Forecasting VIX using ARIMA GARCH

Jiaqi Jiang
Master of Science in Analytics
Georgia Institution of Technology
Atlanta, GA 30332
jjiang334@gatech.edu

Summary

The CBOE Volatility Index (VIX) is a real-time market index that represents the market's expectation of 30-day forward-looking volatility. The VIX is also known as the "Fear Index", which is used by many investors to measure market risk before making investment decision. Recently, the VIX jumped and fell dramatically due to the Covid-19 pandemic. My goal is to test whether ARIMA and GARCH model can fit the VIX. I used the VIX daily price from January 2nd, 2004 to April 24th, 2020, experimented ARIMA model both with seasonality and without seasonality, and found the best model is ARIMA (1,1,2). The test set lies on the lower bound of the 90% confidence interval of my forecast but did not explain the volatility well. Therefore, I further tested ARIMA GARCH and found ARIMA (2,1,2) – GARCH (1,1) is better than ARIMA (1,1,2) in predicting VIX.

1 Introduction and related work

Previous research on fitting time series model using ARIMA by Katja Ahoniemi (2006) [1] indicates that an ARIMA (1,1,1) model enhanced with exogenous regressors (such as major stock index returns) has predictive power regarding the directional change in the VIX index. GARCH terms are statistically significant, but do not improve forecasts. On the other hand, Farah Hayati Mustapa and Mohd Tahir Ismail (2019) [2] conclude that the ARIMA (2,1,2)-GARCH (1,1) has the best predictive power on S&P 500, which is VIX's underlying asset. In addition, I used the ARIMA-GARCH hybrid model with rugarch package in R to predict the S&P 500 index in the course homework.

Since ARIMA with exogenous factors is out of the scope of this project, my goal is to determine whether a simple ARIMA model or a ARIMA-GARCH model has stronger predicting power in VIX. I will first test how VIX can be fitted with a simple ARIMA model, and then test with an ARIMA-GARCH hybrid model.

2 Methods

2.1 Data and Software

The data set I used includes the VIX daily price from January 2nd, 2004 to April 24th, 2020, sourced from CBOE website. I used the last 10 data points as the test set and the remaining as the training set. I used Python 3 to perform all the tasks except for ARIMA-GARCH model, which I used Rugarch package in R. The output from R is then coded back to Python 3 notebook for plot purposes.

2.2 Data Clean and Analysis

Figure 1 shows the log price for the VIX index. There is an obvious trend in time series data, but no certain seasonality in the data. However, the Dick-Fuller test (Appendix Table 1) shows rejection of the null hypothesis, which suggests the data does not have a unit root and is stationary. Further research and analysis show that the test is affected by data frequency and sample size. The p-value is usually small for the test when the data points are denser. Figure 2 shows decomposition of trend using STL package from Python 3.

Figure 1: Log price for VIX

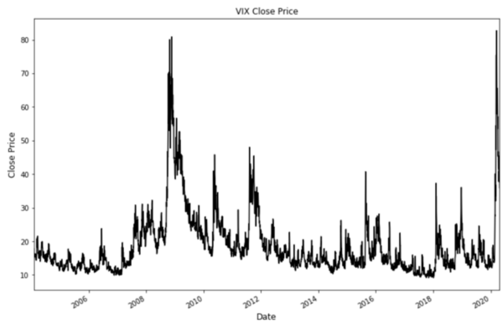


Figure 2: Trend of log price for VIX

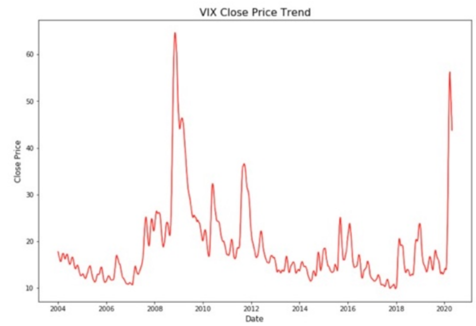
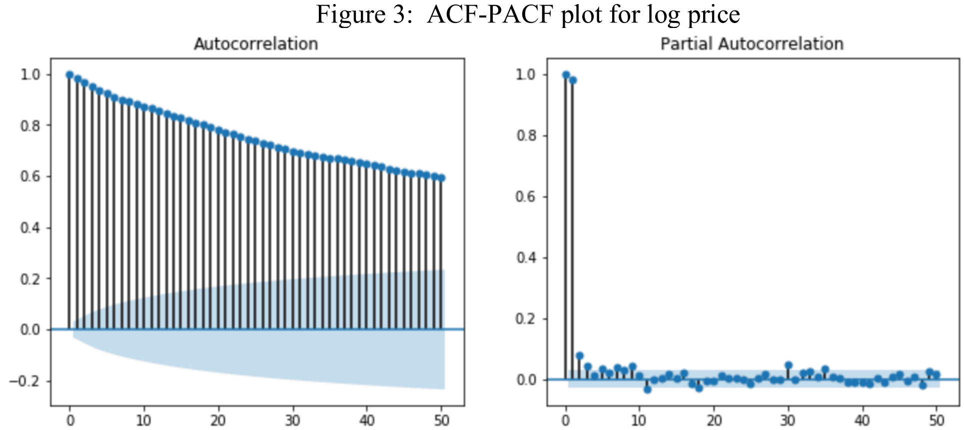


Figure 3 shows the trend for log price of VIX. As seen in Figure 3, the ACF shows a very slow decay, which suggests the time series is not stationary. Additionally, The PACF with its first lag almost equal to one suggests the time series is not stationary.

Figure 3: ACF-PACF plot for log price



Next, I took the first order difference which is equivalent to the log returns of the VIX. As shown in Figure 4, the log return shows constant mean. This is also suggested by Dick-Fuller test (Appendix Table 2). For the ACF and PACF graph in Figure 5 without lag zero, the ACF decays fast after one or two lags, which might suggest MA of order one or two; PACF truncates fast after first lag, which might suggest AR of order 1. In addition, I continued my analysis with log returns (first order difference of the Log price of VIX).

Figure 4: First order difference

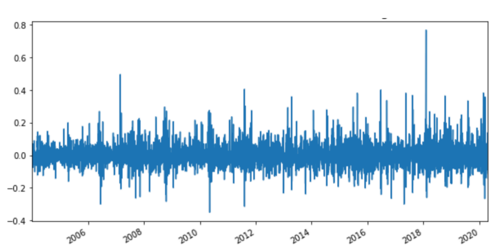
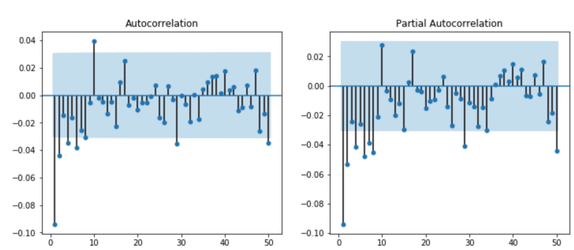


Figure 5: ACF-PACF for first order difference

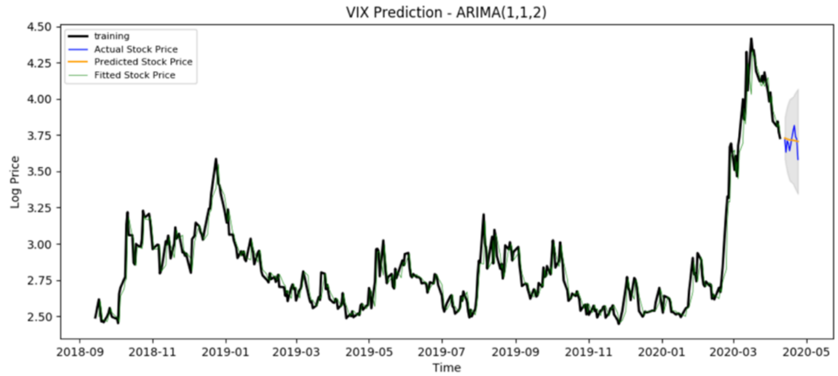


2.3 Model One - ARIMA

In order to search for best orders for the ARIMA (p, 1, q) model, I used the Auto Arima function from the pyramid package in Python3. It grid-searched p from 1 to 3 and q from 0 to 5 and found the model with highest BIC score is ARIMA (1,1,2), which matches our initial guess by inspecting ACF and PACF plots. All AR and MA lags are statistically significant (Appendix Table 3).

The QQ plot and the histogram (Appendix Figure 1) show the residual is likely to be normal distribution but with a long right tail. ACF and PACF of the residual suggest there is no correlation in the residual. I also used the Auto Arima (p, 1, q) (P, D, Q) model with seasonal trend, which generates a BIC higher than the model without seasonal trend. The result in Figure 6 shows the actual price fell on the lower bound of 90% confidence interval of the prediction.

Figure 6: ARIMA(1,1,2) prediction



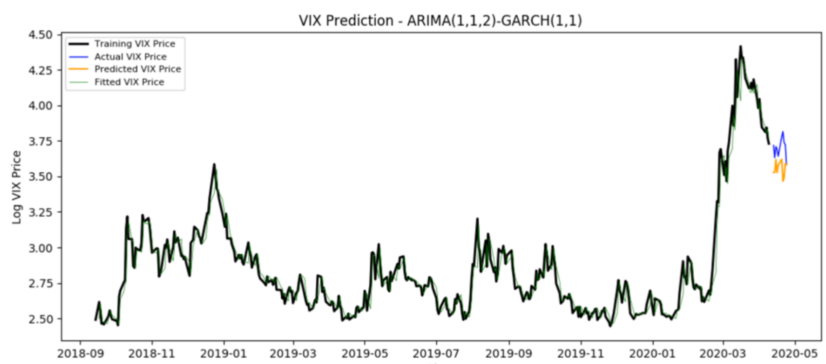
2.3 Model Two - ARIMA-GARCH Hybrid

Next, I inspected how ARIMA-GARCH performed on predicting the volatility of log VIX price. After research, there is no package to perform this task in Python 3. Therefore, I shifted to R and used Rugarch package. I started with the ARIMA model (1,1,2) (or the ARMA (1,2) of the first order difference) from the previous section.

I searched for a GARCH model which will minimize the BIC score, the result shows GARCH (1, 1). Using the GARCH (1, 1) I used the same method and refined my ARIMA to ARIMA (2, 1, 2). I compared the original ARIMA-GARCH (1, 1, 2) - (1,1) model with ARIMA-GARCH (2, 1, 2) - (1,1) model. I found the ARIMA-GARCH (1, 1, 2) - (1,1) model has a lower AIC and BIC score. The statistics of the models are attached in Appendix Table 4.

The result in Figure 7 shows the ARIMA-GARCH (1,1,2) - (1,1) model predicts both the trend and volatility better than the original ARIMA (1,1,2) model.

Figure 7: ARIMA-GARCH (1,1,2) -(1,1) model



3 Conclusion

In conclusion, the ARIMA-GARCH model has better predicting power in the trend and volatility of the VIX index, which is slightly different the finding from the previous research. Further analysis could be focused on bringing in extra exogeneous factors such as S&P index or indicator variables for financial crisis.

4 Reference

[1] Katja Ahoniemi Helsinki School of Economics, FDPE, and HECER 2006 Modeling and Forecasting Implied Volatility – an Econometric Analysis of the VIX Index
[2] Farah Hayati Mustapa and Mohd Tahir Ismail 2019 J. Phys.: Conf. Ser. 1366 012130

5 Appendix

Appendix Table 1: Dick-Fuller test summary log price

ADF Statistic:	-4.27876
P-value:	0.000483

Appendix Table 2: Dick-Fuller test summary first order difference

ADF Statistic:	-22.951931
P-value:	0.000000

Appendix Table 3: ARIMA (1,1,2) result

ARIMA Model Results						
Dep. Variable:	D.Log.VIX Close		No. of Observations:	4084		
Model:	ARIMA(1, 1, 2)		Log Likelihood:	4880.155		
Method:	css-mle		S.D. of innovations	0.073		
Date:	Sun, 26 Apr 2020		AIC	-9750.310		
Time:	20:35:55		BIC	-9718.736		
			HQIC	-9739.130		
	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0003	0.001	0.449	0.654	-0.001	0.002
ar.L1.D.Log.VIX Close	0.8516	0.063	13.561	0.000	0.729	0.975
ma.L1.D.Log.VIX Close	-0.9605	0.065	-14.682	0.000	-1.089	-0.832
ma.L2.D.Log.VIX Close	0.0453	0.024	1.908	0.056	-0.001	0.092

Appendix Table 4: ARIMA GARCH model comparison

	ARIMA GARCH (1,1,2)-(1,1)	ARIMA GARCH (2,1,2)-(1,1)
Akaike	-2.656901	-2.655675
Bayes	-2.644561	-2.641793
Shibata	-2.656908	-2.655684
Hannan-Quinn	-2.652532	-2.650760

Appendix Figure 1: ARIMA (1,1,2) diagnostic plot