ORIE 5640 Statistics for Financial Engineering

Project 2

Modeling the Long-memory Property of Volatility in Gold Futures Returns

Group members:
Rui Dai (rd576)

May 9, 2021

1 Introduction

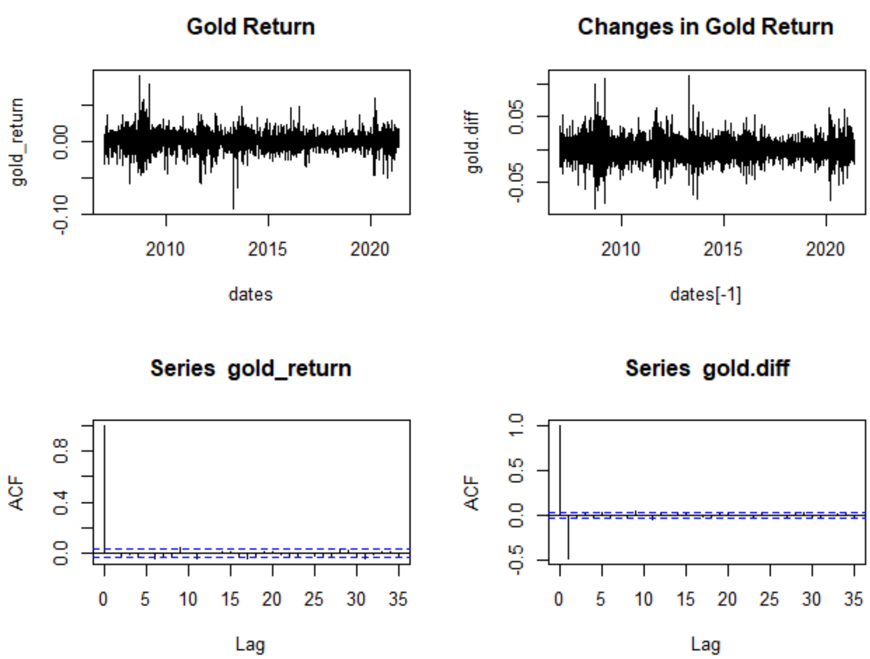
The gold market plays an important role in global finance since it exhibits low correlation with other assets. As a result, investors often consider it as an instrument to diversify portfolios and avoid risks, especially during crisis periods [1]. During such periods, gold returns may also have unstable volatility, which illustrates the importance of understanding and modeling volatility. Long-memory is a common property of financial time series that attracts interests for a long time because of its financial implications, and a number of literature have shown that the gold volatility is a long-memory process [2] [3], implying volatility persistence over long horizons. For simplicity, the financial implications are not discussed here, but including the long-memory property into the volatility model is likely to increase the predictive power and achieve more satisfactory performance when forecasting returns. Motivated by the above idea, this project examines the fit of GARCH(1,1), Integrated GARCH(1,1) and Fractionally Integrated GARCH(1,d,1) models to the volatility of gold futures returns and compares the out-of-sample forecasting ability on returns of the corresponding time series models.

2 Data

Due to the difficulty to obtain a general representation of spot gold prices, this project uses gold futures contracts traded on the COMEX division of the New York Mercantile Exchange to represent the gold market. The prices from Jan 1, 2007 to May 7, 2021 are acquired from Yahoo Finance via the quantmod package in R, and daily returns are calculated using the built-in function of quantmod.

2.1 Initial Analysis

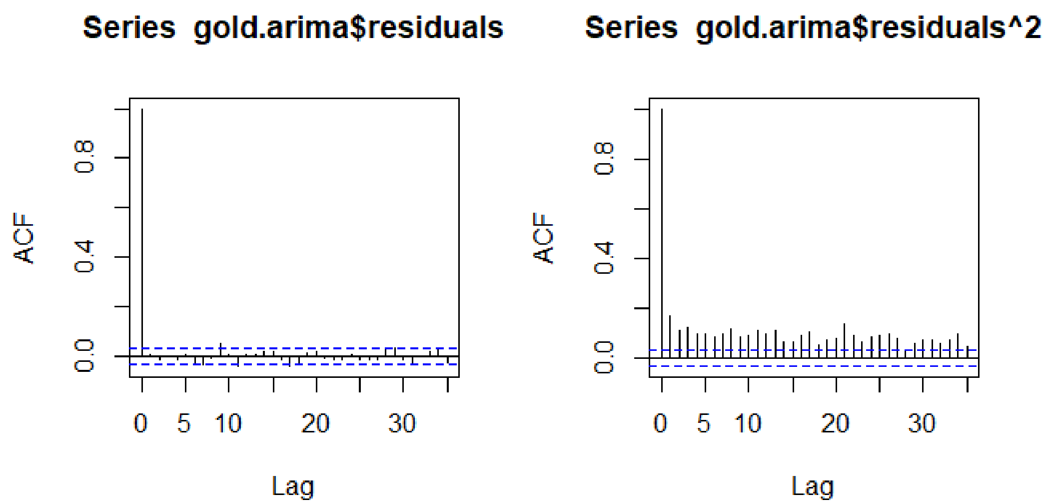
The following plots show the returns of gold futures, the differenced returns of gold futures and the corresponding ACF plots.



ADF and KPSS tests are performed on both series and the results show that both series are stationary at the 95% level of significance, so there is no need to difference the returns. From the plots, there is volatility clustering in the returns.

2.2 Auto ARIMA Residuals

An ARIMA(0,0,0) model with non-zero mean is selected by auto ARIMA, and the ACF plots of its residuals and the squared residuals are as follows.



Although the ARIMA(0,0,0) model seems to fit the conditional mean well, the ACF plot of the squared residuals shows significant non-zero autocorrelation, indicating the existence of conditional heteroskedasticity and the necessity of GARCH model.

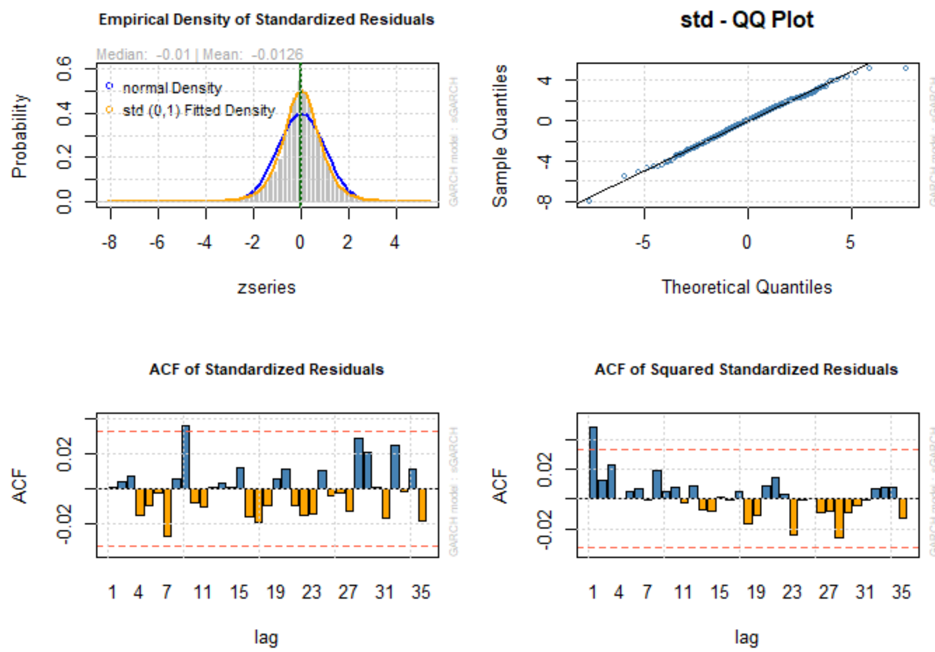
3 GARCH Model Fitting

Following the results of auto ARIMA, the mean model used in GARCH will be ARIMA(0,0,0). The distribution model for the noise is ”std” as financial data usually have heavier tails than implied by the conditional heteroscedasticity.

3.1 GARCH(1,1)

GARCH(1,1) is the most basic and most commonly used GARCH-class model in volatility modeling. The estimated parameters and related plots are as follows:

	Estimate	Std. Error	t value	$\Pr(>\lvert t\rvert)$
mu	0.0005	0.0001	3.3909	0.0007
omega	0.0000	0.0000	0.7258	0.4679
alpha1	0.0349	0.0042	8.2560	0.0000
beta1	0.9620	0.0025	380.6237	0.0000
shape	4.4943	0.6030	7.4536	0.0000



The shape parameter is significant, indicating the std distribution model is a good choice. The ACF plot of the squared standardized residuals shows little autocorrelation, and the adjusted Pearson goodness-of-fit test shows large p-values for all groups, indicating the GARCH(1,1) model fits well. Details of the model fit can be found in appendix.

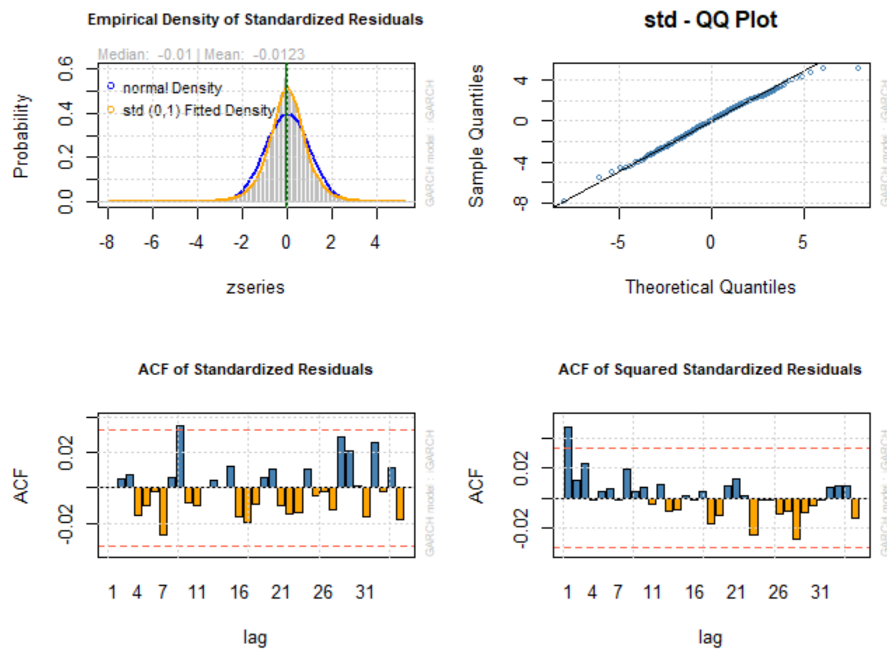
3.2 IGARCH(1,1)

Similar to the idea of stationary ARMA processes, the standard GARCH model only considers shocks that decay to zero at an exponential rate. Therefore, it is only suitable for short-memory processes [4]. Consider the standard GARCH model as an I(0) type of process, Integrated GARCH (IGARCH) gives a way to model I(1) type of process, similar to the difference between ARMA(p,q) and ARIMA(p,1,q). This model is proposed by Engle and Bollerslev [5] by introducing a unit root condition into the GARCH(p,q) model. That is,

$$\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_i = 1$$

following the notations in the lecture notes.
The estimated parameters of an IGARCH(1,1) fitting and related plots are as follows:

	Estimate	Std. Error	t value	Pr(> $\|t\|$)
mu	0.0005	0.0001	3.4065	0.0007
omega	0.0000	0.0000	1.2900	0.1970
alpha1	0.0373	0.0030	12.3435	0.0000
beta1	0.9627	NA	NA	NA
shape	4.2785	0.2351	18.2003	0.0000

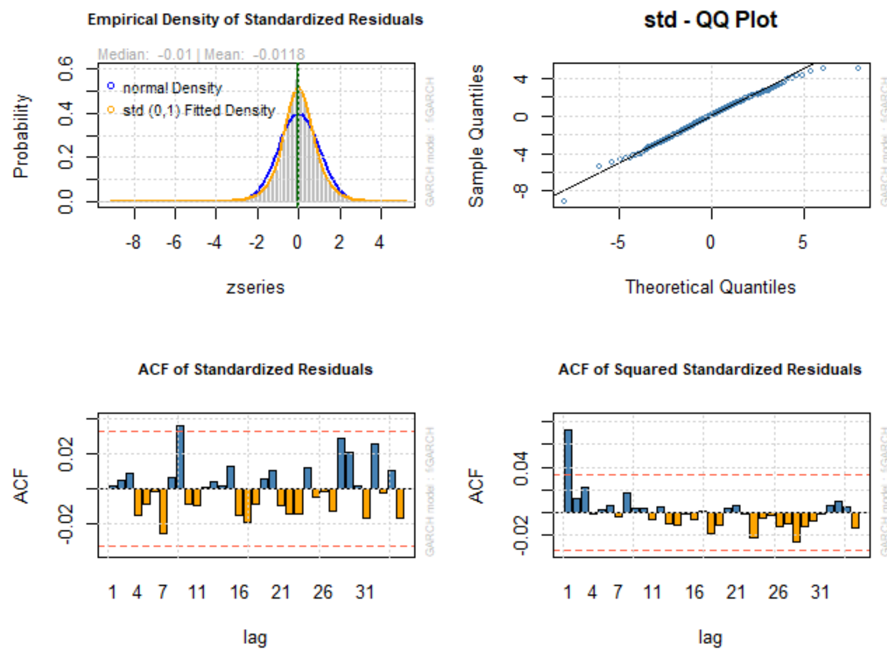


The shape parameter is significant, indicating the std distribution model is a good choice. The ACF plot of the squared standardized residuals shows little autocorrelation, and the adjusted Pearson goodness-of-fit test shows large p-values for all groups, indicating the IGARCH(1,1) model fits well. Details of the model fit can be found in appendix.

3.3 FIGARCH(1,d,1)

IGARCH model is still not satisfactory since it assumes that shocks will die out, which is unrealistic. To model a long-memory volatility process, Baillie et al. [6] introduced Fractionally Integrated GARCH (FIGARCH) model, allowing fractional integration. The shocks are assumed to decay at a slow hyperbolic rate, which is reasonable for real data. Detailed mathematical discussion of the model is omitted here due to page limits.
The estimated parameters of a FIGARCH(1,d,1) fitting and related plots are as follows:

	Estimate	Std. Error	t value	$\Pr(>\lvert t\rvert)$
mu	0.0005	0.0001	3.4033	0.0007
omega	0.0000	0.000000	2.6981	0.0070
alpha1	0.0000	0.0136	0.0001	0.9999
beta1	0.9602	0.0031	308.5111	0.0000
delta	0.9644	0.0106	91.2573	0.0000
shape	4.2818	0.2376	18.0207	0.0000



The shape parameter is significant, indicating the std distribution model is a good choice. The ACF plot of the squared standardized residuals shows little autocorrelation, and the adjusted Pearson goodness-of-fit test shows large p-values for all groups, indicating the FIGARCH(1,d,1) model fits well. Details of the model fit can be found in appendix.

4 Prediction Performance

The one-step-ahead out-of-sample forecasting are performed with rolling window on gold futures returns in 2021, and the estimated Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) are used in performance evaluation. The results are as follows:

	MAE	RMSE
GARCH(1,1)	0.676235	0.01067288
IGARCH(1,1)	0.6762321	0.01067285
FIGARCH(1,1)	0.6762241	0.01067276

Due to the difficulty in measuring volatility forecast performance, the volatility forecasting ability is implicitly represented by the forecasting performance of conditional means. See appendix for the plots of volatility prediction.

5 Conclusion

GARCH(1,1), IGARCH(1,1) and FIGARCH(1,d,1) all give satisfactory fit to the volatility of gold futures returns. In terms of forecasting ability on returns, both MAE and RMSE selected FIGARCH(1,d,1) as the most preferrable model, which is under expectation and aligns with results from previous literature [4]. However, the difference is small and more analysis should be done using data in different time periods to determine the best fit. To conclude, as an important property of the volatility of gold returns, long-memory should be carefully considered when investors intend to make investments with relevant models.

References

[1] F. A. O'Connor, B. M. Lucey, J. A. Batten, and D. G. Baur, "The financial economics of gold — a survey," *International Review of Financial Analysis*, 2015.

[2] B. K. Uludag and Z. Lkhamazhapov, "Long memory and structural breaks in the returns and volatility of gold: Evidence from turkey," *Applied Economics*, 2014.

[3] S. R. Bentes, "Long memory volatility of gold price returns: How strong is the evidence from distinct economic cycles?," *Physica A: Statistical Mechanics and its Applications*, 2016.

[4] S. R.Bentes, "Forecasting volatility in gold returns under the garch, igarch and figarch frameworks: New evidence," *Physica A: Statistical Mechanics and its Applications*, 2015.

[5] R. F. Engle and T. Bollerslev, "Modelling the persistence of conditional variances," *Econometric Reviews*, 1986.

[6] R. T. Baillie, T. Bollerslev, and H. O. Mikkelsen, "Fractionally integrated generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 1996.

6 Appendix

6.1 GARCH(1,1) Fit

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,0)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu      0.000484    0.000143   3.39089 0.000697
omega   0.000001    0.000001   0.72582 0.467948
alpha1  0.034856    0.004222   8.25598 0.000000
beta1   0.962040    0.002528 380.62365 0.000000
shape   4.494327    0.602970   7.45364 0.000000

Robust Standard Errors:
        Estimate  Std. Error   t value Pr(>|t|)
mu      0.000484    0.000158  3.069033 0.002148
omega   0.000001    0.000006  0.098945 0.921182
alpha1  0.034856    0.026017  1.339745 0.180328
beta1   0.962040    0.014503 66.333111 0.000000
shape   4.494327    3.662737  1.227041 0.219807

LogLikelihood : 11414.06

Information Criteria
------------------------------------

Akaike       -6.3756
Bayes        -6.3669
Shibata      -6.3756
Hannan-Quinn -6.3725

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                  0.0008197  0.9772
Lag[2*(p+q)+(p+q)-1][2] 0.0342031  0.9688
Lag[4*(p+q)+(p+q)-1][5] 0.5705394  0.9463
d.o.f=0
```

```
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                               statistic  p-value
Lag[1]                             8.239 0.004099
Lag[2*(p+q)+(p+q)-1][5]            9.800 0.010361
Lag[4*(p+q)+(p+q)-1][9]           10.585 0.037508
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]     1.845 0.500 2.000  0.1744
ARCH Lag[5]     1.892 1.440 1.667  0.4957
ARCH Lag[7]     2.001 2.315 1.543  0.7172

Nyblom stability test
------------------------------------
Joint Statistic:  493.154
Individual Statistics:
mu       0.7931
omega   80.8909
alpha1   0.9864
beta1    0.7433
shape    1.4145

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.28 1.47 1.88
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           0.3031 0.76186
Negative Sign Bias  1.6407 0.10095
Positive Sign Bias  1.8757 0.06078   *
Joint Effect       10.2144 0.01683  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     23.45       0.2180
2    30     29.88       0.4202
3    40     40.06       0.4231
4    50     55.38       0.2466
```

```
Elapsed time : 0.3586559
```

6.2 IGARCH(1,1) Fit

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : iGARCH(1,1)
Mean Model      : ARFIMA(0,0,0)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000483    0.000142   3.4065 0.000658
omega   0.000000    0.000000   1.2900 0.197047
alpha1  0.037294    0.003021  12.3435 0.000000
beta1   0.962706          NA       NA       NA
shape   4.278498    0.235078  18.2003 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000483    0.000142  3.41254 0.000644
omega   0.000000    0.000001  0.47351 0.635849
alpha1  0.037294    0.015191  2.45497 0.014090
beta1   0.962706          NA       NA       NA
shape   4.278498    0.232768 18.38098 0.000000

LogLikelihood : 11413.45

Information Criteria
------------------------------------

Akaike       -6.3758
Bayes        -6.3689
Shibata      -6.3758
Hannan-Quinn -6.3733

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                   0.001022  0.9745
Lag[2*(p+q)+(p+q)-1][2]  0.044005  0.9610
```

```
Lag[4*(p+q)+(p+q)-1][5]   0.597352   0.9417
d.o.f=0
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                          statistic  p-value
Lag[1]                        7.940 0.004836
Lag[2*(p+q)+(p+q)-1][5]       9.426 0.012904
Lag[4*(p+q)+(p+q)-1][9]      10.174 0.046044
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]     1.819 0.500 2.000  0.1774
ARCH Lag[5]     1.862 1.440 1.667  0.5025
ARCH Lag[7]     1.949 2.315 1.543  0.7280

Nyblom stability test
------------------------------------
Joint Statistic:  227.5046
Individual Statistics:
mu      0.79347
omega  92.06364
alpha1  0.05848
shape   1.21059

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:          1.07 1.24 1.6
Individual Statistic:     0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           0.3067 0.75910
Negative Sign Bias  1.4707 0.14145
Positive Sign Bias  2.0045 0.04509  **
Joint Effect       10.0776 0.01792  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     21.90       0.2893
2    30     26.26       0.6118
3    40     39.03       0.4686
4    50     53.51       0.3053
```

```
Elapsed time : 0.4153769
```

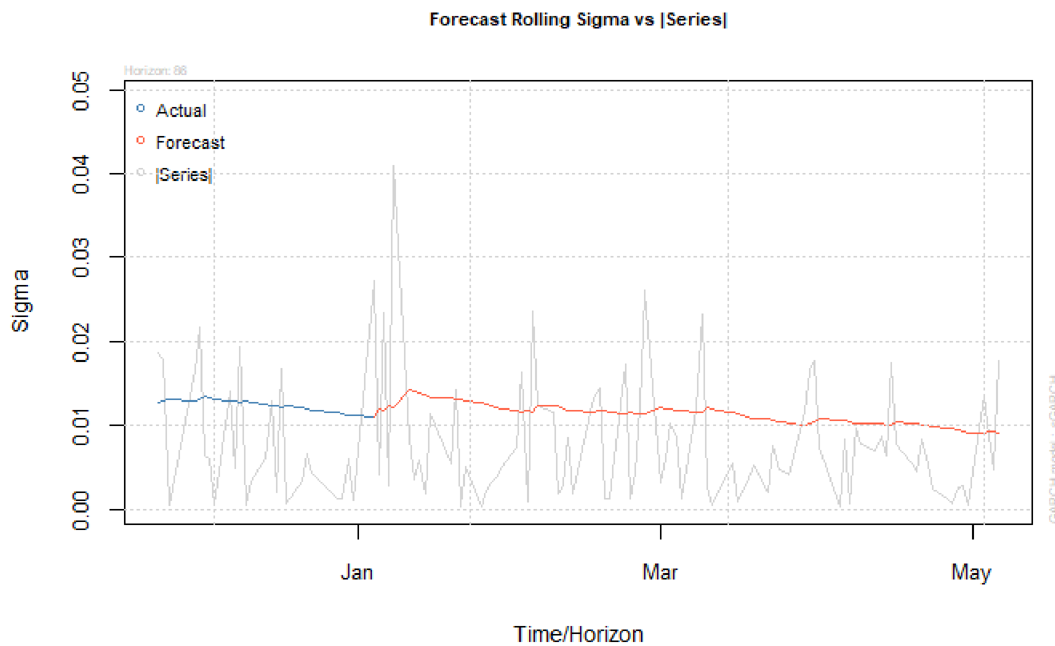
6.3 FIGARCH(1,d,1) Fit

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : fiGARCH(1,1)
Mean Model      : ARFIMA(0,0,0)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error    t value Pr(>|t|)
mu      0.000482    0.000142   3.403327 0.000666
omega   0.000001    0.000000   2.698055 0.006975
alpha1  0.000002    0.013647   0.000111 0.999912
beta1   0.950210    0.003080 308.511144 0.000000
delta   0.964432    0.010568  91.257323 0.000000
shape   4.281767    0.237602  18.020735 0.000000

Robust Standard Errors:
        Estimate  Std. Error   t value Pr(>|t|)
mu      0.000482    0.000143  3.363754 0.000769
omega   0.000001    0.000000  2.097760 0.035926
alpha1  0.000002    0.015020  0.000101 0.999920
beta1   0.950210    0.010825 87.780491 0.000000
delta   0.964432    0.027289 35.341480 0.000000
shape   4.281767    0.211757 20.220236 0.000000

LogLikelihood : 11414.89

Information Criteria
------------------------------------

Akaike       -6.3755
Bayes        -6.3651
Shibata      -6.3755
Hannan-Quinn -6.3718

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
```

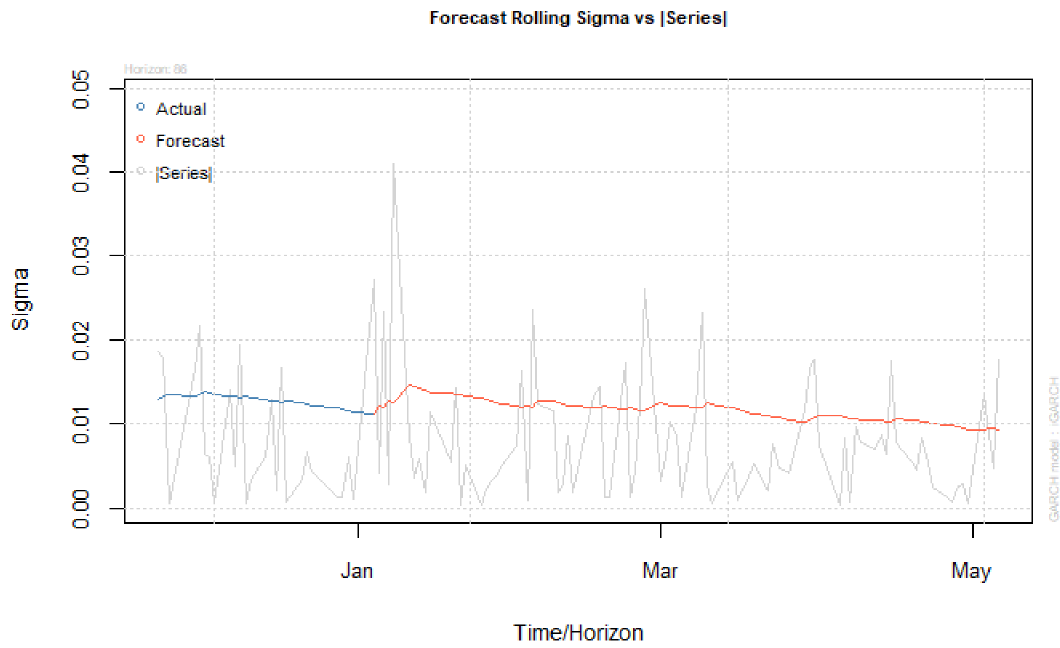
```
                              statistic p-value
Lag[1]                         0.009277  0.9233
Lag[2*(p+q)+(p+q)-1][2]        0.042368  0.9623
Lag[4*(p+q)+(p+q)-1][5]        0.635816  0.9349
d.o.f=0
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                              statistic   p-value
Lag[1]                            19.03 1.287e-05
Lag[2*(p+q)+(p+q)-1][5]           20.48 1.429e-05
Lag[4*(p+q)+(p+q)-1][9]           21.17 9.979e-05
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]     1.727 0.500 2.000  0.1887
ARCH Lag[5]     1.743 1.440 1.667  0.5308
ARCH Lag[7]     1.836 2.315 1.543  0.7520

Nyblom stability test
------------------------------------
Joint Statistic:  218.8641
Individual Statistics:
mu      0.75471
omega  60.83710
alpha1  0.20005
beta1   0.08226
delta   0.21263
shape   1.21420

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:          1.49 1.68 2.12
Individual Statistic:     0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value     prob sig
Sign Bias            0.671 0.502236
Negative Sign Bias   2.702 0.006934 ***
Positive Sign Bias   1.380 0.167562
Joint Effect        12.610 0.005561 ***


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
```

```
  group  statistic  p-value(g-1)
1    20      22.89        0.2420
2    30      26.54        0.5965
3    40      34.49        0.6756
4    50      42.84        0.7200


Elapsed time : 1.622824
```

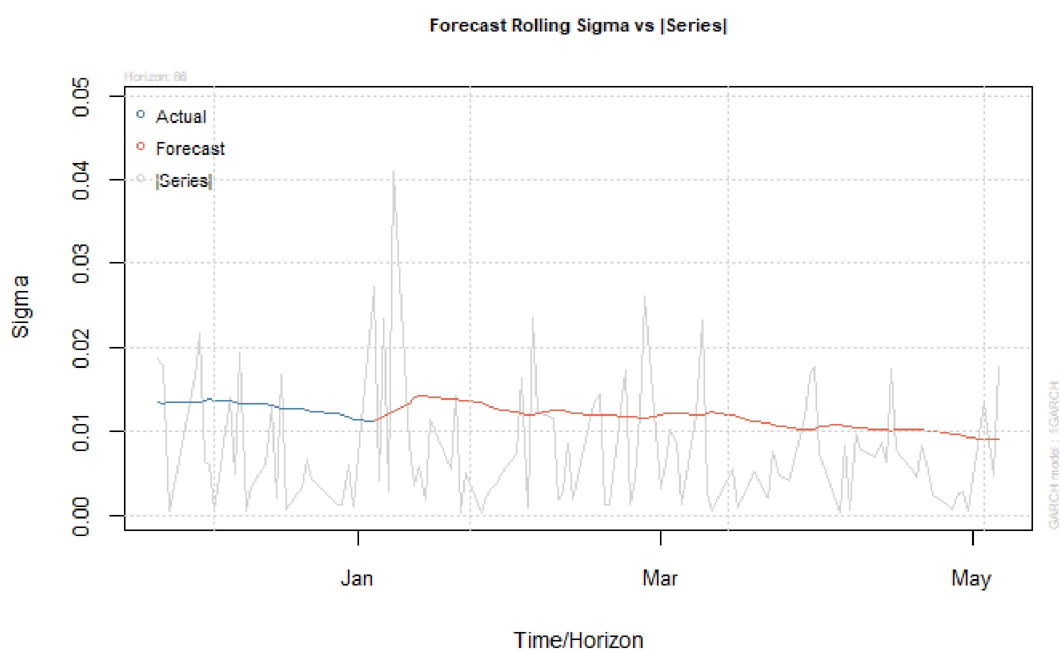
6.4 Sigma Prediction Plot - GARCH(1,1)



6.5 Sigma Prediction Plot - IGARCH(1,1)



6.6 Sigma Prediction Plot - FIGARCH(1,d,1)



6.7 R codes

A .R file is submitted as supplementary which contains all R codes for this project.