Application of ARIMA-GARCH-S model to short-term stock forecasting

Yi Wang

Tandon school of engineering, New York Univeristy, Brooklyn, New York, 11201, USA yw7742@nyu.edu

ABSTRACT:

The performance of the stock market in the financial domain profoundly impacts the economic well-being of many individuals. Thus, accurately predicting stock prices is an essential task. Although traditional financial time series models such as ARIMA and GARCH play a crucial role in predictions, they may fail to capture all the market dynamics. This study explores a composite model combining ARIMA, GARCH, and Stacking techniques (ARIMA-GARCH-S) to enhance the accuracy of predictions. The research data are derived from the stock closing price time series data of "Amazon" from June 29, 2020, to April 12, 2022, with 653 entries and "Caterpillar" from February 8, 2021, to October 21, 2022, with 621 entries. The model's fitting performance is evaluated by comparing the fitting residual plots and variance graphs, while predictive performance is determined by comparing the MAPE, RMSE, and EC statistical metrics. The results indicate that the ARIMA-GARCH-S composite model has a significant predictive advantage over the ARIMA model. This finding not only offers a new avenue for model innovation but also provides financial market participants with a more precise and stable prediction tool, aiding them in making more informed investment decisions.

KEYWRODS: Stock Market Forecasting; ARIMA-GARCH-Stacking Model; Time Series Analysis; Fitted Residuals; MAPE/RMSE/EC Indicators

1. Introduction

The stock market has always been the focus of financial research and investment strategies. Since its inception, the stock market has not only been a place for the exchange of capital but also a nexus for information, expectations, and market confidence. For most people, the performance of the stock market directly or indirectly affects their financial health, including retirement funds, education savings, or other long-term investments. Therefore, predicting stock prices and market trends has always been a core issue of concern for investors, scholars, and policymakers. Enhancing the accuracy of predictions for stock closing price time series data is not only vital for investors and financial institutions but also profoundly impacts the daily lives of ordinary people. Financial time series models have a long history of application and are often considered classic methods in this domain. However, with the increasing complexity of financial markets, a single prediction model might no longer adequately capture market dynamics. Improving the accuracy of predicting stock closing price time series data has become an urgent issue. This study aims to explore the integration of existing models in search of higher predictive accuracy. Through the developed composite model, we hope to provide a new perspective and method for time series forecasting.

2. Literature Review

Zhang Yingchao et al. (2019) used the Shanghai Stock Exchange Index data and applied the ARIMA(4,1,4) model for predictions, demonstrating that the ARIMA model could provide accurate forecasts in the short term for the financial forecasting field[1]. Ariyo et al. (2014) predicted stock prices for Nokia and Zenith Bank based on the ARIMA model, achieving satisfactory forecast results, with a regression standard error of 3.5808 for Nokia and 0.7872 for Zenith Bank[2]. Hassan Mohammadi et al. (2010) employed various GARCH models to study the behavior of oil returns and their conditional variance, emphasizing the advantages of the MA(1)–EGARCH(1,1) and MA(1)-APARCH(1,1) models in out-of-sample predictions[3]. Ray Chou (1988) used the GARCH(1,1) estimation technique to study the persistence of stock return volatility, proving that stock return volatility has persistent characteristics[4]. Yu Yaning et al. (2018) and Xu Shuya et al. (2019) respectively used the ARIMA-GARCH model to predict stock prices for Guizhou Maotai and Yutong Bus, both proving that compared to a single model, the ARIMA-GARCH model has superior predictive performance[5][6]. Farah Hayati Mustapa et al. (2019) predicted the S&P 500 index stock prices, finding the ARIMA(2,1,2)-GARCH(1,1) model most suitable, which showed higher accuracy in dynamic predictions[7]. Bohdan Pavlyshenko et al. (2018) utilized stacking techniques to combine models and forecast sales of over 1000 products covered by more than 800,000 stores, proving that combining models using stacking significantly enhances forecasting accuracy[8]. Building on this foundation, this study introduces the stacking concept into the ARIMA-GARCH composite model. While both ARIMA and GARCH models have applications in financial research, combining the two and then introducing stacking to establish a meta-model for time series analysis is relatively rare. This composite approach captures market volatility and trends from different angles, and by using deep learning concepts, it can filter out the optimal combination pattern, further enhancing the model's predictive performance. This research will use stock closing price time series data from listed companies on Yahoo Finance (2000-2020). To ensure the comprehensiveness of the study, two groups of listed company stock closing price data will be randomly selected, and two independent experiments will be conducted.

3. Method

3.1 Principle of the ARIMA Model:

The ARIMA model, an acronym for "Autoregressive Integrated Moving Average Model," is a commonly used model in time series analysis. Introduced in the early 1970s by Box and Jenkins, this renowned method for forecasting time series is also known as the Box-Jenkins model. It combines elements of Autoregressive (AR) models, Integration (I) - representing differencing - and Moving Average (MA) models. An ARIMA(p,d,q) model is formulated as:

$$X_t^{(d)} = c + \sum\nolimits_{i=1}^{p} \varphi_i X_{t-i}^{(d)} + \epsilon_t + \sum\nolimits_{j=1}^{q} \theta_j \epsilon_{t-j}$$

Where pdq denote the order of the autoregressive, differencing, and moving average parts of the model, respectively, $X_t^{(d)}$ represents the time series after d times of differencing,[c is a constant,[ϵ_t is the random error term at time t,[φ_i and $\theta_j$correspond to the coefficients of the AR and MA parts, respectively.

In essence, the AR part predicts current observations using past values, the MA part uses past forecast errors, and the I part stabilizes the time series through differencing.

This study initially applies the ARIMA model for preliminary fitting of stock closing prices. Through this model, we aim to capture the main trends, patterns, autocorrelations, and seasonalities present in the time series.

3.2 Principle of the ARCH Model:

The ARCH (Autoregressive Conditional Heteroskedasticity) model, introduced by Engle in 1982, is designed to account for conditional heteroskedasticity within financial time series. The core idea is that the volatility at a given time is based on the shocks from previous periods. This implies that a large shock in the past indicates potentially greater volatility in the future and vice versa. This model structure allows for the capture of volatility clustering commonly observed in financial time series, where periods of high volatility tend to be followed by high volatility, and periods of low volatility follow low volatility. Mathematically, it is expressed as:

$r_t = \mu + \epsilon_t$ and $\sigma_t^2 = \alpha_0 + \sum\nolimits_{i=1}^{q} \alpha_i \epsilon_{t-i}^2$

Where r_t represents the return at time t, ϵ_t is the shock,[σ^2 is the conditional variance

The parameters of the model are estimated using the Maximum Likelihood Estimation method.

3.3 Principle of the GARCH Model:

The GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model is a robust tool for characterizing volatility in financial time series. It is a model for describing and predicting the phenomenon of conditional heteroskedasticity in time series. It extends the ARCH model by assuming that current variance depends not only on the squared errors from the previous q periods but also on the variances from the previous p periods, thus capturing the phenomenon of volatility clustering more effectively. The GARCH model is formed by adding lagged squared residuals to the ARCH model:

$$\sigma_t^2 = \alpha_0 + \sum\nolimits_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum\nolimits_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

Where $\sum\nolimits_{i=1}^{q} \alpha_i \epsilon_{t-i}^2$ represent the ARCH terms, $\sum\nolimits_{j=1}^{p} \beta_j \sigma_{t-j}^2$ represent the GARCH terms, $\beta_j$is the coefficient of the GARCH term, which represents the effect of the variance in period t-j on the variance at time t. Overall, the GARCH model describes how the volatility of a time series depends on past errors and past volatility. This model is particularly useful in the analysis of financial time series since the returns on financial assets often exhibit volatility clustering which means large fluctuations tend to follow large fluctuations and small fluctuations tend to follow small fluctuations..

Upon the forecasts generated by the ARIMA model, this study employs the GARCH model to fit the residuals of the ARIMA model, which helps capture the underlying volatility in the time series and enhances the precision of the forecasts.

3.4 Principle of Stacking Technique:

Stacking is an ensemble learning technique where multiple models are used as base models to construct a meta-model. This meta-model learns how to best combine the predictions of the base models to further improve forecast accuracy. In this study, we first train the ARIMA and GARCH models independently to obtain forecast values for the time series:

$$\hat{y}_{ARIMA} = ARIMA(X_t)$$

$$\epsilon_{ARIMA} = X_t - \hat{y}_{ARIMA}$$

$$\hat{\epsilon}_{GARCH} = GARCH(\epsilon_{ARIMA})$$

Where X_t is the time series data used for training, ϵ_{ARIMA} represents the fitting residuals from the ARIMA model.

Subsequently, the predictions from ARIMA and the forecasts of residuals from GARCH are combined and analyzed through linear regression to act as a meta-model:

$$\hat{y}_{meta} = \alpha + \beta_1 \hat{y}_{ARIMA} + \beta_2 \hat{\epsilon}_{GARCH} + \epsilon$$

Where α is the intercept, β_1 and β_2 are regression coefficients, ϵ denotes the error term.

By minimizing the sum of squared residuals, the values for α, β_1, β_2 are estimated. The constructed meta-model selects the optimal linear combination of the ARIMA and GARCH models to further improve the model's precision and performance.

4. Experiments

4.1 Experimental Procedure:

(1) Randomly select a segment of stock closing price time series data for a particular company from a database.

(2) Perform the Ljung-Box test on the data for white noise. If the stock series is a white noise series, it indicates that the data set has no statistical predictive value, and a new set of data must be chosen.

(3) If the stock series is not a white noise series, indicating that the data has statistical predictive value, continue with the Durbin-Watson (DW) test for autocorrelation. If the stock data shows weak autocorrelation, the data set is not suitable for building an ARIMA model, and another data set must be selected.

(4) If the stock data exhibits sufficient autocorrelation, an ARIMA model can be constructed for analysis, and the experimental process continues.

(5) Divide the data into a prediction set and a training set; the prediction set is for evaluation, and the training set is for training the model.

(6) Conduct an Augmented Dickey-Fuller (ADF) test on the stock series to check for stationarity. If the series is not stationary, differencing is applied until it becomes stationary.

(7) Select the optimal ARIMA model order (p, q) by using the Akaike Information Criterion (AIC), with d confirmed by the ADF test mentioned above.

(8) Construct the ARIMA model and train it with the training set to obtain the forecast results and fitting residuals of the ARIMA model.

(9) Establish a GARCH model based on the fitting residuals of the ARIMA model, and use the AIC to select the most appropriate GARCH model order.

(10) Combine the fitting results of the ARIMA and GARCH models as base models and input them into the meta-model established by Stacking to perform regression analysis, thus obtaining the optimal combination method for the ARIMA and GARCH models and, consequently, the ARIMA-GARCH-S model.

(11) Obtain the forecast results and fitting residuals of the ARIMA-GARCH-S model.

(12) Perform a white noise test on the fitting residuals of the ARIMA-GARCH-S model. If the result is not white noise, further optimization of the model is needed. If the result is white noise, proceed to evaluate and compare the performance of the ARIMA model and the ARIMA-GARCH-S model.

(13) Repeat the above steps to complete two sets of independent experiments.

4.2 Experiment 1:

4.2.1 Original Data

Randomly selected from Yahoo Finance, the stock closing price time series data for "Amazon" from June 29, 2020, to April 12, 2022, includes a total of 653 entries. The original time series data is shown in Figure 1.

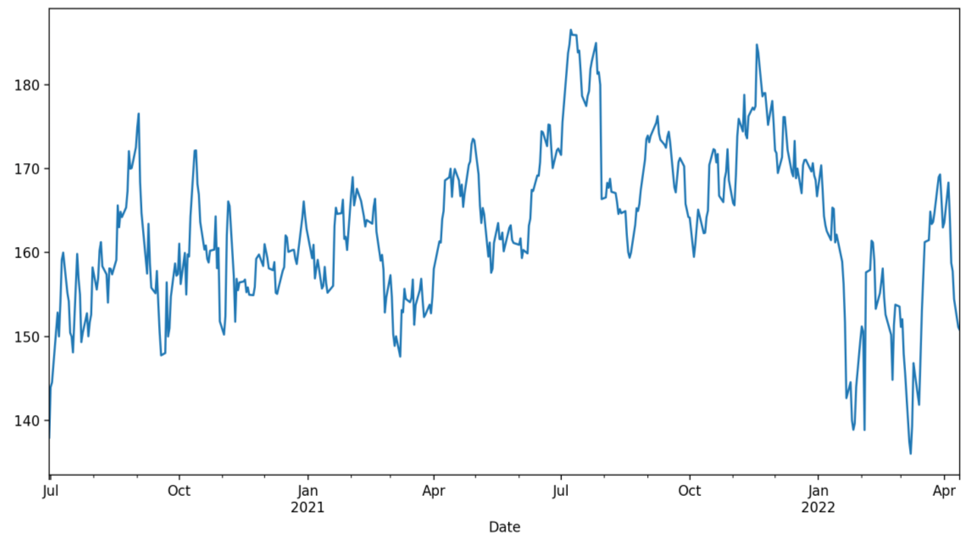

Figure 1. Amazon 2020/6/29-2022/4/12 Stock Closing Price Time Series Data

4.2.2 Data Processing

First, perform the Ljung-Box test for white noise on the stock series. Assuming the time series is white noise, the Ljung-Box test result P-Values = 0.000000, $P < 0.05$, rejects the null hypothesis, indicating that the stock series is not white noise and has statistical analytical value.

Continue with the DW test on the stock series. The DW test statistic is 1.075158, indicating a relatively strong negative autocorrelation, thus an ARIMA model can be established for analysis.

Finally, conduct an ADF test for stationarity on the stock series. Assuming the time series is non-stationary. Without differencing, the ADF test statistic is -4.036741, with a P-Value of 0.001230, $P < 0.05$, rejecting the null hypothesis, indicating that the time series is already stationary and meets the requirements for constructing an ARIMA model.

4.2.3 Dataset Division

Divide the stock series data, using the first 649 entries for the training set to train the model, and the last 4 entries for the prediction set to evaluate the model's performance.

4.24 ARIMA Model Parameter Selection

With d = 0 confirmed by the ADF test, construct several ARIMA models with orders ranging from (1,5), and select the optimal ARIMA model order (p, q) by comparing the Akaike Information Criterion (AIC). The optimal model confirmed is ARIMA(3,0,4), and the AIC results for each model are shown in Table 1.

Table 1. AIC values for ARIMA models of different orders

ARIMA order (p,q)	AIC Value	ARIMA order (p,q)	AIC Value
(3, 4)	2.155569	(4, 1)	5.803636
(5, 2)	2.477351	(2, 3)	5.891981
(1, 2)	3.096501	(2, 5)	6.395334
(3, 5)	3.560064	(4, 4)	6.528158
(3, 1)	3.954647	(1, 4)	7.050423
(5, 3)	3.963581	(2, 4)	7.198567
(1, 1)	3.984731	(1, 5)	7.272507
(1, 3)	5.080768	(2, 1)	7.468202
(2, 2)	5.082294	(5, 5)	7.503022
(5, 4)	5.365728	(5, 1)	7.700762
(4, 5)	5.567539	(3, 3)	7.783244
(3, 2)	5.797856	(4, 2)	8.390644

4.25 ARIMA Model Construction and GARCH Model Parameter Selection

For the obtained ARIMA(3,0,4) model, use the training set data to train the ARIMA model and acquire its fitting residual time series data. Use the residuals to establish a GARCH model, similarly using AIC as the criterion to select the most appropriate GARCH model order. The optimal model obtained is GARCH(1,5).

4.26 Construction of ARIMA-GARCH-S Model

The fitting results of the ARIMA and GARCH models are used as the base models, input into the meta-model established by Stacking, and regression analysis is conducted to derive the optimal combination of the ARIMA and GARCH models. The coefficients obtained for the final model are

β_1 = -5.02686, β_2 = 0.47987, α = 926.39523

The ARIMA-GARCH-S model obtained is:

$$\hat{y}_{meta} = 926.39523 - 5.02686\hat{y}_{ARIMA} + 0.47987\hat{\epsilon}_{GARCH} + \epsilon$$

4.27 Model Prediction

Use the established ARIMA(3,0,4) and ARIMA-GARCH-S models to predict Amazon's stock closing price for the subsequent four days. Compare the results with the prediction set data to get the preliminary residuals. The specific results are shown in Table 2.

Table 2. results of ARIMA(3,0,4) and ARIMA-GARCH-S models prediction

	Day 1	Day 2	Day 3	Day 4
Actual	153.34766	152.23483	151.12199	150.78751
ARIMA(3,0,4)	154.27303	154.21621	154.19899	154.19378
ARIMA-GARCH-S	153.33748	152.26185	151.06718	150.82547
Residuals of ARIMA-GARCH-S	0.01018	-0.02702	0.05481	-0.03796

Performing the Ljung-Box white noise test on the ARIMA-GARCH-S model gives a result of P-Values = 0.999816, P > 0.05, which means the null hypothesis cannot be rejected, indicating that the stock series is white noise and that all underlying patterns has already been successfully captured, so there is no need for further optimization of the model.

4.28 Accuracy Comparison and Analysis

To compare the fitting effects and model performance of ARIMA(3,0,4) and ARIMA-GARCH-S models in multiple dimensions, this study first used residual time series plots, residual frequency distribution histograms, and variance plots to visually compare the fit and fitting residuals of the two models.

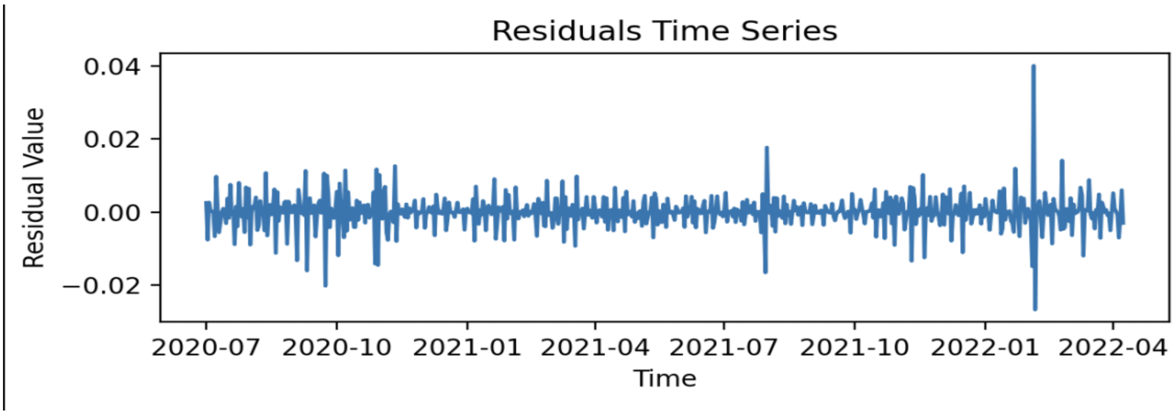

Figure 2. ARIMA-GARCH-S fitted residual time series plot

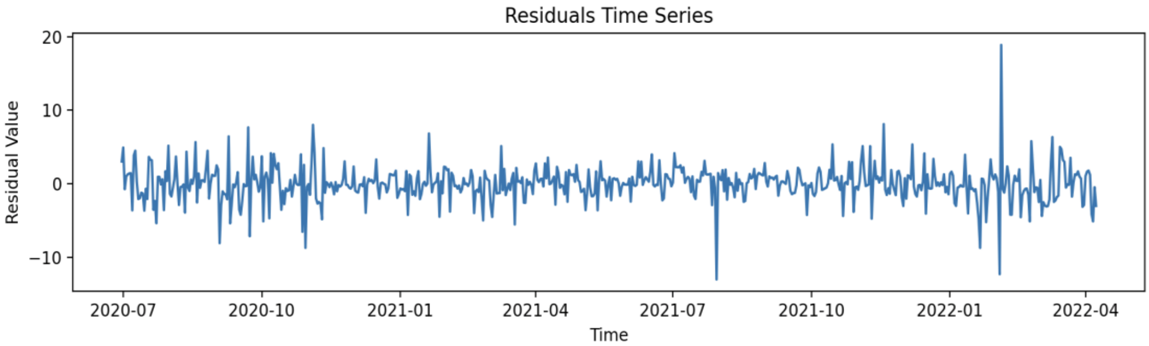

Figure 3. ARIMA(3,0,4) fitted residual time series plot

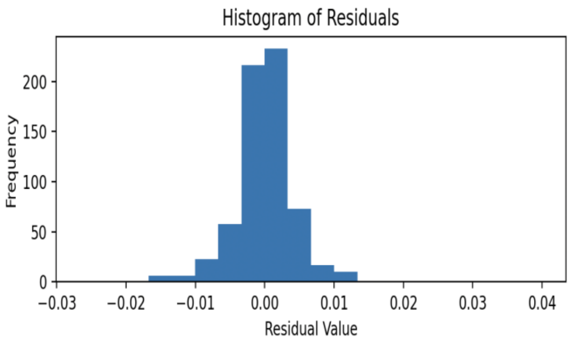

Figure 4. Histogram of the frequency distribution

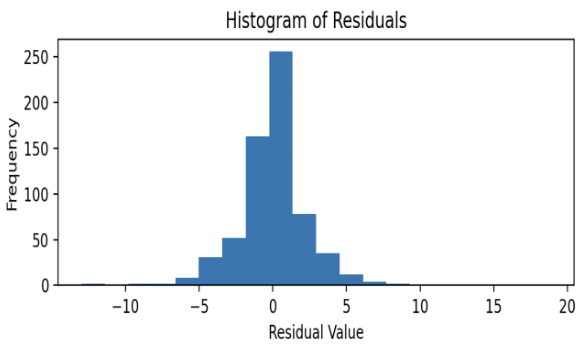

Figure 5. Histogram of the frequency distribution of the fitted residuals of ARIMA-GARCH-S of the fitted residuals of ARIMA(3,0,4)

From the time series plot of the residuals and the histogram of the frequency distribution of the residuals, it can be clearly seen that the fitted residuals of the ARIMA-GARCH-S model are much smaller than the fitted residuals of the ARIMA model, which is about two orders of magnitude less, and thus the fitting effect of the ARIMA-GARCH-S model is better than that of the ARIMA model.

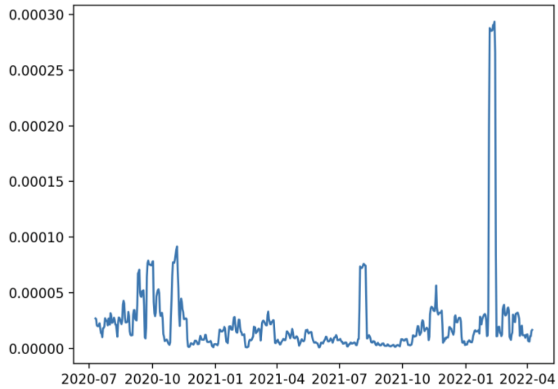

Figure 6. Variance plots of ARIMA-GARCH-S

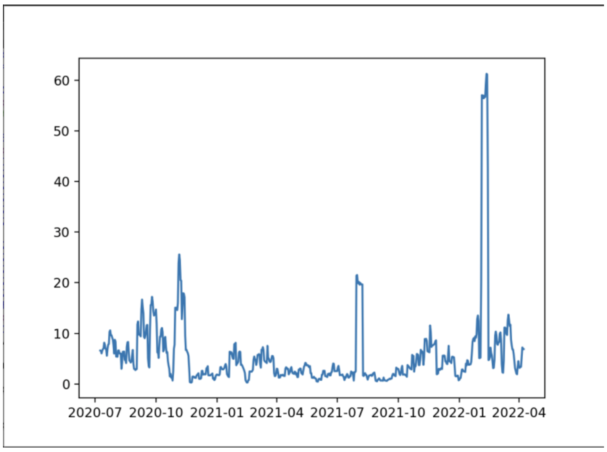

Figure 7. Variance plots of ARIMA(3,0,4)

By comparing the variance plots, it can be clearly seen that the ARIMA-GARCH-S model has a smaller degree of bias and a smaller range of fluctuation than the ARIMA model, and therefore the fitting performance of the ARIMA-GARCH-S model is better than the ARIMA model.

In order to further examine the predictional performance and accuracy of the models, this study quantitatively compares the models' prediction residuals using the Mean

Absolute Percentage Error(MAPE), Root Mean Square Error(RMSE) and the Equal Coefficient(EC). The formula for calculating the statistics is as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{At}\right| \times 100\%$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2}$$

$$EC = \frac{\sqrt{\sum_{t=1}^{n}(F_t - A_t)^2}}{\sqrt{\sum_{t=1}^{n}A_t^2} + \sqrt{\sum_{t=1}^{n}F_t^2}}$$

Where n is the number of observations, A_t is the t-th actual value, F_t is the t-th predicted value

The final calculations are shown in Table 3.

Table 3. Predictive test of the closing price of the stock "Amazon" for 4 days

test statistic	ARIMA(3,0,4)	ARIMA-GARCH-S
MAPE	1.55015%	0.02146%
RMSE	2.54228	0.03633
EC	0.94028	0.99999

According to the test results, in the 4-day stock closing price prediction, the ARIMA model has a MAPE of 1.55015%, an RMSE of 2.54228, and an EC of 0.94028. The ARIMA-GARCH-S model has a MAPE of 0.02146%, an RMSE of 0.03633, and an EC of 0.99999.This shows that the ARIMA -GARCH-S model's prediction error as a whole is substantially smaller than that of the ARIMA model, which is about two orders of magnitude, so the prediction results and model performance of the ARIMA-GARCH-S model are better than that of the ARIMA model.

4.3 Experiment 2:

4.3.1 Original Data

Randomly selected from Yahoo Finance, the stock closing price time series data for "Caterpillar" from February 8, 2021, to October 21, 2022, includes a total of 621 entries. The original time series data is shown in Figure 8.

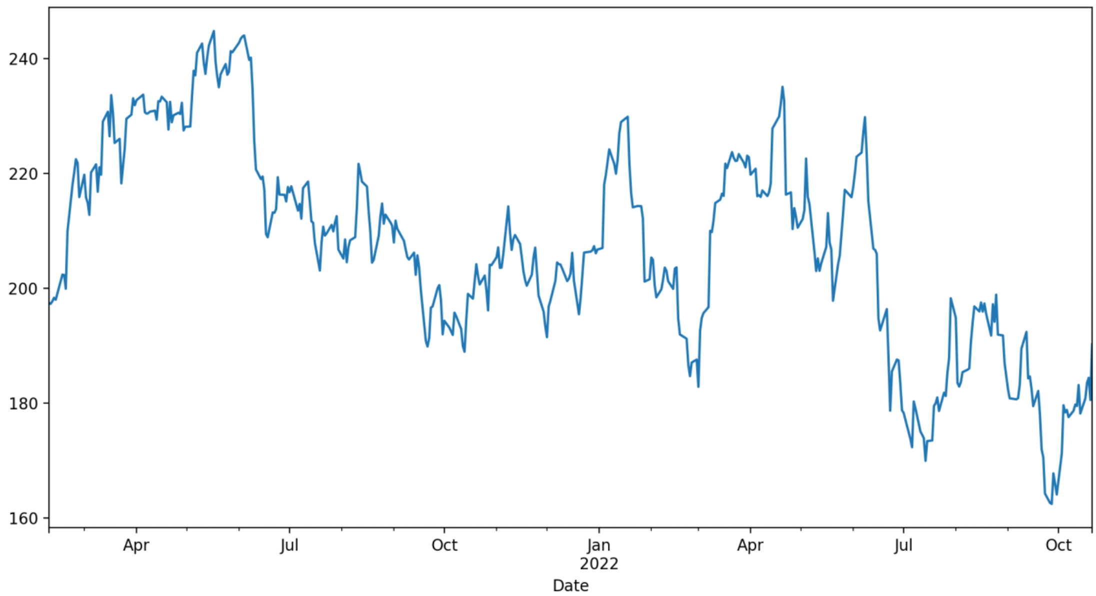

Figure 8. Caterpillar 2021/2/8-2022/10/21 Stock Closing Price Time Series Data

4.3 2 Data Processing

First, perform the Ljung-Box test for white noise on the stock series. Assuming the time series is white noise, the Ljung-Box test result P-Values = 0.000000, $P < 0.05$, rejects the null hypothesis, indicating that the stock series is not white noise and has statistical analytical value. Continue with the DW test on the stock series. The DW test statistic is 1.18376, indicating a relatively strong negative autocorrelation, thus an ARIMA model can be established for analysis.

Finally, conduct an ADF test for stationarity on the stock series. Assuming the time series is non-stationary. Without differencing, the ADF test statistic is -2.294219, with a P-Value = 0.173839,[P > 0.05, fail to reject null hypothesis, indicating that the time series is non-stationary and needs to be differentiated.

After the time series was first-order differenced, the ADF test was performed again, and the ADF test statistic is -22.390577, with a P-Value of 0.000000, P < 0.05, rejecting the null hypothesis, indicating that the time series is now stationary and meets the requirements for constructing an ARIMA model. The time series after the first-order differencing is shown in Figure 9.

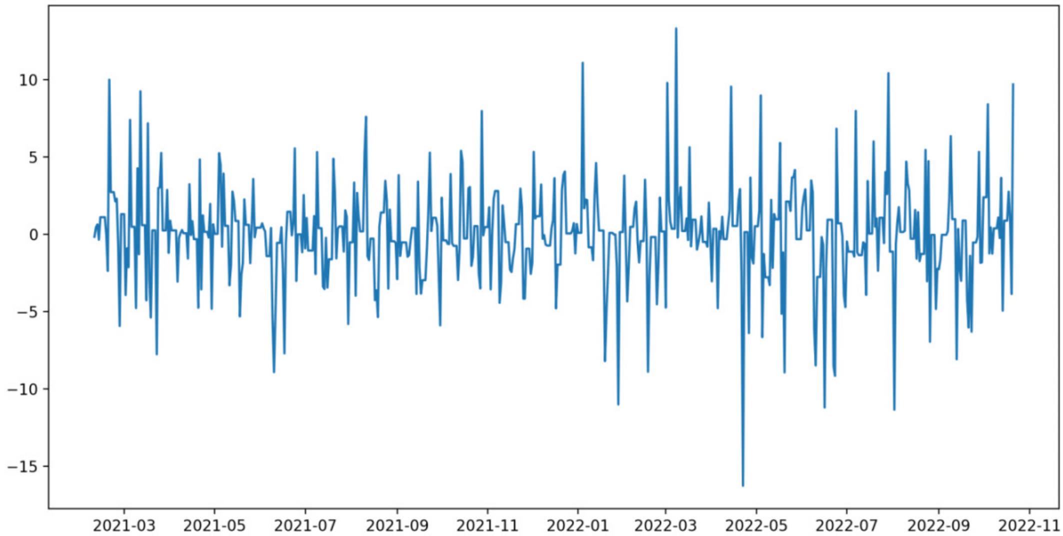

Figure 9. Caterpillar 2021/2/8-2022/10/21 First Order Differential Stock Closing Price

4.3.3 Dataset Division

Divide the stock series data, using the first 617 entries for the training set to train the model, and the last 4 entries for the prediction set to evaluate the model's performance.

4.34 ARIMA Model Parameter Selection

With d = 1 confirmed by the ADF test, construct several ARIMA models with orders ranging from (1,5), and select the optimal ARIMA model order (p, q) by comparing the Akaike Information Criterion (AIC). The optimal model confirmed is ARIMA(3,1,2)

4.35 Construction of ARIMA-GARCH-S Model

For the obtained ARIMA(3,1,2) model, use the training set data to train the ARIMA model and acquire its fitting residual time series data. Use the residuals to establish a GARCH model, similarly using AIC as the criterion to select the most appropriate GARCH model order. The optimal model obtained is GARCH(1,1). The fitting results of the ARIMA and GARCH models are used as the base models, input into the meta-model established by Stacking, and regression analysis is conducted to derive the optimal combination of the ARIMA and GARCH models.

4.36 Model Prediction

Use the established ARIMA(3,1,2) and ARIMA-GARCH-S models to predict Caterpillar's stock closing price for the subsequent four days. Compare the results with the prediction set data to get the preliminary residuals. The specific results are shown in Table 4.

Table 4. results of ARIMA(3,1,2) and ARIMA-GARCH-S models prediction

	Day 1	Day 2	Day 3	Day 4
Actual	183.55000	184.41000	180.53999	190.22000
ARIMA(3,1,2)	182.00733	183.64589	184.75632	186.79243
ARIMA-GARCH-S	183.28622	184.41951	180.87853	190.13575
Residuals of ARIMA-GARCH-S	0.26378	-0.00951	-0.33854	0.08425

Performing the Ljung-Box white noise test on the ARIMA-GARCH-S model gives a result of P-Values = 0.998972, P > 0.05, which means the null hypothesis cannot be rejected, indicating that the stock series is white noise and that all underlying patterns has already been successfully captured, so there is no need for further optimization of the model.

4.37 Accuracy Comparison and Analysis

MAPE, RMSE and EC are calculated for ARIMA(3,1,2) and ARIMA-GARCH-S models and the final results are presented in Table 5.

Table 5. Predictive test of the closing price of the stock "Caterpillar" for 4 days

test statistic	ARIMA(3,1,2)	ARIMA-GARCH-S
MAPE	1.71867%	0.09513%
RMSE	2.84997	0.21873
EC	0.91792	0.99963

It can be seen that the overall prediction error of the ARIMA-GARCH-S model is substantially smaller than that of the ARIMA model, so the prediction results and model performance of the ARIMA-GARCH-S model are better than those of the ARIMA model.

5. ARIMA-GARCH-S Model Strengths and Potential Application

5.1 Advantages analysis:

The ARIMA-GARCH-S model showcases a set of notable strengths, as evidenced by our experiments. Its precise fitting capability is highlighted by significantly smaller residuals compared to the ARIMA model alone, indicating a more accurate capture of the underlying time series trends. This precision is visually supported by the fitted residual time series plots, residual frequency distribution histograms, and variance plots.

The model also exhibits enhanced stability, especially when modeling the volatility of residuals, which is crucial for capturing dynamic changes in stock prices. As volatility is a principal concern in stock market analysis, the ability to provide a stable outlook on price movements is invaluable.

A significant improvement in predictive accuracy is evident through the comparison of statistical metrics such as MAPE, RMSE, and EC. The ARIMA-GARCH-S model's predictive residuals are reduced by approximately two orders of magnitude, showcasing a substantial enhancement over the traditional ARIMA approach.

5.2 Application Prospects:

The ARIMA-GARCH-S model, as demonstrated in this study, holds significant advantages for forecasting stock prices, offering extensive potential for application within financial sectors that demand high accuracy and stability in predictions.Its efficacy in risk management allows for a refined approach to predicting and managing market volatility, thereby enhancing investment risk strategies. In the realm of algorithmic trading, the model's accurate predictions are invaluable, potentially boosting trade strategy effectiveness and profitability. Additionally, for asset pricing, the model provides precise forecasting inputs, crucial for investors seeking to make informed decisions. Collectively, these applications highlight the transformative potential of the ARIMA-GARCH-S model in financial analysis and trading.

5.3 Future Developments:

In future research, combining deep learning and ensemble learning methods with existing time series analysis models represents a particularly enticing direction. This integration is likely to significantly enhance the model's ability to process complex data structures and capture nonlinear patterns, thereby leading to groundbreaking progress in forecasting accuracy and model robustness. Specifically, deep learning models are capable of automatically learning and extracting features from data, while ensemble learning methods improve prediction accuracy and robustness by consolidating multiple models. By leveraging these advanced machine learning techniques, we can anticipate the development of more powerful combined forecasting models. These models will not only be adept at capturing complex relationships within time series data but will also exhibit greater adaptability and precision when confronted with new, unknown market conditions.

6. Conclusion

The experimental results indicate that the ARIMA-GARCH-S model has significant advantages over the ARIMA model in multiple aspects. By examining the fitted residual time series plots, residual frequency distribution histograms, and variance plots, it is evident that the fitted residuals of the ARIMA-GARCH-S model are substantially smaller than those of the ARIMA model, suggesting a more precise fitting capability. Moreover, the volatility of the hybrid model is considerably less than that of the ARIMA model, indicating its enhanced stability in capturing the dynamic changes of stock prices. Additionally, the comparison of three statistical metrics——MAPE, RMSE, and EC—further confirms the superiority of the ARIMA-GARCH-S model in forecasting. Relative to the ARIMA model, the overall predictive residuals of the hybrid model are reduced by approximately two orders of magnitude, significantly improving the model's predictive accuracy.

The findings of this study are valuable academically, as they provide new combinatory ideas for model innovation, and they also have profound implications for practical applications in the financial market. A more accurate and stable model for predicting stock closing prices means that market participants, such as investors, fund managers, and financial analysts, can assess future stock price trends more accurately, thereby making wiser investment decisions.

7. References

[1] YingChao, Z., YingJuan, S. (2019) An empirical study on the analysis and forecasting of SSE index based on ARIMA model. J. Journal of Economic Research., 11: 131-135.

[2] Ariyo, A. A., Adewumi, A. O., Ayo, C. K. (2014) Stock price prediction using the ARIMA model. 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation., 106-112.

[3] Mohammadi, H., Su, L. (2010) International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models. Energy Economics., 32: 1001–1008.

[4] Chou, R. (1988) Volatility persistence and stock valuations: Some empirical evidence using. Journal of Applied Econometrics., 3: 279–294.

[5] ShuYa, X., XiaoYing, L. (2019) Research on Stock Forecasting Based on ARIMA-GARCH Modeling. J. Journal of Henan Institute of Education., 28: 20-24.

[6] YaNing, Y., LiuXiao, G. (2018) Exploration and study of stock data processing process based on ARIMA-GARCH modeling. J. Science Education Guide., 18: 1-7.

[7] Mustapa, F. H., Ismail, M.T. (2019) Modelling and forecasting s&p 500 stock prices using hybrid arima-garch model. Journal of Physics: Conference Series., 1-13.

[8] Pavlyshenko, B. (2018) Using Stacking Approaches for Machine Learning Models. IEEE Second International Conference on Data Stream Mining & Processing. Lviv. pp. 255-258.