Frontiers in Business, Economics and Management
ISSN: 2766-824X | Vol. 13, No. 3, 2024

Stock Closing Price Prediction Based on the ARIMA-GARCH Model

Shang Li[1, *]

[1]College of William & Mary, Virginia, 23187-8795, United States of America
[*]Corresponding author's email: sli36@wm.edu

Abstract: Amidst a burgeoning stock market, a plethora of predictive models for stock prices have been steadily surfacing, with a wide array of Autoregressive models finding extensive application. This study sets out to evaluate the efficacy of the ARIMA-GARCH model in the domain of stock price prediction, taking as its dataset the stock information of Jinan Hi-Tech Development (600807). This paper begins by transforming the closing prices of Jinan Hi-Tech Development (600807) from July 13, 2023, to July 27, 2023, into a time series for the purpose of model fitting, identifying the ARIMA model parameters, and examining ARCH effects alongside the normality and independence of residuals. Subsequently, GARCH model parameters are discerned and integrated with the ARIMA model to establish the ARIMA-GARCH model, which is then subjected to residual testing. Concluding the study, a comparative error analysis between the predicted and actual closing prices reveals that the ARIMA-GARCH model boasts substantial accuracy in short-term stock price forecasting.

Keywords: ARIMA-GARCH; Stock Prediction; Time Series; Closing Price.

1. Introduction

With the advancement of the internet and technology, a myriad of investment avenues has blossomed. Among them, the stock market, one of the most well-established markets, has seen rapid growth, with buying stocks becoming a principal investment method for many. Consequently, the prediction of stock prices and the volatility of returns has garnered significant attention. To accurately forecast stock prices, the traditional econometric models have been set aside in favor of the ARCH model proposed by Engle in 1982.[1] After over four decades of evolution, the ARCH model and its derivative, the GARCH model, have been extensively applied. When combined with artificial neural networks, they provide more precise predictions. Accurate stock price predictions not only assist investors with risk management but also reflect the state of the economy, offering insights into the administration of the stock market.

2. Literature Review

There is a substantial body of research on stock price prediction, providing solid theoretical foundations and empirical evidence. Wu found that the traditional log-normal models struggled to fit the abnormal fluctuations in the stock prices of Beijing Tianqiao and proposed a source model for stock price volatility.[2] YANG et al. building on the classic ARCH model, fitted the intraday high-frequency trading data of China National Trade stocks and discovered that the confidence intervals obtained were more accurate than those from traditional econometric models.[3] They suggested that utilizing various extended models of ARCH could further improve prediction accuracy. Farman et al. applied GARCH (1,1) and FIGARCH (fractionally integrated GARCH) models to analyze the volatility of the Indian stock market over the period from 2008 to 2021, concluding the volatility models based on the GARCH equation were effective in yielding precise predictions, essential for investment portfolio distribution, assessment of financial performance, and pricing of financial options.[4] Chimrani et al. assessed stock price volatility on the Pakistan Stock Exchange from 2009 to 2016 using ARCH, GARCH, and EGARCH models. They found that the decision criteria Akaike Information Criterion (AIC) and Schwarz Criterion (SC) were most favorable for GARCH and EGARCH models, making them more suitable for stock price prediction. [5]Adebayo et al. utilized ARIMA (3,1,1) and ARIMA (1,1,4) models to predict the stock markets of Botswana and Nigeria, respectively, demonstrating that these models could predict future stock markets with reasonable accuracy in the short term.[6] Xu and Liang predicted the stock price of Yutong Bus using the ARIMA-GARCH time series model and conducted an error analysis, finding that this model fitted and predicted the closing stock price effectively. [7]TOMA identified volatility patterns using the GARCH model and predicted future stock prices for Microsoft Corporation using the ARIMA model based on historical data, assessing the effectiveness of the proposed method and showing very positive results compared to similar methods.[8]Qi et al. combined Gated Recurrent Unit (GRU) neural networks with Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) to predict the S&P 500 and CSI 300 stock indices with higher accuracy, providing insights for future implementation of more precise stock predictions combining neural networks.[9]

Conventional econometric methodologies, including regression analysis and time series analysis, generally presuppose steady variance in elucidating the volatility of stock market returns. However, a wealth of empirical research indicates that this assumption needs to be more comprehensive and objectively describe financial data. Moreover, economic time series data exhibit instability, with significant fluctuations in some periods and minor fluctuations in others, forming a phenomenon known as volatility clustering. Therefore, it is necessary to apply first or higher-order differencing to make non-stationary raw data stationary.

As a representative of emerging stocks, Jinan Hi-Tech Development's stock has attracted much attention since its listing. Hence, this paper will use the ARIMA-GARCH model to analyze the historical closing prices of Jinan Hi-Tech stocks and predict their future closing prices.

3. Theoretical Foundation of the ARIMA-GARCH Model

3.1. ARIMA Model

The Autoregressive Integrated Moving Average (ARIMA) model, introduced by Box and Jenkins in the early 1970s, is a method for forecasting time series data. [10]The model is comprised of three components: the Autoregressive (AR) model, the Differencing process (I), and the Moving Average (MA) model. The ARIMA model is constructed through a process where a non-stationary time series is converted into a stationary form. Following this transformation, the model predicts the dependent variable using only its previous values and the present and past values of the random error term.

$$Y_t = c + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \theta_1 \delta_{t-1} + \theta_2 \delta_{t-2} + \cdots + \theta_q \delta_{t-q} + \varepsilon \quad (1)$$

In the time series data this article is considering, denoted by Y_t, the parameters from φ_1 to φ_p belong to the AR model. These parameters delineate the connection between the present value and the values from the preceding p time points. The parameters from θ_1 to θ_q are part of the MA model, and they depict the association between the current value and the errors from the prior q time points. ε represents the error term, and c is a constant. This equation is based on the premise that the time series in question is stationary, thereby enabling the direct application of the AR and MA models. If the time series is non-stationary, this article must consider the 'I' component of the ARIMA model, which involves differencing the data.

3.2. GARCH Model

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, an advancement of the ARCH model, was formulated by Bollerslev in 1986. [11] The GARCH model is formed by adding a pth-order autoregressive term to the variance function, building upon the foundation of the ARCH model. It can effectively fit heteroskedasticity functions that exhibit long memory characteristics.

$$\sigma_t^2 = \omega + \alpha_1 \delta_{t-1}^2 + \alpha_2 \delta_{t-2}^2 + \cdots + \alpha_q \delta_{t-r}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \cdots + \beta_p \sigma_{t-s}^2 \quad (2)$$

The GARCH model is able to accurately simulate the changes in the volatility of time series variables, enabling a more precise grasp of risk (volatility), especially when applied in the theory of Value at Risk. The knowledge derived from these models is instrumental in setting prices and assessing which assets might yield greater returns. Furthermore, they contribute to forecasting the returns on existing investments, thus facilitating decisions related to asset allocation, hedging strategies, risk management, and portfolio optimization.

3.3. ARIMA-GARCH Model

Combining the ARIMA(p,d,q) model with the GARCH(s,r) model, Y_t follows:

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \cdots + \sum_{j=1}^{q} \beta_j \delta_{t-j} \quad (3)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{r} \alpha_i \delta_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \quad (4)$$

$$\delta_t = \sigma_t \varepsilon_t, \varepsilon_t \sim N(0,1) \quad (5)$$

4. Constructing the Stock Price Forecast Model

To forecast stock prices using the ARIMA-GARCH model, the following steps are typically involved:

(1) Conduct a correlation test on the time series data to check for autocorrelation between the data points.

(2) Difference the non-stationary series to stabilize it and then perform robustness checks.

(3) Split the data into training and test samples, determine the order of the ARIMA model, and identify its parameters.

(4) Test the residuals for independence and normality, and check for ARCH effects.

(5) Identify the parameters of the GARCH model using the residuals from the ARIMA model.

(6) Carry out diagnostic evaluations on the residuals of the developed ARIMA-GARCH model to ensure they conform to a white noise process.

(7) Use the constructed ARIMA-GARCH model to forecast the stock prices of Jinan Hi-Tech Development.

(8) Evaluate the model's performance.

5. Empirical Study

5.1. Data Source

This paper establishes a stock price forecasting model for the closing price data of Jinan High-Tech Development (600807) from January 4, 2021, to July 12, 2023, and selects the closing prices from July 13, 2023, to July 27, 2023, as the test data to forecast their closing prices.

5.2. Data Processing

5.2.1. Stationarity Test

The ADF (Augmented Dickey-Fuller) test for stationarity is conducted on the time series data.[12] If the series is found to be non-stationary, a differencing method is used to eliminate unit roots to achieve stationarity. From the ADF stationarity test depicted in Figure 1, it can be concluded that the original series is unstable. Therefore, the differencing method should be applied to process the original series.

```
        Augmented Dickey-Fuller Test

data:  t1
Dickey-Fuller = -2.3237, Lag order = 8, p-value = 0.4413
alternative hypothesis: stationary
```

Figure 1. ADF Test Results

This article performs first-order differencing on the original series, and upon testing, as shown in Figure 2, the first-order differenced time series is found to be stable.

```
        Augmented Dickey-Fuller Test

data:  dlnt1
Dickey-Fuller = -10.52, Lag order = 8, p-value = 0.01
alternative hypothesis: stationary
```

Figure 2. ADF Test Results after First-Order Differencing

5.2.2. Model Parameter Identification and Order Determination

Based on the AIC criterion, the smaller the AIC value, the more accurate the model fit. [13]This article uses R language for parameter identification and order determination. As shown in Figure 3, the model with the highest degree of fit is ARIMA (0,1,3).

```
ARIMA(0,1,3)

Coefficients:
         ma1      ma2      ma3
      0.0820  -0.0401  -0.1873
s.e.  0.0406   0.0396   0.0473

sigma^2 = 0.01176:  log likelihood = 488.6
AIC=-969.2   AICc=-969.13   BIC=-951.56
```

Figure 3. ARIMA Model Fitting and Automatic Parameter Optimization Results

For the established ARIMA(0,1,3) model, this article conducts an ARCH effect test on its standardized residuals. The results, as shown in Figure 4, indicate that the model exhibits ARCH effects, suggesting the suitability of proceeding with a GARCH model implementation.

```
        ARCH LM-test; Null hypothesis: no ARCH effects

data:  train_data
Chi-squared = 561.18, df = 12, p-value < 2.2e-16
```

Figure 4. The ARCH LM-Test Results

When determining the order of the GARCH model, while the GARCH(1,2) model has the smallest AIC value, the difference compared to the GARCH(1,1) model is negligible. Moreover, having a higher number of parameters in the GARCH model can lead to increased instability in the model. Therefore, this article opts for the GARCH(1,1) model.

5.2.3. Model Fitting and Estimation

Using the ARIMA(0,1,3)-GARCH(1,1) model, this article fitted the data and estimated parameters, as illustrated in Figure 5.

The p-values for all these coefficients are less than 0.001, demonstrating a robust fit for the model.

	Estimate	Std. Error	t value	Pr(>|t|)
mu	2.833881	0.000045	6.3598e+04	0
ma1	0.513601	0.019621	2.6176e+01	0
archm	2.538482	0.000354	7.1778e+03	0
omega	0.000386	0.000060	6.3907e+00	0
alpha1	0.152559	0.003596	4.2430e+01	0
beta1	0.846441	0.000004	2.4023e+05	0
skew	0.611063	0.022162	2.7573e+01	0
shape	7.139228	0.955051	7.4752e+00	0

Figure 5. ARIMA-GARCH Model Fitting Results

5.2.4. Residual Autocorrelation and Normality Test

The ARIMA-GARCH model was subjected to tests for residual autocorrelation and independence. As indicated in Figures 6 and 7, the residuals satisfy the conditions for autocorrelation and normality tests.

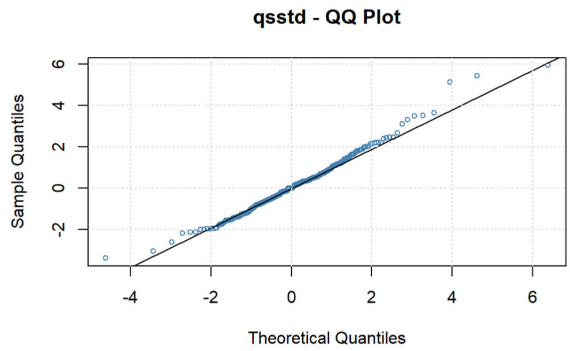

Figure 6. Standardized Sample Data QQ Plot

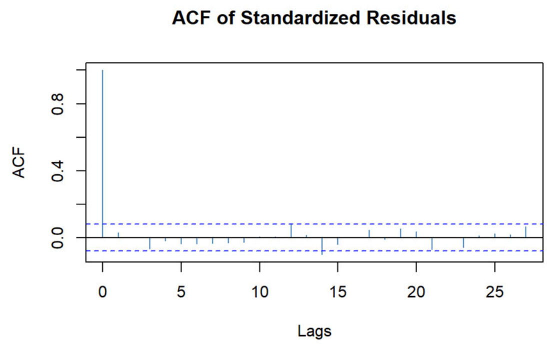

Figure 7. ACF of Standardized Residuals

5.2.5. Forecasting Results

The established ARIMA(0,1,3)-GARCH(1,1) model was utilized to predict the closing stock prices for the next 15 days. The outcomes are presented in Figure 8.

	Series	Sigma
T+1	3.007	0.06616
T+2	3.009	0.06898
T+3	3.016	0.07169
T+4	3.023	0.07430
T+5	3.029	0.07682
T+6	3.035	0.07926
T+7	3.041	0.08162
T+8	3.047	0.08392
T+9	3.053	0.08615
T+10	3.058	0.08832
T+11	3.063	0.09044
T+12	3.069	0.09250
T+13	3.074	0.09452
T+14	3.079	0.09650

Figure 8. Forecasting Results

5.2.6. Accuracy Evaluation

Upon assessing the model's accuracy, as illustrated in Figure 9, the model's forecasted values exhibit an average deviation of approximately 10.34% from the actual values. This deviation being a positive number may indicate a slight tendency of the model to overestimate. Typically, the magnitude of the forecasting error is about 12.42%. The mean prediction error of the model stands at 3.23%. A positive Mean Percentage Error (MPE) suggests that the model might systematically overestimate the actual values.

From these data, it is discernible that the trained ARIMA-GARCH model is capable of forecasting the closing prices of Jinan Hi-Tech with a fair degree of accuracy.

Table 1. Accuracy Evaluation Results

	ME	RMSE	MAE	MPE
Test Set	0.1034146	0.1242084	0.1034146	3.229931

6. Conclusion

By employing the ARIMA-GARCH model to fit the time series of the closing stock prices of Jinan Hi-Tech and to predict future closing prices, followed by an error analysis, it is evident from the forecast results that the ARIMA-GARCH model is capable of predicting short-term stock closing prices with a relatively high degree of precision.

References

[1] Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50(4), 987–1007.

[2] Wu, W., & Wu, C. (2000). Discussion on Stock Price Fluctuation Models. Systems Engineering Theory and Practice, 2000(04), 63-69.

[3] Yang, K., Ma, Y., Zhang, X., et al. (2019). Empirical analysis of high-frequency stock trading data based on ARCH(q) model. Jilin Normal University Journal (Natural Science Edition), 40(02), 68-72.

[4] Ali, F., Suri, P., Kaur, T. (2022). Modelling time-varying volatility using GARCH models: evidence from the Indian stock market. F1000 Research, 11, 1098.

[5] Chimrani, C. R., Ahmed, F., & Panjwani, V. K. (2018). Modeling Sectoral Stock Indexes Volatility: Empirical Evidence from Pakistan Stock Exchange. International Journal of Economics and Financial Issues, 8(2), 319-324.

[6] Adebayo, F. A., Sivasamy, R., & Shangodoyin, D. K. (2014). Forecasting Stock Market Series with ARIMA Model. Journal of Statistical and Econometric Methods, 3(3).

[7] Xu, S., & Liang, X. (2019). Research on Stock Price Prediction Based on ARIMA-GARCH Model. Journal of Henan Institute of Education (Natural Science Edition), 28(04), 20-24.

[8] Toma, L. R. (2023). Exploring the Effectiveness of ARIMA and GARCH Models in Stock Price Forecasting: An Application in the IT Industry. Informatica Economica, 27(3), 61-72.

[9] Qi, C., Ren, J., & Su, J. (2023). GRU Neural Network Based on CEEMDAN–Wavelet for Stock Price Prediction. Applied Sciences, 13(12), 7104.

[10] Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.

[11] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. Journal of Econometrics, 31(3), 307-327.

[12] Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. Econometrica, 55(2), 251-276.

[13] Sugiura, N. (1978). Further Analysis of the Data by Akaike’s Information Criterion and the Finite Corrections. Communications in Statistics—Theory and Methods A,7, 13-26.