Introduction

This article compares several time series model to forecast the daily realized volatility of SP 500 index. The benchmark is ARMA-EGARCH model for SPX daily return series. It is compared to the realized GARCH model of Hansen, Huang and Shek(2012). Finally, an esemble forecasting algorithm is developed.

Assumption

The realized volatility is invisible so we can only estimate it. This is also the hard part for volatility modeling. It is difficult to judge the forecasting quality if the true value is unknown. Nevertheless, researchers develop estimators for the realized volatility.Andersen, Bollerslev Diebold (2008) and Barndorff-Nielsen and Shephard (2007) and Shephard and Sheppard (2009) proposed a class of high frequency based volatility (HEAVY) models.

Assumption: The HEAVY realized volatility estimator is unbiased and efficient. There is no model misspecification.

In the following, HEAVY estimator is taken as *observed realized volatility* to determine the forecasting performance.

Source of Information

- SPX daily data (close-close return)
- SPX intraday high frequency data (HEAVY model estimation)
- VIX
- VIX derivatives (VIX futures)

In this article, I mainly focus on the first two.

Data Collection

Realized Volatility Estimation and Daily Return

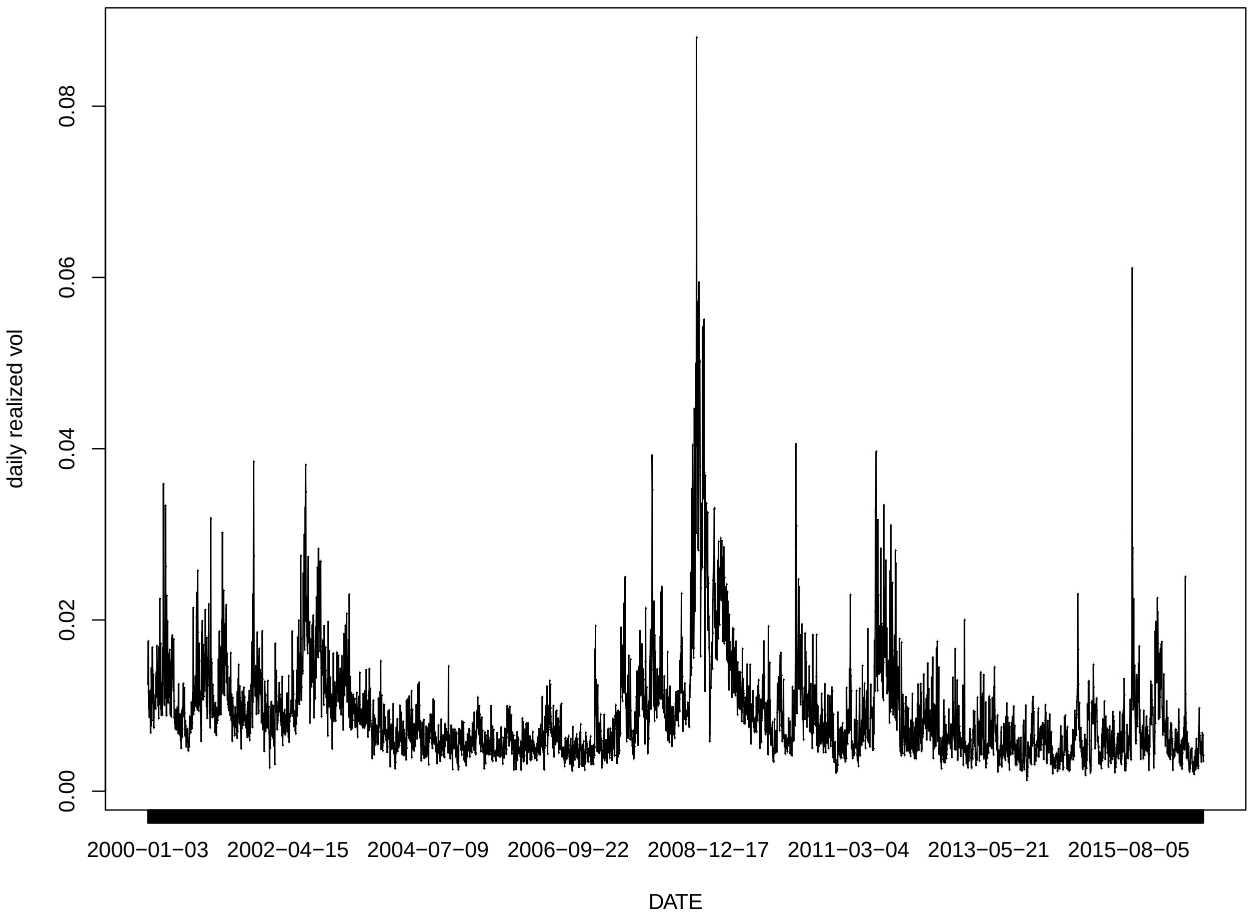
Oxford-Man Institute of Quantitative Finance maintains a Realized Library which publishs the real-time daily realized volatility estimation for equity indices and commodities. I take their publishment as the source of SPX realized volatility estimation and daily return.

```r
library(lubridate)
SPXdata <- read.csv("SPX_rvol.csv")
rownames(SPXdata) <- ymd(SPXdata$DATE)
SPXdata$SPX2.rvol <- sqrt(SPXdata$SPX2.rv)
head(SPXdata)
```

```
               SPX2.rv       SPX2.r     SPX2.rs SPX2.nobs SPX2.open
2000-01-03 0.000157240 -0.010103618 0.000099500      1554  34191.16
2000-01-04 0.000298147 -0.039292183 0.000254283      1564  34195.04
2000-01-05 0.000307226  0.001749195 0.000138133      1552  34196.70
2000-01-06 0.000136238  0.001062120 0.000062000      1561  34191.43
2000-01-07 0.000092700  0.026022074 0.000024100      1540  34186.14
2000-01-10 0.000117787  0.010537636 0.000033700      1573  34191.50
           SPX2.highlow SPX2.highopen SPX2.openprice SPX2.closeprice
```

```
2000-01-03   0.02718625   0.005937756        1469.25         1454.48
2000-01-04   0.04052226   0.000000000        1455.22         1399.15
2000-01-05  -0.02550524   0.009848303        1399.42         1401.87
2000-01-06  -0.01418039   0.006958070        1402.11         1403.60
2000-01-07  -0.02806616   0.026126203        1403.45         1440.45
2000-01-10  -0.01575486   0.015754861        1441.47         1456.74
                 DATE   SPX2.rvol
2000-01-03 2000-01-03 0.012539537
2000-01-04 2000-01-04 0.017266934
2000-01-05 2000-01-05 0.017527864
2000-01-06 2000-01-06 0.011672103
2000-01-07 2000-01-07 0.009628084
2000-01-10 2000-01-10 0.010852972
```

`SPXdata$SPX2.rv` is estimated realized variance. `SPXdata$SPX2.r` is the daily return (close-close). `SPXdata$SPX2.rvol` is the estimated realized volatility



`SPXdata$SPX2.rvol` plot.

Benchmark: SPX daily ret modeling

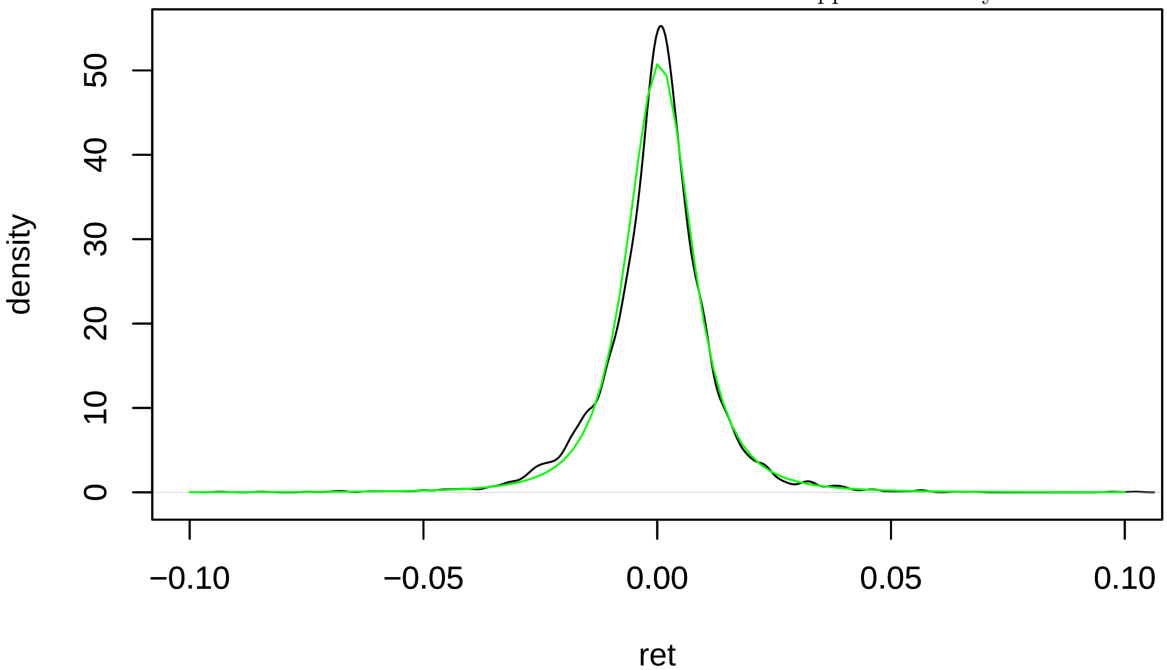
ARMA-eGARCH

Given the daily return with the belief of heteroskedasticity in conditional variance, GARCH model can be the benchmark for fitting and forecasting.

First, the return series is stationary

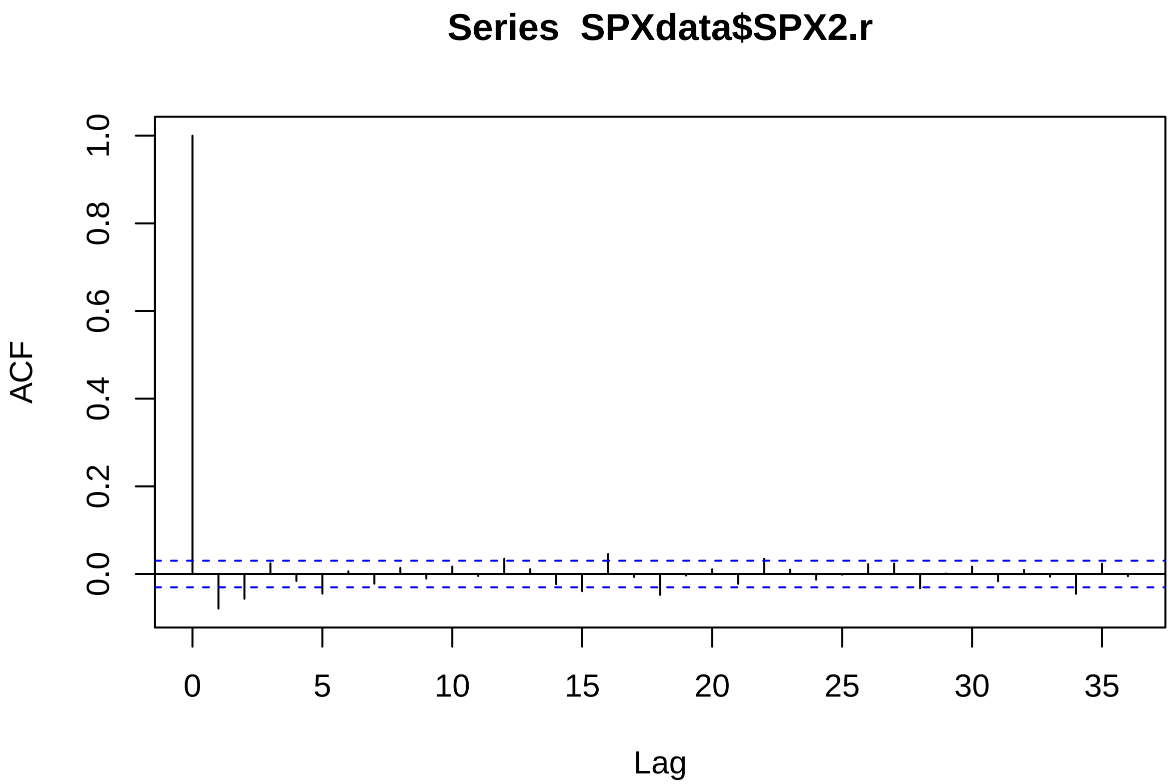
```
	Augmented Dickey-Fuller Test

data:  SPXdata$SPX2.r
Dickey-Fuller = -15.869, Lag order = 16, p-value = 0.01
alternative hypothesis: stationary
```

The return distribution shows an extra kurtosis and fat tail. It can be approximated by a scaled t-distribution



Return distribution density plot. Black line is the kernal-smoothed density and green line is the scaled t-distribution density.

```
acf(SPXdata$SPX2.r)  ## acf plot
```




```
    Box-Ljung test

data:  SPXdata$SPX2.r
X-squared = 26.096, df = 1, p-value = 3.249e-07
```

The autocorrelation plot shows some week correlation in return series. The Ljung-Box test confirms the suspect.

```r
library(forecast)
auto.arima(SPXdata$SPX2.r)
```

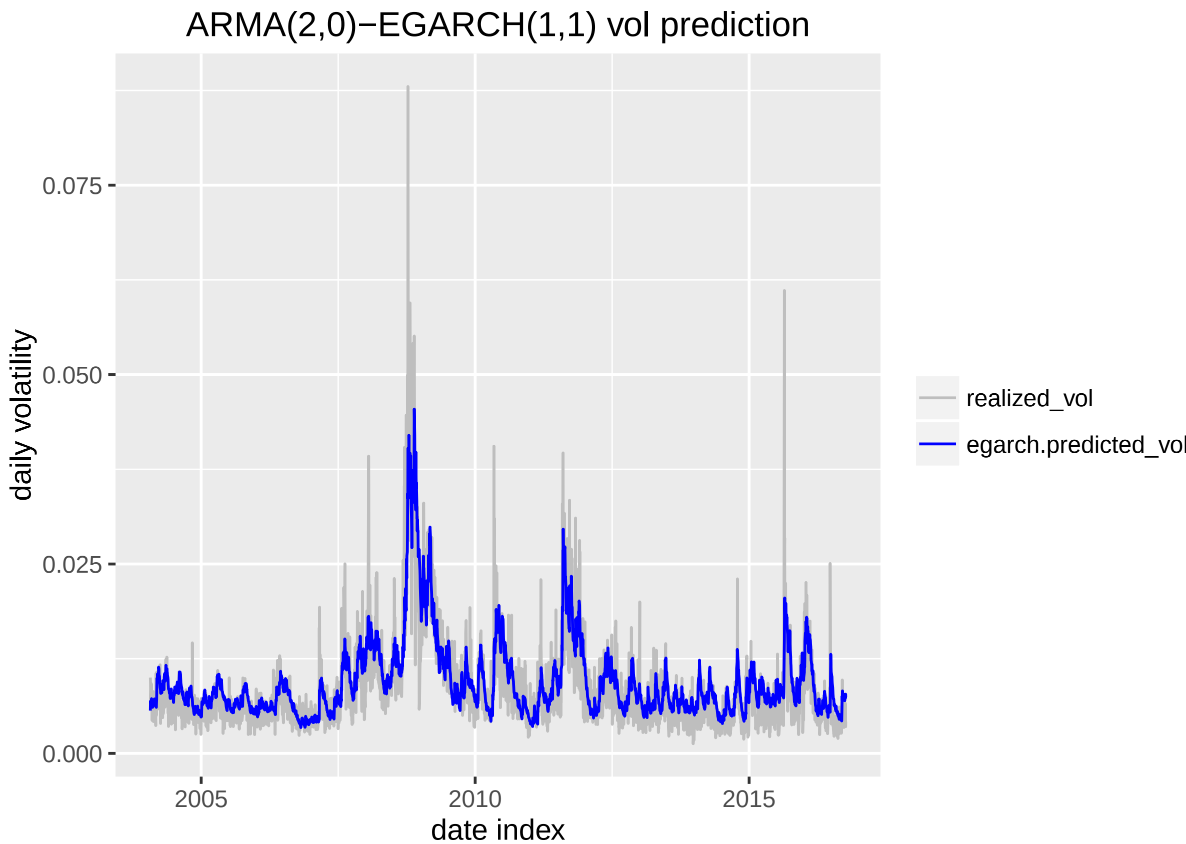
```
Series: SPXdata$SPX2.r
ARIMA(2,0,0) with zero mean

Coefficients:
          ar1      ar2
      -0.0839  -0.0633
s.e.   0.0154   0.0154

sigma^2 estimated as 0.0001412:  log likelihood=12624.97
AIC=-25243.94   AICc=-25243.93   BIC=-25224.92
```

`auro.arima` indicates ARIMA(2,0,0) to model the autocorrelation in return series, and eGARCH(1,1) is popular for volatility modeling. So I choose the ARMA(2,0)-eGARCH(1,1) with t-distribution error, as the benchmark model.

```
egarch_model$spec
```

```
*---------------------------------*
*       GARCH Model Spec          *
*---------------------------------*

Conditional Variance Dynamics
------------------------------------
GARCH Model     : eGARCH(1,1)
Variance Targeting  : FALSE

Conditional Mean Dynamics
------------------------------------
Mean Model      : ARFIMA(2,0,0)
Include Mean        : TRUE
GARCH-in-Mean       : FALSE

Conditional Distribution
------------------------------------
Distribution    :  std
Includes Skew   :  FALSE
Includes Shape  :  TRUE
Includes Lambda :  FALSE
```

With 4189 observations for return (from 2000-01-03 to 2016-10-06), I train the model with the first 1000 observations, then rolling-forecast one ahead each time, and re-estimate the model every 5 observations (roughly 1 week in calendar). The **out-of-sample** forecasting and corresponding realization is in the following plot.



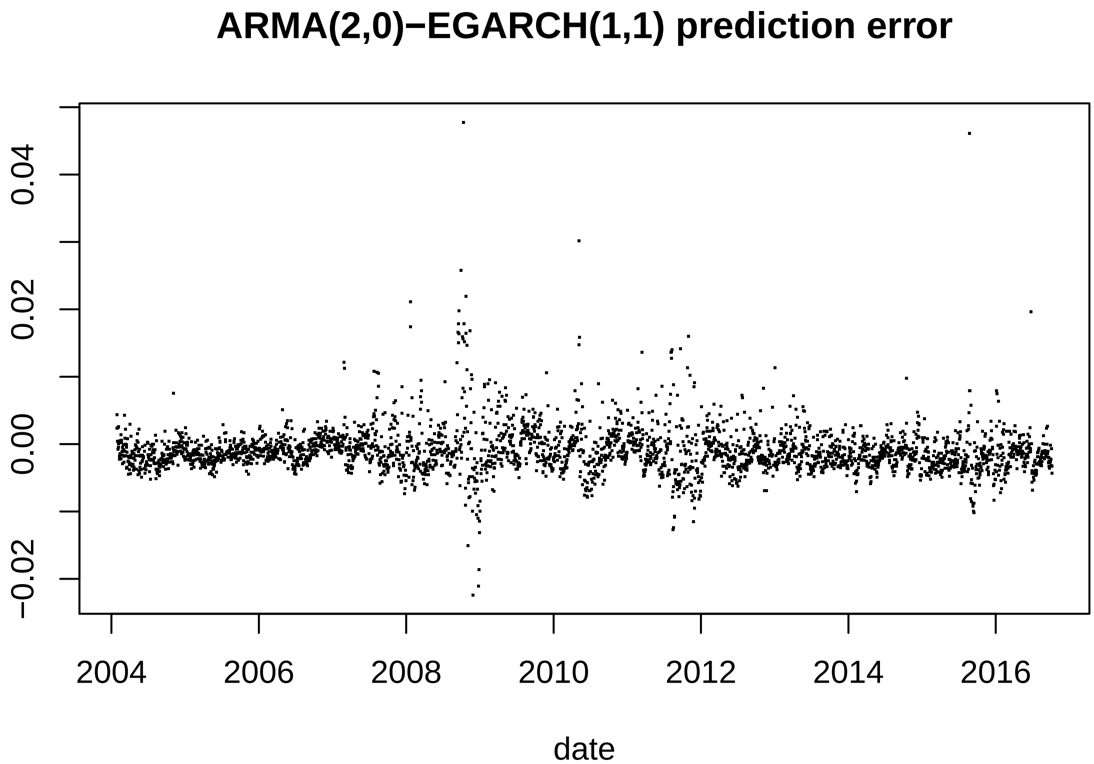
The prediction shows a strong correlation to realization, more than 83%.

```r
cor(egarch_model$roll.pred$realized_vol, egarch_model$roll.pred$egarch.predicted_vol)
```

```
[1] 0.8304458
```

The error summary and plot

```
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.0223800 -0.0027880 -0.0013160 -0.0009501  0.0003131  0.0477600
```



The mean squre of error (MSE):

```
[1] 1.351901e-05
```

For details of the R code, check `GARCH.R`

Improvement: Realized GARCH Model and Long Range Dependence(LRD) Modeling

Realized GARCH

`realGARCH` model is proposed by Hansen, Huang and Shek (2012) (HHS2012) which relates the realized volatility measure to the latent *true volatility* using a representation with asymmetric dynamics. Unlike the standard GARCH model, it is a joint modeling of returns and realized volatility measure(HEAVY estimator in this article). The asymmetric reaction to shocks also makes for a flexible and rich representation.

Formally:

$$y_t = \mu_t + \sigma_t z_t, z_t \sim iid(0,1) log\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i log r_{t-i} + \sum_{i=1}^{p} \beta_i log\sigma_{t-1}^2 log r_t = \xi + \delta log\sigma_t^2 + \tau(z_t) + u_t, u_t \sim N(0,\lambda)$$

It defines the dynamics of return y_t, the latent conditional variance σ_t^2 and realized variance measure r_t. The asymmetric reaction comes via $\tau(.)$

$$\tau(z_t) = \eta_1 z_t + \eta_2(z_t^2 - 1)$$

which has nice property $E\tau(z_t) = 0$. This function also forms the basis for the creation of a type of news impact curve $\nu(z)$

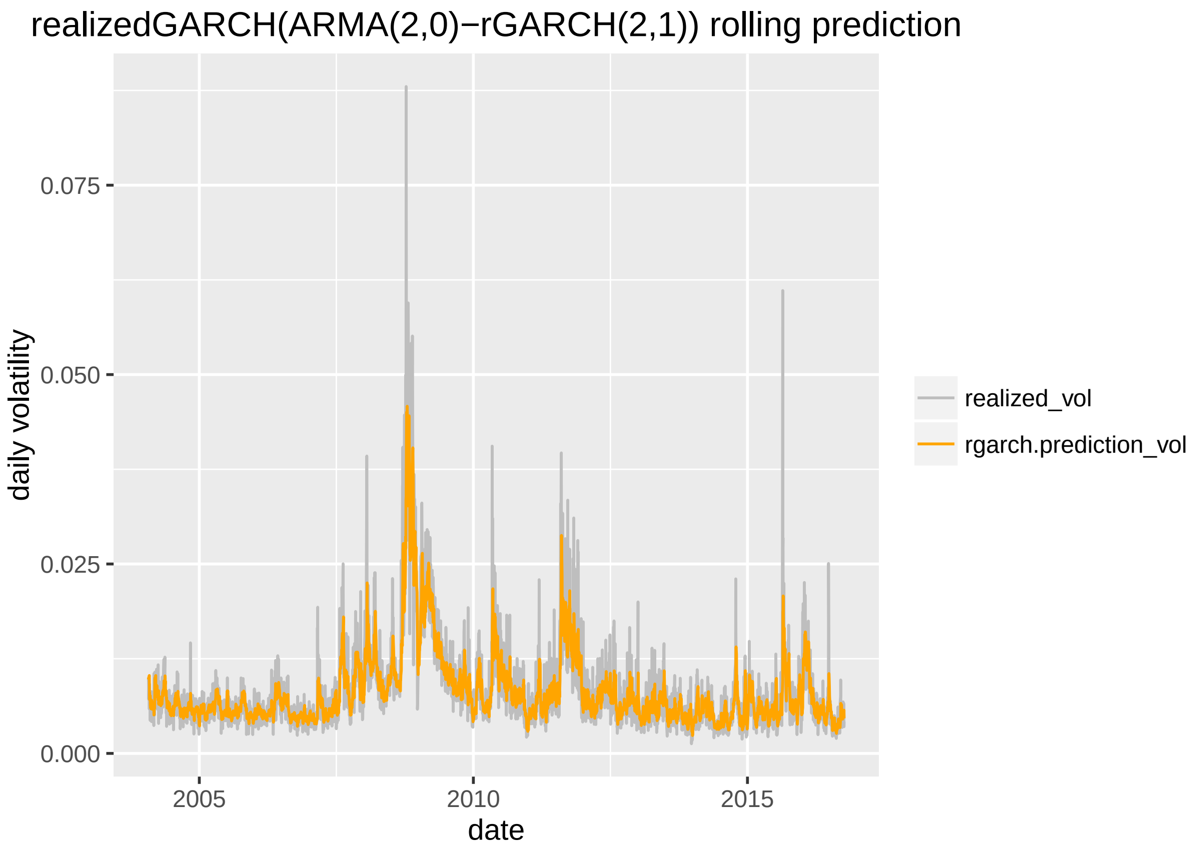
$$\nu(z) = E[log\sigma_t | z_{t-1} = z] - E[log\sigma_t] = \delta\nu(z)$$

so $\nu(z)$ is the change in volatility as a function of the standartized innovations.

The model specification:

```
*---------------------------------*
*       GARCH Model Spec          *
*---------------------------------*

Conditional Variance Dynamics
------------------------------------
GARCH Model     : realGARCH(2,1)
Variance Targeting  : FALSE

Conditional Mean Dynamics
------------------------------------
Mean Model      : ARFIMA(2,0,0)
Include Mean        : TRUE
GARCH-in-Mean       : FALSE

Conditional Distribution
------------------------------------
Distribution    :  norm
Includes Skew   :  FALSE
Includes Shape  :  FALSE
Includes Lambda :  FALSE
```

The rolling-forecast procedure is the same as that of ARMA-EGARCH model above. The **out-of-sample** forecasting and corresponding realization is in the following plot.



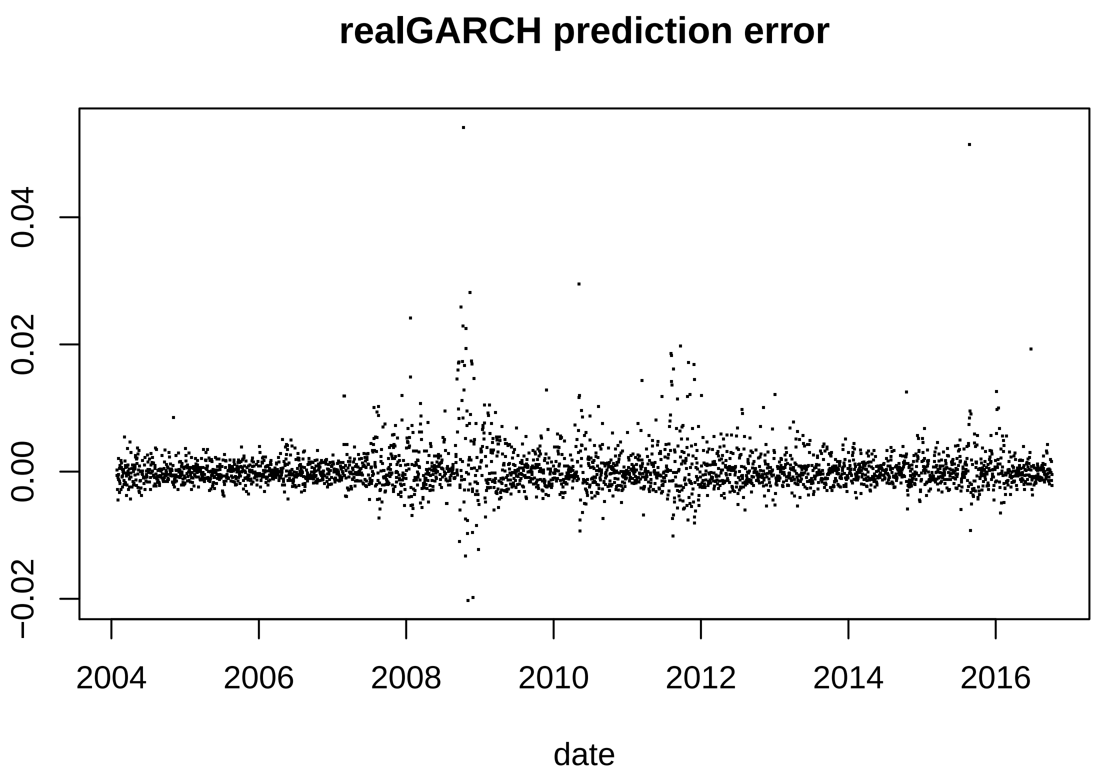
The correlation of forecasting and realization is more than 84%

```
cor(rgarch_model$roll.pred$realized_vol, rgarch_model$roll.pred$rgarch.prediction_vol)
```

```
[1] 0.8424984
```

The error summary and plot:

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```



The mean square of error (MSE):

```
[1] 1.20419e-05
```

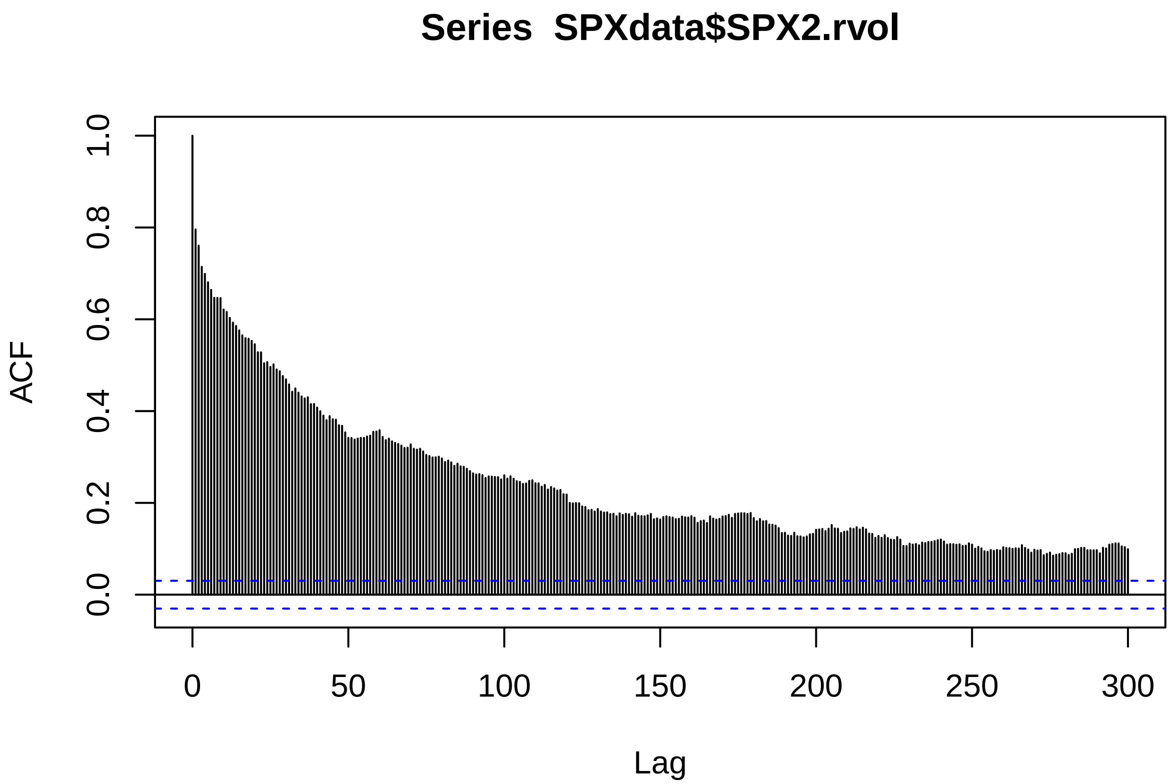
For more details of the R code, check ***`rGARCH.r`***

The LRD modeling: ARFIMA(0,d,0)-eGARCH(1,1)

Since the realized volatility is *"known"*, another idea is to model the realized volatility directly.

The realized volatility acf plot shows a very slow decay in autocorrelation.

```
acf(SPXdata$SPX2.rvol, lag = 300)
```



The double rejection of `adf.test` and `kpss.test` suggests a significant long range dependence (LRD) in the realized volatility series.

```
    Augmented Dickey-Fuller Test

data:  SPXdata$SPX2.rvol
Dickey-Fuller = -5.9652, Lag order = 16, p-value = 0.01
alternative hypothesis: stationary


    KPSS Test for Level Stationarity

data:  SPXdata$SPX2.rvol
KPSS Level = 1.8541, Truncation lag parameter = 14, p-value = 0.01
```

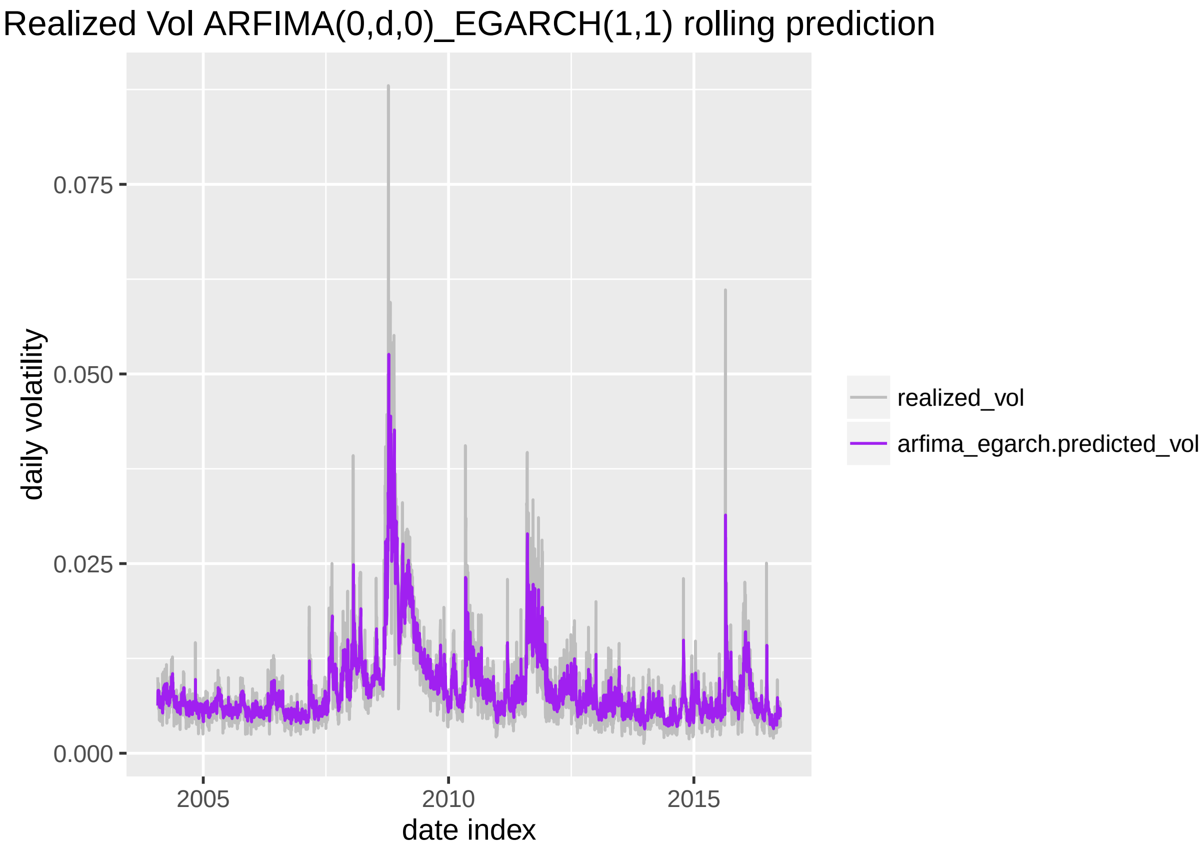
To model the characteristics of LRD, fractional-ARIMA(ARFIMA) model would be a good choice. The model selection based on AICc criteria suggests ARFIMA(0,d,0). So I model the realized volatility by ARFIMA(0,d,0)-eGARCH(1,1).

The model specification:

```
*---------------------------------*
*       GARCH Model Spec          *
*---------------------------------*

Conditional Variance Dynamics
------------------------------------
```

```
GARCH Model     : eGARCH(1,1)
Variance Targeting  : FALSE

Conditional Mean Dynamics
------------------------------------
Mean Model      : ARFIMA(0,d,0)
Include Mean        : TRUE
GARCH-in-Mean       : FALSE

Conditional Distribution
------------------------------------
Distribution    :  norm
Includes Skew   :  FALSE
Includes Shape  :  FALSE
Includes Lambda :  FALSE
```

The rolling-forecast procedure is the same as that of ARMA-EGARCH model above. The **out-of-sample** forecasting and corresponding realization is in the following plot.



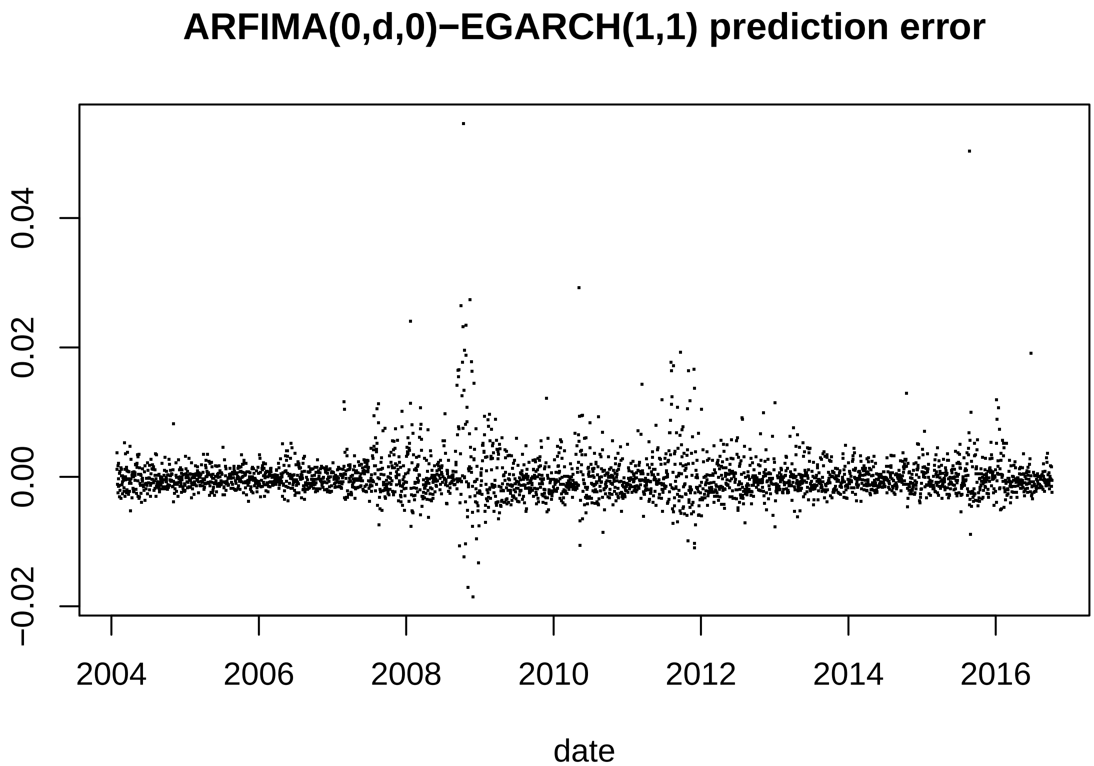
The correlation of forecasting and realization is more than 84%

```r
cor(arfima_egarch_model$roll.pred$realized_vol, arfima_egarch_model$roll.pred$arfima_egarch.predicted_vo
```

```
[1] 0.8440872
```

The error summary and plot:

```
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max. 
-1.851e-02 -1.665e-03 -4.912e-04 -1.828e-05  9.482e-04  5.462e-02 
```



The mean square of error (MSE):

```
[1] 1.18308e-05
```

For more details about the R code, check `rVol_fARIMA.R`

Remark:

- The ARMA-eGARCH model for daily return series and ARFIMA-eGARCH model for realized volatility utilize different information sources. ARMA-eGARCH model only involves the daily return, while the ARFIMA-eGARCH model is based on HEAVY estimator, which is computed from intraday tick data. RealGARCH model combines them.
- ARFIMA-eGARCH model is slightly better performed than realGARCH model, measured by mean squared error. It is probably due to the feature of LRD of ARFIMA-eGARCH model.

Ensemble Model

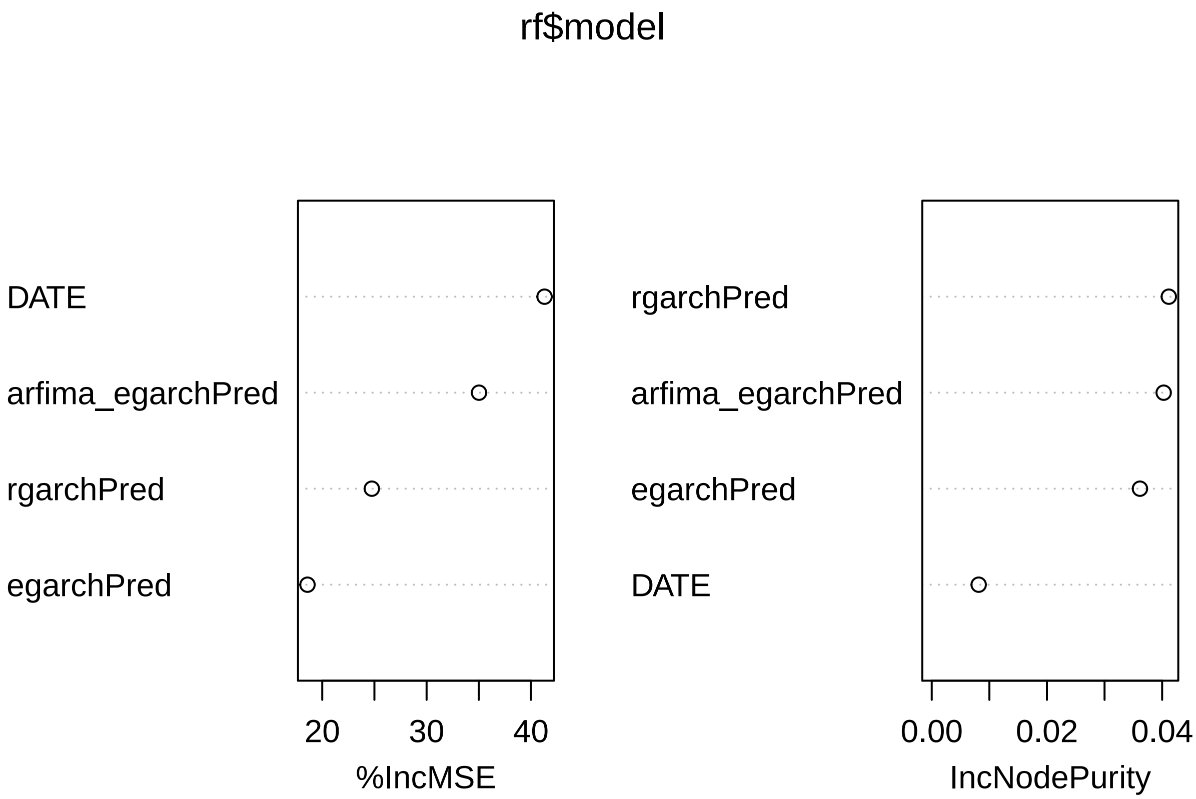
Random Forest Ensemble

Now three forecasting have been constructed

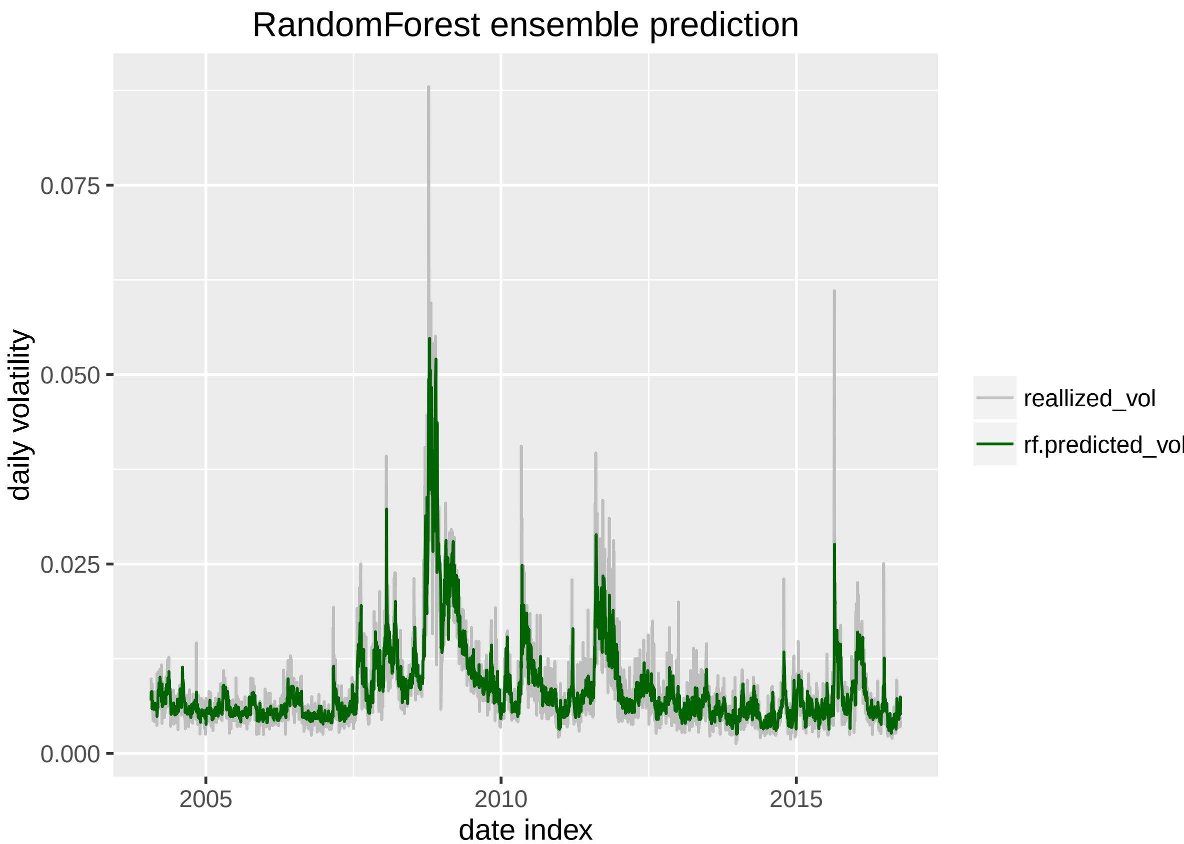
- ARMA-eGARCH `egarch_model`
- realGARCH `rgarch model`
- ARFIMA-eGARCH `arfima_egarch_model`

The model average is expected to reduce forecasting variance, so to improve accuracy, though these three forecasting shows high correlation. The random forest ensemble is employed.

```
randomForest(formula = real ~ ., data = forecasting, ntree = 1000,
    mtry = 2, importance = TRUE)
```



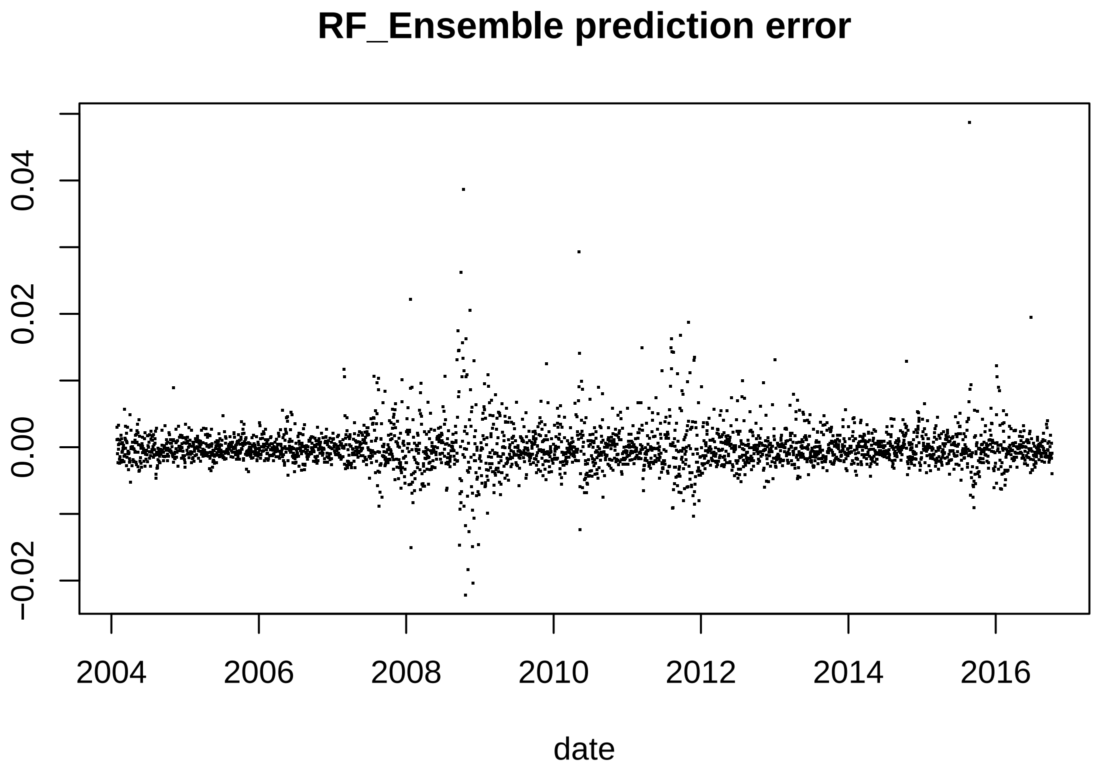
The forest consists of 500 trees, and each tree randomly select 2 forecasting to fit the realizatoin. The following plot is the out-of-bag fitting and realization.



The correlation of forecasting and realizatoin:

```
[1] 0.8510685
```

The error plot:



The mean square error:

```
[1] 1.123454e-05
```

The ratio of MSE to the variance of realized volatility

```
[1] 0.2805191
```

Remarks

The realGARCH model and ARFIMA-eGARCH model which involve the information of realized measure out-performs the standard ARMA-eGARCH model of return series. The MSE of random forest ensemble shrinked by more than 17% compared to the benchmark.

From the view of information source, the realGARCH model and ARFIMA-eGARCH model capture the incremental information in the intraday high frequency data (by model the HEAVY realized volatility estimator)

Further Development: the Implied Volatility

The above methods do not involve the implied volatility data. Implied Volatility is computed from SPX European options. A natural perception is to treat implied volatility as a predictor to forward realized volatility. However, much research shows that VIX, the model free implied volatility is a biased estimator and not as efficient as the forecasts based on past realized volatility. Torben G. Andersen, Per Frederiksen and Arne D. Staal (2007) agree with this view. Their work shows that the introduction of implied volatility to time series analysis frame work gives no significant benefit. However the authors point out the possibility of incremental information in the implied volatility, and suggest a combination model.

So the further development may be an ensemble model which combines the time series forecasting and the prediction information in implied volatility (if there is).

For the codes and related file, please check my GitHub repository (https://github.com/ericwbzhang/Vol_prediction)