Prediction of CSI China Mainland Low Carbon Economy Index Based on ARIMA-GARCH Model

Yijie Lei*

School of Economics and Management, Nanjing University of Science and Technology, Nanjing, China

*Corresponding author: 931643563@qq.com

Abstract. China is entering a new stage of development. It is a hot topic to play the driving role of green finance and help the realization of carbon neutralization. The change of stock price of green finance related industries can directly reflect the development trend of low-carbon economy and the development prospect of carbon neutralization in the future. Therefore, this paper selects the closing price of CSI China Mainland Low Carbon Economy Index (Mainland L-C Index) from July 1, 2010 to June 18, 2021 as the research data, establishes ARIMA-GARCH model to analyze and predict the future mainland low-carbon index, so as to effectively grasp the development direction of low-carbon economy and provide basis for the government to formulate relevant policies.

Keywords: Carbon Neutrality; Green Finance; Mainland L-C Index; ARIMA-GARCH Model.

1. Introduction

Global warming, ozone layer destruction and other climate problems are common, and have attracted great attention all over the world. Facing the increasingly serious climate problem, the Chinese government has always taken positive action. In 2020, the Chinese government announced that China's carbon dioxide emissions would reach a peak by 2030 and be carbon neutral by 2060. To achieve carbon neutrality, we must promote the transformation of industrial structure to low-carbon direction, which requires green finance to guide funds into low-carbon field. Yi Gang, governor of the people's Bank of China, proposed to vigorously develop green finance to help achieve the goal of carbon neutralization and carbon peak as scheduled. Green finance mainly includes stocks, bonds and other businesses related to green industries. the research on green finance covers a wide range. Wang Xin and Wang Ying used the double difference model to investigate the impact of green credit policy on green innovation and believed that green finance can promote green innovation [1]. Chen Danning used the event study method to explore the impact of green bonds issued by Chinese listed companies on stock prices [2]. Zhu Yijie analyzed the development of green stock index and proposed that social funds should flow into the green industry through index products to boost the development of green industry [3]. The research on green finance is diverse, and the stock market is an important part of the financial system. The stock price changes related to green finance can directly reflect the development trend of low-carbon economy. In studying the stock price trend, Zhang Yingchao and Sun Yingjun used ARIMA model to predict and analyze the future Shanghai stock index [4]. Wu Yuxia and Wen Xin predicted the law and trend of stock price change in gem by establishing ARIMA model. The results show that the static prediction effect of this model is good [5]. Luo Zhidan et al. established a mixed prediction model to predict the trend of stock price in the future [6]. Li Li established the ARMA-GARCH combination prediction model to describe the internal relationship between stock price and trading volume [7].

To sum up, in order to achieve the goal of carbon neutralization, it is particularly important to predict the price trend of green finance related stocks. This paper analyzes the historical closing price of Mainland L-C Index, and constructs a stock price prediction model based on ARIMA-GARCH, so as to effectively grasp the development direction of green finance and provide a basis for the government to formulate relevant policies.

2. Model Introduction

2.1 ARIMA Model

The full name of ARMA model is autoregressive moving average model, which is one of the representative time series prediction methods proposed by Box and Jenkins in the early 1970s, so it can also be called Box Jenkins model [8]. ARIMA (p, d, q) model is an extension of ARMA (p, q) model, which is called differential autoregressive moving average model. The basic idea of ARIMA model is to treat the data sequence formed by the object to be predicted over time as a random sequence, and use the corresponding mathematical model to approximately describe this series. The model can predict the future value from the past value and current value of the time series.

ARIMA model can be converted into ARMA model by finite difference of non-stationary time series, which can be converted into stationary series. The basic expression of ARMA model is:

$$Y_t=\alpha_1 Y_{t-1}+\alpha_2 Y_{t-2}+\ldots+\alpha_p Y_{t-p}+\varepsilon_t+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-1}+\ldots+\theta_q\varepsilon_{q-1} \quad (1)$$

2.2 GARCH Model

The full name of ARCH model is autoregressive conditional heteroscedasticity model, which solves the problem caused by the second assumption of time series variables in traditional econometrics, that is, constant variance. GARCH model, proposed by bollerslev, is an extension of arch model, also known as generalized autoregressive conditional heteroscedasticity model [9]. GARCH model is suitable for volatility analysis and prediction, and can provide some reference for investors.

If variance is represented by ARMA model, ARCH model can be transformed into GARCH model. Generally, the expression of single variable GARCH (p, q) model is:

$$\sigma_t^2=\alpha_0+\alpha_1\varepsilon_{t-1}^2+\ldots+\alpha_p\varepsilon_{t-p}^2+\beta_1\sigma_{t-1}^2+\ldots+\beta_q\sigma_{t-q}^2 \quad (2)$$

3. Empirical Research

3.1 Data Sources

This paper uses the closing price of Mainland L-C Index (index code: 000977) excluding holidays from July 1, 2010 to June 18, 2021 as the sample data of this study, with a total of 2666 data. Among them, the data of Mainland L-C Index comes from NetEase Finance.

The Mainland L-C Index is a sample stock composed of 50 low-carbon economy theme companies with high daily average total market value selected from Shanghai and Shenzhen A shares to reflect the overall trend of low-carbon economy company stocks. By establishing ARIMA model to analyze and predict the low-carbon index, it is helpful to grasp the development direction of China's low-carbon economy in the future, analyze the activity of green financial market and the psychological expectations of the public for the future development of low-carbon economy, and provide reliable basis for policy makers.

As shown in Table 1, we first make descriptive statistics on the closing price of Mainland L-C Index. The average value is 1185.039, the maximum value is 2400.817, the minimum value is 607.616, and the standard deviation is 332.648, indicating that the fluctuation range of the closing price series is large.

Table 1. Descriptive statistics

	Min	Max	Mean	Std	Sample
Closing Price	607.616	2400.817	1185.039	332.648	2666

3.2 Seasonal Analysis and Stability Analysis

As shown in Fig 1, we use the statistical software EViews to make the timing diagram of the original data. Firstly, we can see from the figure that there is no obvious seasonal component in the closing price time series of Mainland L-C Index, so there is no need for seasonal decomposition in this modeling. It can be seen from the figure that the Mainland L-C Index showed a sharp upward trend from 2014 to 2015, and decreased rapidly during 2015, which to some extent reflects that the stock market is unstable and needs to be stabilized.

Figure 1. Timing chart of closing price of Mainland L-C Index

As shown in Fig 2, we perform logarithmic first-order difference on the sample data and made autocorrelation and partial autocorrelation diagrams after data processing by using statistical software. We find that most of the P values were greater than 0.05, indicating that the sequence of logarithmic first-order difference is more stable. In order to explain the stationarity of the current sequence more accurately, we need to carry out the unit root test. If there is no unit root, the model can be established. In the result of unit root test, the T value of ADF is -48.76707. Under the three significant levels of 1%, 5% and 10%, the critical values of the unit root test are -3.432612, -2.862425, - 2.567286, and the P value is 0.0001. Therefore, we can reject the original hypothesis and consider that the sequence after the logarithmic first-order difference has no unit root and is a stationary sequence. After transforming non-stationary series into stationary series, we can establish ARIMA model.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.056	0.056	8.3384	0.004
		2	-0.022	-0.025	9.5856	0.008
		3	0.007	0.009	9.7005	0.021
		4	0.014	0.013	10.224	0.037
		5	0.010	0.009	10.480	0.063
		6	-0.023	-0.024	11.933	0.063
		7	0.026	0.030	13.810	0.055
		8	0.018	0.013	14.676	0.066
		9	0.023	0.022	16.046	0.066
		10	-0.021	-0.023	17.269	0.069
		11	0.012	0.015	17.626	0.091
		12	0.006	0.001	17.714	0.125

Figure 2. AC and PAC diagrams after logarithmic first-order difference

3.3 Model Selection

We establish ARIMA (p, d, q) model for the original closing price series, where p is an autoregressive term, q is a moving average term, and d = 1. We find that p and q of ARIMA model can choose 1 and 2 respectively. Therefore, we preliminarily judge that the model can be ARIMA (1,1,1), ARIMA (1,1,2), ARIMA (2,1,1) or ARIMA (2,1,2). According to the AIC and SC minimum criteria, we list and compare the AIC and SC values of ARIMA models of second order and below, as shown in Table 2. Through the table, we find that the AIC value of ARIMA (2,1,1) model is the smallest, which is -5.34933, and the SC value of ARIMA (1,1,1) model is the smallest, which is -5.338564. Therefore, the final model is selected from ARIMA (2,1,1) model and ARIMA (1,1,1) model. We observe the regression results of the two models. For ARIMA (1,1,1) model, all parameters are significant at the 5% level. All parameters in ARIMA (2,1,1) are also significant at the 5% level. However, the p value of AR (2) in ARIMA (2,1,1) model is larger than other p values. To sum up, this paper chooses ARIMA (1,1,1) model to model the closing price series of China mainland low carbon index.

Table 2. Comparison of AIC and SC values of different ARIMA models

Model	AIC	SC
ARIMA(1,1,1)	-5.347438	-5.338564
ARIMA(1,1,2)	-5.349329	-5.338238
ARIMA(2,1,1)	-5.349333	-5.338241
ARIMA(2,1,2)	-5.348597	-5.335287

According to the above analysis, we believe that ARIMA (1,1,1) model is the best prediction model. We further find that the coefficients of AR (1) and MA (1) in the parameter estimates of ARIMA (1,1,1) are statistically significant, but the constant term C is not significant. Next, we discard the constant term C [10] and re estimate and test the ARIMA (1,1,1) model. The results show that each parameter has significant statistical significance, and the new AIC value and new SC value are smaller than those when the constant term is included, indicating that the model is more in line with the original data. Therefore, we choose ARIMA (1,1,1) as the final mean equation, which can be expressed as:

$$\Delta \ln Y_t = -0.3951 \Delta \ln Y_{t-1} + 0.4517 \alpha_{t-1} + \alpha_t \quad (3)$$

3.4 Residual Test

After we get the expression of ARIMA (1,1,1) model, we then test whether the residual sequence of the model is a white noise sequence. If the test results show that the residual sequence is not a white noise sequence, it means that some useful information has not been extracted by the residual sequence, and the previous model needs to be further improved. It is found that the sample autocorrelation diagram and partial autocorrelation diagram of the residual sequence fluctuate up and down near 0 in the random region, and the p value is much greater than 0.05, which shows that the residual sequence of the model established in this paper is a white noise sequence. As shown in Fig 3, we then observe the timing chart of the residuals. It is obvious that there are often large fluctuations after large fluctuations, and the fluctuations show aggregation. Therefore, we consider that there may be conditional heteroscedasticity in the residuals.

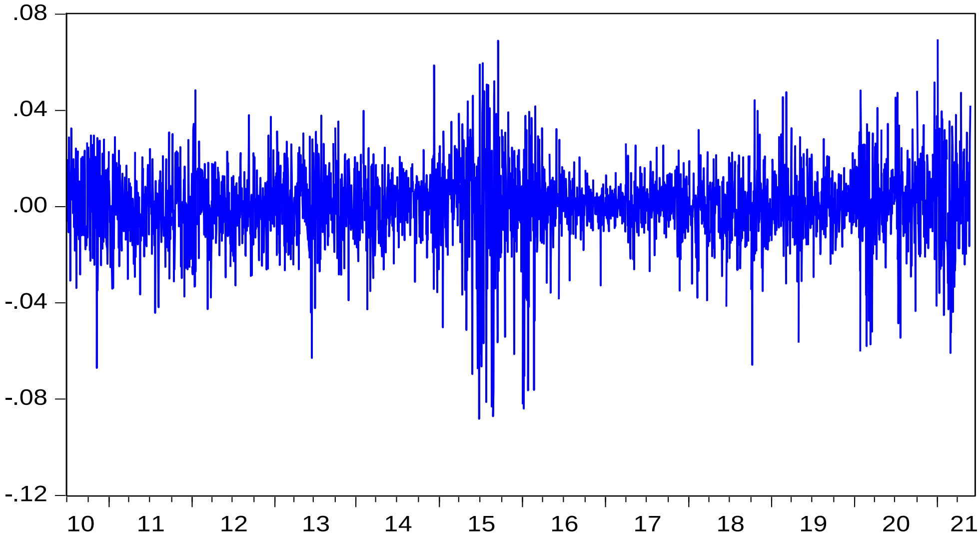


Figure 3. Sequence diagram of residuals

3.5 ARCH Effect Test

In order to determine whether there is conditional heteroscedasticity in the residuals, we conducted ARCH-LM test on the residuals. The results show that the p value is 0, so the residuals sequence of ARIMA (1,1,1) model does have ARCH effect. In order to carry out further research, we also make JB statistical chart after logarithmic difference of Mainland L-C Index. As shown in Fig 4, it can be seen from the observation chart that, firstly, after the logarithmic difference of the closing price of the Mainland L-C Index, there is a large kurtosis = 6.425, while the K value of the standard normal distribution is 3 (6.425 > 3), and the bulge of the distribution is greater than the normal distribution, so it indicates that there is a form of "sharp peak and thick tail". Secondly, the sequence distribution of JB statistical chart has a long-left tail, indicating that the mean is not zero, which violates the conclusion of normal distribution. Finally, after data processing, the JB statistic is 1500.324, which is far greater than the critical value of 5.99 at the significance level of 5%, indicating that the sequence is not normally distributed and can be considered to show the "thick tail" feature [11].

In summary, the establishment of ARIMA-GARCH models will have a significant meaning than a separate ARIMA model.

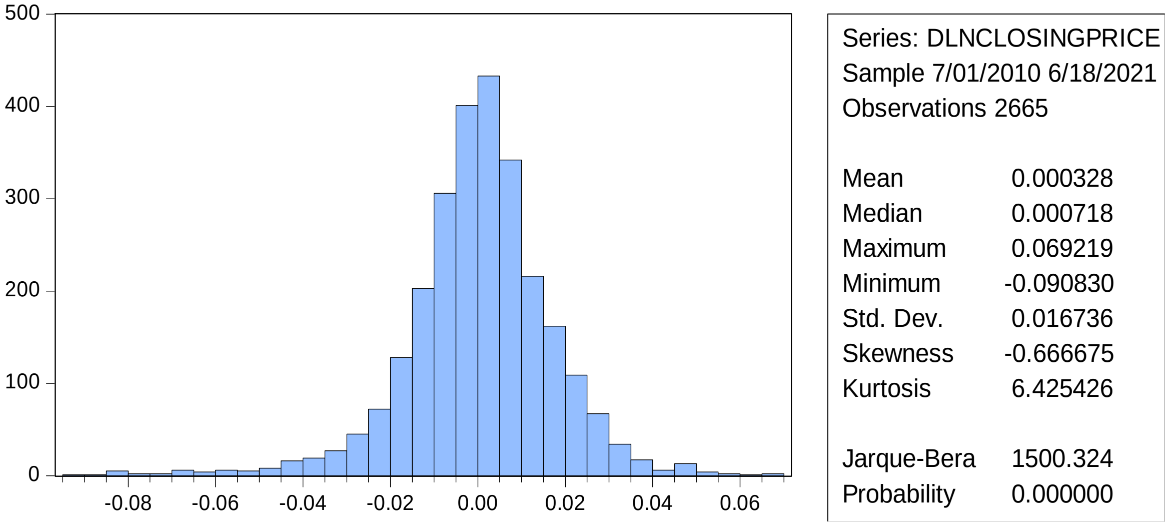


Figure 4. JB statistical chart

3.6 Establishment of ARIMA-GARCH Model

Through ARCH test, we find that there is conditional heteroscedasticity in the residuals of the model, so next, we establish the volatility equation for the residuals of the model. Referring to previous studies, we chose to establish GARCH (1,1) model to explain the conditional heteroscedasticity [12]. We analyze the results of errors under t distribution, biased t distribution, generalized error distribution and biased generalized error distribution. After observing whether the estimation of each parameter is remarkable, follow the AIC and SC minimum principles, and we finally determine the ARIMA (1, 1, 1) -GARCH (1, 1) -t model, so the improved model is:

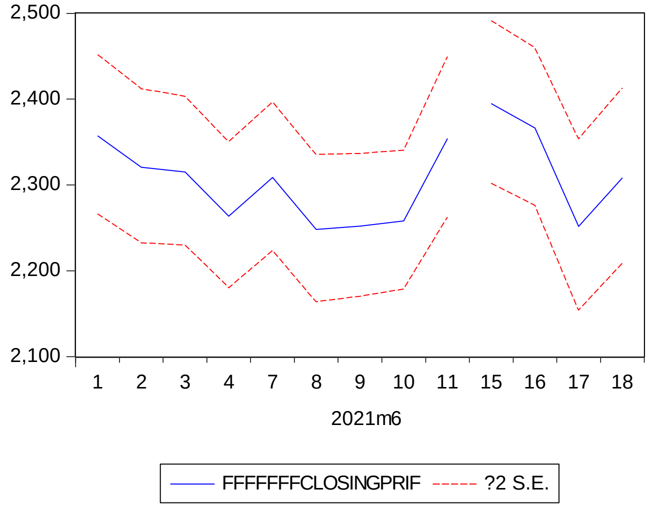
$$\Delta \ln Y_t = -0.8135\Delta \ln Y_{t-1} + \alpha_t + 0.8528\alpha_{t-1} \tag{4}$$

$$\alpha_t = \sigma_t \varepsilon_t \tag{5}$$

$$\sigma_t = 0.00000218 + 0.0667\alpha_{t-1}^2 + 0.9279\sigma_{t-1}^2 \tag{6}$$

3.7 Model Forecast

We use the closing price trading data from July 1, 2010 to May 31, 2021 as the original data to build ARIMA (1,1,1) - GARCH (1,1) model to predict the closing price of Mainland L-C Index from June 1, 2021 to June 18, 2021. We use statistical software to make a static forecast of the closing price, as shown in Fig 5.



Forecast: FFFFFFFCLOSINGPRIF	
Actual: CLOSINGPRICE	
Forecast sample: 6/01/2021 6/18/2021	
Included observations: 13	
Root Mean Squared Error	52.24388
Mean Absolute Error	43.41816
Mean Abs. Percent Error	1.880496
Theil Inequality Coefficient	0.011317
Bias Proportion	0.000001
Variance Proportion	0.000442
Covariance Proportion	0.999557
Theil U2 Coefficient	1.013174
Symmetric MAPE	1.878320

Figure 5. Static forecast chart of closing price

It can be seen from the figure that the fitting effect of the whole model is good, which is basically consistent with the actual value, so the model has good prediction effect. In order to show the fitting effect of the model in more detail, we list the comparison table between the actual value and the predicted value of the closing price of Mainland L-C Index from June 1, 2021 to June 18, 2021. As shown in Table 3, it can be seen from the table that all error ratios have been stable below 5%, indicating that the prediction effect of ARIMA (1,1,1) - GARCH (1,1) model is good. At the same time, we find that the average error ratio before June 9, 2021 is less than that after June 9, that is, with the increase of time, the prediction error of the model is gradually increasing, which shows that ARIMA-GARCH model has a good prediction effect on Mainland L-C Index in the short term, but it is not suitable for long-term prediction.

Table 3. Comparison table of actual value and predicted value

Date	Actual value	Predicted value	Error	Error ratio
2021.06.01	2325.398	2357.200	31.802	1.37%
2021.06.02	2311.702	2320.633	8.931	0.39%
2021.06.03	2268.288	2315.081	46.793	2.06%
2021.06.04	2303.315	2263.680	-39.635	1.72%
2021.06.07	2254.639	2308.749	54.11	2.40%
2021.06.08	2246.862	2248.143	1.281	0.06%
2021.06.09	2261.806	2251.951	-9.855	0.44%
2021.06.10	2347.282	2258.177	-89.105	3.80%
2021.06.11	2398.611	2354.177	-44.434	1.85%
2021.06.15	2364.683	2394.961	30.278	1.28%
2021.06.16	2257.494	2366.247	108.753	4.82%
2021.06.17	2301.764	2251.761	-50.003	2.17%
2021.06.18	2358.036	2308.581	-49.455	2.10%

4. Conclusion and Enlightenment

By establishing ARIMA (1,1,1) model for the closing price of Mainland L-C Index from July 1, 2010 to May 31, 2021, we find that there is heteroscedasticity in the residual sequence of the model. In order to make the established model fit the sample data better, we further improve the model and finally obtain ARIMA (1,1,1)-GARCH (1,1) model. Although the final prediction data of the model is not consistent with the real value to a great extent, the error ratio between the predicted closing price and the real closing price has been kept below 5%, indicating that the prediction effect of the new model is good and has a certain reference value. At the same time, we find that the prediction effect of the new model decreases with the increase of time, indicating that the model is only suitable for short-term prediction. If you want to make accurate long-term prediction, you need to improve the model, which is also the deficiency of this paper.

We find that in the future, Mainland L-C Index will fluctuate steadily, with an average of about 2307.56, which directly reflects that the development trend of China's low-carbon economy will move forward steadily in the future. This also shows that the market is full of expectations for low-carbon industries. In the future, it will attract more funds to invest in low-carbon industries, help industrial transformation and help achieve the goal of carbon neutralization. For the government, in terms of green finance, it can adopt stable welfare policies to guide funds to invest in green industries, pay attention to the dispersion and control of possible risks in the process of financing activities, establish sound policies to protect investors and enterprises, and give full play to the role of green finance in promoting carbon neutralization.

References

[1] Wang Xin, Wang Ying. Research on green credit policy promoting green innovation [J]. Management World, 2021, 3 7(06): 173 - 188.

[2] Chen Danning. Stock price effect of green bond issuance of Chinese Listed Companies [J]. Journal of Shanxi University of Finance and Economics, 2018, 40 (S2): 35 - 38.

[3] Zhu Yijie. Research and Reflection on the development of green stock index in China [J]. Financial Perspectives Journal, 2018 (07): 59 - 64.

[4] Zhang Yingchao, Sun Yingjun. An empirical study on the analysis and prediction of Shanghai stock index based on ARIMA model [J]. Economic Research Guide, 2019 (11): 131 - 135.

[5] Wu Yuxia, Wen Xin. Short term stock price forecasting based on ARIMA model [J]. Statistics & Decision, 2016 (23): 83 - 86.

[6] Luo Zhidan, Liu Ying, Guo Wei. Short term prediction of stock price based on Taylor expansion and ARIMA hybrid model based on tracking differentiator[J]. Mathematics in Practice and Theory, 2019, 49 (23): 67 - 77.

[7] Li Li. Research on the dynamic relationship between volume and price of stock market based on ARMA-GARCH model [J]. Statistics & Decision, 2011 (04): 144 - 146.

[8] Box G, Jenkins G, Reinsel G. Time series analysis forecasting and control - Rev. ed [J]. Journal of Time, 1976, 31 (2): 238 - 242.

[9] Bollerslev T. Generalized autoregressive conditional heteroskedasticity [J]. Eeri Research Paper, 1986, 31 (3): 307 - 327.

[10] Feng Pan, Cao Xianbing. An empirical study on stock price analysis and prediction based on ARMA model [J]. Mathematics in Practice and Theory, 2011, 41 (22): 84 - 90.

[11] Fang Yan, Gen Xueyang, Qin Shanshan. Research on price prediction of media sector index in Shanghai and Shenzhen-Analysis Based on ARIMA-GARCH model [J]. Price: Theory & Practice, 2018 (01): 102 - 105.

[12] Yang Qi, Cao Xianbing. Stock price analysis and prediction based on ARMA-GARCH model [J]. Mathematics in Practice and Theory, 2016, 46 (06): 80 - 86.