# A comparison of different procedures for combining high-dimensional multivariate volatility forecasts

Alessandra Amendola and Giuseppe Storti

**Abstract** Aim of this paper is to investigate the effect of model uncertainty on multivariate volatility prediction. This effect is expected to be particularly relevant in applications to vast dimensional datasets since it is well known that, in this case, the need for tractable model structures requires the imposition of severe and often untested constraints on the volatility dynamics. By means of an application to the optimization of a vast dimensional portfolio of stock returns, the paper compares the performances of different models and combination procedures. The main finding is that results are highly sensitive not only to the choice of the model but also to the specific combination procedure being used.

**Key words:** multivariate volatility, forecast combination, weights estimation

## 1 Introduction

In multivariate volatility prediction model uncertainty is a relevant problem to be faced by researchers and practitioners. The risk of model misspecification is particularly sizeable in large dimensional problems where highly restrictive assumptions on the volatility dynamics are usually required (see e.g. Pesaran, Schleicher & Zaffaroni, 2009). In order to reduce the impact of misspecification at the forecasting stage, a typical approach is to consider the combination of forecasts from different competing models. Although some recent papers have been focused on the evaluation of forecast accuracy of MGARCH models (Patton & Sheppard, 2008; Laurent,

Alessandra Amendola
Department of Economics and Statistics, University of Salerno, Via Ponte don Melillo, 84084 Fisciano, Salerno (Italy) e-mail: alamendola@unisa.it

Giuseppe Storti
Department of Economics and Statistics, University of Salerno, Via Ponte don Melillo, 84084 Fisciano, Salerno (Italy) e-mail: storti@unisa.it

Rombouts & Violante, 2011) less attention has been paid to the combination of volatility forecasts from different models as a strategy for improving the predictive accuracy (Amendola & Storti, 2008). Also, it has to be considered that in theory different combination strategies could be implemented but, for a given application, only one must be chosen. A combination strategy is defined by the identification of two different elements: a combination rule, which is a function of the alternative forecasts available, and an estimator of the weights assigned to each model. As a consequence of these choices, an additional source of uncertainty, related to the choice of the combination strategy, is introduced into the analysis.
Aim of this work is to discuss some alternative forecast combination strategies for (possibly HD) multivariate volatility forecasts and compare their empirical performances. Section 2 introduces the reference model used for the analysis while some alternative estimator of the combination weights are discussed in Section 3. The statistical properties of the estimators have been assessed by a Monte Carlo simulation whose results are not presented here but are available upon request. Section 4 concludes illustrating the results of an application to the optimization of a portfolio of stock returns.

## 2 The reference model

The data generating process is assumed to be given by

$$\mathbf{r}_t = \mathbf{S}_t \mathbf{z}_t \qquad \text{t=1,...,T, T+1,...,T+N}$$

where T is the end of the in-sample period, $\mathbf{z}_t \overset{iid}{\sim} (\mathbf{0}, \mathbf{I}_k)$ $\mathbf{S}_t$ is any $(k \times k)$ positive definite (p.d.) matrix such that $\mathbf{S}_t\mathbf{S}_t^{'} = \mathbf{H}_t = Var(\mathbf{r}_t|\mathbf{I}^{t-1})$, $\mathbf{H}_t = C(\mathbf{H}_{1,t}, \ldots, \mathbf{H}_{n,t}; \mathbf{w})$ with $\mathbf{H}_{j,t}$ being a symmetric p.d. $(k \times k)$ matrix. In practice $\mathbf{H}_{j,t}$ is a conditional covariance matrix forecast by a given 'candidate model'.
The function $C(.)$ is an appropriately chosen *combination function* and $\mathbf{w}$ is a vector of combination parameters. The weights assigned to each candidate model depend on the values of the elements of $\mathbf{w}$ but do not necessarily coincide with them. Different combination functions $C(.)$ can in principle be used and there is no *a priori* valid procedure for selecting the optimal function. Among all the possible choices of $C(.)$, the most common is the *linear* combination function

$$\mathbf{H}_t = w_1\mathbf{H}_{1,t} + \ldots + w_n\mathbf{H}_{n,t} \qquad w_j \geq 0$$

where $\mathbf{w}$ coincides with the vector of combination weights. The assumption of non-negative weights is required in order to guarantee the positive definiteness of $\mathbf{H}_t$ but can be too restrictive. Alternatively, in order to get rid of the positivity constraint on the $w_j$, two different combination functions can be selected: the *exponential* and *square root* combination function. The exponential combination is defined as

$$\mathbf{H}_t = \text{Expm}\left[w_1\text{Logm}(\mathbf{H}_{1,t}) + \ldots + w_n\text{Logm}(\mathbf{H}_{n,t})\right]$$

where Expm(.) and Logm(.) indicate matrix exponential and logarithm respectively. Differently from the other two functions, the square root combination (for $\mathbf{S}_t$) is not directly performed on the $H_{j,t}$ but on the $S_{j,t}$

$$\mathbf{S}_t = w_1 \mathbf{S}_{1,t} + \ldots + w_n \mathbf{S}_{n,t}$$

with $\mathbf{H}_t = \mathbf{S}_t \mathbf{S}_t^{'}$ and $\mathbf{H}_{j,t} = \mathbf{S}_{j,t} \mathbf{S}_{j,t}^{'}$.

## 3 Weights estimators

For the estimation of the combination parameters we consider three different estimation approaches: Composite Quasi ML (CQML), Composite GMM (CGMM) and 'Pooled' Mincer-Zarnowitz (MZ) regressions. All the estimators considered share the following features: i) do not imply any assumption on the conditional distribution of returns ii) can be applied to large dimensional problems. In the CQML method the estimated $w_i$ are obtained by performing the following optimization:

$$\hat{\mathbf{w}} = \underset{\mathbf{w}}{\text{argmax}} \sum_{i \neq j} L(\mathbf{r}^{(ij)} | \mathbf{w}, \mathbf{I}^N),$$

where $\mathbf{r}_t^{(ij)} = (r_{i,t}, r_{j,t})'$, $\hat{\mathbf{w}} = (\hat{w}_1, \ldots, \hat{w}_k)'$ and

$$L(\mathbf{r}^{(ij)} | \mathbf{w}, \mathbf{I}^N) = -0.5 \sum_{h=1}^{N} \log(|\mathbf{H}_{T+h}^{(ij)}|) - 0.5 \sum_{h=1}^{N} \mathbf{r}_{T+h}^{(ij)} \mathbf{H}_{T+h}^{(ij)} (\mathbf{r}_{T+h}^{(ij)})'$$

is the (bivariate) quasi log-likelihood for the couple of assets (i,j) computed over the prediction period [T+1,T+N].
The CGMM estimator extends the same framework to a GMM setting. The $\hat{w}_i$ are obtained by performing the following optimization:

$$\hat{\mathbf{w}} = \underset{\mathbf{w}}{\text{argmin}} \sum_{i \neq j} m(\mathbf{r}^{(i,j)}; \mathbf{w})' \boldsymbol{\Omega}_N^{(i,j)} m(\mathbf{r}^{(i,j)}; \mathbf{w})$$

$\mathbf{r}_t^{(i,j)} = (r_{i,t}, r_{j,t})$, for $t = T+1, \ldots, T+N$. $m(\mathbf{r}^{(i,j)}; \mathbf{w}) = \frac{1}{N} \sum_{t=T+1}^{T+N} \mu(\mathbf{r}_t^{(i,j)}; \mathbf{w})$ and $\mu(\mathbf{r}_t^{(i,j)}; \mathbf{w})$ is a $(p \times 1)$ vector of moment conditions $\boldsymbol{\Omega}_N^{(i,j)}$ is a consistent p. d. estimator of

$$\boldsymbol{\Omega}^{(i,j)} = \lim_{N \to \infty} N E(m(\mathbf{r}^{(i,j)}; \mathbf{w}^*) m(\mathbf{r}^{(i,j)}; \mathbf{w}^*)')$$

with $\mathbf{w}^*$ being the solution to the moment conditions i.e. $E(m(\mathbf{r}^{(i,j)}; \mathbf{w}^*)) = \mathbf{0}$.
Finally, in the 'Pooled' MZ regressions the $\hat{w}_i$ are the OLS estimates of the parameters of the pooled regression model

$$\text{vech}(\tilde{\mathbf{\Sigma}}_{T+h}) = w_1 \text{vech}(\tilde{\mathbf{H}}_{1,T+h}) + \ldots + w_n \text{vech}(\tilde{\mathbf{H}}_{n,T+h}) + \mathbf{e}_{T+h}$$

for $h = 1, ..., N$, where, depending on the type of combination chosen, $\tilde{\Sigma}_t$ and $\tilde{\mathbf{H}}_{i,t}$ are appropriate transformations of $\mathbf{H}_{i,t}$ and $\mathbf{\Sigma}_t = \mathbf{\Sigma}_t = \mathbf{r}_t \mathbf{r}_t^{'}$.

## 4 An application to stock returns

We consider an application to the optimization of a portfolio of stocks using data from Chiriac and Voev (2011). Data refer to 2156 open to close daily returns on 6 NYSE stocks from 3/1/2000 to 30/7/2008. Six different candidate models and five combination strategies are considered. For each of this we compute the associated minimum variance portfolio and compare the empirical volatilities of the optimized portfolios (table 1). The CGMM gives the lowest variance but the results appear to be very sensitive to the choice of the model or combination strategy used.

| **Model** | **Portfolio Variance**$^*$ | **Comb. strategy** | **Portfolio Variance**$^*$ |
|---|---|---|---|
| DCC | 2.33188 | REG(rv) | 2.08441 |
| CC | 2.37658 | REG | 2.08733 |
| ES | 2.33857 | CGMM | *2.07337* |
| MCOV(22) | 2.67185 | CQML | 2.10192 |
| MCOV(100) | 2.10778 | EW | 2.08147 |
| VECH | 2.09339 | | |

**Table 1** Realized portfolio variances ($(^*)$: $\times\ 10^4$) for different models, constant conditional correlation (CC), dynamic conditional correlation (DCC), exponential smoothing (ES), k-days moving covariance(MCOV(k)), and weights estimators, CGMM, QML, equally weighted (EW), MZ regression (REG), MZ regression using realized covariance as dependent variable (REG(rv)). In all cases a linear combination function is used.

## References


1. Amendola, A., Storti G.: A GMM procedure for combining volatility forecasts. Computational Statistics & Data Analysis, **52**(6), 3047–3060 (2008).
2. Patton, A., Sheppard, K. Evaluating volatility and correlation forecasts. Oxford Financial Research Centre, OFRC Working Papers Series, (2008).
3. Laurent, S., Rombouts, J.V.K., Violante, F.: On the Forecasting Accuracy of Multivariate GARCH Models. Forthcoming in Journal of Applied Econometrics, (2011).
4. Chiriac, R., Voev, V. : Modelling and forecasting multivariate realized volatility, Journal of Applied Econometrics, **26**(6), 922–947 (2011).
5. Pesaran, M. H., Schleicher, C., Zaffaroni, P.: 2009. Model averaging in risk management with an application to futures markets. Journal of Empirical Finance, **16**(2), 280–305 (2009)