Coal plays an important role in China’s energy structure and its price has been continuously decreasing since the second half of 2012. Constant low price of coal affected the profits of coal enterprises and the coal use of its downstream firms; the precision of coal price provides a reference for these enterprises making their management strategy. Based on the historical data of coal price and related factors such as port stocks, sales volume, futures prices, Producer Price Index (PPI), and crude oil price rate from November 2013 to June 2016, this study aims to forecast coal price using vector autoregression (VAR) model and portray the dynamic correlations between coal price and variables by the impulse response function and variance decomposition. Comparing predicted and actual values, the root mean square error (RMSE) was small which indicated good precision of this model. Thus this short period prediction can help these enterprises make the right business decisions.

1. Introduction

Coal is a kind of fundamental energy source which is uppermost in China. In 2015, the total energy consumed in China was 4.3 billion tons of standard coal, of which the consumed coals took a percentage of 64%. Bhattacharya et al. suggested that improving overall efficiency in coal sector will continue to play a significant role in maintaining sustainable growth in China in the long run [1]. On the supply side, China government made amendment to Law of the People’s Republic of China on the Prevention and Control of Atmospheric Pollution on the 16th session of the 12th Standing Committee of the National People’s Congress. It stipulates that the relevant ministry of the State Council and local people’s governments at different levels have the responsibility to take proper measures to adjust the energy structure and promote the production and utilization of clean energies; and the related departments should optimize using method of coals, promote the clean and effective utilization, and consequently reduce the proportion of coals in the consumption of primary energy. Meanwhile, coals have the priority in the implementation of production capacity reduction based on the 13th five-year plan. The production capacity is planned to be reduced by about 0.5 billion tons in order to advance the revolution of energy supply side. On the demand side, China energetically promotes the utilization of clean energies and reduces thermal power generation. In 2015, the absolute generating capacity of thermal power was 4210.2 billion KWH which fell 2.8% year on year. The equipment utilization hours of thermal power averaged 4329 hours which reduced by 410 hours year on year. The Action Plan of the Clean and Effective Utilization of Coals in Industrial Circle aims to save more than 0.16 billion tons of coal consumption volume by 2020. Yuan et al. found that coal used for power generation would reach the peak at around 1280 million ton coal equivalent under the clean coal power plan declared by the Chinese government [2]. Within the revolution of both supply side and demand side, both sides in the business will pay much more attention to the price of coals, so the coal price forecasting seems more important to both of them.

In 2013, China government canceled the key contract on coal supply and the double-track pricing system of electrical coals, which marked the notion that the coal market totally stepped into marketization period and the coal price started to entirely be determined by market. However, since the second half of 2012 electrical power industry and steel industry which normally have high consumption of coals had
reduced the coal consumption because of the influence by
the decelerated economic growth. At the same time, impelled
by the previous high price, many companies started to enter
into coal market and snatch resources and finally resulted
in the situation that coal production capacity is excessively
more than demand amount. This unbalanced supply-demand
situation caused the price of coals to start to drop down and
maintained the falling trend. The price in 2015 was reduced
by 60% compared with the price at high level in 2011. At
the beginning of 2016, the price of coals showed some slight
rise because of the coal consumption peak in winter and the
resource tension at port. Afterwards, the price tends to be
stable.

For both supply and demand on coals and related energy
policy makers, accurate prediction of coal price will help
them to make more rational decisions. So it is useful to
explore a prediction model of coal price under market
economy. Based on the available literature, Liu and Ma
predicted coal price by using ARMA and Grey-forecasting
model [3]. Liu et al. analyzed the major reasons of coal price
in China [4]. Xu et al. built a FS-SVM model for coal price
[5]. These papers have some certain practical values but more
common types of energy price analysis articles are in oil
price, natural gas price, and so forth. Kanamura proposed
a SDV model for energy prices using the supply-demand
relationship [6]. García-Martos et al. built a multivariate
model for energy prices and compared its results with those
of the univariate model [7]. Yuan et al. studied the relations
between Chinese energy consumption and energy prices [8].
Wei and Guo studied the relationship between oil prices and
the Chinese macroeconomy [9]. Zhang and Yao employed
the state-space model and log-periodic power law model to
explore the dynamic bubbles of oil prices and predict their
crash time [10]. van Goor and Scholtens investigated daily
natural gas prices in the UK [11]. The essential thoughts and
forecasting methods of these papers have good inspirations
for our research; in this paper we choose PPI, port stocks,
sales volume, futures prices of steam coals, and crude oil
price to represent, respectively, economic situation, supply
and demand, financial instruments, and substitute to predict
the coal price in China. Since the coal future in China started
running since September 2013, this factor is used less in other
articles.

Considering the superiorities of the VAR, this model is
used in this paper. The VAR model allows us to explore how
variable values depend on their own lags and the lags of
other variables, thus offering us a rich structure to evaluate
integrated data. The characteristics of this model make it
useful to predict data. Wang et al. analyzed the trend of
RMB equilibrium exchange rate in the VAR model system
[12]. Zheng et al. used the VAR model to forecast pork
consumption in China [13]. Sing et al. used VAR model to
estimate the value of annual construction work carried out by
main contractors, which was useful for forecasting private-
sector construction completions [14]. Shah and Ghonasi
developed a VAR model to find the determinants of price level
for the period 1994–2008 in India and tested the predictive
ability of the model [15]. Aydin and Cavdar compared the
prediction performances of Artificial Neural Network and
VAR model [16].

2. Materials and Methods

2.1. Vector Autoregression Model. The VAR model can be
used to analyze the dynamic interaction of time series and
the dynamic impact of random disturbances on the variable
system and for forecasting. As such, it explains the influence
of various impacts on variables, and it is one of the relatively
extensive applications in multivariate time series models. It
can easily handle multiple variables and predict the time
series. Compared with some neural network models or
those with complex kernel functions, this model could more
straightforward analyze the economic questions and observe
the influence of variables on prices. We used the VAR model
to predict the coal price with the port stocks, sales volume,
 futures prices, PPI, and price of WTI.

The mathematical expression of a general VAR (P) model is

\[ y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + Bx_t + \varepsilon_t, \]

\[ t = 1, 2, \ldots, T, \]

where \( y_t \) is a \( K \times 1 \) vector of time series \( t = 1, 2, \ldots, T; A \) is
a \( K \times K \) parametric matrix; \( x_t \) is a \( D \times 1 \) vector of exogenous
variables; and \( B \) is the \( K \times D \) coefficient matrix to be estimated.
\( \varepsilon_t \) represents the random error term, while \( p \) represents the
lag period.

Akaike Information Criterion (AIC) and Swartz Criteria
(SC) are used to select the lag periods of the VAR model in
this study. The AIC and SC are computed as follows:

\[ \text{AIC} = -2l + \frac{2n}{T}, \]

\[ \text{SC} = -2l + \frac{n \ln T}{T}, \]

where \( n \) represents the total number of estimated parameters
and \( T \) represents the sample length. \( l \) is determined using

\[ l = -\frac{Tk}{2} (1 + \ln 2\pi) - \frac{T}{2} \ln \left| \sum \right|. \]

In practice, the unit root test is used to measure the
stationary case, while the cointegration test is used to check
whether any correlations exist. If the VAR model is stable,
then our analysis can continue to predict the coal price and
error test.

2.2. Data. This paper makes some price estimations for coals
with the five variables of port stocks, sales volume, futures
prices of steam coals, PPI, and crude oil price selected,
respectively, in the aspects of supply and demand, financial
instruments, economic situation, and the price of substitute.
The major factor that influences prices is the degree of mar-
etization, among which the relationship between demand
and supply plays a significant role, with the alternatives also
making difference. For the coal price in China, the data of
coal capacity remains inaccurate due to the existing extensive mining and small mines. Usually the port inventory could be taken as an important indicator of the demand and supply relationship in the circulation stage.

(i) Port Stocks. All together with the participation of both upstream and downstream enterprises, the port inventory of coals is able to intuitively reflect the supply and demand situation on coals. In 2014, the nationwide port stocks monthly averaged 41.82 million tons. In 2015, it dropped to 37.44 million tons which was decreased by 4.38 million tons compared with 2014. In 2016 port stocks keep the declining trend with the amount less than 30 million tons for each of the first 3 months. It is observed that the coal quantity supplied at port has been relatively lower compared with demand quantity since 2014. The low price of coals caused the trading companies of coals to be unwilling to deposit a large storage at port. If such a situation would sustain for a while, it may turn into an inducement for the modest rise of large storage at port. If such a situation would sustain for a while, it may turn into an inducement for the modest rise of large storage at port.

(ii) Futures Prices. Chinese steam coal futures were officially listed for running in September of 2013. It is found by Wei that steam coal futures have shown the function of price discovery [17]. Therefore, the futures prices of steam coal are able to give an index for the price trend of spot market.

(iii) Sales Volume. The sales volume influences the sale price of coal enterprises. When consumers anticipate the price rising or strained transport capacity, they will correspondingly increase the quantity of coals that they will purchase in the same month. Meanwhile, the coal enterprises may reduce the coal price for boosting the sales volume.

(iv) PPI. PPI means a price indicator to reflect the variation in the price level of goods and service commodity in different periods from the viewpoint of producers, which also has significantly referential value for national economic accounting. Because coal price is a monthly data, so PPI is used in this paper which is also a monthly data as economic situation in China. In recent two years, Chinese economy is running in a poor state with a declining PPI and serious problems of weak demand and excess capacity. The sustained economic downturn resulted in great price stress on coals.

(v) Crude Oil Price. As a kind of important fossil energy source, crude oil is a major substitute to coal. The price fluctuation of crude oil will correspondingly influence the demand and price for coal. When crude oil rises in price, the demand on coal will have some corresponding increase as the substitute. The fluctuation could also happen to the price. China's oil price depends on the Chinese policies and lags the market reaction. The WTI price could influence not only the import of coals but also the domestic oil price somehow. And there are some papers about the WTI and Chinese coal price. Guo et al. analyzed the relationship between the world oil price and China's coke price [18]. Liu et al. did the comparative analysis between the price of WTI, BJI steam coal price, and Shanxi blending coal price [4]. So, in this paper, we employ price of WTI as crude oil price.

As to eliminate heteroscedasticity and make the variables obey linear relation, we take natural logarithm of port stocks (LOGSS), sales volume (LOGS), futures prices (LOGF), and price of WTI (LOGW). Since PPI is a dimensionless variable, its original time series (PP) is used. We selected 29 months from November 2013 to March 2016 as time units in the model. For all variables, we obtained data in this paper from the China coal market network (http://www.cctd.com.cn/) and Zhengzhou Commodity Exchange net (http://www.czce.com.cn/portal/index.htm).

3. Results

3.1. Unit Root Test Results. Before using the VAR model, it was necessary to guarantee the stationary data to prevent spurious regression. Table 1 shows the results of augmented Dickey–Fuller unit root tests. Our results show that the null hypothesis was not rejected. In fact, while all the variables are first-order difference stationary, rejection of the null hypothesis occurs under the 10% level. Thus, it can be

<table>
<thead>
<tr>
<th>Variable</th>
<th>ADF test statistic</th>
<th>Prob.</th>
<th>1% level</th>
<th>5% level</th>
<th>10% level</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOGF</td>
<td>−1.108679</td>
<td>0.6980</td>
<td>−3.689194</td>
<td>−2.971853</td>
<td>−2.625121</td>
<td>Nonstationary</td>
</tr>
<tr>
<td>DLOGF</td>
<td>−3.699871</td>
<td>0.0119**</td>
<td>−3.699871</td>
<td>−2.976263</td>
<td>−2.627420</td>
<td>Stationary</td>
</tr>
<tr>
<td>PP</td>
<td>−0.326606</td>
<td>0.5582</td>
<td>−2.653401</td>
<td>−1.953838</td>
<td>−1.609571</td>
<td>Nonstationary</td>
</tr>
<tr>
<td>DPP</td>
<td>−1.923290</td>
<td>0.0533*</td>
<td>−2.653401</td>
<td>−1.953838</td>
<td>−1.609571</td>
<td>Stationary</td>
</tr>
<tr>
<td>LOGP</td>
<td>−0.664299</td>
<td>0.8388</td>
<td>−3.711457</td>
<td>−2.981038</td>
<td>−2.629906</td>
<td>Nonstationary</td>
</tr>
<tr>
<td>DLOGP</td>
<td>−4.070135</td>
<td>0.0043***</td>
<td>−3.711457</td>
<td>−2.981038</td>
<td>−2.629906</td>
<td>Stationary</td>
</tr>
<tr>
<td>LOGS</td>
<td>−5.806618</td>
<td>0.0001***</td>
<td>−3.699871</td>
<td>−2.976263</td>
<td>−2.627420</td>
<td>Stationary</td>
</tr>
<tr>
<td>DLOGS</td>
<td>−7.044197</td>
<td>0.0000***</td>
<td>−3.711457</td>
<td>−2.981038</td>
<td>−2.629906</td>
<td>Stationary</td>
</tr>
<tr>
<td>LOGSS</td>
<td>−4.601816</td>
<td>0.0011***</td>
<td>−3.689194</td>
<td>−2.971853</td>
<td>−2.625121</td>
<td>Stationary</td>
</tr>
<tr>
<td>DLOGSS</td>
<td>−8.448770</td>
<td>0.0000***</td>
<td>−3.699871</td>
<td>−2.976263</td>
<td>−2.627420</td>
<td>Stationary</td>
</tr>
<tr>
<td>LOGW</td>
<td>−1.183059</td>
<td>0.6667</td>
<td>−3.699871</td>
<td>−2.976263</td>
<td>−2.627420</td>
<td>Nonstationary</td>
</tr>
<tr>
<td>DLOGW</td>
<td>−3.536514</td>
<td>0.0149**</td>
<td>−3.711457</td>
<td>−2.981038</td>
<td>−2.629906</td>
<td>Stationary</td>
</tr>
</tbody>
</table>

Notes: D represents the first difference of variables. * indicates 10% significant level. ** indicates 5% significant level. *** indicates 1% significant level.
concluded that all the variables are first-order difference stationary and can proceed with cointegration test.

3.2. Cointegration Test Results. Since most time series of variables are nonstationary and the transformed time series often do not have direct economic significance, Engle and Granger proposed the cointegration theory and methods, providing another way for nonstationary series modeling [19]. The cointegration test is used on multiple variables. Although they may have independent long-term variation, if they are cointegrated, then there exists a long-term and stable relationship between these variables. We use the method of multivariate cointegration proposed by Johansen and Juselius [20], based on the VAR model, and our cointegration tests are carried out on data, using the first-order lag. Test results are shown in Table 2.

Max-Eigen Statistic show that at the confidence level of 95%, there is only one long-term cointegration relationship between coal price, port stocks, sales volume, futures prices of steam coals, PPI, and crude oil price.

3.3. Construction of Our Model. A VAR model is built based on the statistical properties of the data. It is constructed by taking each endogenous variable in the system as the lag value function of all the endogenous variables in the system. In this way, a single variable autoregression model is expanded to a "vector" autoregression model, comprising a multivariate time series variable.

An important aspect of the VAR model is the determination of lag order. The bigger the lag period, the greater the need for estimated parameters, and the greater the reduction in the degrees of freedom of the model, while an insufficient lag period will not reflect the dynamic characteristics of the model. AIC are used to evaluate lag order in this study. Lag order for our VAR model is given in Table 3, reflecting AIC.

As shown in Table 3, the selected lag order is 2 in accordance with the AIC information rule.

The simulation result of LOGP (price) is shown as follows:

\[
\text{LOGP} = 0.33921 \times \text{LOGF}(-1) + 0.08415 \\
\quad \times \text{LOGF}(-2) + 0.91306 \times \text{LOGP}(-1) \\
\quad - 0.40670 \times \text{LOGP}(-2) - 0.07897
\]

To ensure our model was well specified, it was necessary to conduct a stability test. If the VAR model is not stable, then the prediction cannot be carried out. Stability was assessed using an autoregressive characteristic polynomial. When all the characteristic roots are less than 1, that is, they are located within the unit circle, then the model is stable. VAR roots of this characteristic polynomial are shown in Figure 1.

3.4. Impulse Response Functions. Because the VAR model is not a theoretical model, no apriority constraints are made on the variables. Thus, the dynamic effect on the system is analyzed as the VAR model is impacted. The impulse response function is an analysis tool used by many researchers to describe this causality, that is, the shock generated by the change of one variable in the VAR model on another variable (e.g., Xu and Lin [21]; Zhu et al. [22]). Analyses are presented dynamically using graphs, which show the response direction, amplitude, and persistence of the variables in the model related to a shock. Such changes can be observed over time. Its advantage is to highlight variable changes over time at a system-scale; that is, when the whole system undergoes external shocks, it will be temporarily unstable but will achieve balance with adjustments over time. In impulse response analysis, the VAR model is transformed into a vector process of infinite order. All other conditions are unchanged. Thus, the error term impacted by a unit at some time point

\[
\begin{align*}
\text{LOGS}(-1) &+ 0.00431 \times \text{LOGS}(-2) \\
- 0.01296 \times \text{LOGSS}(-1) - 0.01240 \\
+ 0.09014 \times \text{LOGW}(-1) + 0.01221 \\
- 0.01588 \times \text{PP}(-2) + 1.75760.
\end{align*}
\]
Table 3: Lag length test results.

<table>
<thead>
<tr>
<th>Lag</th>
<th>LogL</th>
<th>LR</th>
<th>FPE</th>
<th>AIC</th>
<th>SC</th>
<th>HQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>28.54615</td>
<td>NA</td>
<td>7.58e−09</td>
<td>−1.670085</td>
<td>−1.382121</td>
<td>−1.584458</td>
</tr>
<tr>
<td>1</td>
<td>165.9928</td>
<td>203.6247</td>
<td>4.44e−12</td>
<td>−9.184654</td>
<td>−7.168908 *</td>
<td>−8.585267</td>
</tr>
<tr>
<td>2</td>
<td>217.3081</td>
<td>53.21584 *</td>
<td>2.24e−12</td>
<td>−10.31912</td>
<td>−6.575590</td>
<td>−9.205971 *</td>
</tr>
</tbody>
</table>

Notes: * indicates lag order selected by the criterion.

Figure 2: Responses of the coal price to influencing factors.

3.5. Variance Decomposition. Variance decomposition analyzes the contribution degree of each impact on the endogeneous variable, highlighting the importance of different structural shocks. Therefore, variance decomposition can illustrate the relative importance of a given factor in the VAR model. To quantitatively describe the contributions of port stocks, sales volume, futures prices, PPI, and crude oil price on coal price, variance decomposition of our VAR model is given in Table 4. Here, the contributions to changes in coal price standard error are shown for these five factors.

As shown in Table 4, in the future 15 periods, coal price has the greatest impact during the first six periods with its own contribution rate above 45%. Since then, rates in the rest of the period level off at about 22%. The contribution rate of PPI to coal price increase steeply from the fifth period and then exceeds all other variables at the seventh period and then remains stable and keeps the high level between 62% and 65%. Futures price shocks rank third between fifth and eighth periods and in the remaining periods, and sales volume shocks rank third after eighth period. Contribution rates of port stocks and crude oil price remain relatively static. Therefore, in the long run, economic conditions have the biggest contribution to the changes of coal price; in a short time, coal price has the biggest contribution to its own changes.

3.6. Results and Error Test. Figure 3 shows the obtained results. We used 29 months (November 2013–March 2016) to predict the coal price between January 2014 and April 2016. The blue line represents the predicted coal price within 28 months since January 2014 and the red line represents the actual coal price. Clearly, the simulated price curve has the same downward trend as the practical situation in the 28 months (January 2014–April 2016). The curves are comparatively close to each other, which proved the accuracy of the VAR model.

We applied the root mean square error (RMSE) to further verify precision. The formula of RMSE is expressed as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_{ci} - y_i)^2}{n}},$$

where $y_{ci}$ is the coal price in the simulated period, $y_i$ is the coal price in the actual period, and $n$ is the number of periods.
Table 4: Variance decomposition analysis.

<table>
<thead>
<tr>
<th>Period</th>
<th>S.E.</th>
<th>LOGF</th>
<th>PPI</th>
<th>LOGP</th>
<th>LOGS</th>
<th>LOGSS</th>
<th>LOGW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.041035</td>
<td>0.729212</td>
<td>9.299317</td>
<td>89.97147</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0.052049</td>
<td>7.701295</td>
<td>9.956613</td>
<td>78.18305</td>
<td>0.080677</td>
<td>2.045599</td>
<td>2.032769</td>
</tr>
<tr>
<td>3</td>
<td>0.05851</td>
<td>12.06564</td>
<td>10.90689</td>
<td>70.85491</td>
<td>0.191851</td>
<td>5.118254</td>
<td>2.472727</td>
</tr>
<tr>
<td>4</td>
<td>0.073097</td>
<td>13.37388</td>
<td>10.90689</td>
<td>66.4985</td>
<td>0.588931</td>
<td>6.204526</td>
<td>2.427272</td>
</tr>
<tr>
<td>5</td>
<td>0.094675</td>
<td>12.9068</td>
<td>19.39546</td>
<td>59.33157</td>
<td>0.530491</td>
<td>5.646364</td>
<td>2.189317</td>
</tr>
<tr>
<td>6</td>
<td>0.11771</td>
<td>10.38902</td>
<td>36.39546</td>
<td>45.83592</td>
<td>1.490347</td>
<td>4.238021</td>
<td>1.651235</td>
</tr>
<tr>
<td>7</td>
<td>0.138359</td>
<td>7.231468</td>
<td>51.05857</td>
<td>34.6953</td>
<td>2.74231</td>
<td>3.178658</td>
<td>1.093694</td>
</tr>
<tr>
<td>8</td>
<td>0.155671</td>
<td>4.972065</td>
<td>59.05059</td>
<td>28.80485</td>
<td>3.467079</td>
<td>2.8815</td>
<td>0.823913</td>
</tr>
<tr>
<td>9</td>
<td>0.169861</td>
<td>3.748737</td>
<td>62.73128</td>
<td>25.76903</td>
<td>3.893932</td>
<td>2.944914</td>
<td>0.912112</td>
</tr>
<tr>
<td>10</td>
<td>0.180673</td>
<td>2.72994</td>
<td>64.80983</td>
<td>23.17257</td>
<td>4.364957</td>
<td>3.310999</td>
<td>1.616162</td>
</tr>
<tr>
<td>11</td>
<td>0.191171</td>
<td>2.526387</td>
<td>64.76071</td>
<td>22.79765</td>
<td>4.468586</td>
<td>3.496792</td>
<td>1.949877</td>
</tr>
<tr>
<td>12</td>
<td>0.192665</td>
<td>2.422941</td>
<td>64.49293</td>
<td>22.75454</td>
<td>4.511523</td>
<td>3.641249</td>
<td>2.185818</td>
</tr>
<tr>
<td>13</td>
<td>0.193143</td>
<td>2.376791</td>
<td>64.20818</td>
<td>22.83548</td>
<td>4.51951</td>
<td>3.738696</td>
<td>2.310902</td>
</tr>
<tr>
<td>14</td>
<td>0.193264</td>
<td>2.361875</td>
<td>63.99327</td>
<td>22.99525</td>
<td>4.496838</td>
<td>3.797168</td>
<td>2.355594</td>
</tr>
</tbody>
</table>

Table 5: Comparison results.

<table>
<thead>
<tr>
<th>Prediction method</th>
<th>RMSE</th>
<th>MAE</th>
<th>MRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>April 2016 (VAR)</td>
<td>1.39%</td>
<td>4.8</td>
<td>1.07%</td>
</tr>
<tr>
<td>May 2016 (VAR)</td>
<td>1.48%</td>
<td>5.6</td>
<td>1.26%</td>
</tr>
<tr>
<td>June 2016 (VAR)</td>
<td>1.46%</td>
<td>5.5</td>
<td>1.23%</td>
</tr>
<tr>
<td>July 2016 (VAR)</td>
<td>1.44%</td>
<td>5.4</td>
<td>1.21%</td>
</tr>
<tr>
<td>FS-SVR [5]</td>
<td>2.23%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVM [5]</td>
<td>7.27%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: MAE indicates the mean absolute error and MRE indicates the mean relative error.

where \( y_{ci} \) means predicted coal price, \( y_i \) means actual coal price, and \( n = 1, 2, 3, \ldots, 28 \).

We obtained “RMSE = 1.39%.” And we predict the coal price of May 2016 (using the data between November 2013 and April 2016), June 2016 (using the data between November 2013 and May 2016), and July 2016 (using the data between November 2013 and June 2016) by the same processes; the RMSE are 1.48%, 1.46% and 1.44%, respectively, which indicate good precision of the model.

According to the prediction results, the comparison results between the VAR model in this paper and other methods are shown as Table 5.

4. Conclusions and Discussion

This study performed a VAR model for predicting the coal price in China; port stocks, sales volume, futures prices, PPI, and crude oil price were used as influence factors in this model. After verifying the stability of the model, impulse response, variance decomposition, and the prediction of coal price were carried out. Results of the impulse response function indicated that the economic growth and increased coal price will become the driving force of coal price. The variance decomposition results showed that the contribution of state of the economy account is over 50% since the seventh period. RMSE were applied to verify precision and the value are 1.39%, 1.48%, 1.46%, and 1.44%, respectively, which indicated good precision of the model and indicated that it is instructive to coal related corporations and government departments.

Competing Interests

The authors declare that they have no competing interests.

Authors’ Contributions

Xiaopeng Guo designed the study and defined VAR model. Jiaxing Shi prepared the data, participated in designing the study, wrote the manuscript, and revised it. Dongfang Ren provided good advice for the model.

Acknowledgments

Project is supported by the Fundamental Research Funds for the Central Universities of China (no. MS201439).
References


