An Empirical Study on Stock Returns of Chemical Industry Based on ARMA-GARCH Model--The Case of NIKKE Chemicals

Lechi Shen, Ning Yang *

School of Economics and Management, China University of Mining and Technology, Xuzhou, China, 221116

* Corresponding author: 15365882598@163.com

Abstract. This paper uses the ARMA-GARCH model to make predictions on the volatility of stock returns in China, selects the time series of daily returns of the representative enterprise of the chemical industry, NIKKE CHEMICAL, as the object of study, analyzes the stock returns of the enterprise over the past 6 years, and uses the ARMA model to make predictions of stock returns, and at the same time, joins the effect of volatility, and models the risk rate by using the GARCH model, and empirical evidence The results show that the chosen ARMA-GARCH model has a good fit to the time series of daily returns. This paper examines stock prices and analyses them by constructing a theoretical model that can provide advice for investment decisions in the stock market.

Keywords: Time Series, ARMA-GARCH Model, Stock Return Forecasting.

1. Introduction

In the stock and even the futures market, the daily price of each period of time put together to consider can analyze the price changes on a particular day, while observing and analyzing the continuous changes in one of the prices can be roughly understand the daily price changes, how to analyze and predict the price of the stock in the latter time changes, up and down to how much, is very important [1-3].

In the quantitative calculation and prediction, the stock price itself is often sensitive to the time factor, so the establishment of a time series model for price analysis and forecasting is a more ideal method [4].

The basic idea of time series forecasting method is: when predicting the future change of a phenomenon, the past behavior of the phenomenon is used to predict the future, i.e., the historical data through the time series reveals the law of the phenomenon over time, and extends this law to the future, so as to make a prediction of the future of the phenomenon. Time series analysis method is a quantitative analysis method to study the development and change law of things by analyzing the correlation of variables at different moments and revealing their correlation structure [5-6].

Therefore, this paper for the Shenzhen Stock Exchange Venture A share NIKKE CHEMICAL (300214) in the stock daily return data from May 2, 2017 to May 1, 2023, constructed the stock return volatility of ARMA-GARCH model, based on which the empirical analysis to explore.

2. Introduction to ARMA and GARCH models

2.1. ARMA model

The ARMA(p,q) model is a method dedicated to the analysis and forecasting of non-stationary time series and is one of the linear models often used in finance for forecasting over short periods of time.

AR is autoregressive, p is a regression term stating the existence of a correlation between the predicted historical data and the data at that time, using the observed historical time data to predict itself, MA is moving average, q is the number of moving average terms, emphasizing on the accumulation of the error term in the autoregression, ARMA is autoregressive mean moving model, ARIMA model is the difference calculation on the ARMA model, namely Any unsmoothed series

that can be smoothed after differencing using a certain order of differencing, at which point an ARMA model can be fitted to realize the differenced series [7].

2.2. ARCH model and GARCH model

BP neural network is back propagating, mainly composed of three parts: input layer, middle layer and output layer. The number of nodes in the input and output layers is relatively easy to determine, but the determination of the number of nodes in the hidden layer is a very important and complex problem [8].

In order to find the variance of the returns, it is necessary to model the volatility of the return series. The volatility is the conditional standard deviation of the rate of return. The volatility of a yield series is not directly observable. However, there are some observable characteristics: 1. There is a volatility aggregation effect: volatility is higher in one consecutive period and lower in another consecutive period. 2. Volatility reacts differently to large increases and decreases in yields, that is to say, it reacts differently to positive and negative market sentiments, which is known as leverage. the ARCH model also has significant drawbacks. It assumes that volatility is subject to positive and negative disturbances to the same extent. However, a large number of studies have shown that the prices of assets in financial markets respond differently to positive and negative disturbances. the parameters of the ARCH model are constrained to a certain range, which is demanding for the model to satisfy the conditions. In ARCH (1), the range of values is limited to a certain range, and for higher-order ARCH models, the constraints on the parameters will be more complicated [9-10].

To address the shortcomings of the ARCH model, the constraints on the parameters are more stringent. In view of this, Bollerslev proposed a generalized ARCH model. It is also known as the classical GARCH model. Set is the mean-corrected new interest series, the same as the ARCH same. It is usually assumed to obey the standard normal distribution, t-distribution, and GED distribution. are the parameters of the ARCH and are known as the parameters of the GARCH. The coefficients need to satisfy, as a way to ensure that the unconditional variance is finite and its conditional variance is time varying. When n=0, at this point the GARCH(m,n) model degenerates into an ARCH(m) model [11].

At this point still start with the GARCH (1,1) model. We can see that the larger or will lead to the larger, which also indicates that the larger will be accompanied by another large, which is the phenomenon of "volatility aggregation" that is common in financial time series. Similarly, the kurtosis in the GARCH model is sharper than the normal distribution, and the tails are thicker than the tails of the normal distribution. The phenomenon of "sharp peaks and thick tails" of the return series is satisfied.

Similar to the ARCH model, the GARCH model also has the same defects. It has the same portrayal of positive and negative disturbances, which obviously does not match with the response of the asset return series in the financial market to good or bad market conditions.

3. Results

3.1. Data sources

In this paper, the stock daily return data of NIKKE CHEMICAL (300214), an A-share of Shenzhen Stock Exchange Venture Stock, from May 2, 2017 to May 1, 2023 are selected to analyze and forecast. The company belongs to the leading enterprises in the plastic modifier industry and is a national high-tech enterprise. The large data prediction model for the user's electricity consumption is implemented in the Clementine software.

3.2. Analysis of experimental results

The data acquisition and cleaning are done by using R software, and the natural logarithmic treatment is done on the returns in order to make the data more accurate and reasonable. First look at the image distribution of the data, as shown in Figure 1. It is obvious from Figure 1 that the data shows

an overall steady state over time, fluctuating around a certain number, and its smoothness is tested using the ADF unit root test.

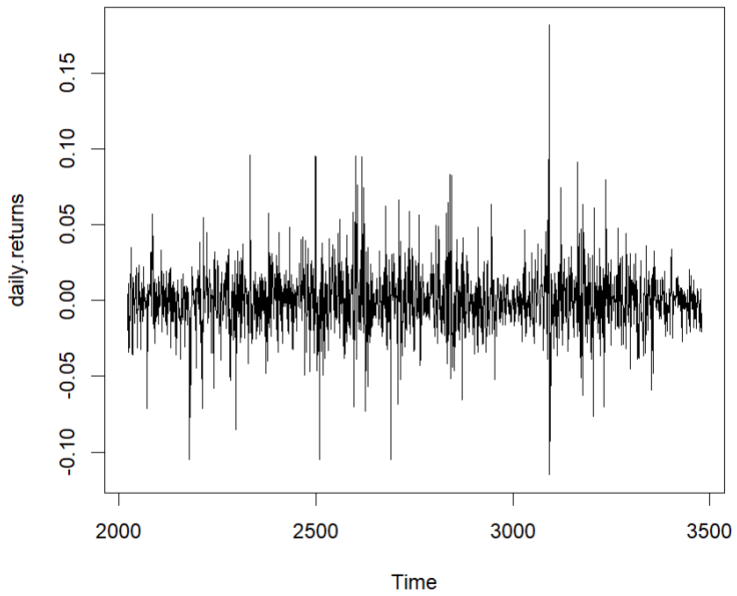

Figure 1. Logarithmic yield series

The P-value is 0.01, which is less than the significance level α. The original hypothesis is rejected and the logarithmic yield series is considered to have no unit root, i.e., the series is smooth and model construction can be carried out.

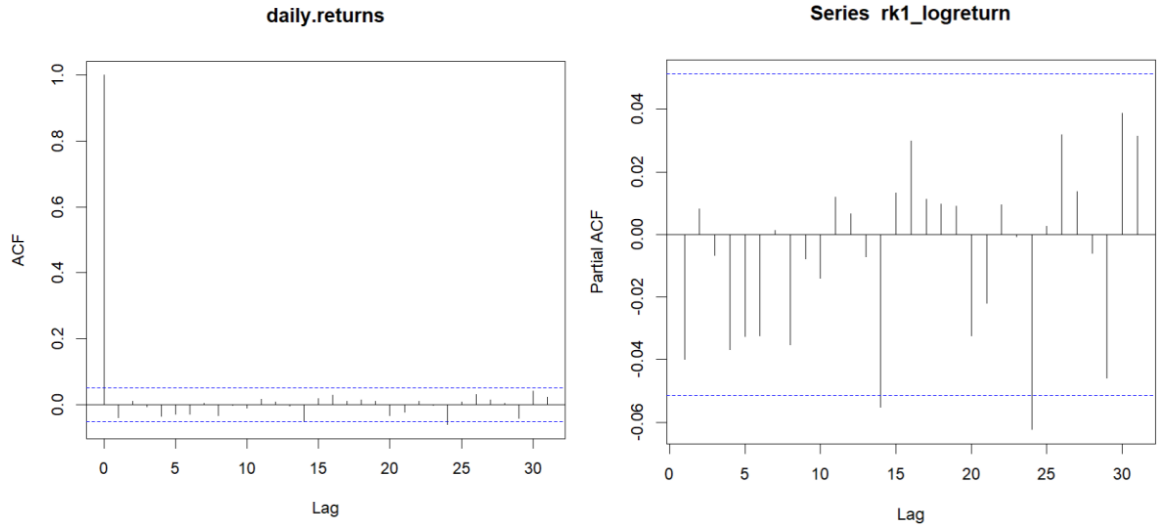

Figure 2. ACF, PACF plots of logarithmic yield series

The autocorrelation coefficient plots and partial autocorrelation coefficient plots are obtained based on the logarithmic yield series, as shown in Figure 2. The autocorrelation and partial autocorrelation plots in Figure 2 show that the autocorrelation coefficient and partial autocorrelation coefficient are significant at the 14th and 24th orders, and the ACF and PACF plots do not trail, so that the model cannot be simply designated as an AR or MA model, and the ARMA model is considered. The basic method of order setting for the ARMA model is to set the order based on the order in which the ACF and PACF plots are significant, but directly setting the model at the 14 orders, the order is too high, the number of variables is too many, the model is too complex, and it is not easy to follow up the interpretation and prediction of the data. According to the general experience of economics problems, the model is temporarily set as ARMA (5,5). The results of the model based on ARMA (5,5) show information such as the estimated values of the coefficients of each order and the AIC values, where s.e. denotes the standard error of the coefficient estimates. In statistical analysis, it is generally required

that the estimates are greater than three times the standard error for the coefficients to be significant. From the results in the figure, it can be seen that most of the coefficients do not meet the requirement of significance, so it is necessary to reduce the ARMA order. After adjusting the order, ARMA (4,5), ARMA (5,4), ARMA (4,4), ARMA (3, 3) were tried and it was found that the coefficients of all orders of ARMA (3, 3) were significant.

In order to simplify the model, an attempt was made to find a model with lower order and that explains the data information better, and multiple models after downscaling were compared with ARMA (3, 3) to obtain the AIC value for each model, as shown in Table 1. According to Table 1, it can be seen that ARMA (3, 3) has the smallest AIC value and the model fits better.

Table 1. Table of AIC criteria for model fitting

Model	AIC
ARMA (3, 3)	-692
ARMA (3, 2)	-691
ARMA (2, 3)	-691
ARMA (2, 2)	-690

Therefore, this paper chose the ARMA (3, 3) model as the final mean-variance.

After the model is built, white noise test is done on the residuals of ARMA (3, 3) model, if the sequence is white noise indicates that there is still valuable information in the residual sequence that has not yet been extracted, and the model construction is not perfect. Box-Ljung test is done on the residual series to test whether they are uncorrelated or not. p-value is much greater than 0.05, the original hypothesis is accepted and the residual series is a white noise process. Further analyze the residual series and make its QQ plot, as shown in Fig. 3, it can be seen that the residual series does not obey the normal distribution, has a thick-tailed distribution, and appears to be left-skewed. Make the residual time series plot, as shown in Figure 4, the sequence appears fluctuation aggregation phenomenon, there may be conditional heteroskedasticity. From the above analysis, it can be seen that the residual series need to be tested for ARCH effect. The McLeod-Li test function is used to test the residual series, and the results are shown in Figure 5. p-value is much smaller than the level of significance, rejecting the original hypothesis, there is autocorrelation in the residual squared series, i.e., there is conditional heteroskedasticity in the series. In order to solve the phenomenon of conditional heteroskedasticity appearing in the residuals of the model, it is chosen to establish the volatility equation for the residuals with the GARCH model. In economics problems, generally GARCH (1,1) is sufficient to solve most of the problems. Therefore, the GARCH (1, 1) model is prioritized.

ARMA (3, 3) and GARCH (1, 1) are united, and the normality test is done on the united results, the P value is much smaller than the significance level, the original hypothesis is rejected, and the residuals of the model do not conform to the positive too distribution. After constructing the ARMA-GARCH model, the model adequacy is tested using the standardized residuals and their squared series. Figure 14 shows the time series of the model standardized residuals, which is initially judged to be a white noise process, and the ARMA-GARCH model is tested for the ARCH effect, which can be seen in Figure 6, and the model ARCH effect has been eliminated.

Based on the ARMA-GARCH model for interval prediction of log returns, it can be seen from Figure 6 that, except for some outliers, all the returns are within the 95% confidence interval, and the model predicts better.

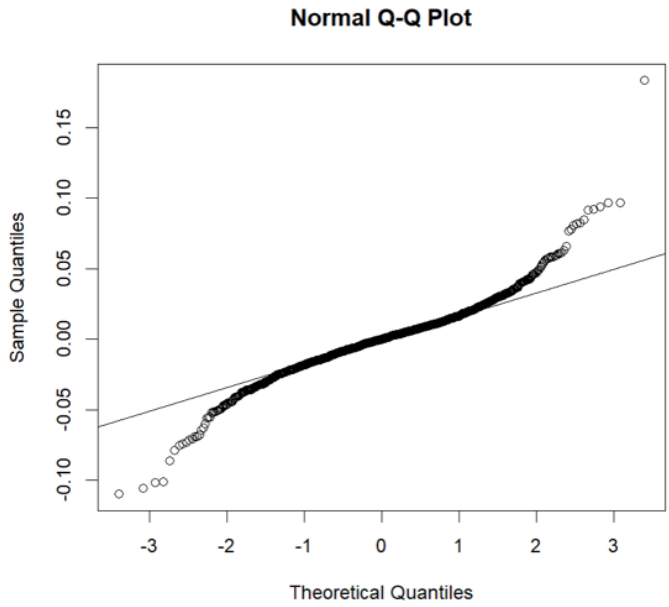

Figure 3. Residual QQ plot

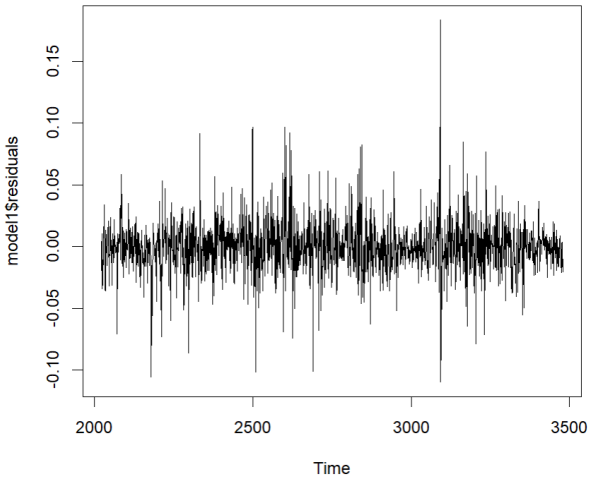

Figure 4. Residual timing diagram

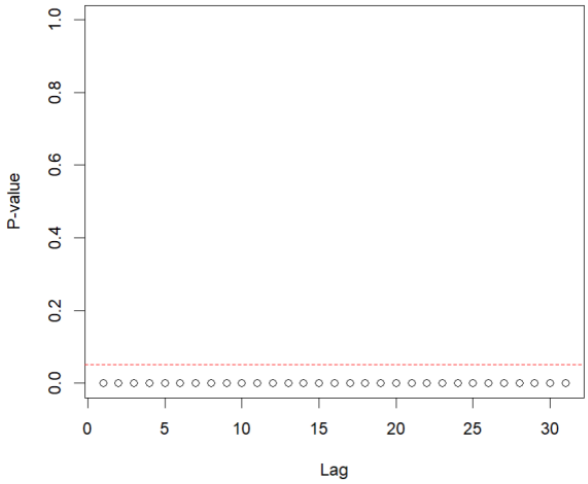

Figure 5. Residual sequence McLeod.Li.test test

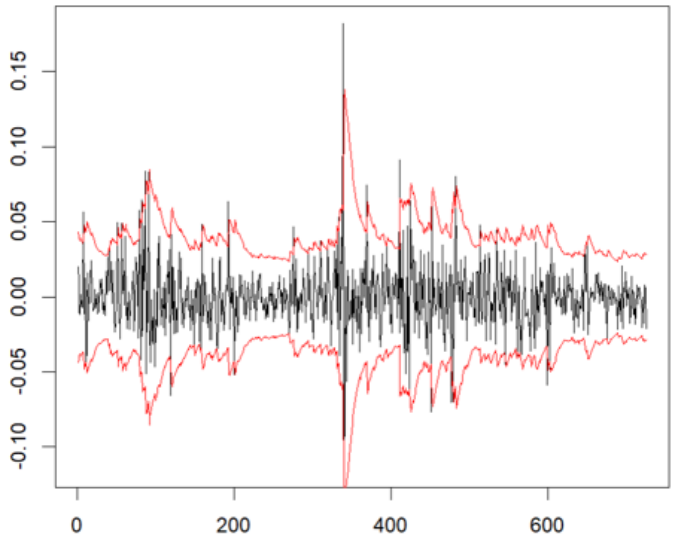

Figure 6. ARMA-GARCH model predictions

4. Conclusion

In this paper, the ARMA-GARCH model is used to analyze and forecast the daily stock return data of NIKKEI CHEMICAL (300214), an A-share of Shenzhen Stock Exchange venture stock, from May 2, 2017 to May 1, 2023. Firstly, the normality test is carried out for the phenomenon of sharp peaks and thick tails of the stock return distribution that often appears in financial time series, and it is found that the stock return shows a clustering effect by plotting the time series of the return, and then the smoothness test and heteroskedasticity test are carried out on the return data because of the establishment of the ARMA-GARCH model, which has three basic requirements for the residual series: zero-mean, purely random, and Heteroskedasticity, the ARCH test results show the existence of heteroskedasticity, which led to the establishment of the ARMA-GARCH model for forecasting return volatility, and the empirical results show that the selected ARMA-GARCH model has a good fit to the daily return time series.

In summary, the ARMA model has a better prediction ability for short-term prices, and the ARCH model can well eliminate conditional heteroskedasticity for volatile data such as stocks, and the combination of the two can make the model more consistent with the reality and provide a better reference for investors.

References

[1] LU Wenxing, REN Huanyu, LIANG Changyong, LI Keqing. Manufacturing PMI forecasting based on wavelet decomposition and ARIMA-GARCH-GRU combination model [J]. Industrial Engineering, 2024, 27 (01): 86 - 95+127.

[2] LIU Qi, SHI Sanzhi, LOU Lei, LIU Lu. Stock price prediction and decision making based on ARIAM-GARCH deep learning[J]. Journal of Changchun University of Science and Technology (Natural Science Edition),2024,47(01):119-130.

[3] Y.X. Zhang, Y. J. Zhang. A study on the accuracy of analyzing S&P 500 index forecasting using ARIMA-GARCH model [J]. Pivot Point, 2023, (S1): 103 - 105.

[4] LIU Zhao, JIA Meng, CHEN Dan, RAO Wenming. Analysis of passenger flow volatility of urban rail transit based on interval estimation[J]. Modern Urban Rail Transportation,2023, (09): 89 - 94.

[5] DONG Min, LOU Feng. Prediction of shared bicycle user volume based on SARIMA and ARIMA-GARCH models--Taking Harrow Outbound as an example [J]. Modern Business, 2023, (17): 46 - 49.

[6] YIN Jia, HUANG Xi, CHEN Xiang, CHEN Chen, CHEN Li, ZHANG Tao, XU Cheng, HUANG Yaping, GUO Pengcheng, WEN Hong. Prediction of safety risk intervals of sauced and marinated meat products based on WPD-ARIMA-GARCH combination model [J]. Food Science, 2024, 45 (03): 176 - 184.

[7] Wang Xiaorui, Wang Xiaorui. Application of ARFIMA-GARCH-LSTM model in the prediction of hand-foot-mouth disease incidence in Shanxi Province [D]. Shanxi Medical University, 2023.

[8] Xie YY. An empirical study on agricultural stock prices based on ARIMA-like model--Taking Lupin Gaoke as an example [J]. Hebei Enterprise,2023, (02): 40 - 42.

[9] Xiong Zheng, Che Wengang. Application of ARIMA-GARCH-M model in short-term stock forecasting [J]. Journal of Shaanxi University of Technology (Natural Science Edition), 2022, 38 (04): 69 - 74.

[10] Jia Qianying. Empirical analysis of domestic carbon financial transaction risk based on financial time series analysis [D]. Shandong University, 2022.

[11] Chen Le, Gao Wenke, Ji Hong, Zhang Lei. Prediction of plunger pump leakage based on HP filtering and ARIMA-GARCH model [J]. Journal of Agricultural Engineering, 2022, 38 (10): 61 - 67.