Comparison of RV, ARV and WRV Based on MEM

Cao Gang, Guo Mingyuan and Zou Wenjun

School of Management and Economics of Tianjin University, Tianjin, 300072
leu2@163.com

Abstract

Realized Volatility (RV) has been widely used since it was put forward. Some scholars put forward the improved form of RV aiming at better performance in describing the volatility. This paper select 2 main form of improvements: Adjusted Realized Volatility (ARV) and Weighted Realized Volatility (WRV) and make a comparison of RV, ARV and WRV of a contract of stock index of futures in China. Considering 3 methods are all non-negative, the MEM is a proper model. The empirical analysis includes 2 parts: one is statistical information, the other is modeling result. In the first part, the paper shows 4 kinds of statistics and the autocorrelation series; in the second part, the paper choose 3 possible distribution of stochastic error term and modeling MEM for RV, ARV and WRV. Then compare the significance of coefficients and likelihood value of different models.

Keywords: *Realized Volatility (RV), Adjusted Realized Volatility (ARV), Weighted Realized Volatility (WRV), stock index futures*

1. Introduction

Realized Volatility have been widely used since it was put forward by Anderson and Bollerslev in 1998 [1]. It estimates the volatility using the sum of square of returns. It is model free and don't need researchers to estimating the parameters of the model.

Some scholars found that there exist errors when using RV in some situations, so they put forward their improved forms. Xu (2004) [2] put forward "Adjusted Realized Volatility" (ARV) and Guo (2006) [3] put forward "Weighted Realized Volatility". Both authors proved their improved form is better than RV.

After that, ARV and WRV are widely used in empirical analysis, while there isn't a literature does comparison of RV, ARV and WRV.

One important reason is – It is difficult to select a proper model to modeling RV, ARV and WRV simultaneously.

While all these 3 statistics is non-negative, so it is proper to use Multiplicative Error Model (MEM). MEM was put forward by Engle (2002) [4]. The core of this model is – express the non-negative process in a product of a time-varying factor and a stochastic variable which is positive.MEM is a expand form of GARCH, it has the similar form to ACD models. It can be used to analyzing high (ultra-high) time series data.

2. The Basic Form of MEM

Multiplicative Error Model was put forward by Engle (2002) [4]. Its basic form is:

$$v_t \left| I_{t-1} = \mu_t \varepsilon_t, \varepsilon_t \propto i.i.d.p(\varepsilon;\pi), t = 1,2,...,T \right.$$
$$\mu_t = \omega + \sum_{i=1}^{q} \alpha_i v_{t-i} + \sum_{j=1}^{p} \beta_j \mu_{t-j} \quad (1)$$

v_t is a non-negative variable, I_{t-1} stands for the information set of τ_{t-1}, $E(v_t | I_{t-1}) = \mu_t$, $\alpha_j, \beta_j \geq 0, \omega > 0,$ $\sum_{j=1}^{\max(p,q)} (\alpha_j + \beta_j) \leq 1$, $p(\varepsilon;\pi)$ is the probability density function of ε_t.

3. Realized Volatility

3.1. Definition of RV

Assume

$$r_{t,n} = p_{t,n} - p_{t,n-1} (t = 1,2,\cdots T, n = 1,2\cdots N)$$

$p_{t,n}$ is n th logarithmic price in t th day, $r_{t,n}$ is the return of logarithmic price of financial asset ,N is the number of sample taken from [t-1,t]

According to Anderson and Bollerslev（1998）,realized volatility is the sum of square of returns in a trading day, so

$$RV_t = \sum_{n=1}^{N} r_{t,n}^2 \quad (2)$$

3.2. The Character of RV

In Anderson and Bollerslev (2000, 2001, 2003) [5-8],realized Volatility of Western countries' financial markets have following characters:

(1)The distribution of RV and $\sqrt{RV}$ are both" skewed to right" and with "high kurtosis" because of the serial correlation and Heteroscedasticity among the high- frequency yield.

(2) $\sqrt{RV}$ was "skewed to right" and with "high kurtosis" while $\ln\sqrt{RV}$ approximately meet normal distribution.

(3) Though the distribution of yield in a trading day didn't meet normal distribution with "high peak and fat tail", the standardized yield is approximately meet standard normal distribution

4. Adjusted Realized Volatility

The theoretical basis of ARV is the measure error of RV.

4.1. The Measure Error of RV

Realized Volatility (RV) is a method to estimating integrated volatility (IV) ,its basic definition is

$$RV_t = \sum_{j=1}^{n} r\left(t-1+\frac{j}{n},\frac{1}{n}\right)^2 \tag{3}$$

$r(t-1,1)=p(t)-p(t-1)$, it represents a process of logarithmic yield. RV is a unbiased estimator of IV, but there exists measure error when using RV to estimate IV, which can be measured by variance of RV. In Xu (2004) [2], the result is

$$V\left[\sum_{j=1}^{n} r\left(t+\frac{j}{n},\frac{1}{n}\right)^2 / F_t\right] = \frac{1}{n} V\left[r(t,1)^2 / F_{t-1}\right] \tag{4}$$

$r(t,1)=p(t)-p(t-1)$ represents the intraday yield， t is integer which measured by day， $r(t+\frac{j}{n},\frac{1}{n}) = p(t+\frac{j}{n}) - p(t+\frac{t-1}{n})$， $j=1,2,\cdots,n$ represents the daily jth sampling.

(4) indicates that the variance of RV is much smaller than the variance of daily yield, Xu (2004) [2] proved: more times of sampling, much lesser of variance of RV and measure error. When $n\to\infty$， measure error $\to 0$.

4.2. Adjusted Realized Volatility

Xu (2004) [2] put forward "Adjusted Realized Volatility (ARV)"to dealing with the measure error of RV.

His method came from the result of Barndoff-Nielsen. In the conclusion of Barndoff-Nielsen (2002) [9], the relationship between IV and RV is

$$\lim_{n\to\infty}(RV_t - IV_t) \xrightarrow{L} \frac{1}{\sqrt{n}} IQ_t^{1/2}\xi_t \tag{5}$$

$\xi_t \sim N(0,1)$ is stochastic error term, $IQ_t \equiv 2\int_{t-1}^{t}\sigma_s^4 ds$， $IV_t = \int_0^1 \sigma_{t-1+s}^2 ds$. σ_t^2 stands for the variance in t, IQ_t can not be observed directly， it is usually estimated by RQ_t ：

$$RQ_t \equiv \frac{2}{3} n \sum_{j=1}^{n} r\left(t-1+\frac{j}{n},\frac{1}{n}\right)^4 \tag{6}$$

Compute variance of equation（6）,the result is：

$$V(IV_t) = V(RV_t) - \frac{1}{n}E(RQ_t) \tag{7}$$

According to equation (7),Xu took the following formula as ARV：

$$ARV_t = \frac{E(RV_t)\left(V(RV_t)-V(IV_t)\right)}{V(RV_t)} + \frac{V(IV_t)}{V(RV_t)} RV_t \tag{8}$$

Xu (2004) [2] proved the mean of ARV is equal to RV and the variance of ARV is smaller than RV. So in his opinion, ARV is more effective than RV.

5. Weighted realized volatility

According to the definition of RV, it gives an equal weight – 1 to each square of intraday return. But there exists "Calendar Effects" in stock market which means the intraday return was higher in the opening and closing period which smaller in the mid-time of trading day. So it's unreasonable to imposing the identical weights. Guo (2006) [3] took "Calendar effects" into consideration and put forward "Weighted Realized Volatility (WRV)", It aimed at to describing the "Calendar Effects" better.

Definition 1

Weighted Realized Volatility (WRV) is a weighed sum of the square of intraday return of financial assets

$$WRV_t = \sum_{n=1}^{N} w_n r_{t,n}^2 \ , \qquad (9)$$

w_n is the weight of square of intraday returns, it meets

$$\sum_{n=1}^{N} w_n = 1 \qquad (10)$$

The most important work to computing WRV is to computing the weights. Guo put forward 2 conditions:

(1)WRV is the unbiased estimator of IV

(2)WRV is the estimator with least variance.

Assume σ_t^2 represents the volatility of tth day, λ_n is the weights of volatilty in nth period contribute to the volatility of tth day.If $r_{t,n}$ is individual,

$$r_{t,n} = \sqrt{\lambda_n}\sigma_t e_{t,n} \ , \quad e_{t,n} \sim i.i.d.N(0,1) \ ,$$

Combine the formula above with equation（10）,

$$WRV_t = \sum_{n=1}^{N} w_n \lambda_n \sigma_t^2 e_{t,n}^2 \ , \quad \text{so } E(WRV_t) = \sum_{n=1}^{N} w_n \lambda_n \sigma_t^2$$

WRV_t is unbiased estimator of RV_t, so $E(WRV_t) = \sigma_t^2 = IV_t$

then $\sum_{n=1}^{N} w_n \lambda_n = 1$

In computing of condition 2,Guo brought Lagrange function to making the least variance,the result is

$$w_n = \frac{\sum_t \sum_{n=1}^{N} r_{t,n}^2}{N \sum_t r_{t,n}^2} \qquad (11)$$

6. The Comparison of Statistical Information among RV, ARV and WRV

6.1. Data Description

The sample came from a contract of stock index futures market of China. It ranged from Apr.16, 2010-Aug 23, 2011 which involved in 330 trading day. The analysis aims at the daily realized volatility of IF01, we compute it with the data of 5-minute index price.

6.2. Statistical Features of RV, ARV and WRV

(1)The series of RV,ARV and WRV

Table 1 shows the statistical information of 3 original series.

Table 1. Statistical Information of RV, ARV and WRV

	RV	ARV	WRV
Maximum	0.000649	0.000647	0.000203
Minimum	0.00000851	0.00000826	0.0000264
Mean	0.0000911	0.0000911	0.0000911
Std.Dev	0.0000832	0.0000829	0.0000315
skewness	2.544017	2.544158	1.088602
Kurtosis	12.32425	12.32501	4.380817
Jarque-Bera	1551.408	1551.643	91.39455
Prob	0.000000	0.000000	0.000000

From Table 1, it is easy to realizing that 3 original series are all "skew to right with high peak" which meets the first item in the character of RV; their mean is equal which is consistent with the conclusion of Xu (2004) and Guo (2006).

(2) The series of ln RV, ln ARV and ln WRV

Table 2 shows the statistical information of 3 logarithmic series.

Table 2. Statistical Information of ln RV, ln ARV and ln WRV

	ln RV	ln ARV	ln WRV
Maximum	-7.339769	-7.342811	-8.502287
Minimum	-11.67374	-11.64003	-10.54273
Mean	-9.620550	-9.617553	-9.358426
Std.Dev	0.792435	0.787634	0.331620
skewness	0.121150	0.133055	0.092899
Kurtosis	2.720391	2.714764	3.128948
Jarque-Bera	1.882238	2.092401	0.703292
Prob	0.390191	0.351270	0.703529

From Table 2, it's clear that 3 logarithmic series improve their result obviously in skewness, kurtosis and J-B statistics. They are much approximate to normal distribution than their original series which meets the second item of character of RV.

(3) The series of $\sqrt{RV}, \sqrt{ARV}, \sqrt{WRV}$

Table 3 shows the statistical information of $\sqrt{RV}, \sqrt{ARV}, \sqrt{WRV}$

Table 3. Statistical Information of $\sqrt{RV}, \sqrt{ARV}$, $\sqrt{WRV}$

	$\sqrt{RV}$	$\sqrt{ARV}$	$\sqrt{WRV}$
Maximum	0.025479	0.025441	0.014248
Minimum	0.002918	0.002968	0.005137
Mean	0.008817	0.008823	0.009415
Std.Dev	0.003668	0.003654	0.001585
skewness	1.186419	1.190761	0.598508
Kurtosis	4.793541	4.805778	3.387773
Jarque-Bera	121.6483	122.8216	21.76920
Prob	0.000000	0.000000	0.000019

These 3 series are all "skew to right with high peak", but perform better than original series, which meets the first item of character of RV.

(4)The series of standardized yield

Set $r_t / \sqrt{RV_t}$ as standardized yield, so as to ARV and WRV. Table 4 shows the statistical information of 3 series of standardized yield.

Table 4. Statistical Information of 3 Series of Standardized Yield

	$r_t / \sqrt{RV_t}$	$r_t / \sqrt{ARV}$	$r_t / \sqrt{WRV}$
Maximum	4.325713	4.324413	7.127733
Minimum	-4.471764	-4.476732	-8.283165
Mean	-0.036628	-0.036461	-0.068528
Std.Dev	1.558360	1.557153	1.753844
skewness	0.046757	0.046388	-0.415055
Kurtosis	3.138024	3.143562	6.280684
Jarque-Bera	0.382187	0.401738	157.4628
Prob	0.826055	0.818020	0.000000

From Table 4, it is easily to getting the following conclusion: the series of $r_t / \sqrt{RV_t}$ and $r_t / \sqrt{ARV}$ is close enough to normal distribution while the series of $r_t / \sqrt{WRV}$ features with "high peak and skew to left".

6.3. The Autocorrelation of RV, ARV and WRV

This part researches the autocorrelation of RV, ARV and WRV. Figures 1-6 shows the series of 100-order autocorrelation coefficients and partial autocorrelation coefficients.

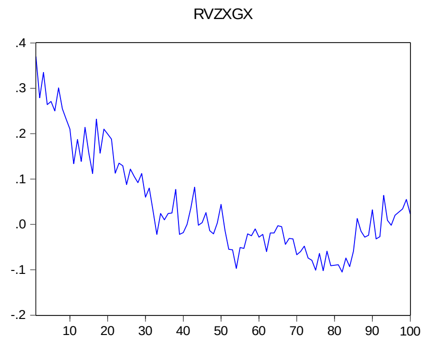

Figure 1. 100-order Autocorrelation Coefficients of RV

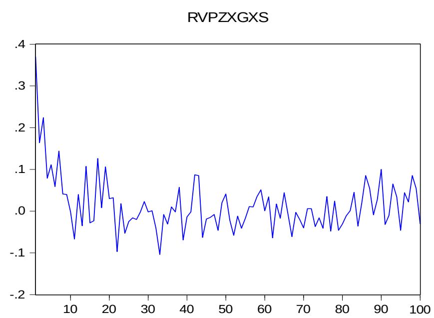

Figure 2. 100-order Partial Autocorrelation Coefficients of RV

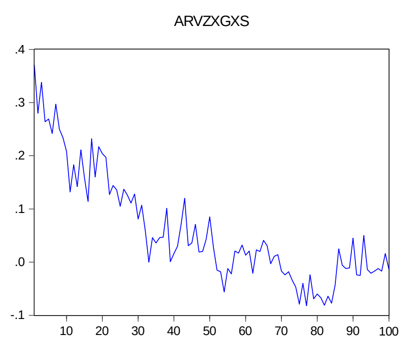

Figure 3. 100-order Autocorrelation Coefficients of ARV

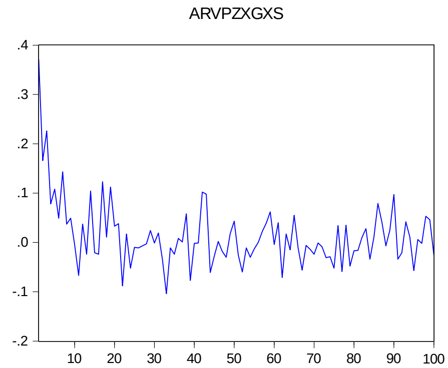

Figure 4. 100-order Partial Autocorrelation Coefficients of ARV

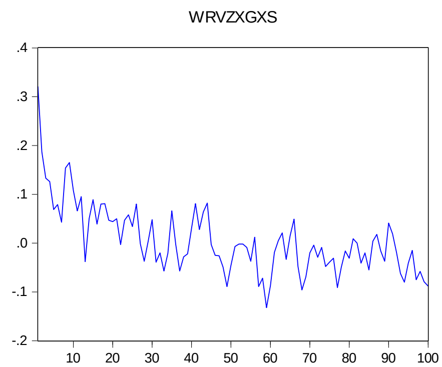

Figure 5. 100-order Autocorrelation Coefficients of WRV

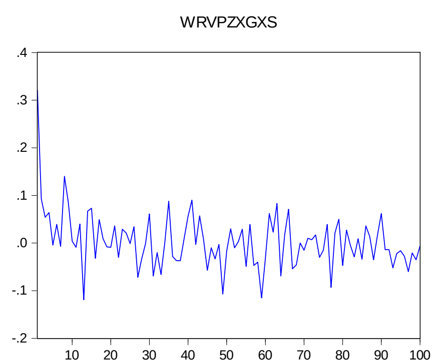

Figure 6. 100-order Partial Autocorrelation Coefficients of WRV

From Figures 1-6 , all of 3 series have character of obvious autocorrelation which meets the first item in character of RV. The trend of autocorrelation and partial autocorrelation coefficients between RV and ARV almost the same While the trend of WRV's is a little different from them. The reason is obvious: the formula of ARV can be divided into 2 parts, the first part is a product of a constant and RV, the second part is a constant. So the trend of autocorrelation and partial autocorrelation coefficients between RV and ARV almost the same. But the formula of WRV break out the structure of RV, so its trend is different from RV and ARV.

6.4. The Summary of Statistical Information of RV, ARV and WRV

Here make a conclusion about the comparison of RV, ARV and WRV :

(1) Both 3 original series are all "skew to right with high peak" while their mean is equal;

(2) 3 logarithmic series improve their result obviously in skewness, kurtosis and J-B statistics. They are much approximate to normal distribution than their original series;

(3) The 3 series of standard deviation are all "skew to right with high peak" ,but perform better than original series;

(4) The trend of autocorrelation and partial autocorrelation coefficients between RV and ARV almost the same While the trend of WRV's is a little different from them.

7. Modeling MEM for RV, ARV and WRV

Before set up the MEM, first problem is the distribution of stochastic error terms.

7.1. The Distribution of Stochastic Error Terms

There are 3 common distributions which meet the expectation of ε_t is 1.

(1) $\varepsilon_t \sim Exp(1)$

Assume $\varepsilon_t \sim Exp(1)$,the density function is

$$\varepsilon_t \sim f(x) = \begin{cases} e^{-x}, x \geq 0 \\ 0, else \end{cases} \quad (12)$$

The density of x_t is ：

$$x_t \sim f(x_t | F_{t-1}) = \frac{1}{\mu_t} e^{-\frac{x_t}{\mu_t}} (x_t \geq 0) \quad (13)$$

(2) $\varepsilon_t \sim Gamma(\lambda, \gamma)$

When X~Gamma (λ, γ)， the density function is：

$$X \sim f(x | \lambda, \gamma) = \begin{cases} \frac{\gamma^{\lambda}}{\Gamma(\lambda)} x^{\lambda-1} e^{-\gamma x}, x > 0 \\ 0, x \leq 0 \end{cases} \quad (14)$$

$\lambda, \gamma > 0$, λ is "shape parameter" and γ is "size parameter", $\Gamma(\bullet)$ is Gamma function. The mean and variance of Gamma (λ, γ) are $E(X) = \frac{\lambda}{\gamma}$, $Var(X) = \frac{\lambda}{\gamma^2}$, because the mean of stochastic error term ε_t is1, so $\lambda = \gamma$ .The density function of ε_t ：

$$\varepsilon_t \left| F_{t-1} \right. \sim Gamma(\lambda, \lambda),\ f(\varepsilon_t \left| F_{t-1} \right.) = \frac{\lambda^{\lambda}}{\Gamma(\lambda)} \varepsilon_t^{\lambda-1} e^{-\lambda \varepsilon_t}, \varepsilon_t > 0, \tag{15}$$

From the linearity of Gamma distribution , if X~Gamma (λ, γ) , a>0 , then Y=aX~ $Gamma(\lambda, \frac{\gamma}{a})$, so $x_t \left| F_{t-1} \right. \sim Gamma(\lambda, \frac{\lambda}{\mu_t})$, the density function of x_t is

$$f(x_t \left| F_{t-1} \right.) = \frac{\lambda^{\lambda}}{\Gamma(\lambda)} x_t^{\lambda-1} \mu_t^{-\lambda} e^{-\lambda \frac{x_t}{\mu_t}} (x_t \geq 0) \tag{16}$$

(3) $\varepsilon_t \sim std.Weibull$

Assume X~ $W(\lambda, \gamma)$, $\lambda, \gamma > 0$, the density function is：

$$X \sim f(x \left| \lambda, \gamma \right.) = \begin{cases} \frac{\lambda}{\gamma^{\lambda}} x^{\lambda-1} e^{-(\frac{x}{\gamma})^{\lambda}}, x \geq 0 \\ 0, x < 0 \end{cases} \tag{17}$$

For meeting the demand of E(X)=1 in MEM, We choose Standard Weibull distribution, then the size parameter γ disappeared, the density function is：

$$\varepsilon_t \sim f(x \left| \lambda, \gamma \right.) = \begin{cases} \lambda[\Gamma(1+\frac{1}{\lambda})]^{\lambda} x^{\lambda-1} e^{-[\Gamma(1+\frac{1}{\lambda})x]^{\lambda}}, x > 0 \\ 0, x < 0 \end{cases} \tag{18}$$

The corresponding density function of x_t is：

$$f(x_t \left| \lambda \right.) = \lambda[\Gamma(1+\frac{1}{\lambda})]^{\lambda} \frac{x_t^{\lambda-1}}{\mu_t^{\lambda}} \exp\{-[\Gamma(1+\frac{1}{\lambda})]^{\lambda} \frac{x_t^{\lambda}}{\mu_t^{\lambda}}\} (x \geq 0) \tag{19}$$

7.2. Modeling Result of RV, ARV and WRV

Table 5 lists the MEM result of RV, ARV and WRV.

Table 5. Modeling Result of RV, ARV and WRV

Variable	Distribution of ε_t	Modeling result	T-statistics	Likelihood value
RV	Exp(1)	None		
	Gamma	$RV_t = \mu_t\varepsilon_t, \varepsilon_t \sim G(1.5341, 1.5341)$ $\mu_t = 0.9877RV_{t-1} + 0.2236\mu_{t-1}$	16.34341 15.71216 11.12122	2732.6005
	Standard Weibull	$RV_t = \mu_t\varepsilon_t, \varepsilon_t \sim W(1.8358)$ $\mu_t = 0.9271RV_{t-1} + 0.2617\mu_{t-1}$	31.85587 14.05512 11.57221	2725.6816
ARV	Exp(1)	$ARV_t = \mu_t\varepsilon_t, \varepsilon_t \sim Exp(1)$ $\mu_t = 0.9964ARV_{t-1} + 0.2169\mu_{t-1}$	12.61212 7.44897	2716.6508
	Gamma	$ARV_t = \mu_t\varepsilon_t, \varepsilon_t \sim G(1.5483, 1.5483)$ $\mu_t = 0.9964ARV_{t-1} + 0.2169\mu_{t-1}$	16.29956 15.85300 10.84792	2733.4485
	Standard Weibull	$ARV_t = \mu_t\varepsilon_t, \varepsilon_t \sim W(1.2019)$ $\mu_t = 0.9319ARV_{t-1} + 0.2574\mu_{t-1}$	31.88473 14.14786 11.40846	2726.3412
WRV	Exp(1)	None		
	Gamma	$WRV = \mu_t\varepsilon_t, \varepsilon_t \sim G(3.3474, 3.3474)$ $\mu_t = 0.9666WRV_{t-1} + 0.2355\mu_{t-1}$	11.71519 10.02686 9.02166	2810.8165
	Standard Weibull	$WRV = \mu_t\varepsilon_t, \varepsilon_t \sim W(1.3636)$ $\mu_t = 0.9813WRV_{t-1} + 0.2044\mu_{t-1}$	23.31570 10.30459 8.15692	2824.2750

Here are the main conclusions of table 5

(1) For RV and ARV, there isn't a proper MEM when $\varepsilon_t \sim Exp(1)$, only ARV is the exception. This may because the exponential distribution of fresh information is far away from the reality.

(2) For RV and ARV, when $\varepsilon_t \sim Gamma$, the coefficient of RV_{t-1} is larger than that in distribution of Standard Weibull, the WRV is also the exception.

(3) For ARV, the coefficient of RV_{t-1} is almost the same when ε_t meets the Exp(1) and Gamma distribution.

(4) For RV, ARV and WRV, when ε_t meets the distribution of Gamma or Weibull, the significance of coefficients and likelihood value performs well.

(5) For each situation of 3 possible distributions, the larger coefficients of RV_{t-1} correspond to the larger likelihood values.

8. Conclusions

Realized Volatility (RV) has been widely used since it was put forward. Some scholars put forward the improved form of RV aiming at better performance in describing the volatility. This paper select 2 main form of improvements: Adjusted Realized Volatility (ARV) and Weighted Realized Volatility (WRV) and make a comparison of RV, ARV and WRV of a contract of stock index of futures in China. Considering 3 methods are all non-negative, the MEM is a proper model .The empirical analysis includes 2 parts: one is statistical information, the other is modeling result. In first part, the paper shows 4 kinds of statistics and the autocorrelation series; in second part, the paper choose 3 possible distribution of stochastic error term and modeling MEM for RV, ARV and WRV. Then compare the significance of coefficients and likelihood value of different models.

References

[1] T. G. Andersen and T. Bollerslev, "Answering the Critics: yes", ARCH Models Do Provide Good Volatility Forecasts, NBER Working Paper, no. 6023, **(1997)**.
[2] Z. G. Xu, "Analysis, Modeling and Application of Financial Market's High Frequency/Ultra High Frequency Time Series", Master Degree Paper of Tianjin University, **(2004)**.
[3] M. Y. Guo, "Analysis, Modeling and Application of volatility of Financial High Frequency Time Series", Doctor Degree Paper of Tianjin University, **(2006)**.
[4] R. Engle, "New Frontier for Arch Models", Journal of Applied Econometrics, vol. 17, Issue 5, **(2002)**, pp. 425-446.
[5] T. G. Andersen and T. Bollerslev, "Exchange Rate Returns Standardized by Realized Volatility are (Nearly) Gaussian", NBER Working Paper, no. 7488, **(2000)**.
[6] T. G. Andersen, "The Distribution of Realized Exchange Rate Volatility", Journal of the American Statistical Association, vol. 96, Issue 453, **(2001)**, pp. 42-55.
[7] T. G. Andersen and T. Bollerslev, "The Distribution of Realized Stock Return Volatility", Journal of Financial Economics, vol. 61, Issue 1, **(2001)**, pp. 43-76.
[8] T. G. Andersen and T. Bollerslev, "Modelling and Forecasting Realized Volatility", Econometrics, vol. 71, Issue 2, **(2003)**, pp. 579 – 625.
[9] O. Bamdorff-Nielsen and N. Shephard, "Econometric analysis of realized volatility and its use in estimating stochastic volatility models", Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 64, Issue 2, **(2002)**, pp. 253- 280.