Research on Coal Price Forecast Based on Exponential Smoothing Forecast and Multiple Linear Regression

Qilian Dai, Ying Fang, Jianxin Yu

Xuhai College, China University of mining and technology, Xuzhou, Jiangsu, 221008, China

Abstract: Aiming at the prediction and estimation of coal price, this paper selects the data related to coal price, and establishes a coal price forecasting model to predict coal price. In view of the existing detection of coal in China and the influence of other factors in the international market, we first consult the analytical literature on coal prices and forecasts, on the basis of which six indicators are summed up, and the factors of coal prices are sorted. Then first of all, according to the feasibility and authority of the collected data, choose Qinhuangdao coal price, import price, Qinhuangdao coal stock raw coal output. In the short term, the prediction selection exponential smoothing method will get the results of inventory and other related indexes through spss operation, and then select the optimal solution to determine the forecast price. Then in the forecast of week and month, we choose to establish multiple linear regression and add a number of indices related to Qinhuangdao thermal coal price for fitting operation, that is, the time forecast of Qinhuangdao thermal coal price.

1. Introduction

Coal is extremely important for modern industry, both heavy industry and light industry are inseparable from the demand for coal, and it is one of the indispensable energy sources in this century. Therefore, the coal price is not only regulated by the state, but also because the globalized resources are bound to be affected by the international market, but there are also many factors to macroscopically interfere with the overall price of coal [1]. In China, as a large industrial country, the excavation and demand for coal has always been huge. First of all, one of the important strategic reserve materials is one of the important links in the economic development of our country. This paper is of great significance to the prediction of coal price [2].

2. Construction of coal prediction model

Exponential smoothing is a common method in production forecasting. It is also used to forecast the trend of economic development in the short and medium term. Among all the forecasting methods, exponential smoothing is the most widely used one. Basic formula of exponential smoothing method:

$$S_t = a \times y_t + (1-a) \times S_t - 1 \tag{1}$$

St-- smoothing value of time t; yt -- the actual value of time t; St-1-- Smoothing value of time t-1;

a -- Smoothing constant, which ranges from 0 to 1[3].

When there is no obvious trend change in time series, exponential smoothing can be used for prediction. The prediction formula is:

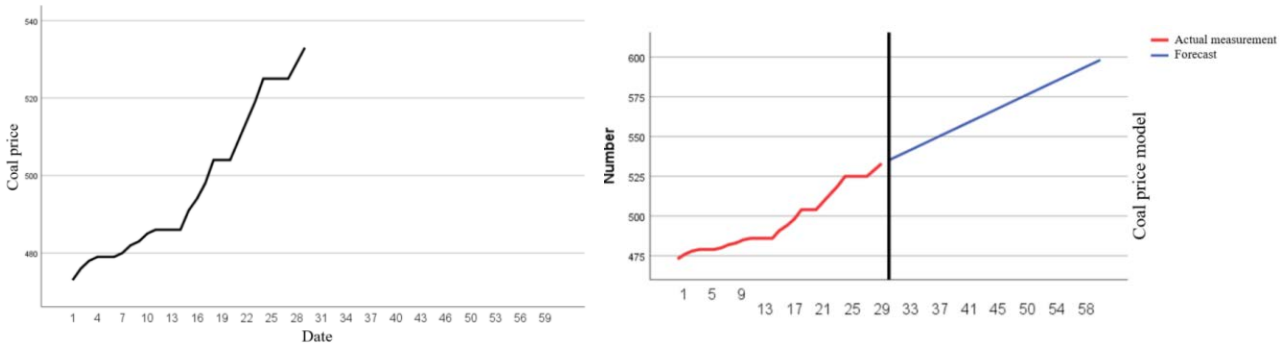$$y_{t+1} = a \times y_t + (1-a) * y_t'$$ (2)

Figure 1: Exponential smoothing prediction result

With the increase of time, Y shows an increasing trend, and the growth range from 22 to 25 is relatively large, showing a linear growth trend, and then the exponential smoothing prediction table can get the above results (take the coal price model as an example).

Table 1: The model of statistical

Model	Number of predictive variables	Holland model statistics				Young-box Q(18)			The number of outliers
		R^2 smoothly	R^2	RMSE	MAE	Statistical	DF	significant	
Y-model_1	0	.204	.988	2.157	1.829	34.404	16	.005	0

As a result, similar Brownian motion and attenuation trend models are obtained. The correlation fitting results are obtained [4]. From the fitting results of the remaining three models, the Holland model and the attenuation trend model R (that is, goodness of fit) = 0.988, followed by the Holt line brown linear, which is 0.985.

The results obtained by using the decay trend and Holland model can be derivative results into the table.) Exponential smoothing method has smoothing effect on the actual sequence, the smaller the weight coefficient (smoothing coefficient) A is, the stronger the smoothing effect is, but it is slow to the change of the actual data.

In the linear variation part of the actual series, the degree of lag deviation of the exponential smoothing value series decreases with the increase of the weight coefficient (smoothing coefficient) a, but when the change of the time series shows a straight line trend, there will still be obvious lag deviation when using the first exponential smoothing method to predict. Therefore, the transformation and optimization of the model, the use of multiple linear regression algorithm for further optimization and calculation by the optimal combination of multiple independent variables to predict or estimate dependent variables is more effective than using only one independent variable to predict or estimate [5]. More in line with reality. Therefore, the practical significance of multivariate linear regression is greater than that of univariate linear regression.

$$Y = \beta + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \cdots\cdots \beta jXj$$ (3)

Then the regression coefficient test is used to test each regression coefficient separately, which is

mainly used to test whether the influence of each independent variable on the dependent variable is significant.

3. Test regression analysis

Set the input quantity to include throughput, transfer quantity, and inventory. R square is one of the important factors to test whether regression is successful or not. DW is residual independence test. It is known that R squared obtained after this data operation is 0.965, indicating a positive correlation

Table 2: ANOVAa

	Sum of squares	Degrees of freedom	The mean square	F	Significant
Regression analysis	3832.069	4	958.017	159.625	.000b
Residual error analysis	138.039	23	6.002		
Total	3970.107	27			

The value of F is the result of F test. $P = 0.000 < 0.05$ in the table supports the null hypothesis and indicates that the linear regression is significant

Table 3: Coefficient

	B	The standard error	Beat	T	Significant
Constant	-144.877	2001.931		-.072	.954
β_2	-3.91.9	.000	-2.720	-.668	.625
α_1	8.213	18/792	.665	.437	.738
α_2	-97.127	98.875	-2.446	-.982	.506
Indonesia thermal Coal Guide price (USD/ton)	1.099	1.827	1.349	.601	.655
Thermal power generation per month (100 million KWH)	.003	.009	.221	.366	.777
End-month inventory of Coal Enterprises (ten thousand tons)	.008	.044	.300	.187	.882

Sig value is the result of T test, and its value < 0.05 indicates the strong significance of this variable.

$$Y = 0.008\alpha_{55} + 0.003\alpha_4 + 1.099\alpha_3 - 97.127\alpha_2 + 8.213\alpha_1 - 3.91.9\alpha - 144.877 \tag{4}$$

The state's regulation and control is essential, the state should strengthen the regulation and control of coal production capacity, actively eliminate backward production capacity, ensure the steady and sustainable growth of coal supply. Although the output and import volume of coal will be affected to some extent under the epidemic situation, the government can implement certain economic regulation on the coal industry to minimize the impact of the epidemic on the coal price. In terms of production technology, China is relatively backward compared with other countries. China's annual scientific research efforts on coal production enterprises need to be improved. There is a lack of professional scientific research personnel in coal enterprises and low investment in research funds, which will directly increase the gap between China and other coal advanced countries
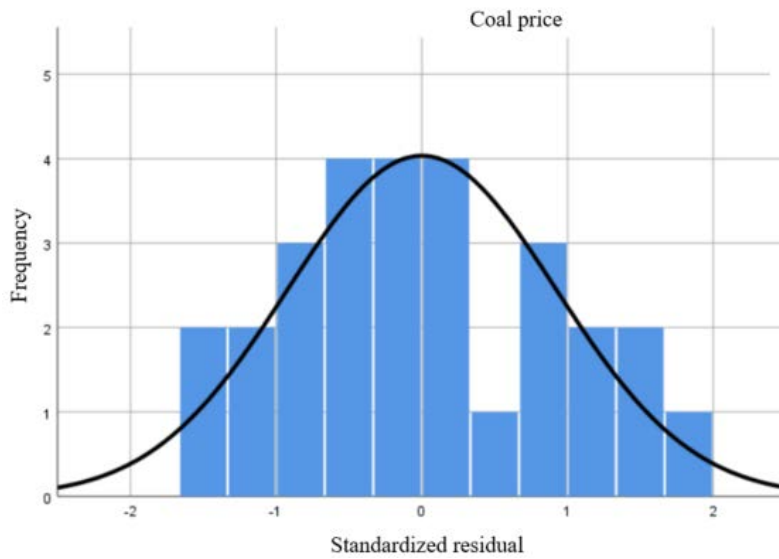
Figure 2: Distribution map

4. Multiple linear regression model

Table 4: Model summary

Model 1	R	R^2	R^2 after adjustment	Errors in standard estimates
1	.999a	.998	.992	9.734

It can be seen that if DW is 9.734 > 2, there is no autocorrelation. After the analysis of the ANOVA results, it is found that the significance of the table is less than 0.036, which indicates that the support hypothesis is significant, that is to say, the linear regression is significant.

The regression equation can be obtained as $Y=-159.523\beta-77.434\beta_1+0.069\beta_2$

5. Conclusion

For the prediction and estimation of coal prices, we first consult the analytical literature on coal prices and forecast, on the basis of which six indicators are summarized. Then select Qinhuangdao coal price, import price and Qinhuangdao coal stock raw coal output according to the feasibility and authority of the collected data. Then based on the exponential smoothing method, the inventory and other related indexes are calculated by spss, and then the optimal solution is selected to determine the forecast price. Finally, considering the current situation in the world and the impact on Qinhuangdao thermal coal prices, we establish a model for Qinhuangdao thermal coal prices rising and falling. For this, we can predict the range of coal price changes in a certain period of time in the future, so we still choose the multivariate linear model to predict the coal price for many times to get the coal price prediction model.

References

[1] Si Lingjun. Study on the influencing factors of coal price fluctuation [J]. Modern Economic Information, 2018 (01): 350.
[2] Zhu Jimao. Analysis and Forecast of influencing factors of Coal Price in China [J]. China Coal, 2017. 43 (11): 20-24.
[3] Wang Wen, Li Guodong. Study on influencing factors of Coal Price in China based on Analytic hierarchy process [J].

Operation and Management, 2016 (12): 87-90.

[4] Wang Yanwei, Jing Xiaozhen. Empirical Analysis of Coal Price trend and its influencing factors in China [J]. Price Theory and practice, 2016 (06): 114117.

[5] Yang Jing. Multi-factor analysis of coal price fluctuation [J]. Knowledge economy, 2015 (14): 81-82.