APPLICATION OF GARCH MODELS IN FORECASTING THE VOLATILITY OF AGRICULTURAL COMMODITIES*[*]

13, september 2004

Tony GUIDA
Université de Savoie
Master in Finance & Banking

and

Olivier MATRINGE
Economic Affairs Officer
UNCTAD, Commodities Branch

Abstract

This paper examines the forecasting performance of GARCH's models used with agricultural commodities data. We compare different possible sources of forecasting improvement, using various statistical distributions and models. We have chosen to confine our analysis on four indices which are the cocoa LIFFE continuous futures, the cocoa NYBOT continuous futures, the coffee NYBOT continuous futures and the CAC 40, the French major stock index. As one may see the sample of indices is containing a genuine stock index also. The implied goal is to find out if the GARCH models are more fitted for stock indices than for agricultural commodities. The forecasts and the predictive power are evaluated using traditional methods such as the coefficient of determination in the regression of the true variance on the predicted one. We find that agricultural commodities time series could not be used with the same methodology than the financial series. Moreover it is interesting to point out that no real "model leader" was found in this sample of commodities. Finally increased forecast performance is not solely observed using non-gaussian distribution in commodities.

JEL classification : C13, C32, C53, G15

Key Word: GARCH, commodities, volatility, forecasting, risk management

[*] The usual disclaimers applies. The conclusions and analysis of this paper did not involve neither UNCTAD nor the Université de Savoie opinions or responsibilities in any cases.

Introduction

The basic version of the least squares model assumes that the expected value of all errors terms, when squared, is the same at any given point. When the variance of the error terms of some time series are not equal, they are said to suffer from heteroskedasticity. That is the main axis of the ARCH and GARCH models. Numerous researches on time series have shown that the volatility of the returns is partially predictable. Furthermore, some well known stylised facts are common to many financial time series. Among them is the *volatility clustering*, i.e. the fact that, high periods of volatility tend to be followed by small period, *fat tails distribution*, i.e. that extreme values are quite numerous in the distribution of the asset returns, and *volatility asymmetric effect*, i.e. price changes are negatively correlated with volatility changes[2].

Engle (1982) at first proposed to model time varying conditional variance with the AutoRegressive Conditional Heteroskeasticity (ARCH) processes, that use past disturbances to model the variance of the series. At that point the major problem was that high ARCH order models were needed to really fit with dynamic process of the conditional variance, and they were pretty hard to quantify. Bollerslev (1986) succeed in finding a solution to that issue, proposing the Generalized ARCH model, adding the forecasts of the variance of the last period in the mean equation.

Since then, lots of derivatives GARCH models had been developed in order to catch the asymmetric effects like Exponential GARCH (EGARCH) by Nelson (1991) for instance or Threshold ARCH (TARCH) by Zakoian (1994).

These days, the main application of heteroskedasticity models is to forecast volatility, which is usually measured by the standard deviation of the returns. One may easily understand how important these models are in financial risk management. However, considering the duality of commodities (such as an underlying asset for futures contract and as a source of income for producing countries) forecasts in volatility could be also helpful in the development of more effective hedge against adverse prices movements. Likewise, it could be very useful for structured finance projects based on commodities. This paper does not seek to find a monolithic answer to the question, are the GARCH models effective to address soft commodities volatility issues, but to justify with a mathematical and statistical frame that soft commodities have to be analysed by more specifics agricultural-oriented GARCH models.
The rest of this papers is organized as follows the first section address the presentation of the models which are used in this paper. The second section presents the data samples and the underlying methodology used. The third one displays the in-sample estimation results of the tested models. The fourth section shows the results of the models forecasts for every models crossed with each statistical distributions. And finally the last section sum up the findings of this paper and point out some assumptions of further research in forecasting volatility in agricultural commodities.

[2] To have further detail and evidence see Black (1976)

Models presentation

Analysis of risk and uncertainty in financial markets has given rise to techniques that allow modelling of temporal dependencies in the variance. The major improvement in the ARCH and GARCH models compared for instance to ARIMA is the distinction between the conditional and the unconditional second order moment.

The GARCH (1,1) model (Bollerslev 1986) is based on the assumption that forecasts of variance changing in time depend on the lagged variance of the asset. An unexpected increase or decrease in the return at time *t* will generate an increase in the expected variability in the next period. The mathematical formula of the model is the following one :

$$\sigma_t^2 = \omega + \alpha\varepsilon^2_{t-1} + \beta\sigma^2_{t-1}$$

Where ω is the mean, ε^2_{t-1} as the news about volatility from the previous period (the ARCH term), σ^2_{t-1} is the last period forecast variance (the GARCH term).

The EGARCH (1,1) (Nelson 1991) model is based on the assumption that the conditional variance is an exponential function of the variables under analysis, which has the advantage to provide against any aberrant negative value of the conditional variance. The mathematical formula of the model is the following one :

$$Log\ \sigma^2_t = \omega + \beta log\ \sigma^2_{t-1} + \alpha \left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}}$$

The TARCH model is based on the assumption that unexpected changes in the return of the index have different effects on the conditional variance of the asset returns. An unexpected increase is presented as a good news and contributes to the variance with the multiplicator α. Instead of an unexpected decrease which is presented as a bad news and contributes to the variance with the multiplicator $\alpha + \gamma$. The mathematical formula of the model is the following one :

$$\sigma^2_t = \omega + \alpha\varepsilon^2_{t-1} + \gamma\varepsilon^2_{t-1}d_{t-1} + b\,\sigma^2_{t-1}$$

Where the good news ($\varepsilon_t > 0$) or the bad news ($\varepsilon_t < 0$) have a different effect on the conditional variance

$$d_{t-1} = 1 \text{ si } \varepsilon_{t-1} < 0$$
$$d_{t-1} = 0 \text{ si } \varepsilon_{t-1} > 0$$

Data and Methodology

Datasets

The sample of data is made of log returns of futures contracts closing prices. All the returns from the global sample are daily trading days returns computed in a continuous way. On a continuous basis, the price return over a given period can be calculated as the logarithm of the ending period price less the logarithm of the beginning period price. The log prices are given by the following formulas : $R_t = \ln (p_t / p_{t-1})$. We clearly assume that storage costs and convenience yields have relatively small effects on the conditional variance of the concerned commodities.

Data consists in 3392 daily observations of the NYBOT cocoa, LIFFE cocoa, NYBOT coffee and CAC 40 (Paris). It provides to span a 13 years period, from the 01 / 01 / 1991 to 31 / 12 / 2003. The full sample is split into two parts : an in sample and an out sample. In sample composed of 2870 observations from 01/01/1991 until the 31/12/2001 in order to estimate the parameters of each models. An out sample composed of 522 observations from 01/01/2002 until the 31/12/2003, in order to make forecasts. Table 1. displays some descriptives statistics about the different indices.

Table 1. Summary Statistics

	Mean	Median	Max.	Min.	Std.dev.	Skewness	Kurtosis	Jarque-Bera
CAC 40	0.00031	3.76E-06	0.2000	-0.1338	0.01512	0.4424	16.7248	26733.63
LIFFE (cocoa)	9.23E-05	0.0000	0.1099	-0.0868	0.01664	0.3707	6.5683	1877.342
NYBOT (cocoa)	8.13E-05	0.0000	0.0996	-0.1001	0.01915	0.2801	5.6001	999.919
NYBOT (coffee)	9.97E-05	0.0000	0.2377	-0.1503	0.02672	0.4114	10.3378	7705.625

The table shows that skewness and kurtosis are clearly observed in the four indices, which is a confirmation of the stylised fact related to fat tails and extreme values with high frequencies data. All the kurtosis exceed 3, which is the normal value and the positive skewness mean that the right tails is particularly extreme. One may notice that the coffee market on the NYBOT seems to be the most volatile on the considered period regarding standard deviation, maximum and minimum values. On the other hand the French stock index seems to be the less volatile compared to the others, even if it presents the higher kurtosis coefficient.

Methodology

Return series are quite seducing for financial statistics because they show some attractive statistical properties like stationarity. Considering that we assume the returns of a financial asset in a continuous fashion, the formula is given by the following equation:

$$r_t = ln\,(P_t / P_{t-1})$$

We also assume that all the four indices returns follow a martingale process, given by the following equation:

$$r_t = \mu + \varepsilon_t$$

Where μ is the mean value of the return, which is expected to be zero , ε_t is a random component of the model, not autocorrelated in time, with a zero mean value. Furthermore ε_t may be considered as a stochastic process. To sum up, the return in the present will be equal to the mean value of r (i.e. the expected value of r based on past information) plus the standard deviation of r (that is the square root of the variance) times the error term for the present period.

Thick tails can be modelled by assuming a conditional normal distribution for returns only if conditional normality implies that returns are normally distributed on each day and the parameters of the distribution are changing from day to day. However we manage to estimate the different GARCH models with non-gaussian distribution in order to find any improvements of that use. In any case we assume that the variance changes with time decay. This characteristic of the variance is called heteroskedasticity. As a matter of fact, the persistence of the volatility is an evidence of autocorrelation in the variance so that is why we use the Q-statistics with 20 lags in order to find some proof of ARCH terms.

The four indices show some evidence of ARCH effect as judged by the autocorrelations of the square returns. The first order autocorrelation is respectively 0.117, 0.146, 0.082 and 0.066 for CAC 40, COFFEE NYBOT, LIFFE and NYBOT COCOA. They gradually decline respectively to 0.027, 0.098, 0.027 and 0.028 after 20 lags. All of these autocorrelations are not so large but they still are significantly different from zero and positive. P values also corroborate the existence of ARCH effect in the four indices.

Tables 2. Autocorrelation statistics in every indices

CAC 40

	AC	Q-Stat	Prob
1	0.117	46.114	0.000
2	0.035	50.195	0.000
3	0.073	68.278	0.000
4	0.042	74.393	0.000
5	0.044	81.033	0.000
6	0.041	86.834	0.000
7	0.043	93.244	0.000
8	0.037	97.888	0.000
9	0.041	103.61	0.000
10	0.036	108.03	0.000
11	0.040	113.56	0.000
12	0.028	116.31	0.000
13	0.032	119.69	0.000
14	0.026	121.92	0.000
15	0.015	122.72	0.000
16	0.027	125.26	0.000
17	0.025	127.39	0.000
18	0.020	128.80	0.000
19	0.021	130.27	0.000
20	0.027	132.84	0.000

COFFEE NYBOT

	AC	Q-Stat	Prob
1	0.146	71.944	0.000
2	0.138	136.98	0.000
3	0.123	187.96	0.000
4	0.069	204.07	0.000
5	0.069	220.28	0.000
6	0.064	234.37	0.000
7	0.079	255.84	0.000
8	0.072	273.61	0.000
9	0.065	287.83	0.000
10	0.234	474.96	0.000
11	0.018	476.06	0.000
12	0.029	479.00	0.000
13	0.063	492.48	0.000
14	0.026	494.72	0.000
15	0.052	503.94	0.000
16	0.021	505.52	0.000
17	0.033	509.26	0.000
18	0.026	511.50	0.000
19	0.010	511.83	0.000
20	0.098	544.43	0.000

COCOA LIFFE

	AC	Q-Stat	Prob
1	0.082	23.106	0.000
2	0.043	29.440	0.000
3	0.090	56.728	0.000
4	0.041	62.448	0.000
5	0.019	63.698	0.000
6	0.050	72.270	0.000
7	0.009	72.576	0.000
8	0.004	72.633	0.000
9	0.018	73.680	0.000
10	0.025	75.795	0.000
11	0.025	77.900	0.000
12	0.041	83.535	0.000
13	0.054	93.310	0.000
14	0.014	93.988	0.000
15	0.058	105.40	0.000
16	0.013	105.95	0.000
17	0.024	107.94	0.000
18	0.044	114.50	0.000
19	0.017	115.48	0.000
20	0.027	117.92	0.000

COCOA NYBOT

	AC	Q-Stat	Prob
1	0.066	14.596	0.000
2	0.026	16.970	0.000
3	0.082	39.884	0.000
4	0.045	46.804	0.000
5	0.036	51.177	0.000
6	0.049	59.426	0.000
7	0.070	75.999	0.000
8	0.065	90.471	0.000
9	0.008	90.703	0.000
10	0.052	99.741	0.000
11	0.079	121.05	0.000
12	0.067	136.18	0.000
13	0.022	137.81	0.000
14	0.024	139.84	0.000
15	0.050	148.29	0.000
16	0.020	149.71	0.000
17	0.023	151.48	0.000
18	0.054	161.29	0.000
19	0.068	177.12	0.000
20	0.028	179.88	0.000

Estimation

The in-sample is composed of 2870 daily observations in order to estimate every models (GARCH, TARCH and EGARCH) crossed with gaussian and non-gaussian (GED and Student t) distributions. We estimate the models by the method of maximum likelihood, making the assumption of a conditionally gaussian and non-gaussian distribution of the errors.

In table 3. we compute different statistics in order to estimate the best model at the in-sample stage for every indices. We use three statistics, (LL) as the Log Likelihood, (AIC) as Akaike Info Criterion and (SC) as Schwarz Criterion. Thus, we ranked every statistics compared to the others models and distributions. As the results, the row rank indicates the sum of the different ranking of every statistics, the lower the sum is the better the model . Moreover the definitive ranking is in parentheses in the row "rank".
Hence considering table 3. the best in-sample results are usually achieved by gaussian distribution. The best models and distributions for every indices are the following ones:

- EGARCH Gaussian for NYBOT COCOA
- GARCH Gaussian for NYBOT COFFEE
- TARCH Gaussian for LIFFE COCOA
- GARCH Gaussian for CAC 40

The results of the estimation and statistical verification of the three models crossed with the three distributions in table 4. indicate that the GARCH components of the variance are statistically significant. Concerning the different GARCH models, they all show a sign of inertia in the development of the conditional variance as long as the sum of the $\alpha + \beta$ coefficient is close to 1, except for the GED NYBOT COFFEE GARCH model for which the sum is close to 0.93. The existence of asymmetric effect is confirmed for the CAC 40 and the NYBOT COFFEE indices whichever distribution concerned. As a consequence, the γ coefficient of the CAC 40 and the COFFEE NYBOT are different from zero in both asymmetric models but remain still very low. Regarding the leverage effect, only the French stock index shows one. In fact CAC 40 is the only index in the sample that shows negative significant γ value for EGARCH and positive significant γ value for TARCH whichever the distribution used. The others indices did not fill the conditions for a leverage effect. Interestingly, CAC 40 is the only index who fully suits the asymmetric ARCH models in the in-sample estimation.

Table 3. Statistical Verification of the Models and relative ranking

NYBOT COCOA, in sample output									
		Gaussian			*GED*			*Student-t*	
	GARCH	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**
LL	7515.744 (2)	7515.994 (3)	7511.577 (1)	7602.247 (5)	7602.384 (6)	7599.656 (4)	7606.557 (8)	7606.602 (9)	7603.399 (7)
AIC	-5.234665 (3)	-5.234142 (2)	-5.231064 (1)	-5.294249 (6)	-5.293647 (5)	-5.291746 (4)	-5.297252 (9)	-5.296586 (8)	-5.294355 (7)
SC	-5.226355 (3)	-5.223755 (2)	-5.220677 (1)	-5.283862 (6)	-5.281183 (5)	-5.279282 (4)	-5.286866 (9)	-5.284122 (8)	-5.281890 (7)
Rank	8 (3)	7 (2)	3 (1)	17 (6)	16 (5)	12 (4)	26 (9)	25 (8)	21 (7)

NYBOT COFFEE, in sample output									
		Gaussian			*GED*			*Student-t*	
	GARCH	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**
LL	6505.572 (1)	6527.146 (2)	6529.891 (3)	6696.152 (7)	6707.665 (8)	6715.186 (9)	6668.295 (4)	6677.070 (5)	6686.375 (6)
AIC	-4.530712 (1)	-4.544049 (2)	-4.546962 (3)	-4.662823 (7)	-4.670150 (8)	-4.675391 (9)	-4.643411 (4)	-4.648829 (5)	-4.655313 (6)
SC	-4.522403 (1)	-4.534663 (2)	-4.536575 (3)	-4.652437 (7)	-4.657686 (8)	-4.662926 (9)	-4.633025 (4)	-4.636365 (5)	-4.642849 (6)
Rank	3 (1)	6 (2)	9 (3)	21 (7)	24 (8)	27 (9)	12 (4)	15 (5)	18 (6)

LIFFE COCOA, in sample output									
		Gaussian			*GED*			*Student-t*	
	GARCH	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**
LL	7843.206 (1)	7843.249 (2)	7843.925 (3)	8044.176 (8)	8044.297 (9)	8043.647 (7)	8039.005 (5)	8039.005 (5)	8037.790 (4)
AIC	-5.462861 (3)	-5.462194 (1)	-5.462662 (2)	-5.602213 (9)	-5.601601 (8)	-5.601148 (7)	-5.598610 (6)	-5.598054 (5)	-5.597066 (4)
SC	-5.454552 (3)	-5.451807 (1)	-5.452279 (2)	-5.591826 (9)	-5.589137 (8)	-5.588684 (7)	-5.588223 (6)	-5.585590 (5)	-5.584602 (4)
Rank	7 (2)	4 (1)	7 (2)	26 (9)	25 (8)	21 (7)	17 (6)	15 (5)	12 (4)

CAC 40 , in sample output									
		Gaussian			*GED*			*Student-t*	
	GARCH	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**
LL	8285.906 (1)	8335.156 (2)	8339.488 (3)	8464.815 (4)	8485.517 (5)	8489.290 (6)	8491.407 (7)	8511.814 (8)	8517.123 (9)
AIC	-5.771363 (1)	-5.804987 (2)	-5.808005 (3)	-5.895342 (4)	-5.909071 (5)	-5.911700 (6)	-5.913872 (7)	-5.927397 (8)	-5.931096 (9)
SC	-5.763054 (1)	-5.794600 (2)	-5.797618 (3)	-5.884955 (4)	-5.896607 (5)	-5.899236 (6)	-5.903486 (7)	-5.914932 (8)	-5.918632 (9)
Rank	3 (1)	6 (2)	9 (3)	12 (4)	15 (5)	18 (6)	21 (7)	24 (8)	27 (9)

Table 4. Quantification of the Models for every distribution

In sample analysis with a GED distribution

NYBOT COCOA			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	-0.000350	-0.000358	-0.000371
ω	2.00E-06	-0.133786	1.86E-06
α	0.031170	0.084709	0.028860
β	0.963526	0.991230	0.964016
γ	-	0.006943	0.005109
GED	1.267591	1.262432	1.268050
LIFFE COCOA			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	-1.02E-05	-1.59E-07	-6.01E-06
ω	3.22E-06	-0.276481	3.20E-06
α	0.031461	0.104036	0.032914
β	0.956668	0.975638	0.957972
γ	-	0.010498	-0.006236
GED	1.037742	1.031950	1.040185

NYBOT COFFEE			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	-1.37E-06	7.09E-06	2.09E-06
ω	5.06E-05	-0.426270	4.02E-05
α	0.108747	0.165920	0.156267
β	0.823991	0.958129	0.848366
γ	-	0.079334	-0.117186
GED	0.998425	1.025798	1.047932
CAC 40			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	0.000375	0.000251	0.000281
ω	6.10E-06	-0.238278	4.29E-06
α	0.074726	0.110618	0.004691
β	0.895905	0.981997	0.927425
γ	-	-0.064335	0.091859
GED	1.186238	1.223252	1.219582

** where the tail parameter r > 0 . The GED is a normal distribution if r = 2 , and fat-tailed if r < 2 .

In sample analysis with a student-t distribution

NYBOT COCOA			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	-0.000637	-0.000633	-0.000649
ω	4.36E-06	-0.131825	2.01E-06
α	0.03068	0.082947	0.029382
β	0.964257	0.991250	0.964475
γ	-	0.007609	0.002988
T-DOF	5.543301	5.485440	5.539174
LIFFE COCOA			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	-0.000678	-0.000657	-0.000659
ω	2.95E-06	-0.193624	2.97E-06
α	0.035771	0.099385	0.037608
β	0.958419	0.984631	0.959412
γ	-	0.010484	0.007533
T-DOF	3.569267	3.567262	3.576330

NYBOT COFFEE			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	-0.000516	-0.000349	-0.000294
ω	2.80E-05	-0.359820	3.80E-05
α	0.088365	0.158197	0.151550
β	0.883460	0.965789	0.860461
γ	-	0.071905	-0.106140
T-DOF	3.828877	3.932012	3.915057
CAC 40			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	0.000530	0.000344	0.000387
ω	5.88E-06	-0.257335	4.44E-06
α	0.08144	0.117055	0.005505
β	0.894311	0.980648	0.923435
γ	-	-0.062043	0.090809
T-DOF	6.036203	6.374411	6.359197

In sample analysis with a Gaussian distribution

NYBOT COCOA			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	-5.22E-05	-8.26E-05	-8.53E-05
ω	2.06E-06	-0.143561	1.89E-06
α	0.034264	0.090591	0.031641
β	0.960633	0.990522	0.961295
γ	-	0.008996	0.005590
LIFFE COCOA			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	5.08E-05	9.92E-05	4.05E-06
ω	4.36E-06	-0.294058	4.38E-06
α	0.030990	0.100489	0.030538
β	0.952923	0.973147	0.952273
γ	-	0.002847	0.002405

NYBOT COFFEE			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	-0.000751	-0.000291	-0.000285
ω	1.74E-05	-0.274476	2.22E-05
α	0.074519	0.134321	0.116827
β	0.903893	0.975828	0.900133
γ	-	0.067397	-0.092766
CAC 40			
Statistics	**GARCH**	**EGARCH**	**TARCH**
μ	0.000437	8.22E-05	0.000263
ω	1.16E-06	-0109499	9.73E-07
α	0.03281	0.076895	-0.001040
β	0.959745	0.993700	0.969959
γ	-	-0.053616	0.055584

Forecasts

Forecasts' performance could be evaluated using the coefficients given by the forecasts output. However, those coefficients could not provide an absolute measure of the predictive power which is the main purpose of this paper. Thus, in order to estimate precisely the accuracy of our forecasts using the GARCH models, we will regress with the least square method the variance realized on the forecasted variance. The R^2, the coefficient of determination, will be the percentage of efficiency. The RMSE denotes the Root Mean Squares Errors, MAE denotes Mean Absolute Error, MAPE denotes the Mean Absolute Percent Error, TIC denotes the Theil Inequality Coefficient and R^2 the determination coefficient from the regression of the true variance by the forecasted variance. In parentheses is the rank of the model used.

Table 5. shows the models performance comparison. We compare the performance of every model for every distribution for every indices, in order to find the best model for a given distribution and a given index. Hence, the relative rankings are in parentheses near the actual value of every statistics. The final rank of a model is given in parentheses in the row "rank".

For any given distributions the asymmetric models and especially the EGARCH model give the best results for the CAC 40, which shows the higher percentage of accuracy in the forecast (24.8 %) on a forecasted period of 2 years. With regards with commodities, we may distinguish two groups of results. The first one is only composed of the NYBOT COFFEE, for which the best result is given by the TARCH models whichever the distribution. The higher percentage of accuracy is shown by the TARCH GED models with 7.26 % on a forecasted period of 2 years. The second group is composed of the two cocoa indices. They show very low percentages of accuracy, the best ones are respectively 1.6 % TARCH student-t and 1.2 % EGARCH Gaussian for the LIFFE and the NYBOT. According to the global sample forecasts results, asymmetric models really help to improve the accuracy of the forecasts compared to the symmetric one. EGARCH and TARCH models are almost always first-ranked in table 5.

Table 6. shows the distribution performance comparison. In order to find out if the use of different distribution may lead to some real improvements, we compare the performance of every distribution for every model for every index. Thus, the method of ranking is the same as table 5., and the final rank is still given in parentheses in the row "rank".

Except for the NYBOT COFFEE index all the indices left seemed to have better results with the Gaussian distribution, especially CAC 40 index. According to the only R^2 statistics, the unbiasedness of the forecasts, the Gaussian distribution shows always better results than the two others distributions. As a matter of fact, the use of non-gaussian distribution did not significantly improve the accuracy of the forecasts.

Table 5 , Forecast Performance, Models Comparison

CAC 40, 2 years Forecasts Performance									
	Gaussian			*GED*			*Student-t*		
	GARCH	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**
RMSE	0.018702 (3)	0.018696 (1)	0.018697 (2)	0.019129 (3)	0.019121 (2)	0.019119 (1)	0.019120 (1)	0.019121 (2)	0.019121 (2)
MAE	0.013501 (3)	0.013499 (1)	0.013500 (2)	0.014045 (3)	0.014039 (2)	0.014037 (1)	0.014035 (1)	0.014039 (2)	0.014039 (2)
MAPE	99.64237 (1)	150.7971 (2)	153.4485 (3)	104.5942 (1)	154.3837 (3)	144.5412 (2)	118.1201 (1)	154.9944 (3)	153.8634 (2)
TIC	0.996303 (3)	0.946867 (2)	0.944635 (1)	0.994128 (3)	0.961089 (1)	0.967144 (2)	0.984373 (3)	0.960717 (1)	0.961406 (2)
R^2	0.141487 (3)	0.221199 (2)	0.247714 (1)	0.142721 (3)	0.214022 (2)	0.234571 (1)	0.141395 (3)	0.214176 (2)	0.241867 (1)
Rank	13 (3)	8 (1)	9 (2)	13 (3)	10 (2)	7 (1)	9 (1)	10 (2)	9 (1)

NYBOT COFFEE, 2 years Forecasts Performance									
	Gaussian			*GED*			*Student-t*		
	GARCH	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**
RMSE	0.023191 (1)	0.023194 (2)	0.023193 (3)	0.023199 (1)	0.023222 (3)	0.023199 (1)	0.020203 (1)	0.023195 (3)	0.023192 (2)
MAE	0.016760 (2)	0.016744 (1)	0.016814 (3)	0.016729 (1)	0.016755 (3)	0.016729 (1)	0.015072 (1)	0.016741 (2)	0.016754 (3)
MAPE	94.39752 (2)	94.14059 (1)	95.78427 (3)	93.84410 (2)	93.13095 (1)	93.86002 (3)	93.48654 (3)	94.08935 (1)	94.30763 (2)
TIC	0.982373 (2)	0.990839 (3)	0.955762 (1)	0.999171 (2)	0.976933 (1)	0.999686 (3)	0.999952 (3)	0.992553 (2)	0.985311 (1)
R^2	0.022435 (3)	0.071626 (1)	0.036770 (2)	0.020768 (3)	0.072572 (1)	0.026606 (2)	0.004041 (3)	0.070187 (1)	0.027753 (2)
Rank	10 (2)	8 (1)	12 (3)	9 (1)	9 (1)	10 (2)	11 (3)	9 (1)	10 (2)

NYBOT COCOA, 2 years Forecasts Performance									
	Gaussian			*GED*			*Student-t*		
	GARCH	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**
RMSE	0.022116 (3)	0.022110 (1)	0.022114 (2)	0.022130 (3)	0.022127 (2)	0.022121 (1)	0.022133 (2)	0.022143 (3)	0.022115 (1)
MAE	0.016293 (1)	0.016294 (3)	0.016293 (1)	0.016294 (2)	0.016294 (2)	0.016293 (1)	0.016295 (2)	0.016296 (3)	0.016293 (1)
MAPE	94.63083 (3)	94.02703 (1)	94.46915 (2)	95.66502 (3)	95.46278 (2)	95.01516 (1)	95.84197 (2)	96.44008 (3)	94.49766 (1)
TIC	0.960599 (1)	0.971228 (3)	0.963220 (2)	0.945344 (1)	0.948036 (2)	0.954461 (3)	0.943192 (2)	0.936038 (1)	0.962756 (3)
R^2	0.006676 (3)	0.009659 (2)	0.012295 (1)	0.006949 (3)	0.008729 (2)	0.010091 (1)	0.002370 (3)	0.007285 (2)	0.009373 (1)
Rank	11 (3)	10 (2)	8 (1)	12 (3)	10 (2)	7 (1)	11 (2)	12 (3)	7 (1)

LIFFE COCOA, 2 years Forecasts Performance									
	Gaussian			*GED*			*Student-t*		
	GARCH	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**	**GARCH**	**TARCH**	**EGARCH**
RMSE	0.019117 (2)	0.019117 (2)	0.019111 (1)	0.019137 (1)	0.019139 (2)	0.019140 (3)	0.019172 (3)	0.019147 (2)	0.019113 (1)
MAE	0.013519 (1)	0.013519 (1)	0.013529 (3)	0.013522 (1)	0.013522 (1)	0.013523 (3)	0.013537 (3)	0.013525 (2)	0.013518 (1)
MAPE	93.15373 (1)	93.16095 (2)	93.57357 (3)	92.76720 (1)	92.78722 (2)	92.80587 (3)	93.82929 (3)	92.97324 (1)	93.34209 (2)
TIC	0.983128 (1)	0.983485 (2)	0.995427 (3)	0.959133 (3)	0.957358 (2)	0.956463 (1)	0.936372 (1)	0.951346 (2)	0.992567 (3)
R^2	0.011538 (2)	0.011734 (1)	0.006495 (3)	0.000748 (2)	0.000489 (3)	0.002264 (1)	0.006066 (3)	0.016388 (1)	0.000947 (2)
Rank	7 (1)	8 (2)	13 (3)	8 (1)	10 (2)	11 (3)	13 (3)	8 (1)	9 (2)

Table 6. Forecast Performance, Distribution Comparison

CAC 40, 2 years Forecasts Performance									
		GARCH			*TARCH*			*EGARCH*	
	Gaussian	**GED**	**Student-t**	**Gaussian**	**GED**	**Student-t**	**Gaussian**	**GED**	**Student-t**
RMSE	0.018702	0.019129	0.019120	0.018696	0.019121	0.019121	0.018697	0.019119	0.019121
	(1)	(3)	(2)	(1)	(2)	(2)	(1)	(2)	(3)
MAE	0.013501	0.014045	0.014035	0.013499	0.014039	0.014039	0.013500	0.014037	0.014039
	(1)	(3)	(2)	(1)	(2)	(2)	(1)	(2)	(3)
MAPE	99.64237	104.5942	118.1201	150.7971	154.3837	154.9944	153.4485	144.5412	153.8634
	(1)	(2)	(3)	(1)	(2)	(3)	(2)	(1)	(3)
TIC	0.996303	0.994128	0.984373	0.946867	0.961089	0.960717	0.944635	0.967144	0.961406
	(3)	(2)	(1)	(1)	(3)	(2)	(1)	(3)	(2)
R^2	0.141487	0.142721	0.141395	0.221199	0.214022	0.214176	0.247714	0.234571	0.241867
	(2)	(1)	(3)	(1)	(3)	(2)	(1)	(3)	(2)
Rank	8 (1)	11 (2)	11 (2)	5 (1)	12 (3)	11 (2)	6 (1)	11 (2)	13 (3)

NYBOT COFFEE, 2 years Forecasts Performance									
		GARCH			*TARCH*			*EGARCH*	
	Gaussian	**GED**	**Student-t**	**Gaussian**	**GED**	**Student-t**	**Gaussian**	**GED**	**Student-t**
RMSE	0.023191	0.023199	0.020203	0.023194	0.023222	0.023195	0.023193	0.023199	0.023192
	(2)	(3)	(1)	(1)	(3)	(2)	(2)	(3)	(1)
MAE	0.016760	0.016729	0.015072	0.016744	0.016755	0.016741	0.016814	0.016729	0.016754
	(3)	(2)	(1)	(2)	(3)	(1)	(3)	(1)	(2)
MAPE	94.39752	93.84410	93.48654	94.14059	93.13095	94.08935	95.78427	93.86002	94.30763
	(3)	(2)	(1)	(3)	(1)	(2)	(3)	(1)	(2)
TIC	0.982373	0.999171	0.999952	0.990839	0.976933	0.992553	0.955762	0.999686	0.985311
	(1)	(2)	(3)	(2)	(1)	(3)	(1)	(3)	(2)
R^2	0.022435	0.020768	0.004041	0.071626	0.072572	0.070187	0.036770	0.026606	0.027753
	(1)	(2)	(3)	(2)	(1)	(3)	(1)	(3)	(2)
Rank	10 (2)	11 (3)	9 (1)	10 (2)	9 (1)	11 (3)	10 (2)	11 (3)	9 (1)

NYBOT COCOA, 2 years Forecasts Performance									
		GARCH			*TARCH*			*EGARCH*	
	Gaussian	**GED**	**Student-t**	**Gaussian**	**GED**	**Student-t**	**Gaussian**	**GED**	**Student-t**
RMSE	0.022116	0.022130	0.022133	0.022110	0.022127	0.022143	0.022114	0.022121	0.022115
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(3)	(2)
MAE	0.016293	0.016294	0.016295	0.016294	0.016294	0.016296	0.016293	0.016293	0.016293
	(1)	(2)	(3)	(1)	(1)	(3)	(1)	(1)	(1)
MAPE	94.63083	95.66502	95.84197	94.02703	95.46278	96.44008	94.46915	95.01516	94.49766
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(3)	(2)
TIC	0.960599	0.945344	0.943192	0.971228	0.948036	0.936038	0.963220	0.954461	0.962756
	(3)	(2)	(1)	(3)	(2)	(1)	(3)	(1)	(2)
R^2	0.006676	0.006949	0.002370	0.009659	0.008729	0.007285	0.012295	0.010091	0.009373
	(2)	(1)	(1)	(1)	(2)	(3)	(1)	(2)	(3)
Rank	8	9	11	7	9	13	7	9	10

LIFFE COCOA, 2 years Forecasts Performance									
		GARCH			*TARCH*			*EGARCH*	
	Gaussian	**GED**	**Student-t**	**Gaussian**	**GED**	**Student-t**	**Gaussian**	**GED**	**Student-t**
RMSE	0.019117	0.019137	0.019172	0.019117	0.019139	0.019147	0.019111	0.019140	0.019113
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(3)	(2)
MAE	0.013519	0.013522	0.013537	0.013519	0.013522	0.013525	0.013529	0.013523	0.013518
	(1)	(2)	(3)	(1)	(2)	(3)	(3)	(2)	(1)
MAPE	93.15373	92.76720	93.82929	93.16095	92.78722	92.97324	93.57357	92.80587	93.34209
	(2)	(1)	(3)	(3)	(2)	(1)	(3)	(1)	(2)
TIC	0.983128	0.959133	0.936372	0.983485	0.957358	0.951346	0.995427	0.956463	0.992567
	(3)	(2)	(1)	(3)	(2)	(1)	(3)	(1)	(2)
R^2	0.011538	0.000748	0.006066	0.011734	0.000489	0.016388	0.006495	0.002264	0.000947
	(1)	(2)	(3)	(2)	(3)	(1)	(1)	(2)	(3)
Rank	8	9	13	10	11	9	11	9	10

Conclusion

This paper has sought to examine the volatility field in the context of agricultural commodities.

First the gain of using GARCH models without introducing more specifics variable in the regressors equation is minimal. The unbiasedness of the forecasts is very dim for agricultural commodities. We may say that perhaps the methodology usually used in financial assets, may not be relevant with agricultural commodities.

Second, the efficiency of the forecasts seems to be bound to the time horizon defined for the forecasts. That is why it directly influences the ranking of the forecasts. Likewise, the choice of the forecast sample, in the case of cocoa if you choose your out sample in the beginning of the 98 you will have very worst forecasts than if you choose it on the beginning of 2000. Because of the trend volatility is currently taking, it depends of the period concerned and moreover if the period is affected by a persistent shock.

Third, as a result the asymmetric models did lead to better forecasts than the symmetric one. Whichever the distribution or the models or the indices, asymmetric models give better results. Obviously, they give much better results when the returns indices show clear evidences of asymmetric and leverage effects (like CAC 40 for instance).

Further research could be pursued on a specification of the convenience yield or a specification of the position of the market or on a shorter period of forecast (less than 1 year).

The predictive ability of GARCH models used with agricultural commodities data is not established. However is not meaning that GARCH models are useless for agricultural commodities forecasts, we are just pondering that they need more specifications in the variance equation to truly capture the trend of the volatility. Plus, one may add that traditionally speaking agricultural commodities are exposed more often to exogenous variables which really disturbed volatility levels more than stock index.

REFERENCES

Andersen, T. G., Bollerslev, T. (1997) Intraday periodicity and volatility persistence in financial markets, Journal of Empirical Finance, 4, 115-158.

Baillie, R. T. and Myers, R. J. (1991) Bivariate GARCH estimation of the optimal commodity futures hedge, Journal of Applied econometrics, 6, 24-109.

Black F. (1976) Studies of stock price volatility changes. ASA, Journal of Business and Economic statistics, 177-181.

Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity, Journal Of Econometrics, 31, 307-327.

Bollerslev, T., Chou, R. Y. and Kroner, K.F. (1992) ARCH modelling in Finance. Journal Of Econometrics 52, 5-59.

Campbell, J., A. Lo, and A. MacKinlay (1997) The econometrics of financial markets. Princeton University Press, Princeton.

Diebold, F. X., and R.S. Mariano (1995) Comparing Predictive Accuracy, Journal of Business and Economic statistics, 13, 253-263.

Engle, R. F. (1982) Autoregressive Conditional Heteroskedasticity with estimates of the variance of U. K. inflation. Econometrica 50, 987-1008.

Engle, R. F. and V.K. Ng, (1993) Measuring and testing the impact of news on volatility, Journal of Finance, 48, 1749-1778.

Fama E. F. (1965) the behaviour of stock market prices, Journal Of Business ,24, 226-241.

French, K. R., G. W. Schwert and R.F. Stambaugh (1987) Expected Stock Returns and Volatility, Journal of Financial Economics, 19, 3-29.

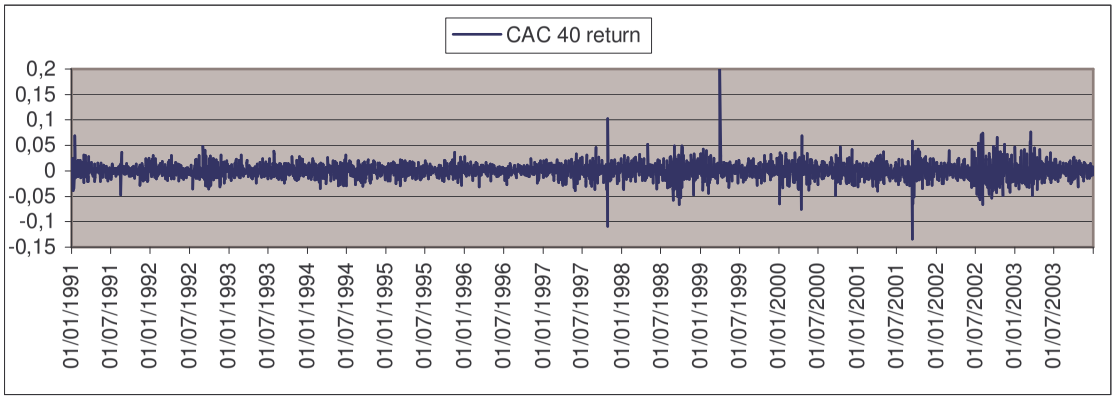
Mandelbrot, B. (1963) The Variation of certain Speculative Prices, Journal of Business, 36, 394-419.

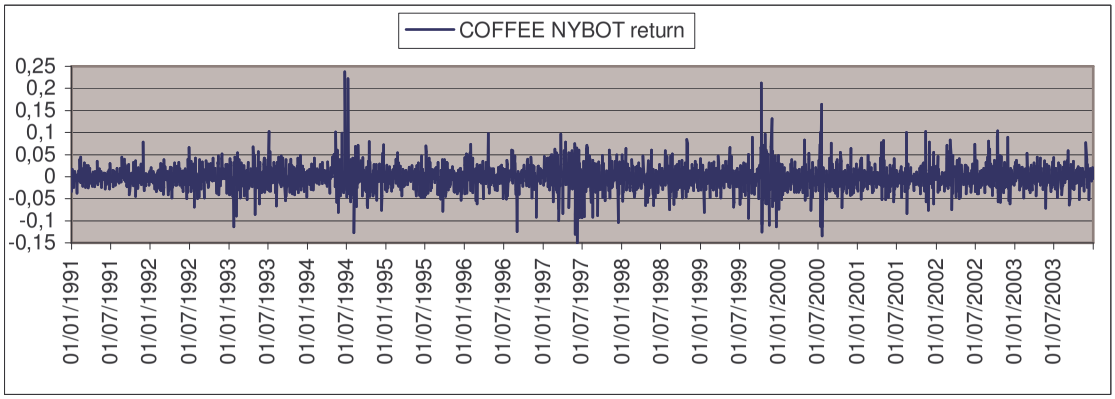
Nelson, D. B., (1991) Conditional Heteroskedasticity in asset pricing : A new approach, Econometrica, 59, 347-370.

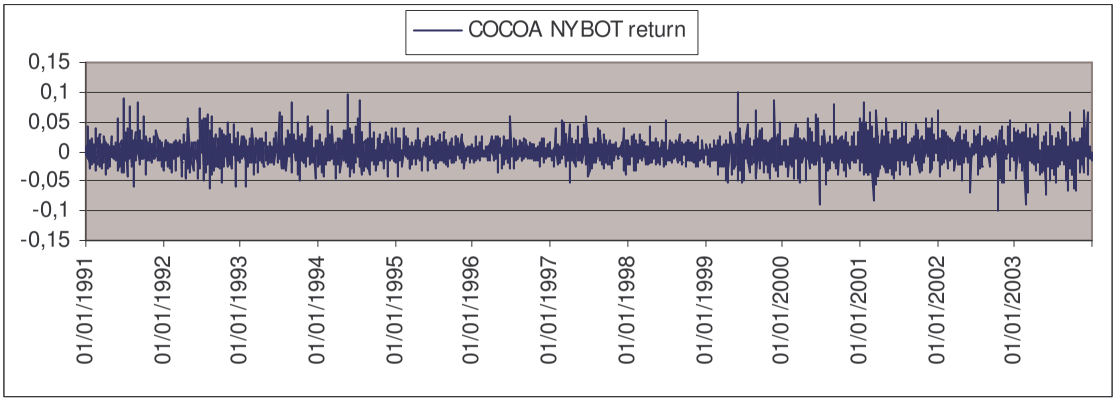
Taylor, S. (1986) Modelling financial time series. Wiley , New York.

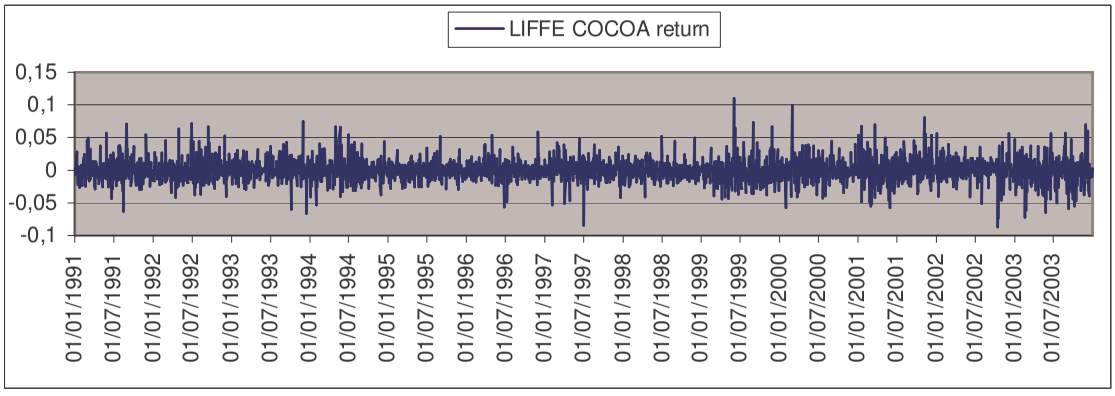
Zakoian, J. M. (1994) Threshold Heteroskedastic Models. Journal Of Economic Dynamics and Control,18, 931-955.

APENDIX A, Volatility Clustering



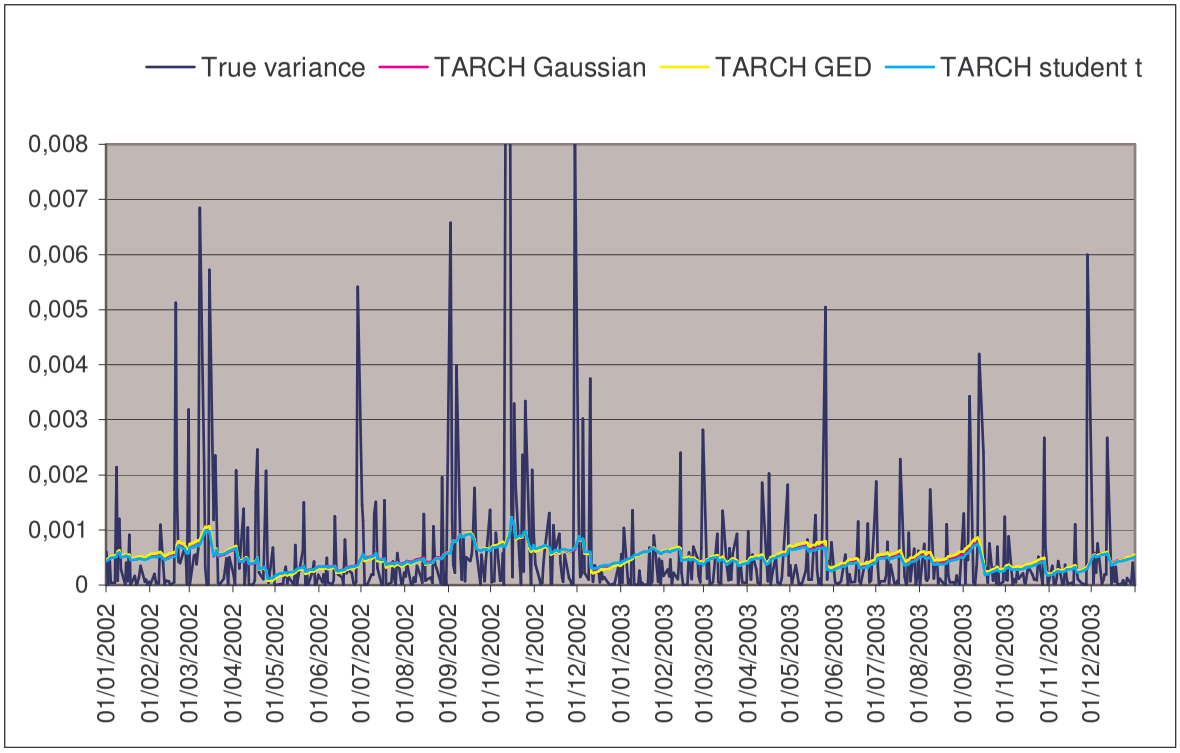




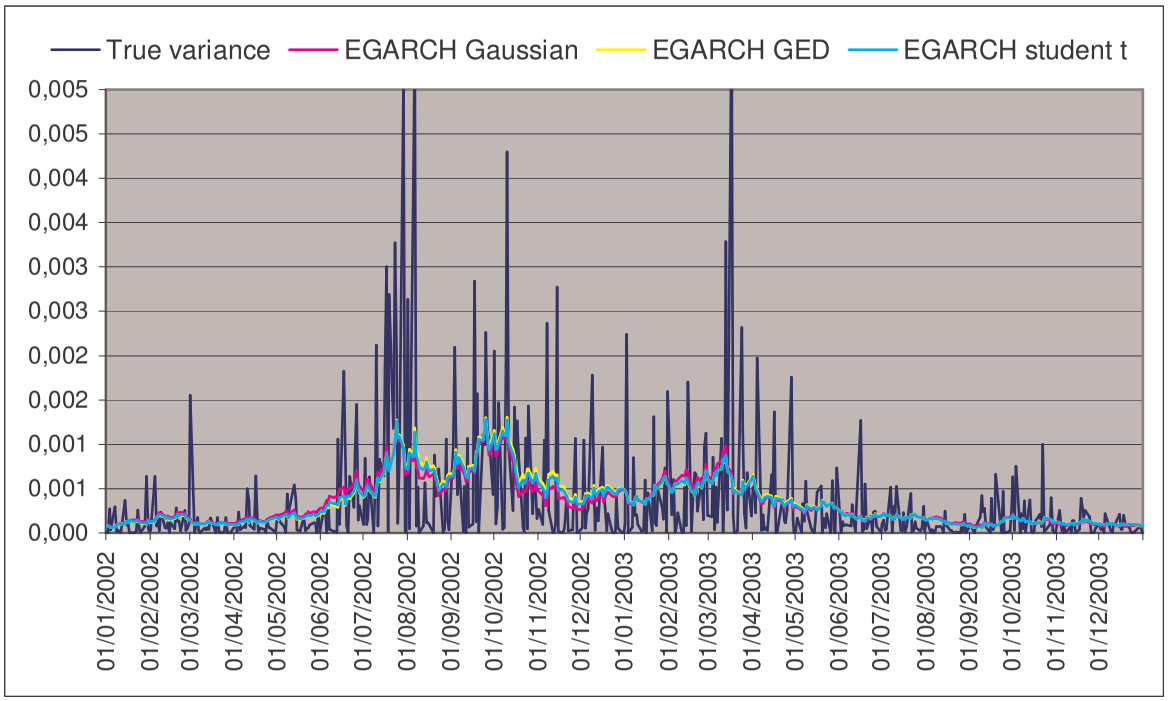


APENDIX B, Best performances forecasts for every indices

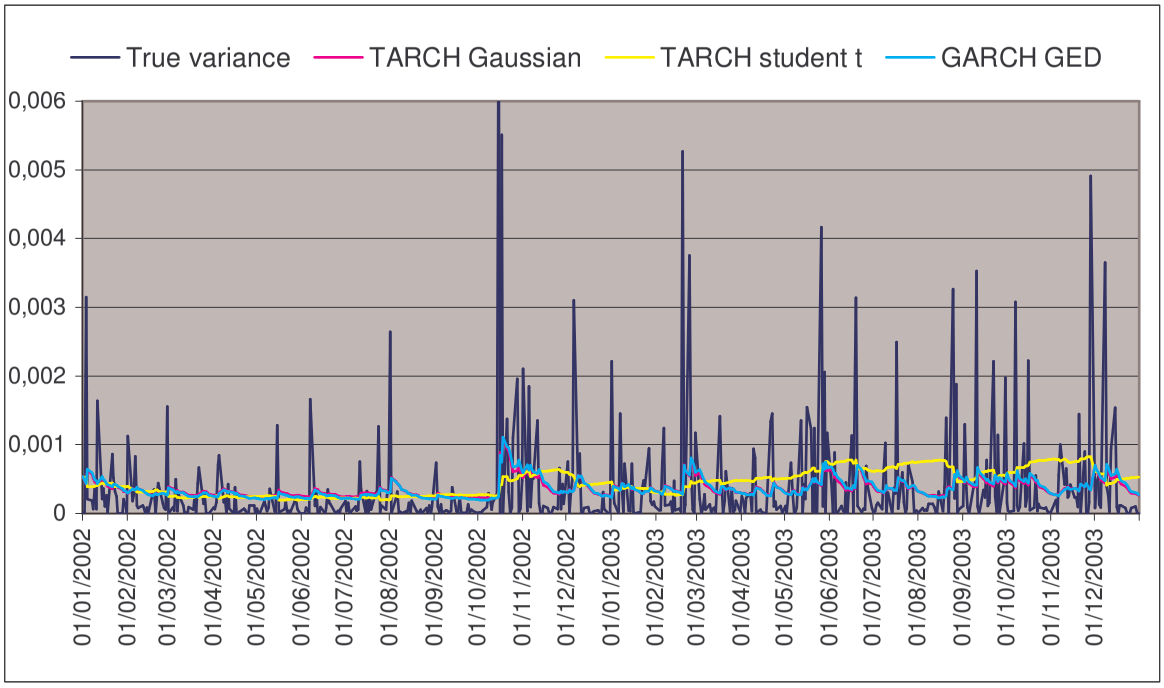
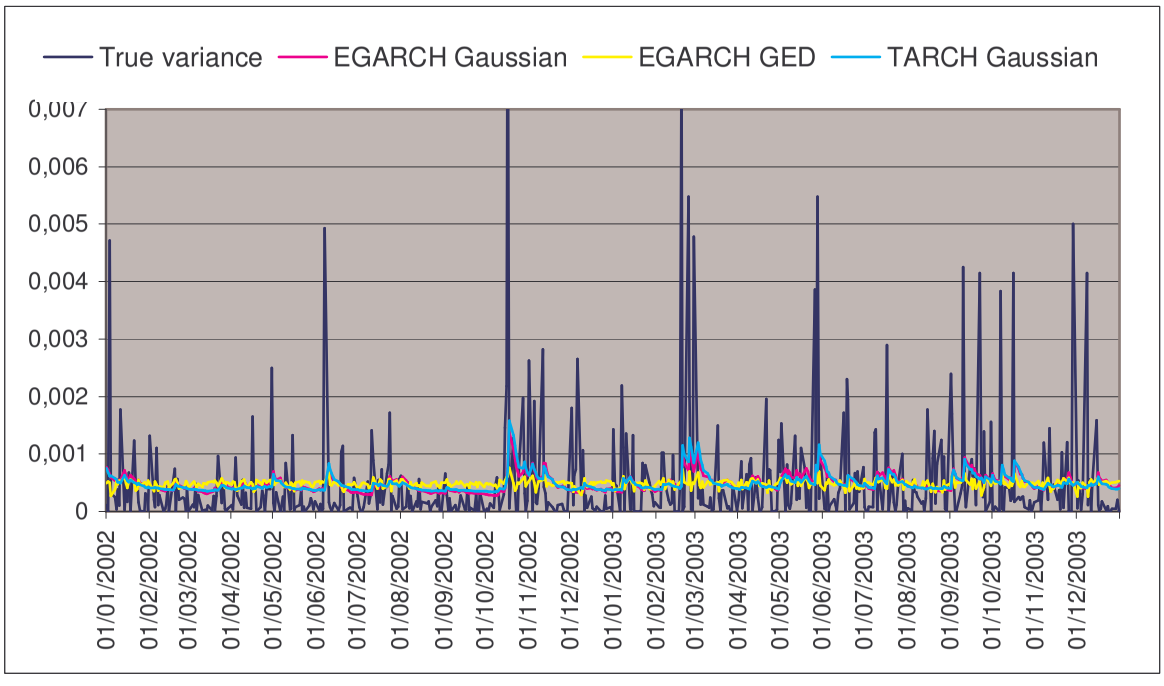
NYBOT COFFEE, 2 years forecasts best performance



CAC 40, 2 years forecasts best performance



LIFFE COCOA, 2 years forecasts best performance



NYBOT COCOA, 2 years forecasts best performance