WORKING PAPER n.03.06

June 2003

The trade-off between complexity and efficiency of VaR measures: a comparison of EWMA and GARCH-type models

M. Caporin[a]

[a] GRETA, Venice

The trade off between complexity and efficiency of VaR measures: a comparison of RiskMetric and GARCH-type models

Massimiliano Caporin
Dipartimento di Scienze Economiche
Università Ca' Foscari di Venezia
Cannaregio - Fondamenta San Giobbe, 873
30121 Venezia - ITALY
and
GRETA Associati
San Marco 3870
30124 Venezia - ITALY
email: mcaporin@greta.it

July 2, 2003

Abstract

This paper compares, on simulated data, the performances of GARCH-type models with the one of the RiskMetric approach in the computation of Value-at-Risk measures. The comparison among the fitted models is based on a set of measures: the tests of Unconditional and Conditional Coverage of Kupiec and Christoffersen and Lopez; the quantiles of the various VaR measures; the moments; the correlation among VaR bounds and the sequencies of exceptions. We show that, even if a long-memory GARCH is the true generator, the RiskMetric model provides a good approximation. It represents the simplest and easier-to-implement solution, the best efficiency at the lower level of complexity.

1 Introduction

Value-at-Risk has become a well known tool for measuring market risk since the implementation of the Basel accord on Capital Requirements (1996). Within this document, banks and financial institutions are required to immobilize resources in an amount adequate to cover their market exposure. The level of needed resources is based on the Value-at-Risk (VaR) and on a set of corrections, increments and exclusions. Many alternative models are available to compute the VaR levels; among them we cite the RiskMetrics (JP Morgan, 1996) and the GARCH-type models. The main difference between these two alternative approaches is on the model structure, very simplistic the first, flexible and with much more complex extensions the second.

Given Basel accord, many subjects face therefore a choice and a trade-off between the complexity of the model and its efficiency. However, no comparative studies have considered this point, which will be the object of this paper. In detail, we compare the fitting of Value-at-Risk measures computed by the RiskMetric model and an alternative set of GARCH-type models on a simulation based study. Moreover, we focus on a particular data generating process, assuming that the market or instrument we are considering shows long-memory behaviour. The analysis includes a comparison among the moments of simulated series, the quantiles of the various VaR models, a group of tests and a correlation study among the VaR bounds and the sequences of exceptions. Our final purpose is to verify if simple and misspecified models (we work on simulated series assuming the data generating process is known) can provide Value-at-Risk bounds that can be considered reliable in the sense they satisfy Basel Accord requirements. This work is structured as follow: next section introduces the various models we consider for the computation of Value-at-Risk; section 3 describes the simulation study while section 4 collects our results; section 5 will conclude.

2 GARCH-type models for Value-at-Risk

Value-at-Risk is defined as the maximum amount of loss a portfolio can incur in with a given level of confidence and in a fixed interval of time. Formally it can be represented as a quantile

$$\int_{-\infty}^{VaR_{m,t}(\alpha,k)} f_{m,t+k}(x) = \alpha \tag{1}$$

where α indicates the confidence level of the quantile, m indicates the model, k refers to the horizon of the possible losses, t is the time to which the Value-at-Risk refers. All these informations are also summarized with the notation $VaR_{m,t}(\alpha, k)$. The horizon is restricted to two cases only, $k = 1$ and $k = 10$, the two levels considered by Basel accord. Various alternative models are available for the evaluation of VaR bounds. Within this paper we focus on a particular class, the GARCH-type models. Within this framework, the conditional volatilities play an essential role in the computation of VaR levels. In fact, the VaR can be represented as a combination of volatilities and residual distribution functions. In particular, assuming also that the standardised residuals (given any GARCH model, these are equal to the mean residuals divided by the conditional volatilities) are normally distributed the VaR can be represented as

$$VaR_{m,t}(\alpha, k) = \Phi^{-1}(\alpha)\,\hat{\sigma}_{t+k,m} \tag{2}$$

where $\Phi^{-1}(\alpha)$ is the quantile of a standardised normal variable and $\hat{\sigma}_{t+k,m}$ represents the forecast of the conditional variance obtained by model m at time t with an horizon k. The differences among the models we consider are included in the term $\hat{\sigma}_{t+k,m}$.

In the following we briefly describe the various models included in this study.

GARCH models were introduced by the seminal works of Engle (1982) and Bollerslev (1986). These models tried to explain several empirical findings of financial market series. The main innovation was in the modelisation of the conditional variances that were structured with a time-dependent relation. The model can be represented with a set of equations. The first and second define the model mean and standardised residual behaviour

$$\begin{gathered} y_t = \mu\left(I^{t-1}\right) + z_t\sigma_t \\ E\left[z_t | I^{t-1}\right] = 0 \qquad E\left[z_t^2 | I^{t-1}\right] = 1 \end{gathered} \tag{3}$$

in this case the standardised residual are coherent with a standardised normal distribution, however other assumptions can be made, including the Student distribution and the GED (Generalised Error Distribution). For the sake of exposition we will assume that the mean is identically equal to zero $\mu(I^{t-1}) = 0$. Finally, the conditional variances are defined:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i z_{t-i}^2 \sigma_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{4}$$

The representation considered is the GARCH(p,q) while in the following we will use the GARCH(1,1) specification. Alternative parameterizations have been suggested in the past years including asymmetric behavior, fat-tailes, heterogeneous behavior and other aspects found in financial market series, however these will not be considered.

The second model we use to evaluate VaR levels is the Integrated GARCH (IGARCH) which adds a unit root to the GARCH model. This parameterisation was suggested by Bollerslev (1996) starting from the ARMA representation of the GARCH model

$$[1-\alpha(L)-\beta(L)]\,\varepsilon_t^2=[1-\beta(L)]\,\upsilon_t \tag{5}$$

where $\alpha(L)=\sum_{i=1}^{p}\alpha_i$, $\beta(L)=\sum_{j=1}^{q}\beta_j$, $\varepsilon_t^2=z_t^2\sigma_t^2$ and $\upsilon_t=\varepsilon_t^2-\sigma_t^2$. Then, the IGARCH is defined introducing a unit root in the autoregressive polinomial

$$(1-L)\,\phi(L)\,\varepsilon_t^2=[1-\beta(L)]\,\upsilon_t \tag{6}$$

The last GARCH-type model is the Fractionally Integrated GARCH (FIGARCH) that represents an extension to the IGARCH case allowing the integration exponent to assume values between 0 and 1. In this case the conditional variances will show a long-memory pattern, a long term correlation. The model was introduced by Baillie, Bollerslev and Mikkelsen (1996) and can be represented as follows

$$(1-L)^d\,\phi(L)\,\varepsilon_t^2=[1-\beta(L)]\,\upsilon_t \tag{7}$$

These "pure" GARCH type models are compared with the RiskMetric model, an exponentially weighted moving average of past squared mean-residuals, defined as

$$\sigma_t^2=(1-\lambda)\sum_{j=1}^{t-1}\lambda^{j-1}\varepsilon_{t-j}^2 \tag{8}$$

The smoothing parameter is normally fixed between 0.94 and 0.97.

Actually, the EWMA is a particular IGARCH model, in fact it can be written in a recursive way as

$$\sigma_t^2=\lambda\sigma_{t-1}^2+(1-\lambda)\,\varepsilon_{t-1}^2 \tag{9}$$

EWMA represents a good approximation of an IGARCH model for an additional reason: in empirical estimations the constant in variance, ω, is likely to be very small, approaching zero, and slightly significant.

In thi study, we do not consider the statistical properties and the problems connected to the parameter estimation. A review of the literature on these topics can be found in Bollerslev, Engle and Nelson (1994). The cited reference provides also an extensive analysis of alternative GARCH-type models.

3 The simulation study

The comparison among the set of alternative models previously described is based on a group of tests and measures of efficiency. At first, we precise that all the simulated series have a FIGARCH structure; this choice is motivated by the numerous findings of long-memory in financial series. These patterns can be misspecified with IGARCH models or by GARCH structures with parameter combination close to the IGARCH case. In these situations the EWMA will necessarily be close to the true GARCH model (many estimated GARCH models on financial series provide a value of the β parameter above 0.9, close to the λ value of EWMA) and we believe that whenever two models behave in a very similar way and whose performances are almost identical our choice must go to the one with the simplest structure. Therefore, when comparing EWMA with IGARCH or GARCH (close to the IGARCH) models, we will always choose the EWMA. In these cases the discrepancies among the various models will be small and combining this point with the availability of GARCH tools in many statistical software make the use of GARCH type models for VaR computation really simple. The trade-off between complexity and efficiency is restricted to the second property, complexity being smoothed by the software packages. Finally, we evidence that the chosen FIGARCH specifications, all FIGARCH(1,d,0), nest the IGARCH model, that can be obtained imposing the equivalence $d = 1$. In such a case the FIGARCH(1,d,0) collapses on an IGARCH, as indicated in the following scheme

$$\begin{gathered} \sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \left[1 - \beta L - (1-L)^d\right]\varepsilon_t^2 \\ \Downarrow \\ d = 1 \\ \Downarrow \\ \sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + (1-\beta)\,\varepsilon_{t-1}^2 \end{gathered} \tag{10}$$

All the considered experiments are based on a set of 1000 replication with two different data generating processes, both FIGARCH(1,d,0) with the following parameter combinations: $\omega = 0.0001$, $\beta = 0.5$ and $d = 0.8$; $\omega = 0.0001$, $\beta = 0.2$ and $d = 0.4$. On the FIGARCH simulated series a group of models is fitted: the true generator, a GARCH(1,1), an IGARCH(1,1) and the EWMA.

For the sake of exposition, we assume that simulated data are daily observations and VaR is computed for 1 and 10 day-ahead. The computation of VaR bounds requires the evaluation of 1 and 10-day-ahead variance forecasts. The derivation of GARCH and IGARCH forecast can be found in Baillie and Bollerslev (1992), while for FIGARCH in Caporin (2002). The computation of 10-day ahead forecasts nests another problem, we should evaluate if the best forecast will be obtained by the square root rule (a 10-day ahead variance forecast is obtained as $\hat{\sigma}^2_{t+10} = \sqrt{10}\hat{\sigma}^2_{t+1}$) or by the sum of a sequence of 1-day ahead forecasts (in this case $\hat{\sigma}^2_{t+10} = \sum_{j=1}^{10} \hat{\sigma}^2_{t+j}$ and $\hat{\sigma}^2_{t+j} = f\left(\left\{\varepsilon^2_i\right\}_{i=1}^{t}, \left\{\hat{\sigma}^2_{t+l}\right\}_{l=1}^{j-1}\right)$), see Caporin (2002).

A second point concerns the length of the series used to compute the VaR levels. Two different sample lengths are considered, 250 and 500 days. The choice is related to the window normally used in the practice that does not exceed two years length. These relative small samples can cause problems in the estimation of FIGARCH parameters. In fact, all long memory models including the ARFIMA of Granger and Joyeux (1980) and the long-memory stochastic volatility of Breidt, Crato and De Lima (1998) share a common characteristic: they depend on the infinite past and long time series are needed to obtain a consistent and robust estimate of the memory parameter. There are no studies that analyse the performances of quasi maximum likelihood estimators in short samples. Something can be obtained in Caporin (2002) limited to samples of length 500, while no other papers consider the length 250. This work will also shed some light in this direction.

The methodology used in the computation of VaR levels requires an additional comment. In the simulation study VaR bounds are computed re-estimating the model with a rolling window on the simulated series. The window amplitude is 250 and 500 days, as specified above. This procedure is used to mimic the extension of the information set that operative subject normally face, every day a new value is provided and the model is re-estimated and/or updated. The Montecarlo covers one year of simulated observations that is assumed to cover 250 days. The choice is motivated by the requirements of the Basel accord that refers to a backtesting period of 250 days.

The comparison is firstly based on the exceptions and on a group of tests. Exceptions are fundamental for fulfilling Basel accord requirements, while the tests of Kupiec (1995) and Lopez and Chistoffersen (1998) measure the unpredictability of the exceptions. The test of Kupiec verifies the null hypothesis of correct unconditional coverage: if the Value-at-Risk model provides accurate risk measures the exceptions (labelled with an e) could be modelled with a binomial distribution with a probability of occurrence equal to the VaR coverage level α. The test compares the theoretical α value with its empiric correspondent $\alpha = e/T$ with a likelihood-ratio-type test

$$LR_{UC} = 2\left[\ln\left(\widehat{\alpha}^{e}\left(1-\widehat{\alpha}^{T-e}\right)\right) - \ln\left(\alpha^{e}\left(1-\alpha^{T-e}\right)\right)\right] \tag{11}$$

In this case the null hypothesis is $\alpha = \widehat{\alpha}$ and the test has an asymptotic $\chi^2(1)$ distribution (under the null). The LR_{uc} test can be considered as the statistical transposition of the Basel Accord requirements.

Christoffersen and Lopez (1998) provide a Conditional Coverage test that improve the previous statistic adding robustness to conditional heteroskedasticity. The test is composed by the sum of the LR_{uc} statistic and of a measure of independence among expections $2\left[\ln\left(L_M\right) - \ln\left(L_I\right)\right]$. This last test verifies the null hypothesis of independence against a null of a first order Markov process. The independence test is again a likelihoo-ratio test where the null hypothesis (independence) is represented by

$$L_I = (1-\pi)^{T_{0,0}+T_{1,0}}\,\pi^{T_{0,1}+T_{1,1}}$$

and the alternative with

$$L_M = \left(1-\pi_{0,1}\right)^{T_{0,0}}\pi_{0,1}^{T_{0,1}}\left(1-\pi_{1,1}\right)^{T_{1,0}}\pi_{1,1}^{T_{1,1}}$$

With $T_{i,j}$ we identify the number of observations in the sample T in state j after having been in state i, where i and j can assume the values 0 or 1 (=exception). Moreover, $\pi_{0,1} = T_{0,1}/\left(T_{0,0}+T_{0,1}\right)$, $\pi_{1,1} = T_{1,1}/\left(T_{1,0}+T_{1,1}\right)$ and $\pi = \left(T_{0,1}+T_{1,1}\right)/T$. Finally, under the null hypothesis the test has an asymptotic $\chi^2(2)$ distribution.

Beside the tests, a descriptive comparison among the various VaR models is considered. We focus on the levels itself and on the sequence of exceptions. A first analysis compares the theoretical VaR quantiles with the exceptions realised with the alternative models. In this case, VaR bounds are computed with the fitted models at the 1-day and 10-day horizon for a range of coverage levels, from 1% up to the 30%. The exception series are then derived and averaged among the 1000 simulation trials. The resulting average exception series provide a VaR model comparison that do not focus only on the queue of the distribution but consider instead a finer distributional comparison.

A second analysis focuses on the moments of the simulated series. In A GARCH-type framework, the VaR bounds are derived using the first and second moment, no attention is given to higher order moments. Our idea is to verify if the estimated models (even if misspecified) adequately explain the distribution properties through higher order moments. Therefore, we simulate series using the estimated conditional volatilities and innovations extracted from a standardised normal distribution (as in the DGP). Alternatively, the innovations are extracted from the standardised residuals of estimated models. In this analysis we proceed in this way: for any simulated FIGARCH(1,d,0) series the various models are estimated and the corresponding conditional variances are saved; we extract the innovation from a standard normal distribution and use it as it is; alternatively, we compute the mean and standard deviation of the standardised residuals of the various fitted models and we apply these values to the casual extractions. This last analysis is used to verify if the misspecification of the model can be in some sense corrected by the lower order moments of the standardised residuals.

Finally, we consider the correlation among the Value-at-Risk bounds and among the corresponding exception sequences. In this case, we can expect a high correlation between two models that adequately and closely explain the behaviour of a series, while different models will show a lower correlation. Similar results can be expected from the correlation among exceptions, where we concentrate the analysis only on the extreme events.

The results of our simulation study are organised in a set of tables. In this paper we report the full set for one experiment, the other are available form the author upon request. The following section highlights the conclusion driven from our simulation study.

4 Simple does not mean bad

A first general remark must evidence that the results we obtained are common over the two parameterisations we considered (d=0.8 and β=0.5, d=0.4 and β=0.2) and the two sample lengths (250 and 500). In all simulations the constant in variance was set equal to 0.0001, to get average volatilities close to the reality. The main point is already contained in the title of this sections: the RiskMetric model, even if simpler than the alternative GARCH and FIGARCH specifications fitted to simulated FIGARCH series provides accurate (in a Basel accord sense) Value-at-Risk estimates. This conclusion derives from the combined analysis of the different tests and efficiency measures we considered. Let us analyse in detail the various behaviour we found in the Montecarlo study.

A starting comment pertains to parameter estimation. Table 1 reports the various Montecarlo averages for the estimated parameters of FIGARCH, GARCH and IGARCH models. The DGP is a FIGARCH(1,d,0) with ω=0.0001, d=0.8 and β=0.5. It is interesting to observe how FIGARCH parameters are close to the true value even with a small sample (estimation is based on 500 observations); however the Montecarlo standard deviation is quite elevate. This is in some sense an expected result; a long memory model requires a long sample to get an unbiased and consistent estimate of the parameters. Whenever this is not the case, the parameter distribution will be much more disperse even if with the correct average. Differently GARCH and IGARCH models provide parameter estimates with smaller standard errors, however the β parameter is far from the EWMA smoothing coefficient. This behaviour could create differences in the VaR bounds obtained by the two models.

[INSERT HERE TABLE 1]

The parameter combination of table 1 is estimated on sample of 500 observations; the same pattern is evidenced with a sample of 250 points but with larger Montecarlo standard deviations. Some differences arise when we consider the second parameter combination we used to simulate returns: the FIGARCH(1,d,0) with ω=0.0001, d=0.4 and β=0.2. In this case the parameter estimates of the fitted FIGARCH model provide disperse and biased values: the Montecarlo average is around 0.6 for d, while at 0.4 for β. We believe this depends on the different memory power of the two FIGARCH models that affect the estimates: in fact with a less persistent memory (d=0.8) a short sample is sufficient to get an unbiased estimate while for more persistent memory (d=0.4) a longer sample is needed to get accurate estimates. Even in this second FIGARCH GDP, the GARCH and IGARCH models provide parameter combinations far from the EWMA smoothing parameter.

We turn now to the analysis of the VaR bounds; we start from the average exception number. Recall that our Montecarlo includes 1000 replications, for each one, 250 1 and 10-step-ahead forecasts are computed based on models estimated on 250 or 500 observation with a rolling approach. This produces 1000 sequences of VaR bounds for each of the fitted models. The VaR bounds are compared with the simulated series and the number of exceptions is computed and averaged. A sample result is included in tables 2 and 3, 1-day VaR and 10-day VaR respectively. What emerges is that the GARCH, IGARCH and FIGARCH models are too conservative; they lead to the computation of larger bounds. In fact, the average exception is really small compared to the VaR confidence level, for the case reported in table 1 we obtain very few exceptions for the 1-day VaR (the percentage average is below 0.01% for a 1% VaR and below 0.5% for a 5% VaR). Differently, the RiskMetric model, even if it represents a completely misspecified model, provides an average percentage of exception of 0.64% and 4.4% respectively. All models clearly satisfy the requirements of Basel accord, however, it must be evidenced that a model which provides larger VaR bound, and by a consequence generate fewer exceptions, cannot be positively considered by financial institutions (unless regulators...). In fact, such models will generate larger capital requirements and will create a bigger opportunity cost for these institutions. It is worth noting that similar results are obtained with a 10-day VaR. In this case the various models provide closer results at both 1% and 5% VaR levels. Clearly, a final conclusion cannot be made on the basis of a single measure, even if some robustness is given by the two sample lengths and the two parameterisations considered.

[INSERT HERE TABLES 2 AND 3]

A subsequent analysis can be seen as a generalisation of the study of average exceptions; it considers a set of VaR quantiles from 1% up to 30% and computes for each the average number of exceptions. An interesting pattern that is evidenced in this case is that the differences among the various model disappear increasing the quantile dimension. This highlights that the various parameterisations generate a different behaviour in the tails while approaching the centre of the distribution they behave in a very similar way. Moreover, while for the 1-day VaR there is a discrepancy among the GARCH-type models and the RiskMetrics, which is closer to the "true" exception number, the difference reduces at the 10-day horizon. It seems that with a shorter horizon the approximation induced by the RiskMetric provides a better measure, while increasing the horizon the GARCH structure become more and more important, influencing the performances of the variance forecasts. These effects can be observed in tables 4 and 5.

[INSERT HERE TABLES 4 AND 5]

Besides the analysis of the exceptions a further measure of model efficiency for VaR models is represented by the correlation among VaR bounds and among the exception series generated by the different fitted models. These correlations are included in tables 6 and 7 which report the Montecarlo average correlation and the corresponding standard errors. The correlations are based on 1% and 5% VaR and we reported the correlation among VaR bounds (the coverage level does not influence the VaR) and the correlation among exceptions at the 1% and 5% level (5% only for 1-day VaR because at the 1% many series reported zero exceptions and correlation could not be computed).

Focus at first on the 1-day VaR. It can be noted that the correlations among FIGARCH, GARCH and IGARCH are close to 1, both for the VaR bounds and the sequences of exceptions. Differently, the correlations of the previous models with the EWMA are smaller, around 80%. This is an expected result, given the behaviour and the patterns evidenced by the models in the average exceptions and in the comparison among VaR quantiles.

Turning to a 10-day horizon creates additional interesting patterns in the correlations. At first the relation among the various models is almost identical to the 1-day horizon if we consider the VaR bounds. If we focus on the sequences of exceptions we note a decrease in the correlations, the minimum being in the cases involving EWMA. The decrease in the correlations is a puzzling fact. It can be explained with two considerations: at first, increasing the horizon in some sense takes into account the long memory pattern of the series, therefore, the FIGARCH provides a better fitting; second, as we noted for the average exception number, at the 10 day level the exceptions increase, this augment the number of points at which correlations are computed and allow for a decrease in these statistics. Let us clarify this point with a simple example. Consider the sequence of exceptions realised by two different models at 1-day and 10-day horizons. For simplicity we focus only on 10 observations.

[INSERT HERE TABLE 8]

While at the 1-day horizon the correlation is around 76% at the 10-day level it decreases to 22%. The correlation analysis evidences that GARCH-type models, indifferently from their memory, determine similar VaR bounds and by a consequence they produce a closer set of exceptions, but only at the 1-day horizon. Turning to the 10-day horizon, the long-memory behaviour of the series become relevant and produces some differences among the GARCH-type models. Moreover, at the 1-day horizon, the correlation among GARCH-type models and the EWMA are significantly different, an expected result determined by the different way the two types of models fit at the lower quantiles: GARCH models provide a poor explanation; they are too much conservative while the EWMA is closer to the theoretical exceptions. At the 10-day level the smaller discrepancies are due both to the long memory of the series, which come into action, and to the model type.

[INSERT HERE TABLES 6 AND 7]

We considered also a different problem. Given the various sequences of conditional variances (used to determine the VaR bounds), we are interested in analysing if these sequences can be used to compare the distribution of the simulated original FIGARCH series with a set of reconstructed returns. To perform this analysis we simulated 1000 return series whose conditional variance was set equal to one of the estimated one. The standardised residuals have been extracted from a standardised normal variable in a case. In a second exercise we computed mean and standard deviation of the standardised residuals of the estimated models and we used these in the simulations. The results are included in table 9. It is evident that in both cases the estimated models are not able to mimic the higher order moments of the simulated returns. The only evident behaviour is the one of GARCH-type models, one close to the other. Moreover, as in the previous analysis, the behaviour of the EWMA is clearly different from the GARCH models.

[INSERT HERE TABLE 9]

A final set of remarks refers to the coverage tests, reported in tables from 10 to 15. It is well known that these tests cannot be used to identify the correct model specification (Lopez 1998, Caporin 2002); therefore we use them only to compare the fitting of various models. The results confirm previous findings: at the 1-day horizon the GARCH-type models behave differently from the EWMA which in turn is the preferred solution; at the 10-day level the long memory of the simulated return turns out to be relevant in variance forecasting and the performances of the FIGARCH models increases. It must be evidenced that since the tests are based on the exceptions they are influenced by their number which gives the number of point on which test statistics are computed. This can explain the clear distinctions between the results on 1-day and 10-day horizon. Summarising, the following set of tables simply indirectly confirm the main result of this work: the EWMA is a closer approximation to the simulated FIGARCH series at the 1-day horizon, while at the 10-day the discrepancy between an exponential smoothing approach and a GARCH-type model become smaller with a preference for long-memory GARCH.

[INSERT HERE TABLES FROM 10 TO 15]

5 Conclusions

In this paper we compared the Value-at-Risk bounds obtained through the estimation of GARCH-type models and the application of the RiskMetrics model (a EWMA). The analysis is developed through a Montecarlo study where the simulated series are generated by a long memory GARCH model. This choice was done in order to mimic the various finding of long term correlation in financial markets. Moreover, the sample length of simulated series was fixed at 250 or 500 observation, the normal ranges used by practitioner, and the VaR bounds were compared on a standard backtesting sample of 250 points. The comparison considered a set of tests and descriptive analysis concentrated both on the VaR measures and on the exceptions realised by the various models.

As a result, we evidence that the EWMA represents the best choice for a financial institution that is interested in satisfying the Basel accord on capital requirements. In fact, even if it represents a misspecified model (this is the case in our simulation study), it provides narrower VaR bounds with an exception number in line with Basel requirements, while GARCH-type models provide too conservative VaR estimates. From a statistical point of view the choice of the EWMA is clearly incorrect while the conclusion is reversed if we focus on the monetary advantages implied by the implementation of the "true" model: narrower VaR measures imply a reduced amount of capital required to cover the market exposure, leaving additional resources that could be used for new investments. This result is evidenced both by the comparison of the quantiles of the various models and by the implied exceptions. Finally, we confirm the results of previous studies (Lopez, 1998) observing that the tests of conditional coverage (Kupiec, 1995) and unconditional coverage (Christoffersen, 1998) cannot be used to identify the best model for Value-at-Risk measurement.

References

[1] Baillie R.T and T. Bollerslev, 1992, Prediction in dynamic models with time-dependent conditional variances, *Journal of Econometrics*, 52, 91-113

[2] Baillie, R.T and T. Bollerslev, H.O. Mikkelsen, 1996, *Fractionally integrated generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics 74, 3-30

[3] Basle Committee on Banking Supervision, 1996, Amendment to the capital market accord to incorporate market risks, Bank for International Settlement

[4] Basle Committee on Banking Supervision, 1996, Supervisory framework for the use of "Baktesting" in conjunction with the internal models approach to market risk capital requirements, Bank for International Settlement

[5] Beltratti A. and C. Morana, 1999, Computing Value at Risk with high frequency data, *Journal of empirical finance*, 6 (5), 431-455

[6] Bollerslev, T., 1986, *Generalized autoregressive conditional heteroskedasticity,* Journal of Econometrics 31, 307-327

[7] Bollerslev, T., R.F. Engle and D.B. Nelson 1994, *ARCH models*, in R.F. Engle and D. McFadden, eds., Handbook of Econometrics, Vol 4, Elsevier Science B.V. Amsterdam

[8] Bollerslev, T and H.O. Mikkelsen, 1996, *Modeling and pricing long memory in stock market volatility*, Journal of Econometrics 73, 151-184

[9] Breidt, F.J., N. Crato and P. de Lima, 1998, *The detection and estimation of long memory in stochastic volatility,* Journal of Econometrics 83, 325-348

[10] Christoffersen P.F., 1998, Evaluating interval forecasts, *International economic review*, 39, 841-862

[11] Christoffersen P.F. and F.X. Diebold 2000, How relevant is volatility forecasting for financial risk management, *Review of Economics and Statistics*, 82-1, 12-22

[12] Diebold F.X., A. Hickman, A. Inoue and T. Schuermann, 1996, Converting 1-Day volatility to h-Day volatility: scaling by $\sqrt{h}$ is worse that you think, University of Pennsilvania working paper

[13] Drost F.C. and T.E. Nijman, 1993, Temporal aggregation of GARCH processes, *Econometrica*, 4, 909-927

[14] Engle, R.F., 1982, *Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation,* Econometrica 50, 987-1008

[15] Engle, R.F. and T. Bollerslev, 1986, *Modelling the persistence of conditional variances,* Econometric reviews 5, 1-50

[16] Granger C.W.J. and R. Joyeux, 1980, *An introduction to long-memory time series and fractional differencing*, Journal of Time Series Analysis, 1, 15-30

[17] J.P. Morgan, 1996, *RiskMetrics Technical Document*, 4th edition, New York

[18] Kupiec P., 1995, Techniques for verifying the accuracy of value-at-risk measurement models, *Journal of Derivatives*, 3, 73-84

[19] Hosking J.M.R., 1981, *Fractional differencing*, Biometrika, 68(1), 165-176

[20] Lopez A.J., 1999, Methods for evaluating value-at-risk estimates, *Federal Reserve Bank of San Francisco Economic Review*, 2

[21] Nelson D.B., 1990, *Stationarity and persistence in the GARCH(1,1) model*, Econometric theory, 6, 318-334

Table 1: Estimated parameters (standard errors)

	FIGARCH(1,d,0)	GARCH(1,1)	IGARCH(1,1)
μ	5.45E-05 (0.00234)	4.47E-05 (0.00235)	5.06E-05 (0.00235)
d	0.82292 (0.15342)		
ω	0.00029 (0.00104)	0.00034 (0.00107)	0.00031 (0.00100)
α		0.29824 (0.05836)	0.32959 (0.08317)
β	0.51247 (0.16009)	0.66877 (0.07820)	

DGP FIGARCH(1,d,0) with ω =0.0001, α=0.5 and d=0.8 - parameter estimates are Montecarlo averages while standard errors are Montecarlo s.e. - Montecarlo is based on 1000 replications with series of 500 observations - the Montecarlo compute VaR levels with a rolling window of amplitude 500 and every step it moves the window by one day re-estimating the model; the process cover one year, 250 observations and the previous table is derived from the estimations of the last step

Table 2: Average Exceptions, standard errors and average percentage of exceptions - 1-day VaR

	VaR	FIGARCH(1,d,0)	GARCH(1,1)	IGARCH(1,1)	EWMA(0.97)
Average exception number (Montecarlo s.e.)	1%	0.00300 (0.05472)	0.00500 (0.08356)	0.00100 (0.03162)	1.60500 (1.25481)
Average exception percent		0.00120	0.00200	0.00040	0.64200
Average exception number (Montecarlo s.e.)	5%	0.63300 (1.31531)	0.67400 (1.25192)	0.35200 (0.89939)	11.01000 (3.12373)
Average exception percent		0.25320	0.26960	0.14080	4.40400

The table reports, for two VaR confidence levels, the average number of exceptions, their standard error and the corresponding average percentage of exceptions. Results are based on the DGP of Table 1. The exceptions are computed for all simulations and for every estimation, that is in one of the 1000 simulations there are 250 records of expections. The Montecarlo averages among all of them.

Table 10: Test of Kupiec - frequency of accepting the null hypothesis - 1-day VaR

Test	VaR	FIGARCH(1,d,0)	GARCH(1,1)	IGARCH(1,1)	EWMA(0.97)
1%	1%	1000	1000	1000	999
5%		3	4	1	802
1%	5%	25	22	8	737
5%		8	5	2	571

The table reports the frequencies of accepting the null hypothesis (correct unconditional coverage) of the test of Kupiec. Two VaR coverage levels are considered as well as two test significance levels. The table is based on 1000 replications. The test is computed on the sequence of exceptions.

Table 12: Test of Independance - frequency of accepting the null hypothesis - 1-day VaR

Test	VaR	FIGARCH(1,d,0)	GARCH(1,1)	IGARCH(1,1)	EWMA(0.97)
1%	1%	997	997	999	238
5%		997	996	999	148
1%	5%	839	821	888	78
5%		736	705	822	36

The table reports the frequencies of accepting the null hypothesis (independence) of the test of independence. Two VaR coverage levels are considered as well as two test significance levels. The table is based on 1000 replications. The test is computed on the sequence of exceptions.

Table 14: Test of Conditional Coverage - frequency of accepting the null hypothesis - 1-day VaR

Test	VaR	FIGARCH(1,d,0)	GARCH(1,1)	IGARCH(1,1)	EWMA(0.97)
1%	1%	1000	1000	1000	321
5%		997	997	999	235
1%	5%	23	19	7	18
5%		7	5	2	3

The table reports the frequencies of accepting the null hypothesis (correct conditional coverage) of the test of Lopez-Christoffersen. Two VaR coverage levels are considered as well as two test significance levels. The table is based on 1000 replications. The test is computed on the sequence of exceptions.

Table 8: exceptions of two different models on the 1-day and 10-day horizons

Horizon	1-day		10-day	
Model	a	b	a	B
t_1	0	0	0	0
t_2	1	1	1	1
t_3	0	0	0	0
t_4	0	1	0	1
t_5	0	0	0	1
t_6	1	1	1	1
t_7	0	0	1	0
t_8	0	0	0	0
t_9	0	0	0	0
t_{10}	0	0	0	1

Table 4: Quantile comparison - 1-day VaR

VaR Quantile	Theoretic Exc.	FIGARCH(1,d,0)	GARCH(1,1)	IGARCH(1,1)	EWMA
0.700	75.000	79.600	79.364	79.045	73.896
0.710	72.500	77.098	76.838	76.466	71.354
0.720	70.000	74.448	74.231	73.721	68.836
0.730	67.500	71.802	71.651	71.083	66.264
0.740	65.000	69.106	68.938	68.376	63.688
0.750	62.500	66.294	66.149	65.520	61.108
0.760	60.000	63.437	63.321	62.651	58.556
0.770	57.500	60.601	60.439	59.683	55.999
0.780	55.000	57.697	57.622	56.759	53.462
0.790	52.500	54.633	54.540	53.524	50.909
0.800	50.000	51.372	51.363	50.350	48.495
0.810	47.500	48.181	48.090	46.973	46.049
0.820	45.000	44.792	44.810	43.521	43.504
0.830	42.500	41.246	41.381	40.026	40.974
0.840	40.000	37.663	37.839	36.341	38.422
0.850	37.500	34.045	34.315	32.643	35.907
0.860	35.000	30.176	30.599	28.839	33.350
0.870	32.500	26.267	26.750	24.902	30.899
0.880	30.000	22.212	22.819	20.869	28.379
0.890	27.500	18.205	18.842	16.892	25.851
0.900	25.000	14.069	14.847	12.839	23.409
0.910	22.500	10.122	10.908	9.024	20.966
0.920	20.000	6.545	7.231	5.534	18.456
0.930	17.500	3.665	4.104	2.865	15.902
0.940	15.000	1.694	1.938	1.176	13.454
0.950	12.500	0.633	0.674	0.352	11.010
0.960	10.000	0.176	0.155	0.072	8.477
0.970	7.500	0.035	0.034	0.014	6.120
0.980	5.000	0.012	0.013	0.005	3.766
0.990	2.500	0.003	0.005	0.000	1.605

The table reports the quantiles of VaR levels in terms of exceptions. We reported the quantile and the theoretical exception value on the common backtesting range (250 onservations) together with the Montecarlo average over the 1000 replications with the four estimated models. The table is based on 1000 replications.

Table 9: moment comparison - 1-day VaR

Order	FIGARCH(1,d,0) simulated	FIGARCH(1,d,0)	GARCH(1,1)	IGARCH(1,1)	EWMA
Residual innovations					
1	-0.00020	-0.00062	-0.00061	-0.00060	-0.00026
2	0.02689	0.11235	0.11218	0.11272	0.24897
3	-0.03651	-0.02204	-0.02278	-0.01541	0.34960
4	1.69892	15.26392	14.73160	14.46808	167.96191
5	-32.66025	4.24275	4.09414	13.37229	818.89366
6	1203.53170	15917.02600	15031.34300	14213.79300	481652.36000
Standardised innovations					
1	-0.00020	1.0E-06	5.4E-06	-1.1E-06	-3.9E-06
2	0.02689	0.02180	0.02185	0.02208	0.02082
3	-0.03651	-0.00034	0.00016	-0.00007	-0.00038
4	1.69892	0.57521	0.60473	0.57605	0.28182
5	-32.66025	-0.26397	0.17059	-0.26734	-0.09492
6	1203.53170	153.95414	169.78073	144.36624	30.37034

The table reports the moments up to order 6 of the simulated series and of the estimated models. The first panel reports the moment obtained using as innovations extractions from a normal variable obtained with the mean and standard deviation of standardised residuals after the model estimation; the second panel is based on standardised normal innovations. The table is based on 1000 replications.

Table 6: Correlation among 1-day VaR and their Exceptions

Models	Var	Exceptions 1%	Exceptions 5%
FIGARCH(1,d,0)-GARCH(1,1)	0.99366 (0.03086)	0.99970 (0.00930)	0.95813 (0.12722)
FIGARCH(1,d,0)-IGARCH(1,1)	0.99384 (0.01713)	1 --	0.96985 (0.10959)
GARCH(1,1)-IGARCH(1,1)	0.99282 (0.03754)	0.99970 (0.00930)	0.97778 (0.08768)
FIGARCH(1,d,0)-EWMA	0.80034 (0.06783)	0.99885 (0.02114)	0.81313 (0.29479)
IGARCH(1,1)-EWMA	0.78661 (0.07786)	0.99795 (0.03797)	0.78871 (0.30855)
GARCH(1,1)-EWMA	0.78820 (0.07345)	0.99957 (0.01344)	0.87539 (0.26101)

The table reports the Montecarlo average correlation among the fitted VaR models and the sequence of exceptions. The table is based on 1000 replications. Percentage coverage of VaR is not indicated since it does not influence the correlation

Table 7: Correlation among 10-day VaR and their exceptions

Models	VaR	Exceptions 1%	Exceptions 5%
F(1,d,0) sum - F(1,d,0) root	0.99603 (0.00580)	0.88252 (0.14602)	0.87818 (0.11429)
G(1,1) sum - F(1,d,0) sum	0.89725 (0.05691)	0.79345 (0.24537)	0.76167 (0.17474)
G(1,1) sum - F(1,d,0) root	0.89660 (0.06077)	0.74527 (0.26284)	0.70022 (0.19418)
G(1,1) root - F(1,d,0) sum	0.90194 (0.05370)	0.77483 (0.24011)	0.78261 (0.15022)
G(1,1) root - F(1,d,0) root	0.90449 (0.05489)	0.74157 (0.24626)	0.76078 (0.15377)
G(1,1) root - G(1,1) sum	0.99601 (0.00688)	0.88323 (0.15122)	0.84682 (0.13603)
IG(1,1) sum - F(1,d,0) sum	0.89015 (0.07928)	0.78663 (0.25917)	0.74057 (0.18230)
IG(1,1) sum - F(1,d,0) root	0.89039 (0.08253)	0.74544 (0.27067)	0.68532 (0.19768)
IG(1,1) sum - G(1,1) sum	0.97643 (0.08043)	0.94936 (0.10814)	0.93071 (0.10628)
IG(1,1) sum - G(1,1) root	0.97856 (0.07934)	0.85871 (0.16674)	0.80765 (0.14991)
IG(1,1) root - F(1,d,0) sum	0.90290 (0.04691)	0.76392 (0.25527)	0.77612 (0.15270)
IG(1,1) root - F(1,d,0) root	0.90573 (0.04884)	0.73572 (0.25505)	0.75799 (0.15605)
IG(1,1) root - G(1,1) sum	0.98624 (0.04108)	0.86734 (0.17232)	0.83202 (0.14080)
IG(1,1) root - G(1,1) root	0.99282 (0.03754)	0.96575 (0.07639)	0.96383 (0.06475)
IG(1,1) root - IG(1,1) sum	0.99009 (0.03088)	0.86926 (0.16171)	0.82094 (0.14736)
EWMA - F(1,d,0) sum	0.80702 (0.07076)	0.75232 (0.24984)	0.71339 (0.18260)
EWMA - F(1,d,0) root	0.78416 (0.07320)	0.70459 (0.27221)	0.67854 (0.18961)
EWMA - G(1,1) sum	0.78502 (0.08106)	0.76850 (0.25708)	0.69965 (0.20432)
EWMA - G(1,1) root	0.78661 (0.07786)	0.72815 (0.26673)	0.70113 (0.18627)
EWMA - IG(1,1) sum	0.77726 (0.09480)	0.76613 (0.26040)	0.67328 (0.20682)
EWMA - IG(1,1) root	0.78820 (0.07345)	0.72171 (0.26786)	0.69089 (0.18949)

The table reports the Montecarlo average correlation among the sequencies of exceptions. Montecarlo standard errors are reported in parenthesis. The table is based on 1000 replications.

Table 5: Quantile comparison - 10-day VaR

VaR Quantile	Theoretic Exc.	FIGARCH(1,d,0) sum	FIGARCH(1,d,0) root	GARCH(1,1) sum	GARCH(1,1) root	IGARCH(1,1) sum	IGARCH(1,1) root	EWMA
0.700	75.000	70.315	73.717	66.549	71.188	64.321	70.608	69.321
0.710	72.500	67.732	71.276	63.804	68.699	61.534	68.140	66.853
0.720	70.000	65.066	68.758	61.153	66.212	58.882	65.598	64.362
0.730	67.500	62.483	66.248	58.563	63.679	56.247	63.014	61.882
0.740	65.000	60.013	63.807	56.005	61.138	53.687	60.491	59.393
0.750	62.500	57.523	61.418	53.446	58.732	51.111	58.093	57.003
0.760	60.000	55.049	58.962	51.013	56.294	48.583	55.677	54.555
0.770	57.500	52.568	56.553	48.616	53.887	46.130	53.186	52.129
0.780	55.000	50.084	54.161	46.081	51.500	43.641	50.790	49.830
0.790	52.500	47.615	51.784	43.664	49.121	41.273	48.376	47.542
0.800	50.000	45.226	49.398	41.244	46.746	38.886	46.129	45.293
0.810	47.500	42.800	47.055	38.918	44.412	36.604	43.794	43.090
0.820	45.000	40.393	44.627	36.665	42.153	34.436	41.508	40.875
0.830	42.500	38.027	42.281	34.407	39.852	32.207	39.312	38.680
0.840	40.000	35.751	39.947	32.280	37.526	30.076	37.016	36.471
0.850	37.500	33.490	37.550	30.240	35.392	27.993	34.861	34.281
0.860	35.000	31.238	35.329	28.102	33.217	25.950	32.658	32.143
0.870	32.500	29.002	32.986	26.026	31.032	23.976	30.469	30.006
0.880	30.000	26.890	30.763	24.068	28.888	22.048	28.297	28.001
0.890	27.500	24.698	28.526	22.124	26.754	20.176	26.236	26.047
0.900	25.000	22.586	26.241	20.194	24.579	18.348	24.061	24.065
0.910	22.500	20.552	24.098	18.345	22.479	16.577	21.995	22.066
0.920	20.000	18.486	21.868	16.497	20.443	14.851	19.984	20.128
0.930	17.500	16.450	19.641	14.733	18.341	13.177	17.897	18.117
0.940	15.000	14.458	17.376	12.938	16.309	11.505	15.916	16.113
0.950	12.500	12.494	15.155	11.168	14.242	9.882	13.887	14.134
0.960	10.000	10.555	13.000	9.372	12.166	8.285	11.911	12.183
0.970	7.500	8.541	10.605	7.664	10.096	6.673	9.853	10.101
0.980	5.000	6.471	8.256	5.824	7.884	5.028	7.657	7.860
0.990	2.500	4.172	5.575	3.854	5.306	3.261	5.137	5.463

Table 3: Average Exceptions, standard errors and average percentage of exceptions - 10-day VaR

	VaR	FIGARCH(1,d,0) sum	FIGARCH(1,d,0) root	GARCH(1,1) sum	GARCH(1,1) root	IGARCH(1,1) sum	IGARCH(1,1) root	EWMA
Average exception number (Montecarlo s.e.)	1%	4.17200 4.18401	5.57500 4.71281	3.85400 4.35289	5.30600 4.91330	3.26100 3.90052	5.13700 4.74539	5.46300 5.26995
Average exception percent		1.66880	2.23000	1.54160	2.12240	1.30440	2.05480	2.18520
Average exception number (Montecarlo s.e.)	5%	12.49400 7.74467	15.15500 8.13776	11.16800 7.85171	14.24200 8.27127	9.88200 7.22885	13.88700 8.02313	14.13400 8.71826
Average exception percent		4.99760	6.06200	4.46720	5.69680	3.95280	5.55480	5.65360

See table 2.

Table 11: Test of Kupiec - frequency of accepting the null hypothesis - 10-day VaR

Test	VaR	FIGARCH(1,d,0) sum	FIGARCH(1,d,0) root	GARCH(1,1) sum	GARCH(1,1) root	IGARCH(1,1) sum	IGARCH(1,1) root	EWMA
1%	1%	813	698	820	720	860	742	698
5%		509	493	476	485	479	502	446
1%	5%	737	740	686	726	679	742	703
5%		571	575	535	569	535	582	528

See table 4.

Table 13: Test of Independance - frequency of accepting the null hypothesis - 10-day VaR

Test	VaR	FIGARCH(1,d,0) sum	FIGARCH(1,d,0) root	GARCH(1,1) sum	GARCH(1,1) root	IGARCH(1,1) sum	IGARCH(1,1) root	EWMA
1%	1%	299	238	367	252	407	260	271
5%		235	148	307	178	344	183	210
1%	5%	78	36	92	45	118	41	50
5%		36	10	57	20	67	15	26

See table 6

Table 15: Test of Conditional Coverage - frequency of accepting the null hypothesis - 10-day VaR

Test	VaR	FIGARCH(1,d,0) sum	FIGARCH(1,d,0) root	GARCH(1,1) sum	GARCH(1,1) root	IGARCH(1,1) sum	IGARCH(1,1) root	EWMA
1%	1%	424	321	471	353	522	357	363
5%		297	235	367	251	407	260	269
1%	5%	18	11	7	7	7	4	8
5%		3	3	1	1	1	1	0

See table 8.