

A GARCH-Volatility Dependent DCC Model

Edoardo Otranto

Abstract Forecasting volatility in a multivariate framework has received many contributions in the recent literature, but problems in estimation are still frequently encountered when dealing with a large set of time series. The Dynamic Conditional Correlation (DCC) modeling is probably the most used approach; it has the advantage of separating the estimation of the volatility of each time series (with great flexibility, using single univariate models) and the correlation part (with the strong constraint imposing the same dynamics to all the correlations). We propose a modification to the DCC model, providing different dynamics for each correlation, hypothesizing a dependence on the volatility of the time series. This new model implies adding only two parameters with respect to the DCC model. Its performance is evaluated in terms of out-of-sample forecasts with respect to other multivariate GARCH models. The results on a real data set seem to favor the new model.

Key words: AR distance, Multivariate GARCH, Out-of-sample forecasts, Statistical analysis of financial data, Time series analysis

1 Introduction

The large production of multivariate volatility models has not had a high level of success in practical applications because of the trade-off between flexibility and parsimony of the models; in practice, the estimation of a multivariate model with a large set of assets is possible only imposing strong restrictions on the number of parameters.

The Dynamic Conditional Correlation (DCC) model of Engle (2002) could be considered an important turning point in the multivariate volatility modelling. In fact, due to its ability to separate the estimation of the volatility and the correlation

Edoardo Otranto
DEIR and CRENoS, Via Torre Tonda 34, 07100 Sassari e-mail: eotranto@uniss.it

parts, it is possible to estimate several univariate GARCH models (one for each time series) to represent the volatility and a single GARCH model to represent the correlation dynamics. In this way the model can be applied to a large set of time series with a small computational effort. On the other hand, imposing the same dynamics on all the correlations is a strong restriction. Greater flexibility could be obtained using the approach of Billio et al. (2006), who consider different dynamics for groups of correlations, but the choice of these groups is well subjective. Otranto (2010) proposed an algorithm to select these groups; anyway some constraint on the matrices of the parameters driving the correlation dynamics of each group (matrices of rank 1) has to be imposed.

The idea proposed in this paper derived from the analysis of stylized facts,¹ in which pairs of time series with similar differences in the volatility structure have similar correlation dynamics. For this purpose we reparameterize the coefficients of the correlation dynamics imposing a dependence on a measure of distance between pairs of GARCH models. Adding only two parameters to the original DCC model, we can estimate a flexible model with different dynamics for all the correlations.

The paper is organized as follows: in the next section we describe the new model proposed, recalling also the existing models belonging to the Conditional Correlation (CC) class; in section 3 a comparison among four multivariate models is performed in terms of out-of-sample forecasts, using a 20-variate Italian data set.

2 CC Models

Let $\mathbf{r}_t$ a $(n \times 1)$ vector of returns of n financial time series ($t = 1, \dots, T$). Let us indicate with μ the mean and $\mathbf{H}_t$ the conditional covariance matrix of $\mathbf{r}_t$.

The time-varying conditional covariance matrix can be decomposed into:

$$\mathbf{H}_t = \mathbf{S}_t \mathbf{R}_t \mathbf{S}_t, \quad (1)$$

where $\mathbf{S}_t$ is a diagonal matrix containing the conditional standard deviations (modeled by univariate GARCH models) and $\mathbf{R}_t$ is a time-varying positive definite matrix of correlations. A general multivariate class of CC models can be written in the following way (Engle, 2002):

$$\begin{aligned} \mathbf{R}_t &= \tilde{\mathbf{Q}}_t^{-1} \mathbf{Q}_t \tilde{\mathbf{Q}}_t^{-1}, \\ \mathbf{Q}_t &= (\mathbf{1}_n \mathbf{1}_n' - \mathbf{A} - \mathbf{B}) \odot \mathbf{R} + \mathbf{A} \odot \mathbf{u}_{t-1} \mathbf{u}_{t-1}' + \mathbf{B} \odot \mathbf{Q}_{t-1} \\ \tilde{\mathbf{Q}}_t &= \text{diag}(\mathbf{Q}_t^{1/2}) \end{aligned} \quad (2)$$

where $\odot$ is the Hadamard product, $\mathbf{1}_n$ a n -vector of ones and $\mathbf{A}$ and $\mathbf{B}$ are $(n \times n)$ matrices of coefficients, $\mathbf{R}$ is the unconditional correlation matrix, $\mathbf{u}_t = \mathbf{S}_t^{-1} \boldsymbol{\varepsilon}_t$, where $\boldsymbol{\varepsilon}_t = \mathbf{r}_t - \mu$. Note that we need to rescale the matrix $\mathbf{Q}_t$, as in the first equation of

¹ To save space, we will not show them, but they are available on request.

(2), because it does not directly provide a correlation matrix; in fact the elements on the diagonal of $\mathbf{Q}_t$ are not constrained to be equal to one.

Of course this representation implies estimating $2n(n-1)$ coefficients and it is impracticable already for $n > 5$; moreover, the $\mathbf{Q}_t$ matrices have to be positive definite. For these reasons we need some simple re-parameterization of matrices $\mathbf{A}$ and $\mathbf{B}$ which provide feasible computations and definite positiveness.

The Conditional Constant Correlation (CCC) model (Bollerslev, 1990) is obtained by (2) putting $\mathbf{A} = \mathbf{B} = \mathbf{0}$. In other words, the correlation matrix is constant along time.

In the DCC model proposed by Engle (2002), $\mathbf{A}$ and $\mathbf{B}$ are two non negative scalar coefficients (call them a and b ; $a + b < 1$).

A consistent estimation of the DCC model can be easily obtained because it is possible to separate the estimation of the variance equations and the one of the correlation matrix, and, under mild conditions, the covariance matrix is positive definite (see Engle and Sheppard, 2001). In particular, using the two-step estimation procedure of Engle and Sheppard (2001), the variance part can be obtained estimating n univariate GARCH models (one for each time series).

It is evident that, in spite of a large flexibility for the estimation of the volatility part (a different GARCH is estimated for each series), the imposition of the same correlation dynamics is a strong constraint. A possibility could be to consider rank 1 matrices obtained as $\mathbf{A} = \mathbf{a}\mathbf{a}'$ and $\mathbf{B} = \mathbf{b}\mathbf{b}'$, where $\mathbf{a}$ and $\mathbf{b}$ are $(n \times 1)$ vectors of coefficients. We can call this model the vector diagonal (Rank 1) DCC (R1DCC) model. If n is too large, it is possible to consider only k different elements in $\mathbf{a}$ and $\mathbf{b}$ (Billio et al., 2006), using the algorithm proposed in Otranto (2010) to detect the series following the same dynamics.

We hypothesize a relationship between the volatility structure and the correlation dynamics, re-parameterizing the matrices $\mathbf{A}$ and $\mathbf{B}$ in (2) in the following way:

$$\mathbf{A} = \{g_A[d_G(i, j)]\}; \quad \mathbf{B} = \{g_B[d_G(i, j)]\}$$

where $g_A[d_G(i, j)]$ and $g_B[d_G(i, j)]$ are functions of the GARCH distance (3), proposed in Otranto (2008). It is an extension of the AR distance (Piccolo, 1990) and, for the GARCH(1,1) case, it is given by:

$$d_G(i, j) = \left[\frac{\alpha_i^2}{1 - \beta_i^2} + \frac{\alpha_j^2}{1 - \beta_j^2} - \frac{2\alpha_i\alpha_j}{1 - \beta_i\beta_j} \right]^{1/2} \quad (3)$$

where (α_i, β_i) and (α_j, β_j) are the GARCH coefficients of the i -th and j -th univariate time series respectively.

The functions $g_A(\cdot)$ and $g_B(\cdot)$ have to respect the stationary constraints $(a_{ij} + b_{ij}) < 1$ for each $i, j = 1, \dots, n$, where a_{ij} and b_{ij} are the elements of matrices $\mathbf{A}$ and $\mathbf{B}$ respectively. Every specification that satisfies this constraint is a valid candidate. We choose the logistic functional form:

$$\begin{aligned}
a_{ij} &= \frac{\exp(\phi_A + \theta_A d_G(i, j))}{1 + \exp(\phi_A + \theta_A d_G(i, j)) + \exp(\phi_B + \theta_B d_G(i, j))} \\
b_{ij} &= \frac{\exp(\phi_B + \theta_B d_G(i, j))}{1 + \exp(\phi_A + \theta_A d_G(i, j)) + \exp(\phi_B + \theta_B d_G(i, j))}
\end{aligned} \tag{4}$$

Note that the DCC model is a particular case of (4), where $\theta_A = \theta_B = 0$, and that the model proposed contains only 2 new parameters with respect to the DCC specification. We call this model GARCH-Volatility Dependent DCC (GVDCC).

3 Comparison of models

One of the main purposes, using financial volatility models, is the ability to forecast the covariance matrix, because it will be used in the investment strategies, such as portfolio allocation, evaluation of the risk, classification of time series with similar co-movements, etc. In this section we use the 20 Italian financial sectors to compare the alternative CC model illustrated in the previous section.

First we estimate 20 univariate GARCH(1,1) models for the common volatility part, whereas for the correlation part we estimate four models: a DCC, a GVDCC, a R1DCC and a CCC. For the R1DCC model, to avoid the heavy estimation of 40 parameters, we apply the algorithm described in Otranto (2010) to identify the series following the same DCC dynamics, obtaining 5 groups of series

The estimates of the correlation part are shown in Table 1; the standard errors are calculated using the *sandwich* covariance matrix illustrated in Engle and Sheppard (2001) and Engle (2002).

We use the standardized residuals $\mathbf{z}_t = \mathbf{H}_t^{-1/2}(\mathbf{r}_t - \boldsymbol{\mu})$ to verify the goodness of fit, using two criteria (see Engle and Sheppard, 2001, and Audrino and Barone-Adesi, 2006). The first criterion calculates the percentage of the 20 series belonging to the data set having variance in a confidence interval of one (we use a confidence level of 95%). In our experiment, as shown in Table 1, all models provide standardized residuals having unit variance. A second criterion calculates the percentage of the 400 series, obtained by the cross-products $\mathbf{z}_t \mathbf{z}_t'$, for which the Ljung-Box test accepts the null hypothesis of no correlation (we use a significance level of 5%). Also the results of such test are shown in Table 1. The Ljung-Box test rejects the null hypothesis in 18% (20% for CCC) of cases. Anyway, in practical cases, this percentage failing is always more than the nominal size of the test and it is consistent with the results obtained by Engle and Sheppard (2001) and Audrino and Barone-Adesi (2006).

The GVDCC model outperforms the other models in terms of likelihood function and AIC and BIC criteria. Moreover, being the DCC model nested in GVDCC, the likelihood ratio test rejects the null of equal models at the conventional significance levels.

We can not compare directly the GVDCC and R1DCC models via a likelihood ratio test because they are not nested; anyway it is possible to verify if each param-

Table 1 Estimates of the correlation part of alternative models belonging to the CC family (standard errors in parentheses) and goodness-of-fit criteria.

DCC				
a	b			
0.016	0.938			
(0.002)	(0.008)			
% of series with variance 1		% Ljung-Box accepted		
100(20/20)		82.50(330/400)		
log-lik	AIC	BIC		
-11411.81	10.147	10.152		
GVDCC				
ϕ_A	θ_A	ϕ_B	θ_B	
-0.867	-0.974	3.093	0.940	
(0.042)	(0.037)	(0.149)	(0.262)	
% of series with variance 1		% Ljung-Box accepted		
100(20/20)		82.25(329/400)		
log-lik	AIC	BIC		
-11395.61	10.128	10.139		
R1DCC				
a_1	a_2	a_3	a_4	a_5
0.118	0.119	0.123	0.128	0.132
(0.006)	(0.011)	(0.009)	(0.006)	(0.009)
b_1	b_2	b_3	b_4	b_5
0.979	0.965	0.964	0.968	0.970
(0.002)	(0.006)	(0.006)	(0.004)	(0.004)
% of series with variance 1		% Ljung-Box accepted		
100(20/20)		82.50(330/400)		
log-lik	AIC	BIC		
-11406.78	10.144	10.169		
CCC				
% of series with variance 1		% Ljung-Box accepted		
100(20/20)		80.25(321/400)		
log-lik	AIC	BIC		
-11968.41	10.634	10.637		

eter in $\mathbf{A}$ of GVDCC is equal to the corresponding parameter in $\mathbf{aa}'$ in R1DCC, and each parameter in $\mathbf{B}$ of GVDCC is equal to the corresponding parameter in $\mathbf{bb}'$ in R1DCC. In practice we derive the covariance matrix of the estimated parameters present in the lower triangular matrices $\mathbf{A}$ and $\mathbf{B}$ using the delta method.² The Wald test is performed using the following statistic:

$$W = \mathbf{m}\delta'(\mathbf{m}\Omega\mathbf{m}')^{-1}\mathbf{m}\delta \quad (5)$$

where δ is a 2×1 vector containing an element of $\mathbf{A}$ ($\mathbf{B}$) and the corresponding element in $\mathbf{aa}'$ ($\mathbf{bb}'$); $\mathbf{m} = [1, -1]$; Ω is a (2×2) diagonal matrix containing the

² The elements on the diagonal are excluded because, rescaling $\mathbf{Q}_t$ as in the first equation of (2), they provide always the value 1 in the conditional correlation.

variance of the parameters in δ . The statistic (5) follows a chi-square distribution with 1 degree of freedom.

Comparing the $n(n-1) = 380$ pairs of coefficients, we obtain that only 37 of them (19 in matrix A and 18 in matrix B) can be considered different at a 5% significance level. In practice, the two models provide matrices that have 90% of similar elements, but for the GVDCC model we have used 4 coefficients, whereas for R1DCC we used 10 coefficients. Moreover, the log-likelihood function and the information criteria favor the more parsimonious model.

3.1 Evaluating the forecasting performance

The GVDCC model seems better with respect to the other CC models in terms of likelihood functions. But what about its forecasting performance? The main purpose of this representation is to use the information derived from the first step estimation (the volatility part) to improve the estimation of the correlation part, hypothesizing a dependence of the correlation dynamics on the similarity of the volatility structure. For this reason we have selected a testing set, constituted by the last $f = 400$ observations of the original data. We verify the performance of the four models on this span, in terms of out-of-sample one-step ahead forecasts.

A first simple comparison can be made in terms of Mean Squared Error (MSE). The results are shown in Table 2. We can note that the GVDCC model has the lowest MSE, whereas DCC seems to have better performance with respect to R1DCC and CCC. Anyway the differences are very small. This is expected, because we are using models belonging to the same family, having the variance part in common, and differing only for the specification of matrices $\mathbf{A}$ and $\mathbf{B}$ in (2). What we have to investigate is if these differences are statistically significant or not. To verify if these results are significantly different, we use the Diebold and Mariano (1995) statistic (which follows a standard Normal distribution) to test the null hypothesis of no difference in the accuracy of two competing forecasts. In Table 3 we show the results. We can note that all the pairs of mean squared forecast errors are significantly different at each size level, so we can conclude that the ranking obtained in terms of MSE can be considered valid and GVDCC presents the best performance.

Table 2 Out-of-sample MSE.

DCC	GVDCC	R1DCC	CCC
12.218	12.203	12.252	12.302

As noted by Audrino and Barone-Adesi (2006), in a volatility framework, it is difficult to judge the best model in terms of forecasts when the differences of MSE are small. In other words, the MSE criterion allows to discriminate among forecasts whose performance is different by orders of magnitude.

Table 3 t-values of the Diebold-Mariano test to verify the equality of means of pairs of the squared forecast errors.

	DCC	GVDCC	R1DCC	CCC
DCC		5.479	-7.853	-8.449
GVDCC	-5.479		-9.531	-11.403
R1DCC	7.853	9.531		-5.825
CCC	8.449	11.403	5.825	

Following Clements et al. (2009), we use the Model Confidence Set (MCS) approach of Hansen et al. (2003) as a useful tool to discriminate among different forecasting methods. This approach compares directly each pair of forecasts, selecting iteratively the best set of models with best forecasting performance.

Briefly, the MCS approach consists in trimming the set of candidate models. First, all the sequences of all loss differentials between model i and j are computed:

$$l_{ij,t} = L_{i,t} - L_{j,t} \quad t = 1, \dots, f; \quad i \neq j$$

where $L_{i,t}$ is a loss function³

At each step the null hypothesis of equal predictive ability:

$$E(l_{ij,t}) = 0 \quad \text{for each } i \neq j \quad (6)$$

is tested for a set of models $\mathcal{M}$. This is made using the range statistic:

$$T_R = \max_{i,j \in \mathcal{M}} \frac{|\bar{l}_{ij}|}{\widehat{\text{var}}(l_{ij})^{1/2}}$$

or the semi-quadratic statistic (less conservative):

$$T_{SQ} = \sum_{i,j \in \mathcal{M}} \frac{\bar{l}_{ij}^2}{\widehat{\text{var}}(l_{ij})}$$

where $\bar{l}_{ij}$ is the mean of $l_{ij,t}$, whereas its variance is obtained from a block-bootstrap procedure (see Hansen et al., 2003, for details). The first test is for the full set of candidate models; if the null (6) is rejected, the worst performing model is eliminated by the set of candidate models. This is identified as the model i such that:

$$i = \max_{i \in \mathcal{M}} \frac{\bar{l}_i}{\widehat{\text{var}}(l_i)^{1/2}}$$

where

$$\bar{l}_i = \frac{\sum_{j \in \mathcal{M}} \bar{l}_{ij}}{m-1}$$

³ Following Clements et al. (2009), we use the Quasi-Likelihood loss function.

and m is the number of models in $\mathcal{M}$. Also the variance of $\bar{l}_i$ is evaluated via block-bootstrap.

Table 4 MCS results for individual out-of-sample forecasts. The first row represents the first model removed, down to the best performing model in the last row.

T_R		T_{SQ}	
Model	p-value	Model	p-value
CCC	0.000	CCC	0.000
R1DCC	0.000	R1DCC	0.000
DCC	0.001	DCC	0.001
GVDCC	1	GVDCC	1

In Table 4 the results of the MCS procedure are shown (we use 10000 bootstrap samples to estimate the variance of $\bar{l}_{ij}$ and $\bar{l}_i$). We can note that both the statistic tests carry out the same results and that the GVDCC model is always the best one. Note also that DCC outperforms R1DCC in terms of out-of-sample forecasts, whereas CCC is the worst, confirming that the correlations follow a certain dynamics.

References

1. Audrino, F., Barone-Adesi, G.: Average conditional correlation and tree structures for multivariate GARCH models. *Journal of Forecasting* **25**, 579–600 (2006).
2. Billio, M., Caporin, M., Gobbo, M. : Flexible dynamic conditional correlation multivariate GARCH models for asset allocation. *Applied Financial Economics Letters* **2**, 123–130 (2006).
3. Bollerslev, T.: Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach. *Review of Economics and Statistics* **72**, 498–505 (1990)
4. Clements, A., Doolan, M., Hurn, S. & Becker, R.: On the efficacy of techniques for evaluating multivariate volatility forecasts. *NCER Working Paper Series* **41**, Financial Economic Program (2009).
5. Diebold, F.X., Mariano, R.S.: Comparing predictive accuracy. *Journal of Business and Economic Statistics* **13**, 253–263 (1995).
6. Engle, R.F.: Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* **20**, 339–350 (2002).
7. Engle, R.F., Sheppard, K.: Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. *Mimeo, UCSD* (2001).
8. Hansen, P., Lunde, A., Nason, J. :Choosing the best volatility models: the model confidence set approach. *Oxford Bulletin of Economic and Statistics* **65**, 839–861 (2003)
9. Otranto, E. : Clustering heteroskedastic time series by model-based procedures. *Computational Statistics and Data Analysis* **52**, 4685–4698 (2008).
10. Otranto, E.: Identifying financial time series with similar dynamic conditional correlation. *Computational Statistics and Data Analysis* **54**, 1–15 (2010).
11. Piccolo, D. : A distance measure for classifying ARIMA models. *Journal of Time Series Analysis* **11**, 153–164 (1990).