

Multivariate Time Series Analysis of China Coal Price Index

Jian Qiao, Xinyi Jia*

School of Science, China University of Mining and Technology (Beijing), Beijing
Email: *jiaxinyi98@163.com

Received: Jul. 27th, 2020; accepted: Aug. 12th, 2020; published: Aug. 19th, 2020

Abstract

Coal is currently the main energy source in my country, and the coal price index is an important representation of the current status of coal prices. Forecasting the coal price index can help companies plan production and adjust national policies. This article predicts through the ARIMAX model that WTI has a significant impact on my country's coal price index. When there are no emergencies, the coal price index may rise slightly in early 2020. Comparison with the ARIMA model shows that both the ARIMA model and the ARIMAX model can make a good prediction of my country's coal price index, but the ARIMAX model is superior to the ARIMA model.

Keywords

Coal Price Index, ARIMAX Model, ARIMA Model

中国煤炭价格指数的多元时间序列分析

乔 舰, 贾欣怡*

中国矿业大学(北京), 理学院, 北京
Email: *jiaxinyi98@163.com

收稿日期: 2020年7月27日; 录用日期: 2020年8月12日; 发布日期: 2020年8月19日

摘 要

煤炭是我国目前的主要能源, 煤炭价格指数是煤炭价格现状的重要呈现。对煤炭价格指数的预测可以帮助企业计划生产、国家调整政策。本文通过ARIMAX模型进行预测发现, WTI对我国煤炭价格指数有显著影响, 当没有突发事件发生时, 2020年初煤炭价格指数可能会出现小幅上涨。与ARIMA模型拟合对比显示, ARIMA模型和ARIMAX模型均能对我国煤炭价格指数进行良好的预测, 但ARIMAX模型优于ARIMA模型。

*通讯作者。

关键词

煤炭价格指数, ARIMAX模型, ARIMA模型

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

我国是煤炭大国, 目前煤炭消费是我国主要的能源消费。据统计, 2019年我国能源消费中, 煤炭消费量约 39.4 亿吨, 占比 57.7%, 超过半数[1]。作为最经济可靠的能源, 在可预见的未来, 煤炭在我国能源结构中的地位不会发生较大改变。

煤炭价格指数反映了煤炭市场的供需关系, 展示了煤炭价格波动的真实情况, 其变化发展不仅影响着煤炭的开采与挖掘, 对全国各行各业的生产经营都起着十分重要的作用。通过对煤炭价格指数的预测, 可以降低生产者的风险, 也可以帮助国家在宏观调控上制定政策、采取措施, 使我国经济保持稳定持续的发展。

中国煤炭价格指数(China Coal Price Index, 简称 CCPI)是一种可以全面、客观、及时地描述全国以及各区域、各品种煤炭市场价格变化的数据, 它创建于 2006 年, 以 2006 年 1 月 1 日为基期(100 点), 在“中国煤炭市场网”按周发布。

申万(2014) [2]等通过对澳大利亚煤炭价格和迪拜石油价格进行回归分析, 证实石油价格与煤炭价格存在正相关关系。王延伟、景晓真(2016) [3]采用灰色综合评价法对我国煤炭价格进行研究, 发现其受经济发展影响大。王帮俊、朱超群(2017) [4]对环渤海动力煤价格进行短期分析, 建立 ARIMA 模型, 得到了良好的预测结果。张同功、赵德志(2018) [5]对秦皇岛煤炭价格运用逐步回归法发现, 对煤炭价格影响较大的是煤炭产量、煤炭进出口量和国际石油价格等。赵修茗、张甜、邹绍辉(2020) [6]对我国煤炭价格指数等交易数据构建 VAR 模型, 发现我国煤炭经济变动与动力煤期货价格呈均衡关系。

煤炭价格指数受多种因素影响, 国际原油价格、原煤产量、原煤进口量、煤炭销量等都会影响其波动, 但由于 ARIMAX 模型要求输入变量为平稳时间序列, 通过检验, 本文最终选取美国西德克萨斯轻质原油(West Texas Intermediate, 简称 WTI)价格为 ARIMAX 模型的输入变量。

使用求和自回归移动平均模型(ARIMA)和传递函数模型(ARIMAX)对 2013~2019 年我国煤炭价格指数和 WTI 价格周数据进行分析, 研究两种模型在分析我国煤炭价格指数方面的优劣性, 并使用两种方法分别对 CCPI 进行短期预测。

2. 模型理论

2.1. ARIMA 模型简介

求和自回归移动平均(ARIMA)模型是在进行一元时间序列分析时的经典模型。其结构为:

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E(\varepsilon_t x_s) = 0, \forall s < t \end{cases}$$

式中, $\nabla^d = (1-B)^d$; $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, 为平稳可逆 ARMA(p,q)模型的自回归系数多项式; $\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$, 为平稳可逆 ARMA(p,q)模型的移动平滑系数多项式。

通过差分将非平稳时间序列转化为平稳, 可以解决绝大多数单变量时间序列的分析问题; 但因为很多时间序列的变化都会受到其他序列的影响, 若仅考虑自身过去对现在的影响, 得出的模型可能并不精确, 这就涉及多元时间序列分析。

2.2. ARIMAX 模型简介

带有输入序列的 ARIMA 模型称为传递函数(ARIMAX)模型, 可以对多元时间序列进行分析。传递函数模型基于假定: 输入序列 x_t 、输出序列 y_t 均为零均值平稳时间序列。若序列不平稳, 可以通过差分运算或方差齐性变换使其平稳, 非零均值序列可通过序列中心变换使序列均值变为零。

实际应用的传递函数模型为:

$$Y_t = \frac{\Omega(B)B^b}{E(B)} X_t + \frac{\Theta(B)}{\Phi(B)} a_t$$

其中, $\Omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s, \omega_0 \neq 1, E(B) = 1 - \varepsilon_1 B - \varepsilon_2 B^2 - \dots - \varepsilon_r B^r$ 。

2.3. ARIMAX 建模步骤

① 对 x_t 、 y_t 进行平稳性检验和纯随机性检验: 若不平稳, 可通过取对数、求差分等方法使其平稳。

② 预白噪化输入序列 x_t : 对 x_t 进行 ARMA 建模, 得到 $\Phi_x(B)x_t = \Theta_x(B)\alpha_t$, 即 $\alpha_t = \frac{\Phi_x(B)}{\Theta_x(B)} x_t$, 此时 α_t 为白噪声序列。

③ 预白噪化输出序列 y_t : 令 $\beta_t = \frac{\Phi_x(B)}{\Theta_x(B)} y_t$ 。

④ 估计脉冲响应函数 \hat{v}_k : 计算 α_t 和 β_t 的互相关系数 $\rho_{\alpha\beta}(k)$, 脉冲响应函数 $\hat{v}_k = \frac{\sigma_\beta}{\sigma_\alpha} \hat{\rho}_{\alpha\beta}(k)$ 。

⑤ 初步判定传递函数模式: 识别延迟参数 b , 衰减参数 r 和参数 s , 基于 $v(B) = \frac{\Omega(B)}{E(B)} B^b$ 估计参数 $\hat{\varepsilon}_j$,

$\hat{\omega}_j$, 从而得到传递函数的估计: $\hat{v}(B) = \frac{\hat{w}(B)}{\hat{\delta}(B)} B^b$ 。

⑥ 估计噪声序列: $\hat{\eta}_t = y_t - \hat{v}(B)x_t = y_t - \frac{\Omega(B)}{E(B)} B^b x_t$ 。

⑦ 噪声序列 ARMA 建模: 对噪声序列 $\hat{\eta}_t$ 构建 ARMA 模型: $\Phi_{\hat{\eta}}(B)\hat{\eta}_t = \Theta_{\hat{\eta}}(B)a_t$, 即 $\hat{\eta}_t = \frac{\Theta_{\hat{\eta}}(B)}{\Phi_{\hat{\eta}}(B)} a_t$,

进而得到传递函数模型: $y_t = \frac{\hat{w}(B)}{\hat{\delta}(B)} B^b x_t + \frac{\Theta_{\hat{\eta}}(B)}{\Phi_{\hat{\eta}}(B)} a_t$ 。

⑧ 参数估计: 基于以上过程所得相应延迟阶数, 对传递函数模型: $y_t = \frac{w(B)}{\delta(B)} B^b x_t + \frac{\Theta_{\hat{\eta}}(B)}{\Phi_{\hat{\eta}}(B)} a_t$ 所有未知参数重新进行最大似然估计。

⑨ 模型检验: 1) 残差序列 a_t 纯随机性 Q_{LB} 检验; 2) 残差序列 a_t 与输入序列 x_t 纯随机性 Q_{LB} 检验。

3. 实例分析

虽然 Eviews 软件可以实现 ARIMAX 模型建模的部分步骤, 但相较之下, SAS 软件能程序化实现整个建模过程, 所以本文选用了 SAS 软件进行实现。

3.1. 数据选择

本文选用 2013 年 1 月至 2019 年 12 月 CCPI 和 WTI 价格进行分析。所用 WTI 价格来源于美国能源信息网站, CCPI 来源于 wind 数据库。数据时序图如图 1 所示, CCPI 均值为 150.2761, 标准差为 14.0821, 最高值出现在 2013 年 1 月 3 日为 171.1, 最低值出现在 2015 年 11 月 13 日为 124.7。

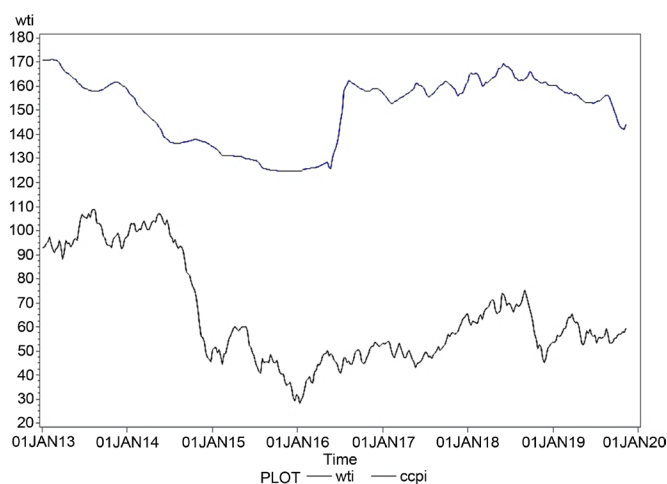


Figure 1. CCPI and WTI price timing chart

图 1. CCPI 与 WTI 价格时序图

3.2. 数据检验

由图 1 可知 WTI 价格和我国煤炭价格指数均为非平稳时间序列。分别对其进行一阶差分, 得到的差分后序列时序图如图 2、图 3 所示。

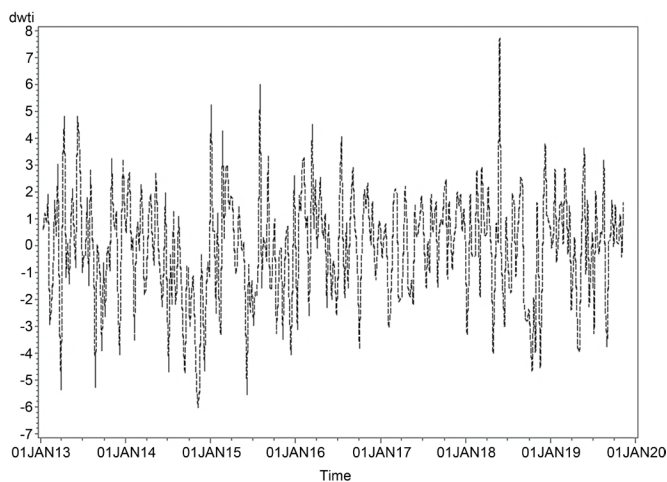


Figure 2. After difference WTI price sequence diagram

图 2. 差分后 WTI 价格时序图

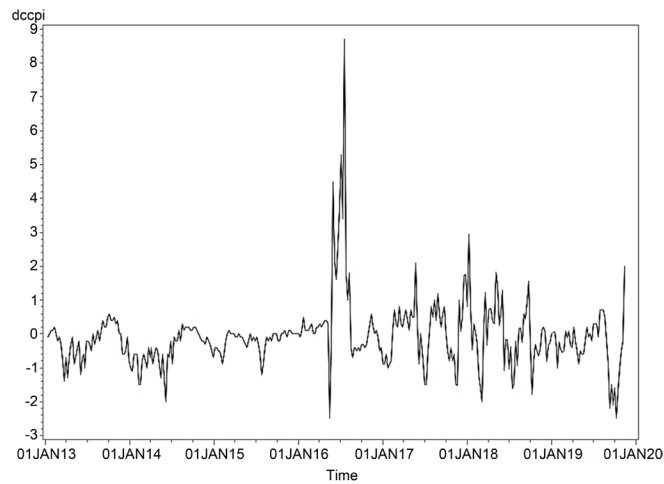


Figure 3. Differential CCPI sequence diagram

图 3. 差分后 CCPI 时序图

由图 2、图 3 可初步判别两数据差分后为平稳序列。进一步通过单位根检验证实了此结论。

二者的白噪声检验结果显示两序列均为非白噪声序列, 可以进行建模。记差分后 WTI 价格序列为 x , 差分后 CCPI 序列为 y 。

3.3. 建立模型

首先, 对 y_t 进行预白化处理, 通过 AIC 准则对比最小信息量结果, 得到最优模型 AR(1), 去掉不显著的常数项后拟合结果为

$$x_t = (1 + 0.2996B)a_t; \quad a_t \stackrel{iid}{\sim} N(0, \sigma_a^2)。$$

预白化 CCPI 差分序列 y_t 得到序列 \tilde{y}_t :

$$\tilde{y}_t = \frac{1}{1 + 0.2996B} y_t。$$

计算 a_t 与 \tilde{y}_t 的互相关系数 $\rho_{a_t, \tilde{y}_t}(k)$, 结果见图 4, 基于该互相关系数初步判定脉冲响应函数的相关参数: $b = 1, r = 1, s = 0$ 。

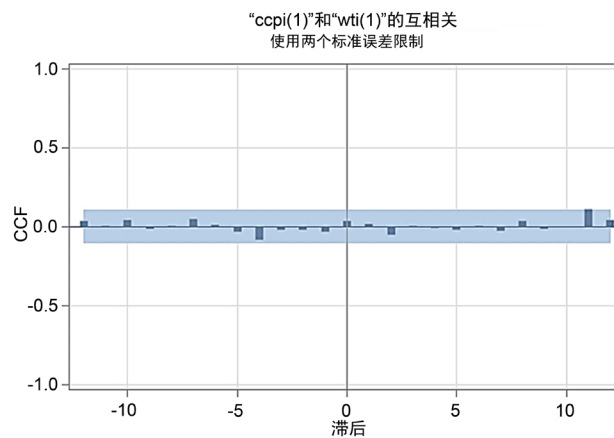


Figure 4. Intercorrelation function of input-output pre-white sequence

图 4. 输入输出预白化后序列的互相关函数

估计结果常数项不显著, 去掉常数项后, 各个参数显著, 估计结果如表 1 所示。

Table 1. Parameter estimation
表 1. 参数估计

参数	估计	标准误差	t 值	近似 Pr > t	滞后	变量	位移
NUM1	0.0119	0.0260	0.46	0.0137	0	wti	0
NUM1,1	-0.0180	0.0279	-0.65	<0.0001	1	wti	0
DEN1,1	-0.6066	1.4986	-0.40	<0.0001	1	wti	0

基于参数估计结果, 脉冲响应函数的估计结果为:

$$\hat{v}(B) = \frac{0.0119B + 0.0180B^2}{1 - 0.7904B}$$

对噪声序列 $\eta_t = y_t - \left(\frac{0.0119B + 0.0180B^2}{1 - 0.7904B} \right) x_t$ 尝试 ARMA(2,0) 模型进行参数估计, 结果如表 2 所示。

Table 2. Parameter estimation
表 2. 参数估计

参数	估计	标准误差	t 值	近似 Pr > t	滞后	变量	位移
AR1,1	0.5520	0.0533	10.35	<0.0001	1	ccpi	0
AR1,2	0.1578	0.0533	2.96	0.0033	2	ccpi	0
NUM1	0.0241	0.0185	1.30	0.1931	0	wti	0
NUM1,1	-0.0298	0.0199	-1.49	0.1361	1	wti	0
DEN1,1	-0.4971	0.6768	-0.73	0.0631	1	wti	0

参数通过了参数显著性检验, 得到最终传递函数模型:

$$y_t = \left(\frac{0.0242B + 0.02986B^2}{1 + 0.4971B} \right) x_t + \left(\frac{1}{1 - 0.5521B + 0.1578B^2} \right) \varepsilon_t$$

该估计结果中的参数都通过了显著性检验, 模型残差通过了白噪声检验, 但模型残差与输入变量之间仍存留一定相关性, 对该模型的预测精度有一定影响。

可以看出, 输入变量 WTI 价格对滞后 1 期 CCPI 有显著影响, 随着时间的发展, 当期国际原油价格对滞后期煤炭价格指数的影响将变得越来越小。

3.4. 预测

通过以上模型拟合, 对我国 2019 年 12 月至 2020 年 2 月煤炭价格指数进行预测, 预测结果如表 3。

Table 3. Forecast of China's coal price index
表 3. 我国煤炭价格指数预测

日期	预测	标准误差	95% 置信限	
2019.12.20	144.6013	0.8813	142.8740	146.3286
2019.12.27	145.6002	1.3540	142.9465	148.2540
2020.01.03	146.7272	1.8795	143.0435	150.4109
2020.01.10	147.7315	2.4778	142.8750	152.5879
2020.01.17	148.5232	3.1224	142.4035	154.6429
2020.01.24	149.0502	3.7768	141.6478	156.4525
2020.01.31	148.0458	4.5114	139.2037	156.8880
2020.02.07	148.2110	5.0625	138.2887	158.1330
2020.02.14	148.3359	5.5864	137.3868	159.2850

结果显示, 我国煤炭价格指数在 2019 年末和 2020 年初有递增趋势, 且浮动不大。这可能是因为进入冬季, 天气寒冷, 供暖消耗大量煤炭, 所以导致价格上涨。

3.5. 对比结论

通过以上模型拟合, 可以看出, 国际原油价格对我国煤炭价格指数有十分重要的影响。输入变量为国际原油价格时, 可以使用 ARIMAX 模型对 CCPI 进行预测。

使用 ARIMA 模型和 ARIMAX 模型的预测效果图如图 5、图 6 所示, 可见两方法的拟合效果均良好。

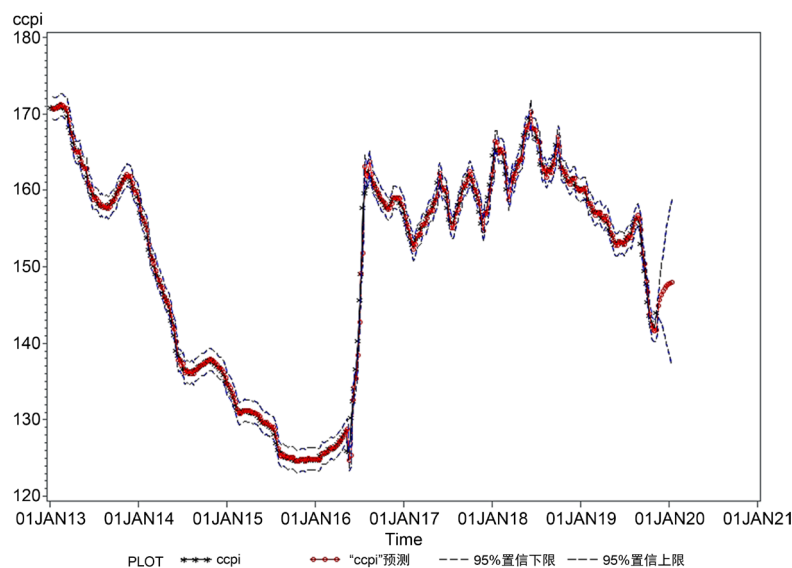


Figure 5. After difference WTI price sequence diagram
图 5. 基于 ARIMA 模型的 CCPI 预测

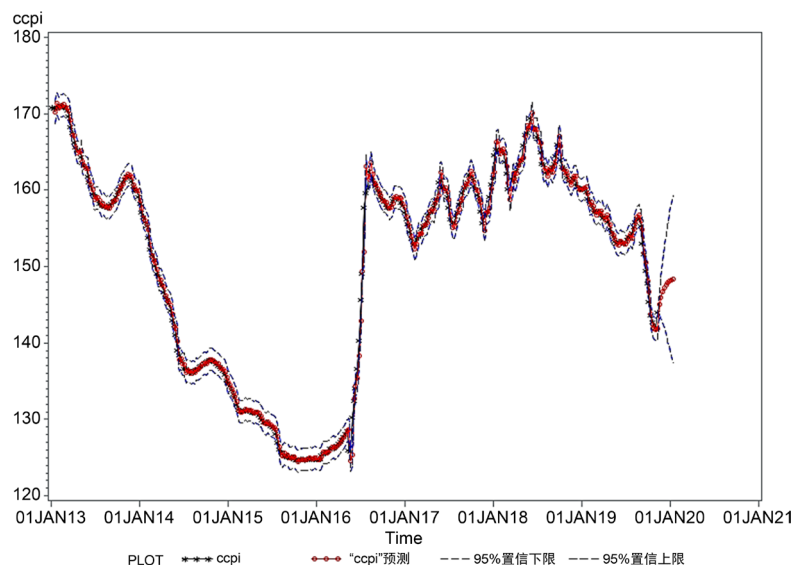


Figure 6. Differential CCPI sequence diagram

图 6. 基于 ARIMAX 模型的 CCPI 预测

将两模型的检验结果对比, 对比结果如表 4 所示。相比之下, ARIMAX 模型的标准误差较小, 不同准则下的最小信息量也较小, 整体优于 ARIMA 模型。

Table 4. ARIMA and ARIMAX model test results

表 4. ARIMA 与 ARIMAX 模型检验结果对比

	ARIMAX 模型	ARIMA 模型
标准误差估计	0.74026	0.74116
AIC	804.952	820.242

4. 小结

4.1. 结论

分别使用 ARIMA 模型和 ARIMAX 模型对我国煤炭价格指数预测的过程中发现, 两种预测方法均能得到较好的拟合效果, 且 ARIMAX 模型的结果优于 ARIMA 模型。

WTI 对我国煤炭价格指数有显著影响。在我国能源结构中, 煤炭和石油是主要消费的一次能源, 二者互为替代品, 当其中一种价格上升时, 对另一种的需求就会增大, 从而也就带动了另一种价格的上涨。

但是由于 ARIMAX 模型对输入变量的平稳性有较高要求, 所以当其无法满足时也可使用 ARIMA 模型。ARIMAX 模型作为一种限制条件较多的模型, 应用受到很大的限制, 但其结果十分可靠。使用更多解释变量的 ARIMAX 模型可能会得到更好的预测效果。

4.2. 不足与展望

我国煤炭价格指数受诸多因素影响, 本文主要研究了 CCPI 与 WTI 价格之间的关系, 还可以尝试寻找其他平稳或差分平稳时间序列对 CCPI 进行 ARIMAX 模型建模。

现有模型普遍在短期预测时效果较好, 后续需根据实际情况对模型进行检验与调整。

虽然通过 ARIMA 模型和 ARIMAX 模型均得到了较好的拟合效果, 且二者预测结果较为一致, 但对

于突发事件的影响是无法预测的。如受今年新冠肺炎的影响, 物流受阻, 人员流动受限, CCPI 的上涨趋势可能也会受到影响。

随着经济全球化趋势不断发展, 国内外联系日趋紧密, 我国煤炭市场与国际原油市场的关联将日益增加, 相关企业与国家在进行决策时都应考虑国际原油价格的长远影响。

基金项目

本文受中国矿业大学(北京)“课程思政”示范课程建设项目(时间序列分析)的支持。

参考文献

- [1] 王健. 新冠疫情对煤炭行业的影响分析[J]. 煤炭经济研究, 2020, 40(4): 42-49.
- [2] 申万, 姜殿虹, 王有勇, 张广军. 国际市场煤炭与石油价格相关性研究[J]. 中外能源, 2014, 19(4): 6-10.
- [3] 王延伟, 景晓真. 我国煤炭价格走势及其影响因素实证分析[J]. 价格理论与实践, 2016(6): 114-117.
- [4] 王帮俊, 朱超群. 去产能背景下煤炭价格走势及预测分析——基于 ARIMA 模型的研究[J]. 价格理论与实践, 2017(5): 73-76.
- [5] 张同功, 赵得志. 我国煤炭价格波动影响因素实证研究[J]. 中国能源, 2018, 40(3): 16-21.
- [6] 赵修茗, 张甜, 邹绍辉. 动力煤期货价格波动对我国煤炭经济影响研究[J]. 中国矿业, 2020, 29(1): 34-40.