**Application and optimization of ARMA model in stock price prediction**


## Abstract

We first study the AR, MA and ARMA model and use them to predict stock prices. It could be found that compared with the actual price, the results of these three models had a certain lag. Then, predicting the trend using ARMA model after removing the residuals from the seasonal decompose to get the predicted price and it could be noticed that this model seems to correctly predict stock price trends at some point although the result was still not satisfactory. This indicated that there might be important information in the residuals, which is the heteroscedasticity of the residuals. Means that the GARCH model could be applied to the residuals and it can be concluded that the model using ARMA for the linear part and GARCH for the residual part could be used to predict the stock price.


## 1. Data Preparation

We used Goldman Sachs (GS) stock price from Jun 2015 to Oct 2017 in our model. For simplification, we chose open price as training data. The following figure shows the division of the data set, the green line is train data while blue line is test data.


Fig 1. The division of the data set

## 2. Autoregressive process

Autoregressive process assumes that the observation at previous several time steps are useful to predict the next step. It depends on autocorrelation. We plotted the observation at the previous time step (t) with the observation at the next time step (t+5) as a scatter plot in the following figure, which clearly shows a relationship or some

correlation.

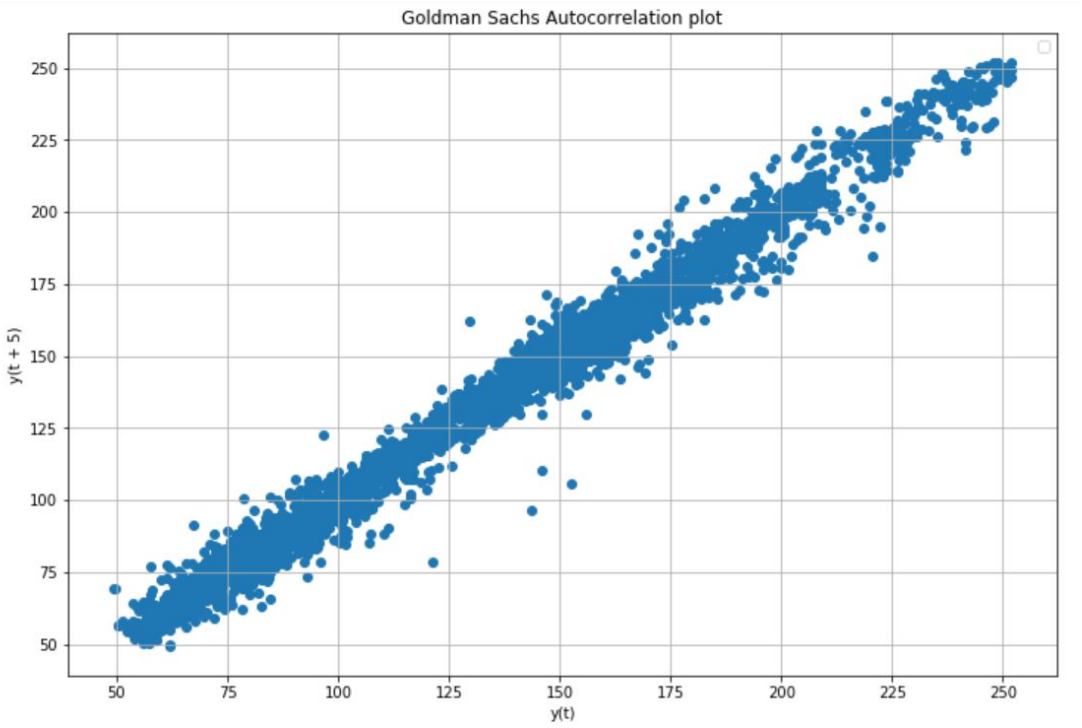



Fig 2. Correlation between $y\,(t)$ and $y\,(t\,+\,5)$

The covariance can also be used to determine whether the data has autocorrelation. From the table below, it can be found that the data is related to the data of the past 1, 5, 10 and 30 days.

| | t | t+1 | t+5 | t+10 | t+30 |
|---|---|---|---|---|---|
| t | 1.000000 | 0.994216 | 0.988282 | 0.956678 | 0.998288 |
| t+1 | 0.994216 | 1.000000 | 0.992890 | 0.963463 | 0.992929 |
| t+5 | 0.988282 | 0.992890 | 1.000000 | 0.971829 | 0.987001 |
| t+10 | 0.956678 | 0.963463 | 0.971829 | 1.000000 | 0.954921 |
| t+30 | 0.998288 | 0.992929 | 0.987001 | 0.954921 | 1.000000 |

The Autoregressive process specifies that the output variable depends linearly on its own previous values and on a stochastic term. The AR process is defined as:

$$X_t = c + \sum_{i=1}^{P} \varphi_i \cdot X_{t-i} + \varepsilon_i$$

Where $\varphi_i$ are the parameters of the model, $c$ is a constant, and $\varepsilon_i$ is white noise.

Figure 3 shows the results of the stock price prediction using the AR model.

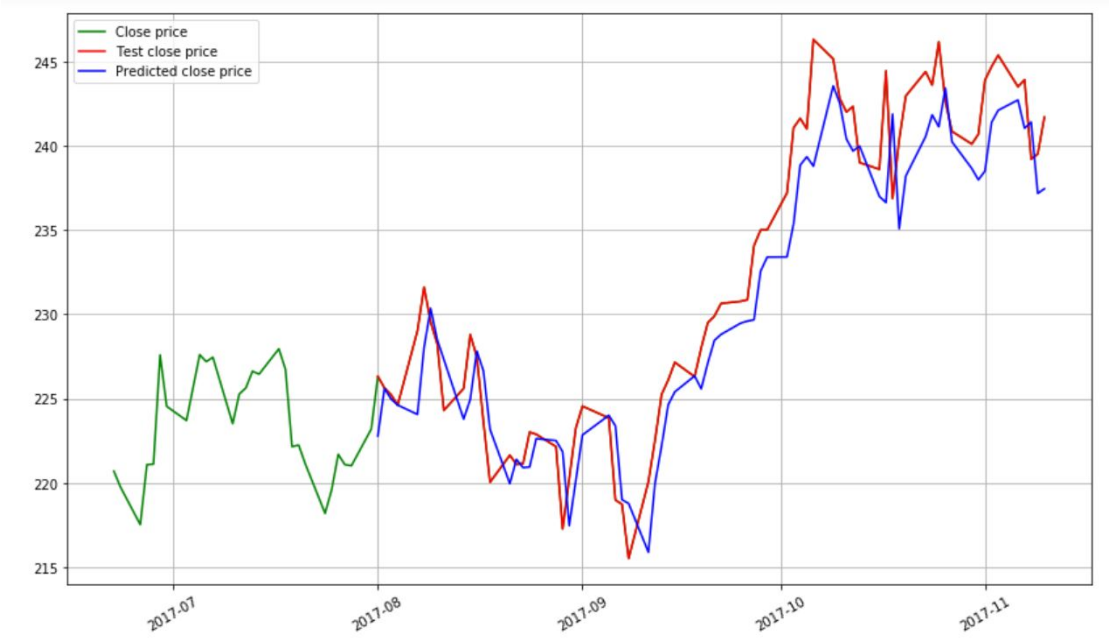



Fig 3. The results of the stock price prediction using the AR model

## 3. Moving-average process

The Moving-average process specifies that the output variable depends linearly on the current and various past values of a stochastic term. The MA process is defined as:

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i};$$

Where $\mu$ is the mean of the series, $\theta_i$ are the parameters of the model and $\varepsilon_i$ are white noise error terms.

Figure 4 shows the results of the stock price prediction using the MA model.

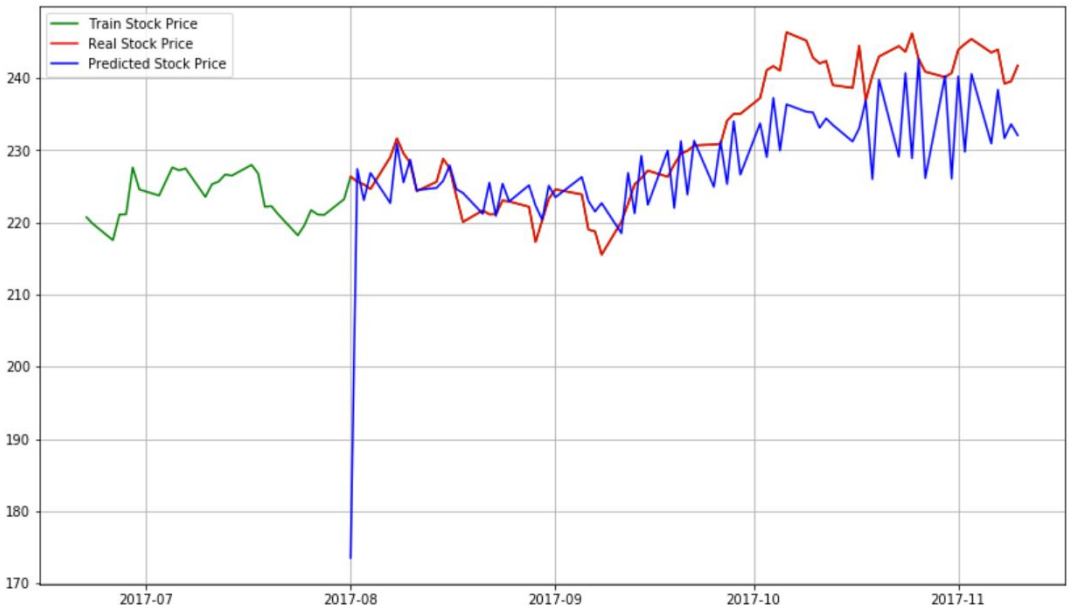


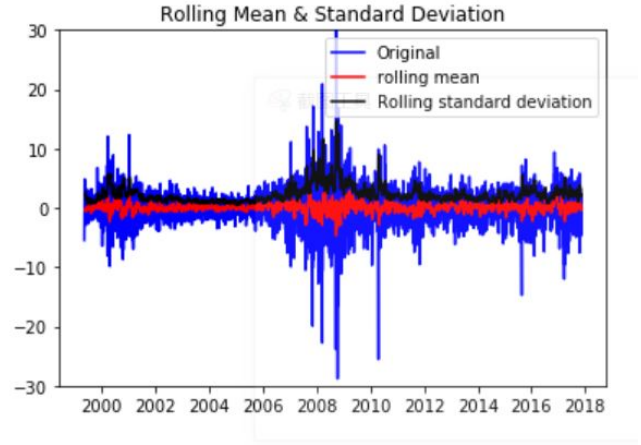

Fig 4. The results of the stock price prediction using the MA model

## 4. ARMA process

ARMA process is simply the merger between AR process and MA process. The ARMA process is defined as:

$$X_t = c + \varepsilon_i + \sum_{i=1}^{P} \varphi_i \cdot X_{t-i} \; + \; \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

The training data of ARMA model should achieve stationary by taking a series of difference. Do Dickey-Fuller Test and it showed that in 90% confidence level, the data after first order difference is stationary, and we can also intuitively see from the figure below that the average change is not large, so the data can be considered stable.




```
Results of Dickey-Fuller Test:
Test Statistic                -1.392256e+01
p-value                        5.263162e-26
#Lags Used                     1.900000e+01
Number of Observations Used    4.640000e+03
Critical value (1%)           -3.431760e+00
Critical value (5%)           -2.862163e+00
Critical value (10%)          -2.567102e+00
dtype: float64
```

Fig 5. Result of Dickey-Fuller test

In ARMA process, Autocorrelation Function and Partial Autocorrelation Function could be used to check the parameter $p$ and $q$ in model. Actually, they all done by computer. And information criterion is also commonly used here.

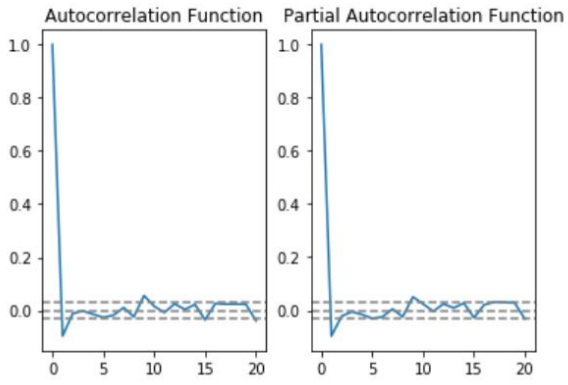



Fig 6. Autocorrelation Function and Partial Autocorrelation Function

Figure 7 shows the results of the stock price prediction using the ARMA model.

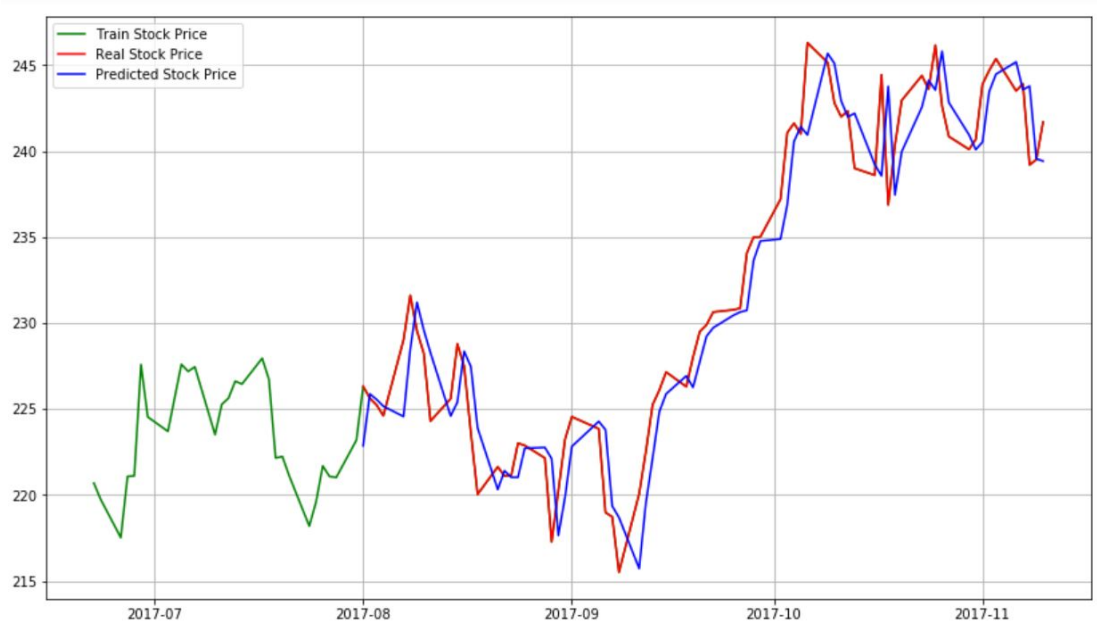



Fig 7. The results of the stock price prediction using the ARMA model

## 5. Weakness of AR, MA and ARMA process

It could be concluded that ARMA model is more accurate than the AR and MA model from the above prediction results. But they all have one common disadvantage, we can notice that prediction had some lags compared to reality in the results of the three models. Since it should be derived from the data of the day before that day, the lag made this model unusable for predictions.

## 6. ARMA process combined with seasonal decomposition

Although the ARMA model has some of the disadvantages mentioned above, the steps of obtaining stationary data during the construction of the ARMA model could give some inspiration. From the difference method it can be noticed that the residuals have strong randomness. We think this is related to external variables. So, would it be better if the residuals were separated from the data? We use seasonal decompose to divide the original data into trends, seasonality and residual part as shown in figure 8.

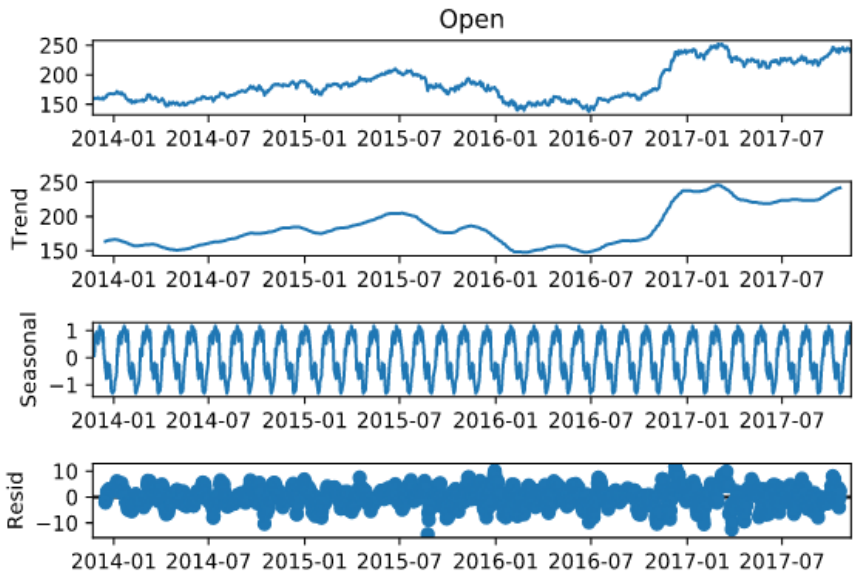



Fig 8. Seasonal Decompose

Then use the ARMA model to predict the trend and add the seasonality to get the prediction result as shown in the figure 9.

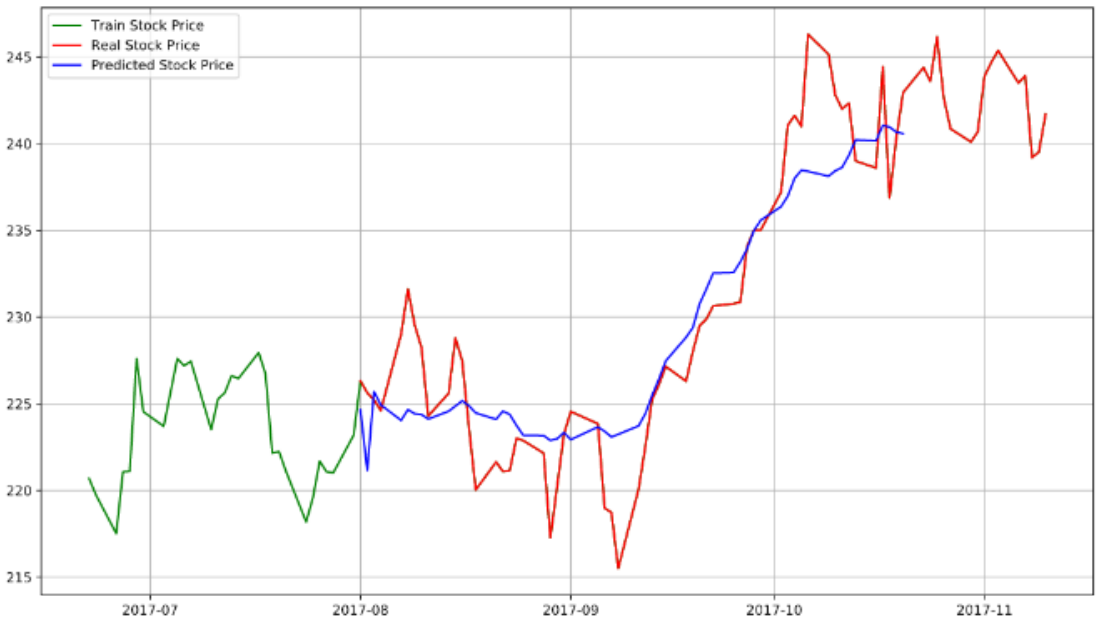



Fig 9. Prediction results of ARMA model after seasonal decompose

Although the result is still not satisfactory, the model successfully predicted the initial decline and intermediate rise. But overall, this cannot be a good predictive model because it is too far from the actual price.

## 7. ARMA-GARCH

The ARMA process combined with seasonal decomposition also could give some inspiration. Since the prediction the model made after removing the residuals was still not ideal, did it mean that there was some important information in the residuals? The answer is yes, we performed an autocorrelation test on the residuals and found that they were not relevant as shown in figure 10.

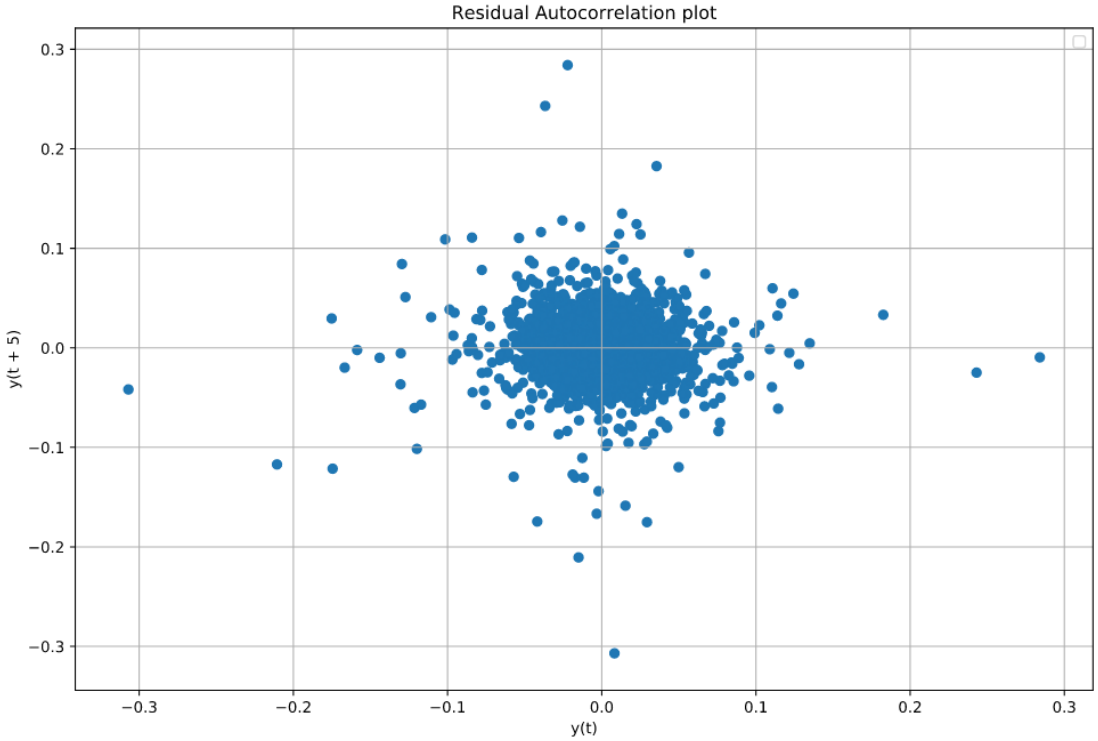



Fig 10. Autocorrelation of residuals

But it could be found that the residuals had significant heteroscedasticity when testing the ARCH effect on the residuals. Here we use the python code in the figure 11 to test the ARCH effect.

```
1 from statsmodels.stats import diagnostic
2 *_, fpvalue = diagnostic.het_arch(resid)
3 print (fpvalue)
4     1.9002719188490938e-263
```

Fig 11. Heteroscedasticity test

The null hypothesis tested is that there is no heteroscedasticity. The $p$ value obtained by the test is almost equal to zero. Means the residuals had heteroscedasticity. That indicates that the variance of the residual will change over time, and variance represents volatility. So, the residual contains information about volatility.

That gives us an idea, maybe we can use ARMA for the linear part and GARCH for the residual part. GARCH models are commonly employed in time series that exhibit time-varying volatility and volatility clustering. We can use GARCH to estimate the variance of the data. The GARCH process is defined as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \ ; \ \varepsilon_t = \sigma_t e_t$$

Where $e_t$ is a random variable with mean 0 and variance 1. Here we consider $\varepsilon_t$ as the residual of return.

And in ARMA-GARCH model:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$$

$$\varepsilon_t = \sigma_t e_t$$

$$\mathrm{X_t} = \mathrm{c} + \varepsilon_\mathrm{i} + \sum_{\mathrm{i}=1}^{\mathrm{P}} \varphi_\mathrm{i} \cdot \mathrm{X_{t-i}} \ + \ \sum_{\mathrm{i}=1}^{\mathrm{q}} \theta_\mathrm{i} \varepsilon_{\mathrm{t-i}}$$

The value of $c, \varphi_i, \theta_i$ are given by ARMA and the value of $\omega, \alpha, \beta$ are given by GARCH.

Before we apply this model to the dataset, we need to convert prices into returns to make it easier. Because we need to determine the distribution of $e_t$. Since the distribution of return has heavy tail phenomenon, it is reasonable to assume that $e_t$ follows the Students t -distribution and it can make the model more accurate. By the way, the transformed data still retains the properties mentioned above such as heteroscedasticity of residuals. So, it makes sense to apply the ARMA-GARCH model to the return.

Figure 12 shows the results of the stock price prediction using the ARMA-GARCH model. It should be noted that unlike the ARMA model, which can only predict one day backwards, the ARMA-GARCH model can predict the results for several days at once. Here we predict the return of next 10 days So, this is more practical. And we also see that this model successfully predicts some trends, and it is very accurate in some places.

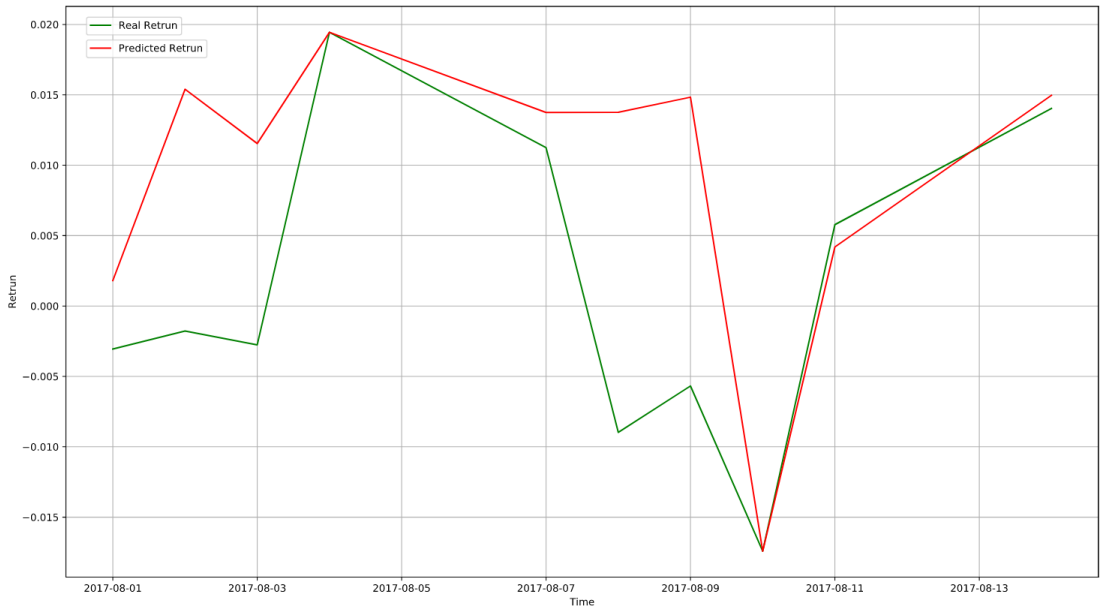
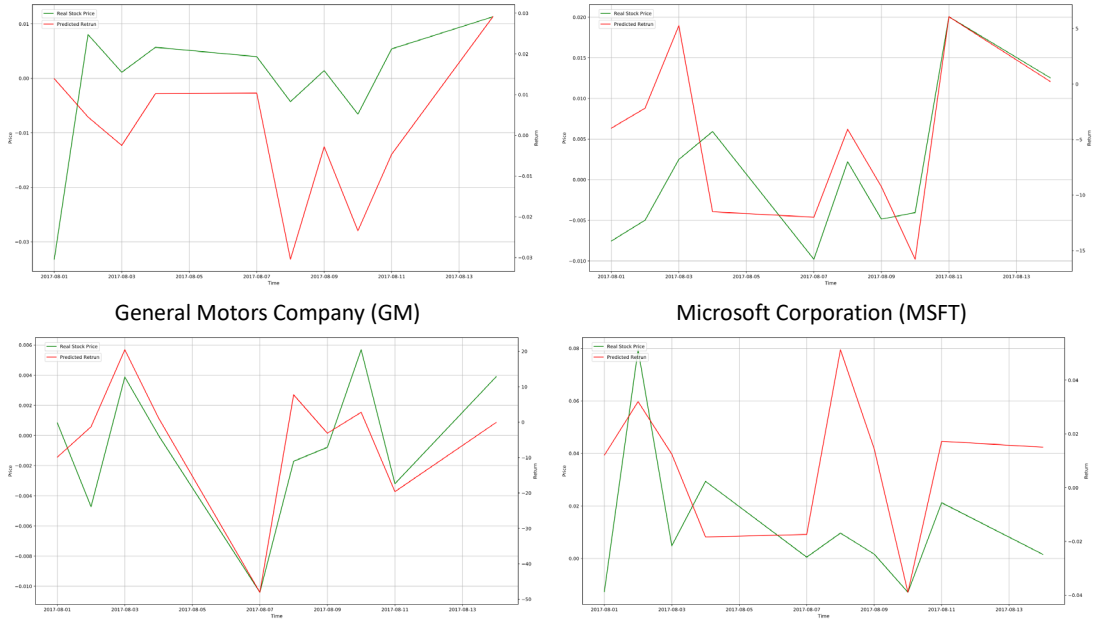



Fig 12. Prediction results of the ARMA-GARCH model

Figure 13 shows the prediction results of Microsoft, Tesla, IBM and GM using ARMA-GARCH model. We can see some defects of this model, that is, in some places the model only predicted the trend, but got wrong predicted value, we can also notice that the prediction results for some stocks will be relatively accurate, like IBM and Microsoft. while the prediction results for other stocks will be comparatively terrible. So, the model may need to be further optimized. But in terms of prediction results, it still surpasses previous models.



General Motors Company (GM)

Microsoft Corporation (MSFT)

International Business Machines (IBM)

Tesla, Inc. (TSLA)

Fig 13. Prediction results of MSFT, TSLA, IBM and GM using ARMA-GARCH model

## 8. Conclusion

The combination of ARMA and GARCH could be used as a tool to predict stock prices, which is better than AR and MA models, but because of the instability of this model, it may only have a reference value. there is still a gap with reality.

In the research process, GARCH was applied on the residuals to further extract the information in the residuals, and turned the irregular information in the residual into a random variable in GARCH, this irregular information is determined by the outside world, such as investor mood, unemployment rate, GDP and other factors. In previous analysis, we assume that the random variable follows the Students t -distribution, but this assumption is still imprecise. But we can get the historical distribution of these random variables from historical prices. Through these random values, it may be possible to analyze which external factors are determining the stock price. Therefore, the ARMA-GARCH model can become a tool for analyzing the past.