Methodology

The percentage rate of return of the stock market index is defined as

$$r_t = 100[log(p_t) - \log(p_{t-1})]$$

The GARCH (p, q) model, proposed by Bollerslev (1986) is given by

$$\mathrm{h_t} = \alpha_0 + \sum_{i=1}^{p} \alpha_i\, \varepsilon_{t-j}^2 + \sum_{j=1}^{q} \beta_j\ \sigma_{t-j}^2 \quad (1)$$

Where p and q are the ARCH and GARCH orders. The mean value of returns is positive α+β<1 (close to unity).

The estimation model consists of a mean equation and a variance equation, identified by the ARCH specification, that is, GARCH, EGARCH, GJR and the MRS-GARCH models.

The conditional variance for the GARCH model is

$$h_t = \alpha_0 + \alpha 1 \varepsilon_{t-1}^2 + \beta h_{t-1} \quad (2)$$

The variance (ht) is a function of an intercept (α_0), a shock from the prior period (α) and the variance from the last period (β).

The standard GARCH model cannot explain an asymmetric response to shocks; the EGARCH model of Nelson (1991) was introduced to account for skewness in the financial returns.

The conditional variance for the EGARCH model is

$$\log(h_t) = \alpha 0 + \alpha 1 + \left|\frac{\varepsilon_{t-1}}{h_{t-1}}\right| + \xi \left|\frac{\varepsilon_{t-1}}{h_{t-1}}\right| + \beta_1 \log(h_{t-1}) \quad (3)$$

The conditional variance for the GJR model that accounts for leverage effect is

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \left[1 - \Gamma_{\{\varepsilon_{t-1} > 0\}}\right] + \xi \varepsilon_{t-1}^2 \Gamma_{\{\varepsilon_{t-1} > 0\}} + \beta_1 h_{t-1} \quad (4)$$

The leptokurtic distribution of the stock market returns can be modelled using Student's t or the GED. If the shape parameter v is equal to 2, the GED is the normal distribution. The distribution has thinner or thicker tails if ν is less or greater than 2.

The sample consists of Alsi from July 1995 to July 2011, that is, 4 012 observations. The Alsi returns are calculated by taking the log differences of price indices, multiplied by 100.

The parameter estimates of the different GARCH models are presented in Tables 3a and 3b. The three distributions considered for the study are the Normal, the Student's t and the GED. The in-sample observations are 3 490. The conditional mean parameters of all the GARCH models are significant. The conditional variance estimates show that almost all the parameters are highly significant.

For the Student's t distribution, the degrees of freedom are always greater than 6, confirming the existence of conditional moments until the sixth order. The estimated conditional variances are relatively high at 0,97 for the EGARCH model, showing a high level of persistence.

Table3a**:** GARCH model: Maximum likelihood estimates

	GARCH		
	N	t	GED
δ	0,0896 (0,016)	0,0877 (0,016)	0,0903 (0,016)
α0	0,051 (0,006)	0,0493 (0,008)	0,0519 (0,008)
α1	0,1463 (0,01)	0,1411 (0,015)	0,1457 (0,014)
β1	0,8231 (0,014)	0,8264 (0,015)	0,8214 (0,015)
ξ	-	-	-
υ	-	8,203 (0.902)	1,492 (0.037)
Log(L)	-5402,27	-5335,174	-5360,226
α+β	0,969	0,9675	0,9671

The conditional kurtosis of the Student's t distribution in GARCH, EGARCH and the GJR model are 4,427, 4.27 and 3,93 respectively, confirming the fat tails of the Alsi returns. The conditional kurtosis of the Student's t distribution is calculated as follows: 3(υ-2)/(υ-4). The shape parameters for models with GED distribution lie between 1 and 2, indicating that GED distribution has fatter tails than gaussan. The maximum likelihood estimates are obtained by maximising the log-likelihood with the BFGS (Broyden, Fletcher,Goldfarb and Shanno) quasi-newton optimisation algorithm, using MATLAB.

Table3b: EGARCH and GJR model: Maximum likelihood estimates

	EGARCH			GJR		
	N	t	GED	N	t	GED
δ	0,0642 (0,016)	0,0736 (0,016)	0,0716 (0,016)	0,0699 (0,017)	0,077 (0,016)	0,078 (0,017)
α0	-0,1667 (0,008)	-0,147 (0,014)	-0,159 (0,012)	0,0513 (0,006)	0,049 (0,008)	0,057 (0,01)
α1	0,222 (0,011)	0,1945 (0,018)	0,210 (0,015)	0,1930 (0,013)	0,180 (0,0139	0,188 (0,022)
β1	0,9760 (0,003)	0,9813 (0,004)	0,9787 (0,004)	0,8315 (0,012)	0,8340 (0,014)	0,814 (0,018)
ξ	-0,074 (0,006)	-0,061 (0,011)	-0,068 (0,009)	0,0764 (0,012)	0,0833 (0,016)	0,091 (0,016)
v	-	8,717 (1,032)	1,548 (0,043)	-	8,83 (1,044)	1,447 (0,04)
Log(L)	-5343,687	-5294,41	-5313,00	-5378,68	-5323,68	-5345,59

Markov regime switching model

In the MRS-GARCH model, the parameters switch between a low volatility regime and a high volatility regime. Empirical analysis demonstrates that MRS-GARCH models outperform the standard GARCH models in forecasting volatility at short horizons (Juri Marcucci, 2005). Financial returns display sudden jumps due to structural breaks or changes in regimes. The four elements of the MRS-GARCH model are the conditional mean, the conditional variance, the regime process and the conditional distribution.

The conditional mean equation is

$$r_t = \mu_t^{(i)} + \varepsilon_t = \delta^{(i)} + \varepsilon_t \qquad (5)$$

where i=1,2 and $\varepsilon_t = \eta_t\left[h_t^{(i)}\right]^{1/2}$ with the zero mean and unit variance process.

The conditional variance of returns is

$$h_t = V\{\varepsilon_t|\tilde{s}_t, \varsigma_{t-1}\} \qquad (6)$$

Where ς_{t-1} is the observable information s_t, the current regime. The GARCH (1,1) term for this variance is

$$h_t^{(i)} = \alpha_0^{(i)} + \alpha_1^{(i)}\varepsilon_{t-1}^2 + \beta_1^{(i)}h_{t-1} \qquad (7)$$

Where h_{t-1} is a state-independent average of past conditional variances and depends on observable information, the current regime and all past states. The maximum likelihood estimation is used to estimate the parameters.

The Markov switching model is governed by a state variable S_t that develops according to a first-order Markov chain.The transition probability is

$$\Pr\big(s_t = (j|s_{t-1})\big) = i = p_{ij} \qquad (8)$$

Parameters of a regime switching model switch across regimes, according to a first-order Markov process and the probabilities of the transition matrix with two regimes are as follows:

$$P= \begin{bmatrix} p11 & p21 \\ p12 & p22 \end{bmatrix} = \begin{bmatrix} p & (1-q) \\ (1-p) & q \end{bmatrix} \qquad (9)$$

The unconditional probability of being in the first state is

$$\pi 1 = (1-q)/(2-p-q) \qquad (10)$$

The estimates of the MRS-GARCH model are presented in Table 4. The conditional mean estimates are significant. The standard deviations of the returns σ of the two states reveal that the first state has low volatility compared to the second state.

Table 4: MRS-GARCH model: Maximum likelihood estimates

	MRS-GARCH-N	MRS-GARCH-t2	MRS-GARCH-t	MRS-GARCH-GED
$\delta 1$	0,11033 (0,108)	0,0619 (0,024)	0,0571 (0,024)	0,1118 (0,018)
$\delta 2$	-0,75617 (0,181)	0,0985 (0,020)	0,0971 (0,021)	-0,213 (0,127)
$\sigma 1$	0,01881 (0,011)	0,1039 (0,0130)	0,0887 (0,022)	0,0099 (0,005)
$\sigma 2$	0,71192 (0,22)	0,03345 (0,010)	0,036 (0,011)	0,5389 (0,177)
$\alpha 1$	0,0902 (0,016)	0,1855 (0,052)	0,1669 (0,038)	0,080 (0,014)
$\alpha 2$	0,0965 (0,013)	0,1021 (0,014)	0,1030 (0,015)	0,322 (0,082)
$\beta 1$	0,8269 (0,016)	0,6597 (0,069)	0,6504 (0,06)	0,9068 (0,014)
$\beta 2$	0,88531 (0,127)	0,8817 (0,016)	0,8820 (0,017)	0,4665 (0,121)
p	0,9596 (0,011)	0,9994 (0,001)	0,9994 (0,001)	0,9896 (0,005)
q	0,5656 (0,131)	0,9997 (0,000)	0,9997 (0,00)	0,9337 (0,035)
$\upsilon 1$	-	4,3459(0,672)	7,98 (0,853)	1,578 (0,055)
$\upsilon 2$	-	10,86 (2.02)	-	
Log(L)	-5338,336	-5294,286	-5300,494	-5314,493
ρ 1	0,9171	0,8452	0,8173	0,9868
ρ 2	0,9818	0,9838	0,985	0,7985
π 1	0,914912	0,333333	0,333333	0,864407
π 2	0,085088	0,666667	0,666667	0,135593

$$\sigma^{(i)} = (\alpha_0^{(i)} \,/\, (-\alpha_1^{(i)} - \beta_1^{(i)}))^{1/2}$$

The two regime states indicate that the first regime is characterised by low volitility and the second regime by high volitility. For the GED, the parameter υ is below the threshold value of 2, implying that the distribution has thicker tails than the normal distributions. The persistance of shocks parameter ρ indicates a high level of persistance in both states or regimes for all distributions in the MRS-GARCH model.The shape parameter υ1 is less than two in the case of the GED distribution for the MRS GARCH model, indicating a thick tail distribution. The unconditional probability of being in state or regime 1 is given by $\pi 1 = (1-q)/(2-p-q)$ and the probability of being in the less volatile state 2 is given by $\pi 2$ in Table 4.

According to the Black-Scholes model, returns of assets follow a normal distribution with constant volatility, which is contrary to the stylised fact of financial time series, namely leptokurtic distributions (Mandelbrot, 1963 and Fama, 1965) and the "leverage effect" (where stock price changes are negatively correlated with changes in volatility).

The normal or Gaussian distribution is symmetric and the density function is as follows: $f_x\,(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$ (11)

Where μ and σ^2 are the expected value and the variance of the series and the standard normal distribution with μ=0 and $\sigma^2 = 1$

The Student's t distribution has a density function

$$f_x(x\colon v) = \frac{\Gamma[\frac{v+1}{2}]}{\sqrt{v\pi}\,\Gamma\left[\frac{v}{2}\right](1+x^2/v)^{(v+1)/2}} \qquad (12)$$

Where v is the degrees of freedom (v> 2). The mean of the distribution is $\mu = 0 \;\; for$ v≥ 2 and the variance $\sigma^2 = \frac{v}{v-2}$for $v \geq 3$ and the kurtosis is γ_2=$\frac{6}{v-4}$ for v ≥ 5

The GED has a symmetric distribution and can be both leptokurtic and platykurtic, based on the degrees of freedom v and has the following density function:

$$f_x(x\colon v) = \frac{v e^{-\frac{1}{2}\left|\frac{x}{\lambda}\right|^{v}}}{\lambda 2^{(v+1)/v}\Gamma[\frac{1}{v}]} \qquad (13)$$

And $\lambda = \left[\frac{2^{-2/v}\Gamma[1/v]}{\Gamma[3/v]}\right]^{\frac{1}{2}}$ (14)

For a GED distribution, the tails are thicker when υ<2 and thinner when υ>2.

Maximum-Likelihood Value Comparison

A method of finding the most appropriate model is indicated by the *negative log-likelihood value* at the *maximum point* for the different models. A significant larger likelihood value for a specific distributional assumption in the MLE indicates that this assumption is most likely the best model.

The switching-regime GARCH models present log-likelihood values that are greatly improved compared with their constant coefficient counterparts. The log-likelihood values for the MRS-GARCH models with different conditional distributions, that is for the Normal, the Student's t and the GED, are presented in Table 4. The best result is represented by the MRS-GARCH with Student's t2 innovations, where -5294,28 represents the highest value for the log-likelihood among regime switching models. However, the EGARCH t is considered best with log-likelihood of -5294,41, although the difference is very marginal.

Evaluation of volatility in-sample forecasts

Six types of loss functions are used in the present study to evaluate the in-sample forecasts of the competing models. Table 1 and 2 in the Appendix report the in-sample goodness of fit statistics for two different samples. Table 1 confirms that the results favour the regime switching model when the sample period is close to a recessionary phase. In Table 2 the largest log-likelihood among the GARCH models is given by the EGARCH model with Student's *t* distribution. The Akaike Information Criteria (AIC) and Schwarz Criteria (SC) indicate that the EGARCH with Student's t errors is the best, followed by the EGARCH with GED errors. Among the MRS-GARCH models, MRS-GARCH-t2 is the best, using the AIC, LOGL and the MAD2 and R2LOG criteria. The in-sample forecast evaluation is conducted by minimising the following statistical loss functions. Different competing models are ranked according the six measures used in the study. The MSE, the MAD, R2LOG and HMSE are used for the analysis.

$$MSE1 = n^{-1}\sum_{t=1}^{n} (\hat{\sigma}_{t+m} - \hat{h}_{t,t+m}^{1/2})^2 \tag{15}$$

$$MSE2 = n^{-1}\sum_{t=1}^{n} (\hat{\sigma}_{t+m} - \hat{h}_{t,t+m})^2 \tag{16}$$

$$QLIKE = n^{-1}\sum_{t=1}^{n} (\log \hat{h}_{t,t+m} + \hat{\sigma}_{t+m}^{2}\hat{h}_{t,t+m}^{-1}) \tag{17}$$

$$R2LOG = n^{-1}\sum_{t=1}^{n} [\log(\hat{\sigma}^2_{t+m}\hat{h}^{-1}_{t,t+m})]^2 \quad (18)$$

$$MAD1 = n^{-1}\sum_{t=1}^{n} \left|\hat{\sigma}_{t+m} - \hat{h}^{1/2}_{t,t+m}\right| \quad (19)$$

$$MAD2 = n^{-1}\sum_{t=1}^{n} \left|\hat{\sigma}^2_{t+m} - \hat{h}_{t,t+m}\right| \quad (20)$$

$$HMSE = n^{-1}\sum_{t=1}^{n} (\hat{\sigma}^2_{t+m}\hat{h}^{-1}_{t,t+m} - 1)^2 \quad (21)$$

Table 2 (Appendix) shows that the AIC and the SC both indicate that the best model among the constant-parameter GARCH is the EGARCH with Student's *t* errors, while among the regime-switching GARCH models and overall the best model is the regime-switching GARCH with Student's *t* (with 2 degrees of freedom) innovations. The MAD criteria in (14) and (15) are generally more robust to the possible presence of outliers than the (MSE criteria in (1) and (2).The R2LOG is the logarithmic loss function and the HMSE is the heteroscedasticity adjusted MSE. Regarding the in-sample fit, it can be gleaned from Table 2 that the regime-switching GARCH model performs better. Actually both the MSE criteria and the mean absolute deviation ones suggest that the smallest values are reached with the GARCH models with Student *t* innovations. Table 2 suggests that EGARCH specification is best because the conditional variance of stock returns indeed responds differently to "positive" and "negative" shocks.

3. Conclusion

This study contributes to the literature on volatility modelling by using the Alsi daily data for South Africa, an emerging-market economy. Estimation and comparison of alternative GARCH-type models (symmetric and asymmetric GARCH Models) were conducted with different distributions.

The largest log-likelihood value among the constant-coefficient GARCH models is given by the EGARCH model with Student's t innovations, while for the regime-switching GARCH models the best value is given by the model with Student's t errors.

This study compares the performance of symmetric GARCH, asymmetric EGARCH and MRS-GARCH models with three error distributions (normal, Student-t and generalised error distribution). Successful volatility model forecasting depends, to a large extent, on the choice of error distribution than the choice of GARCH models. This study can be extended to any volatile time series. The above models could be applied to intra-daily data.

References

Black, F. 1976. Studies of stock market volatility changes, Proceedings of the American Statistical Association, Business and Economic Statistics Section, 177–81.

Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307–27.

Burkhard Raunig, Johann Scharler. Stock Market Volatility and the Business Cycle, MONETARY POLICY & THE ECONOMY Q2/2010.

Engle, R. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50, 987–1007.

Glosten, L, Jagannathan, R, and Runkle, D. 1993. On the relation between expected return on stocks, *Journal of Finance*, 48, 1779–801.

Klaassen, F 2002. Improving GARCH volatility forecasts. *Empirical Economics*, *27(2)*, 363–394.

Marcucci, J. 2005. Forecasting stock market volatility with regime-switching GARCH models. *Studies in nonlinear dynamics and economics,* Vol. 9, Issue 4.

Nelson, D. 1991. Conditional heteroskedasticity in asset returns: a new approach, *Econometrica*, 59, 349–70.

Appendix A Table 1: In-sample goodness of fit statistics(1995-2010)

	k	PERS	AIC	R	BIC	R	LOG L	R	MSE1	R	MSE2	R	R2LOG	R	MAD1	R	MAD2	R
GARCH -N	4	0.96	2.926	13	2.935	13	-3932.139	13	0.713	13	20.371	11	6.95	13	1.514	12	0.621	13
GARCH -t	5	0.954	2.873	7	2.884	6	-3858.622	7	0.701	8	20.186	9	6.909	9	1.492	7	0.615	9
GARCH-GED	5	0.955	2.886	10	2.897	9	-3876.785	10	0.703	10	20.203	10	6.915	10	1.497	10	0.617	11
EGARCH-N	5	0.971	2.891	11	2.902	10	-3883.049	11	0.668	3	19.344	1	6.846	6	1.456	5	0.608	4
EGARCH-t	6	0.976	2.848	1	2.861	1	-3824.630	1	0.662	1	19.656	5	6.822	3	1.438	1	0.603	1
EGARCH-GED	6	0.973	2.861	4	2.874	2	-3842.397	4	0.664	2	19.485	2	6.821	2	1.445	2	0.605	3
GJR-N	5	0.957	2.914	12	2.925	12	-3914.28	12	0.713	12	20.162	8	6.929	12	1.522	13	0.621	12
GJR-t	6	0.953	2.869	6	2.882	5	-3853.025	6	0.699	7	19.998	6	6.89	7	1.495	9	0.614	8
GJR-GED	6	0.954	2.88	8	2.894	8	-3868.097	9	0.703	9	20.008	7	6.897	8	1.503	11	0.616	10
MRS-GARCH-N	10	0.965	2.883	9	2.904	11	-3867.001	8	0.671	5	19.508	3	6.92	11	1.451	4	0.611	7
MRS-GARCH-t2	12	0.976	2.854	**2**	2.880	4	-3826.626	**2**	0.669	4	19.582	4	6.814	**1**	1.447	3	0.604	**2**
MRS-GARCH-t	11	0.985	2.866	5	2.890	7	-3843.953	5	0.694	6	20.753	12	6.84	5	1.483	6	0.611	6
MRS-GARCH-GED	11	0.991	2.856	3	2.880	3	-3829.656	3	0.708	11	23.114	13	6.825	4	1.494	8	0.610	5

K: number of parameters

Persistence of shocks to the conditional variance or volatility

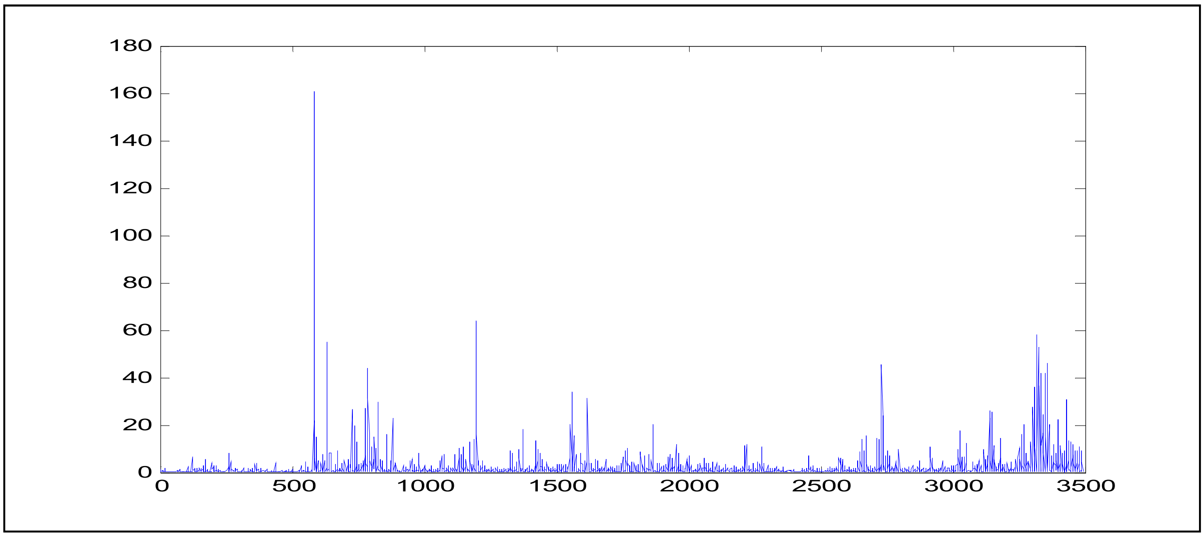
R:RANK; AIC: Akaike Information Criteria, BIC:Schwarz criterion

Table2: In-sample goodness of fit statistics (1995–2011)

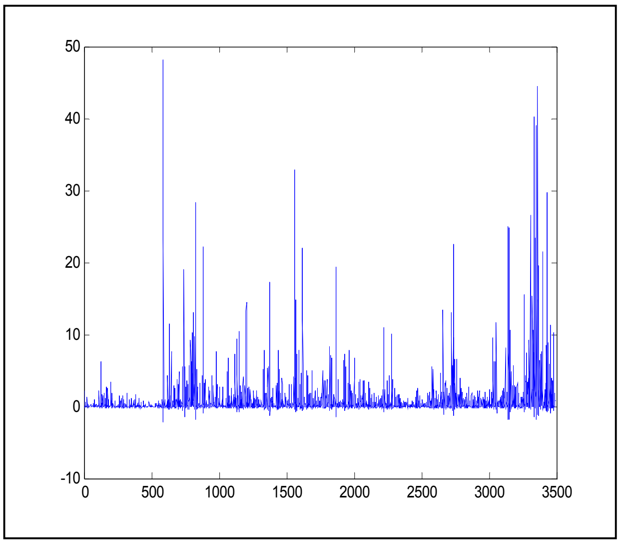
	NUMPAR	PERS	AIC	R	BIC	R	LOGL	R	MSE1	R	MSE2	R	QLIKE	R	R2LOG	R	MAD2	R	MAD1	R	HMSE	R
GARCH-N	4.000	0.969	3.096	12	3.103	12	-5402.272	12	0.847	12	23.322	11	1.255	11	6.832	12	1.857	11	0.677	12	4.046	7
GARCH-t	5.000	0.968	3.059	7	3.067	6	-5335.174	7	0.839	9	23.216	10	1.255	12	6.790	9	1.842	8	0.673	9	4.173	10
GARCH-GED	5.000	0.967	3.069	10	3.078	9	-5353.151	10	0.840	10	23.179	9	1.255	10	6.808	11	1.846	10	0.675	10	4.108	8
EGARCH-N	5.000	0.976	3.063	9	3.072	8	-5343.687	9	0.796	3	21.907	1	1.223	1	6.739	6	1.795	4	0.665	5	3.593	1
EGARCH-t	6.000	0.981	3.033	1	3.044	1	-5289.598	1	0.793	1	22.034	3	1.224	3	6.723	2	1.787	1	0.662	1	3.731	3
EGARCH-GED	6.000	0.979	3.044	4	3.054	2	-5308.072	4	0.794	2	21.959	2	1.223	2	6.726	4	1.791	3	0.663	2	3.661	2
GJR-N	5.000	0.966	3.083	11	3.092	11	-5378.682	11	0.841	11	22.983	8	1.243	6	6.802	10	1.861	12	0.676	11	3.745	4
GJR-t	6.000	0.965	3.053	6	3.063	5	-5323.687	6	0.833	8	22.915	7	1.243	7	6.767	8	1.845	9	0.672	8	3.864	6
MRS-GARCH-N	10.000	0.982	3.063	8	3.081	10	-5338.336	8	0.801	4	22.435	4	1.252	9	6.729	5	1.791	2	0.665	6	4.119	9
MRS-GARCH-t2	12.000	0.984	3.039	2	3.060	3	-5294.286	2	0.808	5	22.548	5	1.240	5	6.726	3	1.804	5	0.664	3	4.501	11
MRS-GARCH-t	11.000	0.988	3.050	5	3.070	7	-5314.493	5	0.831	7	24.006	12	1.236	4	6.746	7	1.838	7	0.670	7	3.764	5
MRS-GARCH-GED	11.000	0.985	3.042	3	3.061	4	-5300.494	3	0.813	6	22.756	6	1.244	8	6.698	1	1.811	6	0.665	4	5.117	12

Appendix B

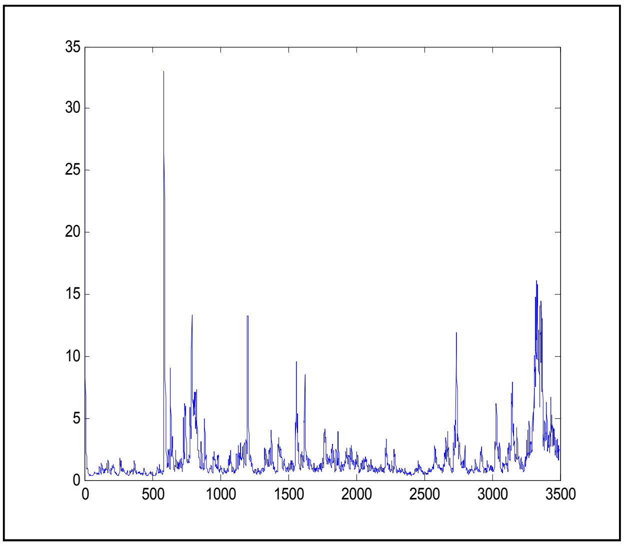
Real errors (squared errors)



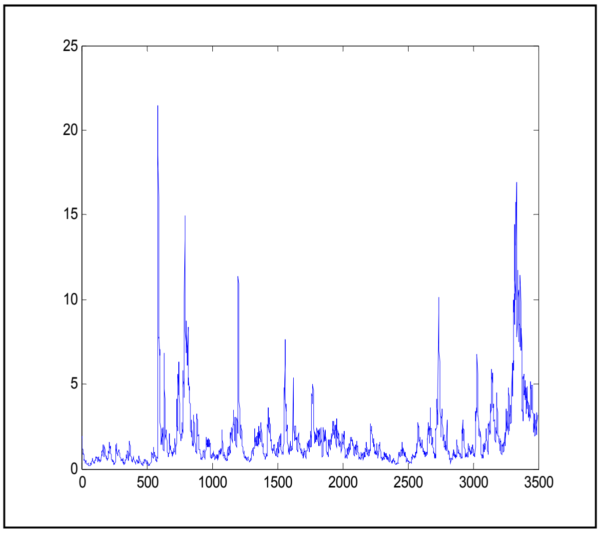
v10-MRS-GARCH-N



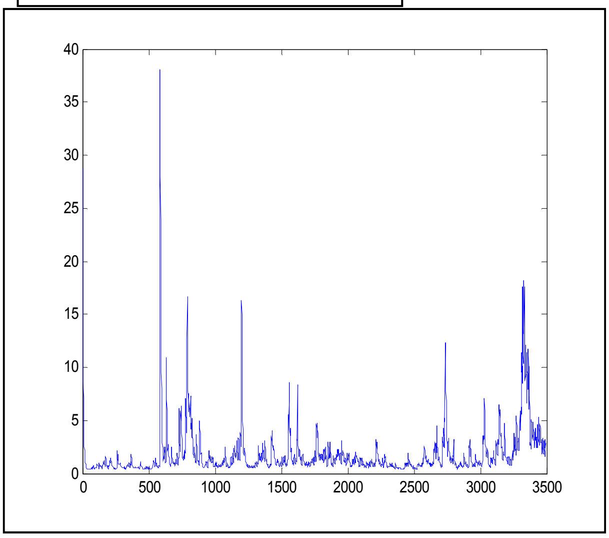
v11-MRS-GARCHt2



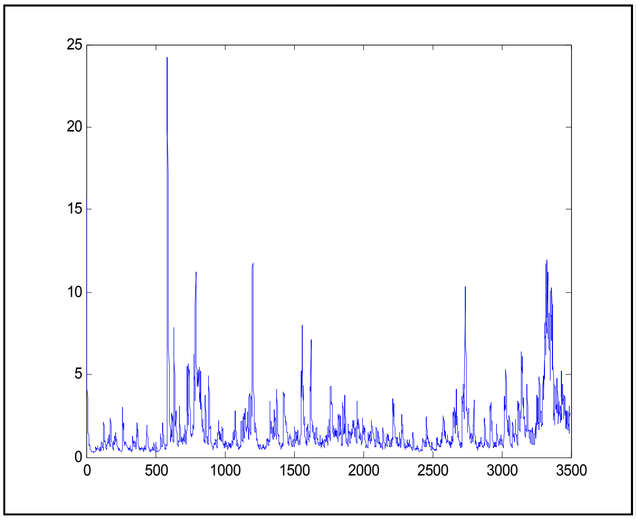
v12-MRS-GARCH-t



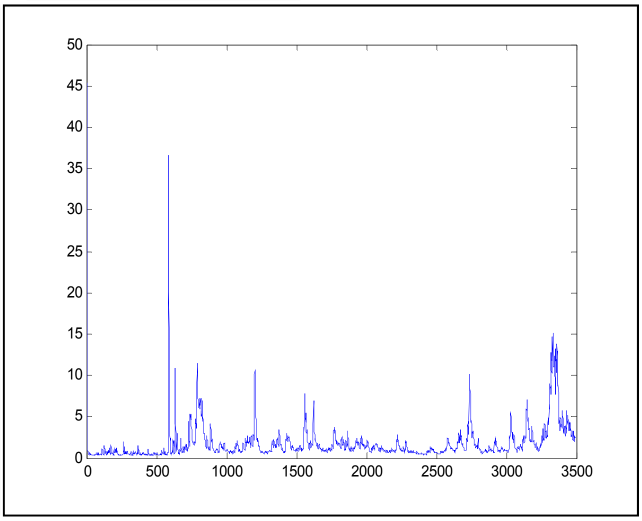
v13-MRS-GARCH-GED



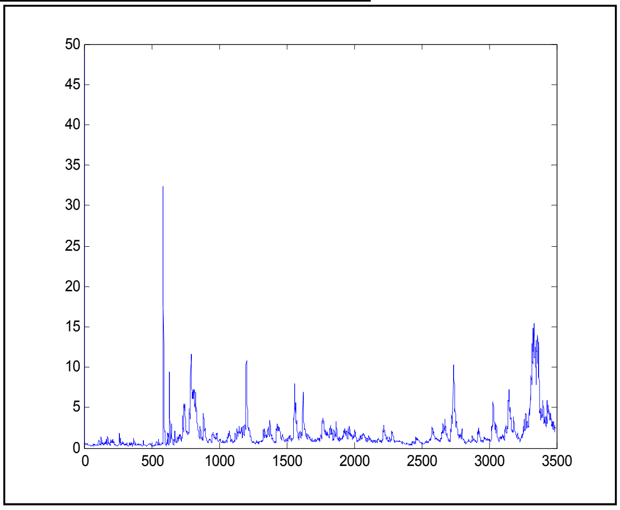
v1:GARCH-N



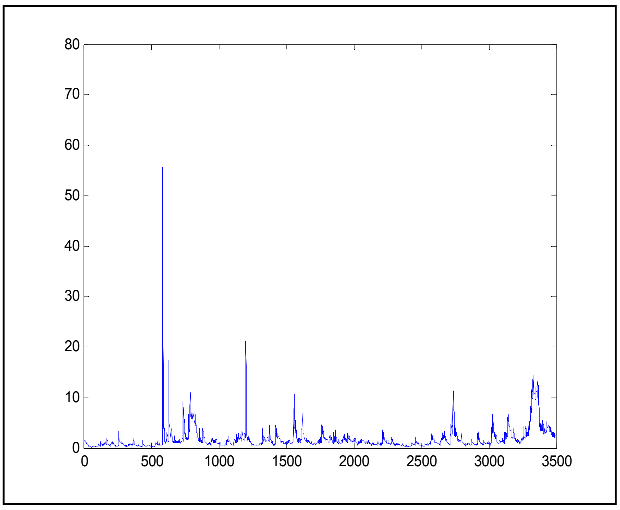
V2-GARCH-t



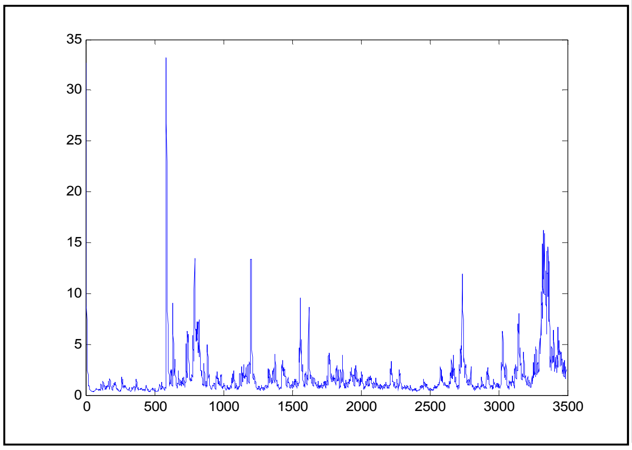
V3-GARCH-GED



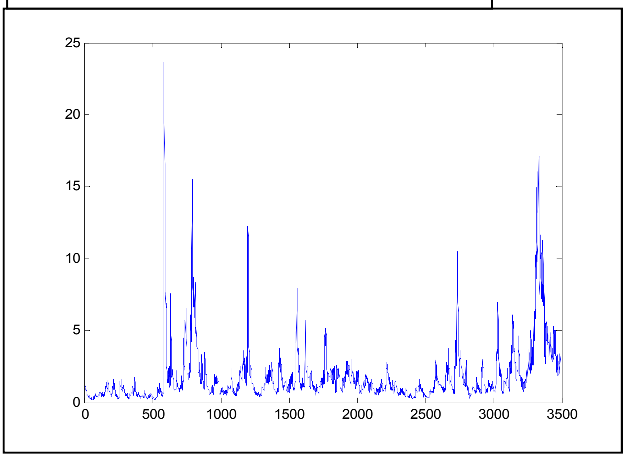
V4-EGARCH-N



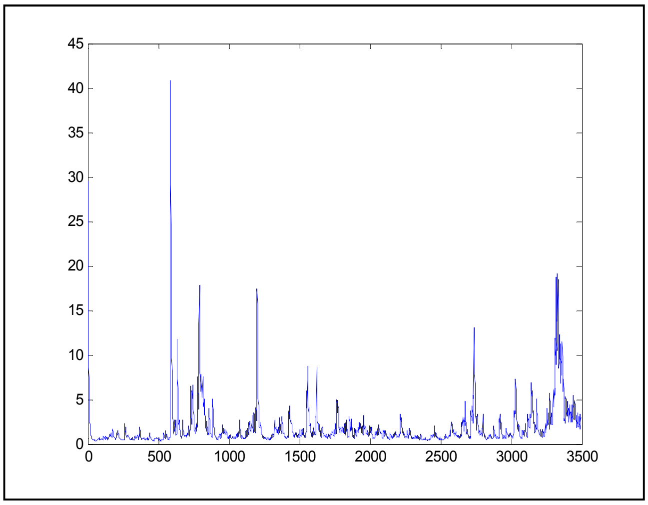
V5-EGARCH-t



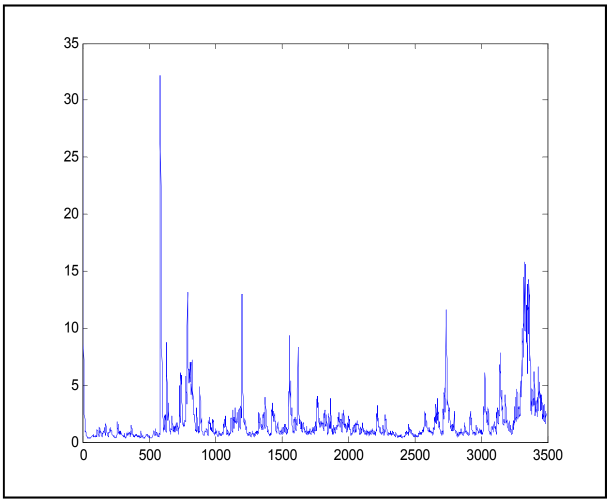
V6-EGARCH-GED



v7-GJR-N



V8-GJR-t



v9-GJR-GED