumer Price Index

(2) PPI: Production Price Index

5. Models

5.1 Analysis and problem one solving

After consulting the literature, we have concluded seven primary factors affecting coal prices, namely coal supply, coal demand, coal trade, coal inventory, transportation methods, macroeconomics and climate[4]. We further divide it into secondary factors.

Taking the thermal coal in Qinhuangdao, Hebei Province as an example, we use the Pearson correlation coefficient and Spearman rank correlation coefficient method to calculate the correlation coefficient between different factors and the price of Qinhuangdao thermal coal.

The Pearson correlation coefficient is used to measure the strength of the linear correlation between two variables X and Y. The formula is:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

Spearman's correlation coefficient is to measure the strength of two correlations, which can be obtained but the scope of application is wider than that of Pearson. The formula is:

$$d = \sum_{i=1}^{N} |R(X_i) - R(Y_i)|^2$$

$$Rs = 1 - \frac{6 \times d}{N \times (N^2 - 1)}$$

Table 1 Impact factor ranking

Ranking	Factor
1	Coal exports
2	Coal yield
3	Generated energy
4	Coal imports
5	Goal consumption
6	Number of ships at Qinhuangdao anchorage
7	Qinhuangdao port railway transfer Volume

According to the correlation coefficient, we generally think that the two variables do not want to be closed. Therefore, we eliminate the factor of correlation coefficient $|r| < 0.5$. According to its correlation coefficient from large to small, we get the final degree.

5.2 Analysis and problem two solving

We combine the historical data of thermal coal prices in Qinhuangdao Port and the main factors affecting coal prices in Problem 1 to establish a coal price prediction model.
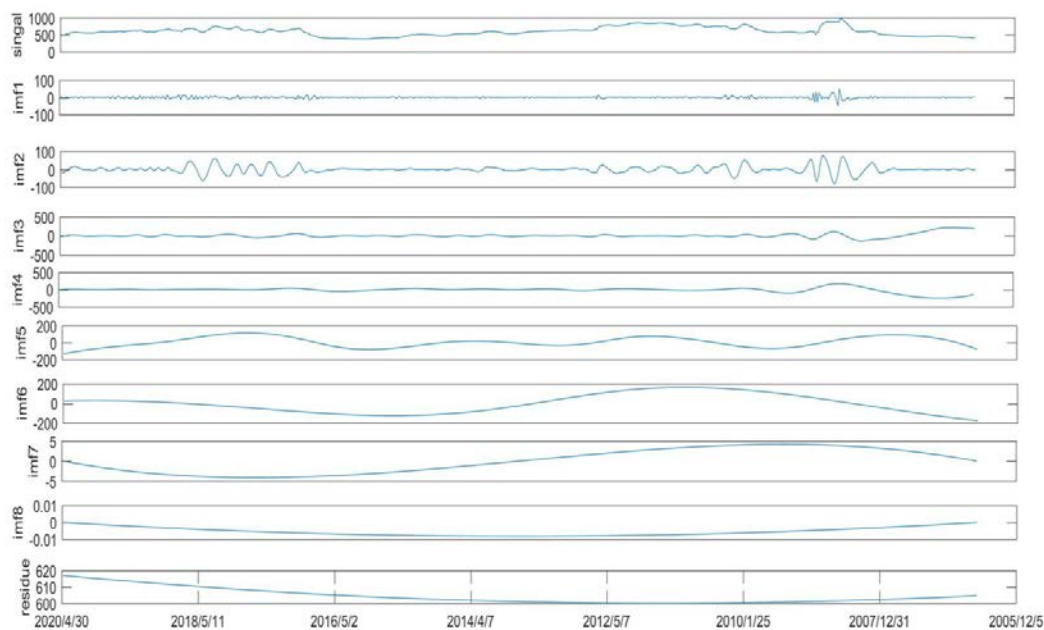

Figure 1. EEMD coal price decomposition

In order to improve the accuracy of the prediction model, the ensemble empirical mode decomposition is used to decompose the carbon price before the prediction. 5.2.1 set the parameter integration times N to 100, and the white noise standard deviation to 0.2. After the carbon price series is decomposed, the respective IMF.

5.2.2 Coal price forecast based on ARIMA model

(1) The stationarity test of the time series after EEMD decomposition

(a) Autocorrelation coefficient test

The time series are basically within the confidence interval after lagging by one period, and the stationarity of the time series after EECM decomposition can be roughly determined.

(b) ADF test

The statistic value of the ADF test is -6.909985, which is less than the critical value of 1% significance, indicating

that the original hypothesis can be rejected on the basis of at least 99% confidence level. Therefore, it can be determined that the time series does not have a unit root.

(2) Model identification and prediction parameters

After the coal price is decomposed by EECM, the time series is a stationary series with zero mean and same variance. We respectively predict the estimation results of the three ARIMAs, so we compare the size of R^2 of the three ARIMA models to determine an optimal coal price prediction model.

Table 2 Evaluation index

Index	ARIMA 01	ARIMA 02	ARIMA 03
R^2	0.77841	0.84022	0.91659

From **Table 2**, it can be seen that the corrected R^2 value of the ARIMA 03 model is the largest, so we first use it as a coal price prediction model.

Table 3 Price forecast

May 2020		May to December 2020		May 2020-April 2023	
Data	Forecast price (yuan/ton)	week (Week of date)	Forecast price (yuan/ton)	month	Forecast price (yuan/ton)
May 1	474.4	May 6	470.4	May 2020	463.8
May 2	473.6	May 13	467.9	June 2020	457.6
May 3	472.8	May 20	466.7	July 2020	453.8
May 4	472	May 27	463.8	August 2020	444.4
May 5	471.2	June 3	462.1	September 2020	432.1
…	…	…	…	…	…
May 29	463.1	December16	413.3	February 2023	308.2
May 30	460.2	December23	409.4	March 2023	312.3
May 31	463.3	December30	404.6	April 2023	320.2

6. Policy recommendations

(1) Pay attention to the balance of coal supply and demand in my country. With the development of society, the proportion and structure of factors affecting the price of coal are also changing. In order to balance the supply and demand of coal and fully allocate coal resources, the government should focus on both supply and demand.

(2) The coal industry should improve the efficiency of comprehensive utilization of coal resources, extend the industrial chain, optimize the structure, promote industrial upgrading, and improve the economic benefits of the industry.

(3) Improving the relevant policy system. Corresponding policies can be introduced to reduce the debt of coal companies and unnecessary tax costs, and alleviate the burden of coal companies. Under special circumstances, the government can appropriately increase coal purchases. While saving coal companies, it can also stabilize coal prices to a certain extent.

7. Error analysis

(1) In the first question, the indicators of the evaluation criterion level are incomplete, which makes the comprehensive evaluation of the target level somewhat lacking, and will produce partial errors in the final result.

(2) The price prediction model of the second question does not take special circumstances into consideration, which will produce slight deviations.

(3) Some indicators in the model are simulated due to the complicated actual measurement conditions, and errors

are unavoidable.

8. Evaluation of the model

8.1 Advantages

(1) This article combines simulation and actual data, and the established model is close to reality and has strong versatility;

(2) The EECM method combined with the ARAMA model can well eliminate the defects caused by the sequence data and make the model results more accurate and reasonable;

8.2 Disadvantages

(1) Due to the limitation of survey data and the lack of corresponding large sample conditions, the evaluation method for calculating the maximum degree of membership is not fully utilized, and its original advantage is reduced;

(2) Although there is a certain degree of representativeness when selecting data indicators, there is no specific basis for distinguishing.

References

1. Sun J, Peng J. Application of temporal sequence analysis technique in coal price prediction. Geology, Technology and Economic Management 2000; (3): 33–40.
2. Ning Y, Zhang D, Li X. Research and application of compound wavelet neural network model for price prediction in international coal market (in Chinese). Coal Economic Research 2001; 6(7): 16–19.
3. Zhang H, Jiang Z. The establishment and forecasting analysis of the ARIMA model of coal prices in China (in Chinese). Industrial Technology Economy 2007; 26(7): 102–105.
4. Ding Z. Research on the influence effect of coal price fluctuation on China entity economy [PhD thesis]. Beijing: China University of Mining and Technology; 2011.