Research on the Improved Neural Network of Coal Price Forecast Based on Co-integration Theory

Luhui Gao¹ and Guoqing Wang²

¹School of Finance and public management, Anhui University of Finance and Economics, Bengbu, Anhui, 233000
²School of finance, Anhui University of Finance and Economics, Bengbu, Anhui, 233000

*Corresponding author: luhuigao@aufe.edu.cn

Abstract. China is a big coal consumer, and changes in coal prices will have a huge impact on the entire market economy system. In this paper, aiming at the prediction of coal price, the grey correlation analysis method is firstly used to describe the correlation degree between the main factors and the coal price, and the main influencing factors are sorted according to the correlation coefficient. Finally, it is determined that a coal price prediction system needs to be established on such indicators as coal reserves, raw coal production, total coal consumption, the proportion of coal in the energy mix, coal road traffic volume, coal imports, etc. Next, the quantitative indicators are combined with the economic cycle and cointegration theory to study, with the help of cointegration theory to improve the traditional neural network, so that it is more suitable for the prediction of coal prices. Finally, this paper obtains an effective coal price prediction model, which can better predict and analyze the future coal price and can play an important reference role in the formulation of government policies.

Keywords: Price forecast, grey relational analysis, cointegration theory, neural network.

1. Introduction

China is the world's largest coal consumer and accounts for more than half of the world's coal consumption. With the continuous changes in the market, China's coal industry has formed a national, competitive and orderly coal trading market system. [1] Changes in coal prices will not only affect the coal industry's own production and operation, but also affect related industries, national economic development and social stability. In December 2016, the National Development and Reform Commission and the National Energy Administration issued the Energy Production and Consumption Revolution Strategy (2016-2030), which calls for the promotion of clean fossil energy to further reduce the proportion of coal consumption and promote the transformation and development of coal. The current national energy revolution strategy will have a far-reaching impact on the coal industry, and coal enterprises will face more severe challenges. In the past, the drastic fluctuation of coal price had a great impact on the production and operation of coal enterprises and their upstream and
downstream enterprises. At the same time, as an important factor of production, the price fluctuation of coal has an important impact on the macro-economy, so the analysis of the factors that affect the price of coal and the prediction of future price fluctuation are very important.

2. Grey Relational Analysis of Coal Price Impact Indicators

Before establishing the coal price model, the first thing to do is to determine the price impact indicators. Considering that there are many factors that affect the coal price, we describe the correlation degree between the main influencing factors and the coal price through the grey correlation degree analysis method, and rank the main influencing factors according to the correlation coefficient.

Based on the data from China’s coal network, this paper obtains that the average price of coal has obvious correlation with coal reserves, coal consumption, the proportion of coal in the energy structure and raw coal production. Because the dimension is not removed and affected by the data unit, the correlation degree between the average price of coal and factors such as coal import and export volume and investment in coal mining and dressing industry cannot be clearly judged, and further grey correlation degree analysis is required.

The reference series (parent factor) is the price of coal; The comparative series (sub-factors) are respectively raw coal production, total coal consumption, coal imports, coal exports, the proportion of coal in the energy mix, investment in coal mining and dressing industry, coal land transportation volume and coal reserves. [2]

After the reference sequence and the comparison sequence are determined, the data are preprocessed. It is difficult to compare the original data due to their different dimensions. In this study, the average method is used to perform dimensionless processing on the experimental data, i.e., the data of each sequence is divided by the average value of the corresponding sequence. For the dimensionless data sequence, the absolute difference between each index sequence (comparison sequence) of the evaluated object and the corresponding elements of the reference sequence is calculated one by one, i.e.:

\[ |X_i - X_0| = |z_{ij} - z_{0j}| \quad i = 1,2, \ldots 10 \quad j = 1,2, \ldots 7 \]

Next, the correlation coefficients of the corresponding elements of each comparison sequence and the reference sequence are calculated:

\[ \xi_{0i}(j) = \frac{\min_{i=1}^{n} \min_{j=1}^{m} (|z_{ij} - z_{0j}|) + \rho \cdot \max_{i=1}^{n} \max_{j=1}^{m} (|z_{ij} - z_{0j}|)}{|z_{ij} - z_{0j}| + \rho \cdot \max_{i=1}^{n} \max_{j=1}^{m} (|z_{ij} - z_{0j}|)} \]

Where \( \rho \) is the resolution coefficient, and its value is in the interval (0, 1). The smaller the value of \( \rho \), the greater the difference between the correlation coefficients, and the stronger the resolution ability, usually \( \rho = 0.5 \). The correlation coefficient \( \xi \) is a positive number not exceeding 1, reflecting the correlation between the i-th comparison sequence \( X_i \) and the reference sequence \( X_0 \) on the j-th attribute.

For each evaluation object (comparison sequence), the average value of the correlation coefficient between each index and the corresponding element of the reference sequence is calculated to reflect the correlation between each evaluation object and the reference sequence.

\[ p_{0i} = \frac{1}{n} \sum_{j=1}^{n} \xi(j), i = 1,2, \ldots 10 \]

The calculated correlation degree is shown in the following table:
Table 1. Correlation Degree between Each Evaluation Object and Reference Sequence.

| Sub-factor                          | Raw coal production (10,000 tons) | Total coal consumption (10,000 tons of standard coal) | Coal imports (10,000 tons) | Coal exports (10,000 tons) | Investment in coal mining and dressing industry (RMB100 million) | Energy consumption structure (coal share) | Land transportation volume of coal (100 million tons) | Coal reserves (100 million tons) |
|------------------------------------|---------------------------------|-----------------------------------------------------|--------------------------|--------------------------|-------------------------------------------------------------------|-------------------------------------------|-------------------------------------------|-------------------------------------|
| degree of association              | 0.6610                          | 0.6088                                              | 0.5485                   | 0.4679                   | 0.4689                                                            | 0.5711                                    | 0.5495                                    | 0.7293                               |
| ranking                            | 2                               | 3                                                   | 6                        | 8                        | 7                                                                | 4                                         | 5                                         | 1                                    |

The greater the correlation degree, the closer the relationship between the comparison series and the reference series, indicating that the greater the influence degree of this influencing factor on the coal price. According to the correlation degree and ranking, the influence degree of each major factor on the coal price from big to small is: coal reserves > raw coal output > total coal consumption > proportion of coal in energy structure > coal land transportation volume > coal imports > investment in coal mining and dressing industry > coal exports.

3. BP neural network approximation model of cointegration function

As a kind of commodity, coal is a kind of consumption energy, its price influencing factors include: coal value, coal supply and demand, alternative energy, economic factors, emergencies, etc. The impact of changes in coal demand on prices has long-term memory, the total supply of coal has a certain impact on prices, but it is not the main reason for the price changes, as a substitute for coal, oil generally has a positive correlation with the impact on coal prices, the increase in oil prices will inevitably lead to a significant increase in coal demand, which in turn will lead to an increase in coal prices. For the complementary products of coal, such as coal and electricity, their impact on coal price is synchronous, and the reduction of electricity price will reduce coal price to a certain extent [3].

To sum up, the influencing factors of coal price are divided into micro-factors, macro-factors, natural factors and factors in non-coal fields, among which micro-factors mainly include coal supply, demand, coal substitutes and supplements, etc, these micro-factors have an important impact on coal price and are fundamental. Macro factors mainly include national economic development, macro policies, political factors and international coal market prices, etc, Macro factors play a guiding and regulating role in coal prices, while other natural factors, such as coal storage, climate and natural conditions, are uncertain in their impact on coal prices, and non-coal factors, as regulating variables, have an impact on coal prices.

The key to the testing and modeling of nonlinear cointegration systems is to estimate the nonlinear cointegration function [4], if the component sequence of vector time series is regarded as the input variable and the nonlinear cointegration function F is regarded as a nonlinear transformation, the output variable is a SMM sequence. Estimating the nonlinear cointegration function is equivalent to mining the relationship between each component sequence from the input vector time series and the output SMM sequence. Due to the nonlinear equilibrium relationship between vector time series and its component series, the general linear cointegration analysis method cannot be effectively solved, Artificial neural networks (ANN) has inherent advantages in the approximation of nonlinear functions.

The basic BP algorithm includes two aspects: the forward propagation of signals and the backward propagation of errors, i.e, the calculation of actual output is carried out in the direction from input to output, while the correction of weights and readings is carried out in the direction from output to input. Assuming that the input variables are x, x, l, x, the output variables are 1, O2L, and the number of samples is p, according to the input-output calculation of the hidden layer node and the input-output calculation of the output layer node, the forward propagation process of the signal can be summarized as follows

$$O_k = \Psi(\text{net}_k) = \Psi \left( \sum_{t=1}^{J} w_{kt} y_t + \alpha_k \right) = \Psi \left( \sum_{t=1}^{J} w_{kt} \Phi \left( \sum_{l=1}^{J} w_{lj} x_t + \theta_l \right) + \alpha_k \right)$$
Using the total error \( E = \sum \sum r_i \) of the training samples, the output error of each layer neuron is calculated layer by layer from the output layer, and then the weights and thresholds of each layer are adjusted according to the error gradient descent method, so that the final output of the modified network can approach the expected value. This feedback process can be summarized as the calculation of the adjustment of each parameter in the network.

\[
\Delta w_{ki} = \eta \sum_{p=1}^{P} \sum_{k=1}^{L} (T_{k}^p - O_{k}^p) \cdot \Psi'(net_k) \cdot y_i
\]

\[
\Delta \alpha_k = \eta \sum_{p=1}^{P} \sum_{k=1}^{L} (T_{k}^p - O_{k}^p) \cdot \Psi'(net_k)
\]

\[
\Delta w_{ij} = \eta \sum_{p=1}^{P} \sum_{k=1}^{L} (T_{k}^p - O_{k}^p) \cdot \Psi'(net_k) \cdot w_{ki} \cdot \Phi'(net_x) \cdot x_j
\]

\[
\Delta \theta_t = \eta \sum_{p=1}^{P} \sum_{k=1}^{L} (T_{k}^p - O_{k}^p) \cdot \Psi'(net_k) \cdot w_{ki} \cdot \Phi'(net_t)
\]

The main idea of estimating nonlinear cointegration function by BP neural network is to train the neural network with vector time series \( \{X_t\} \) whose partial component series are LMM series as input and a certain SMM series \( \{d_t\} \) as network output level teacher value, the output corresponding to \( X_t \) is recorded as \( Y_t = f(X_t; w, \alpha, \theta) \), and the residual is \( u_t = d_t - Y_t \). If the parameter vectors \( w, \alpha, \theta \) exist such that the output sequence \( \{Y_t\} \) is a SMM sequence (i.e., the residual sequence \( \{u_t\} \) is SMM), then a nonlinear cointegration relationship exists between the input component sequences, and the nonlinear cointegration function is \( f(X_t; w, \alpha, \theta) \).

4. Establishment of improved neural network model for coal price prediction

The data in this paper are from China Coal Network, etc, the main data collected are the weekly closing price of Qinhuang Hong Kong Island Steam Coal and the related impact data. In view of the non-uniformity of each variable unit and avoiding the deviation of the data, we preprocessed the data to make it more operational. The processing formula is as follows:

\[
X_i^t = \frac{(X - U_i)}{\sigma_{ij}}
\]

\( Ui \) is the sample mean and \( \sigma_{ij} \) is the sample variance.

The collected data is processed to determine 4 input layers, 1 output layer and 4 hidden layers. After the model is established, input data to train the model, and the result is shown in Figure 1. As can be seen from the results shown in the figure, the Mean Squared Error of the Training data, Validation and Test data rapidly decreases in the horizontal axis direction, and reaches a minimum value at epochs10, i.e, the training result at this time is the most effective.

Figure 1. Mean Square Error Diagram.
Figure 2 shows the fitting effect of the trained BP neural network on the data. Among them, the R2 values of Training data, Validation, Test data and All data are 0.9995, 0.9995, 0.9994 and 0.9995 respectively, all of which are greater than 0.99. The training model has good results in data fitting, which can be used to predict the price trend of Qinhuang Hong Kong Island thermal coal in the next two years.

Through the above analysis, it can be found that the price of Qinhuang Hong Kong Island thermal coal is mainly determined by the supply and demand of the market, although it is subject to the influence of many factors such as railway capacity, disastrous weather and market speculation. Moreover, in the supply and demand of coal, demand factor is still the main influencing factor of the current coal market price in our country, it not only guides each other in the short term, but also has a long-term cointegration relationship between output, demand and coal price in the long term. In addition, according to the forecast results of the variable point model and the neural network, the future coal price will be basically flat with a slight drop. Coal prices will continue to fall without major measures such as large-scale economic stimulus and comprehensive production restrictions. That is, in the foreseeable two-year period, the price of Qinhuang Hong Kong Island thermal coal will decline due to the influence of supply and demand and policies.

5. Conclusion
The improved coal price prediction neural network model based on the cointegration theory obtained in this paper can find that the development of the coal industry in Hong Kong Island and China will have to face various pains brought by low coal prices, declining profits and even losses for a long period of time at present and even in the future, in order to cope with this situation, the government needs to work together with relevant departments, industry associations and coal enterprises to take various measures simultaneously, and gives some suggestions.

A) market-oriented reform and institutional improvement
To further promote the market-oriented reform of coal prices, the government should further promote the market-oriented reform of coal prices, further improve the investment and financing system of the coal industry, straighten out the market relationship between coal and electricity, coal
and steel, and speed up the pace of integration of upstream and downstream development of the coal industry; Focus on building a market price adjustment mechanism between coal and its related industries, fully mobilize the adjustment function of market price between industries, avoid and reduce the direct or indirect mandatory regulation of coal price by government departments, and strive to form a market mechanism of independent convergence, independent negotiation and independent pricing between coal supply and demand, so that the impact of coal price on China's real economy conforms to the relevant theory of market economy.

B) Establish a strategic coal reserve system

Building a strategic coal reserve system and other regulatory measures to avoid frequent fluctuations in coal prices. First, the state should establish and perfect a coal strategic reserve system, which can absorb coal sources when the price of coal is low and the market demand is not strong, and increase the supply of coal when the demand is strong, thus serving as a lever to stabilize the supply and demand of coal, so as to effectively ease the large fluctuation of coal prices, second, the state should make corresponding requirements for the coal inventories of relevant industries, due to the difference between the supply and demand places of coal in our country, most of the coal needs to be transported over a long distance, but the bottleneck and uncertainty in China's transportation add to the volatility of the coal price, and increasing the coal inventory in the major coal-using industries can effectively ease the urgency of the coal demand, thus playing a certain role in promoting the alleviation of the coal price.

C) Improve the market-based pricing mechanism

We should gradually promote and improve the market-oriented pricing mechanism for coal, we should gradually promote and improve the market-oriented pricing mechanism for coal prices, establish a pricing mechanism for coal prices that can reflect the scarcity of resources, the relationship between market supply and demand, coal mine safety and the ecological environment compensation costs in mining areas, break down the artificial barriers between domestic and international coal prices, make full use of both domestic and international resources and markets, and further rationalize the price relationship between coal resources and other resource-based products. At the same time, the relevant taxation mechanism for coal mining should be improved as soon as possible to internalize the external costs of coal production and use, and resource tax and exploration tax should be reflected in the coal price to reflect the real value of coal. However, considering the impact of coal price on China's price, the country should gradually promote the market-oriented reform of coal price in order to realize the steady development of China's economy.

References
[1] Liu Chang, Wang Xuran, China's coal price fluctuations and the choice of policies to reduce production capacity [J/OL], Journal of Xi'an Jiaotong University (Social Science Edition): 1-13 [2020-05-03].
[2] Zhang Jianying, China's coal price impact factors VAR model analysis [J], Economic issues, 2016 (01): 108-112.
[3] Yu Xiaoyu. A Study of Matching Transactions Based on Partial Cointegration Theory [D], China University of Science and Technology, 2020.
[4] Gaili. Research on the Relationship between Economic Cycle Fluctuation and Coal Supply and Demand Based on Co-integration Theory [D], Shandong University of Science and Technology, 2017.