Dynamic Copulas and Long Range Dependence

Beatriz Vaz de Melo Mendes[1]

Silvia Regina Costa Lopes[2]

Abstract

This paper extends the evolution equation of Patton (2006) for the time variation of the copula parameters by specifying an autoregressive fractionally integrated term. For any copula parameter there is a suitable one-to-one transformation so that the maximum likelihood estimation method may be employed. It is suggested an exploratory tool based on the copula data cross-products for detecting the presence of long range dependence on the copula level of real data. We simulate from copula models possessing long range dependence and work out two examples using real data. Modeling long range dependence on the level of dynamic copulas has the potential for providing improved forecasts and are useful for financial and economic applications.

Keywords: Long Memory, Conditional Copulas, Time Series, Financial Applications.
*JEL classification: C*22, *C*51, *G*10.

[1] Corresponding author: Beatriz Vaz de Melo Mendes, IM-Coppead/UFRJ, PO Box 68514, Rio de Janeiro, RJ 21941-972, Brazil. beatriz@im.ufrj.br
[2] Instituto de Matemática - UFRGS, Brazil

1 – Introduction : Condition dependence

The study of dependence in the context of time series is an exciting area of research which has continuously posed new challenges for econometricians, statisticians and probabilists during the last century. Taking care of dependencies becomes important in order to extend standard models towards more efficient ones.

In the *univariate* setting, modeling short and long range serial dependence in the first and second moments of a time series may be successfully accomplished through ARIMA (Box and Jenkins, 1976), ARFIMA (Granger and Joyeux (1980), Hosking (1981)), GARCH (Engle (1982), Bollerslev (1986)), FIGARCH (Baillie, Bollerslev and Mikkelsen (1996), Bollerslev and Mikkelsen (1996)), and FISV (Breidt et al., 1998) models. All these processes rely on the behavior of the autocovariance function and have their statistical properties well established. In addition, computer packages are avaliable for practitioners.

In the *multivariate* setting just a few complex conditional models exist to take care of the dynamics in the moments of the marginal distributions as well as in the covariance matrix. The most popular are the multivariate GARCH models combined with a conditional specification for the mean. However they suffer from the so called *curse of dimensionality*, and are hard to estimate in large dimensions. Lately, the approach via dynamic copulas and pair-copulas has gained increasing popularity.

Attempts to model the time-varying behavior of copula dependence parameters are only a few (Patton (2006), Rockinger and Jondeau (2005), van den Goorbergh, Genest, and Werker (2005), Fantazzini (2007), Pelletier (2006), Hafner and Reznikova (2009), Garcia and Tsafack (2008), and Chollete et al. (2008) among others). Recently, this topic had attracted the attention of researchers in the context of Markov processes (Ibragimov (2005), Beare (2010), Mendes and Aiube (2010)). However, in all these works just short memory was considered, and it is just natural to assume that long memory may also exist on the level of copulas. Modeling long range dependence on the level of dynamic copulas has the potential for providing improved forecasts and are useful for financial and economic applications.

To the best of our knowledge, no model has been proposed so far modeling long range dependence in the dynamic evolution of copula parameters. Accordingly, in this paper we formally define long range dependence on the level of copulas, and extend the evolution equation of Patton (2006) by specifying an autoregressive fractionally integrated term for the time variation of the copula parameters. It is suggested an exploratory

tool based on the copula data cross-products for detecting the presence of long range dependence on the copula level of real data. We simulate from copula models possessing long range dependence and empirically show that long memory may exist in the sequence of copula dependence parameters, independently of the dynamics existing in the marginal processes. Full maximum likelihood estimation may be a difficult computationally expensive problem, and we leave this as a challenge for future research.

The remaining of the paper is as follows. In Section 2 we review the definitions of copulas, of dynamic copulas, of long memory processes, define long memory on the level of copulas, and discuss estimation and exploratory tools for detecting long memory on the level of copulas. In Section 3 we work out two examples with real data on log-returns and realized volatilities. The analyzes show that the dependence parameter indexing the copula associated with the standardized residuals from the ARFIMA and FIGARCH models fitted to the margins may still change with time and show a high degree of persistence coherent with the presence of long range dependence, which might have been generated by the aggregation of common micro units when composing the prices series. We summarize our results in Section 4.

2 - Dynamic copulas and long memory

2.1 Copulas in finance

In the last decade copulas have gained popularity in the areas of finance and insurance because of the flexibility they offer when dealing with multivariate problems. Curiously, the most important theorem in copula theory dates back to the fifties (Sklar, 1959). It states that any multivariate distribution can be expressed by its copula function evaluated at its marginal distribution functions.

Consider a continuous random vector (X_1, X_2) with cumulative distribution function (cdf) F and marginal cdfs F_1 and F_2. Sklar's theorem ensures that there exists a unique copula $C: [0, 1]^2 \rightarrow [0, 1]$ such that

$$F(x, y) = C(F_1(x), F_2(y)). \tag{1}$$

Let $f(x, y)$ represent the joint density, and, $f_i, i = 1,2$, be the marginal densities. When C is absolutely continuous, taking partial derivatives of (1) one obtains

$$f(x,y) \;=\; c(F_1(x), F_2(y))\, f_1(x)\, f_2(y) \tag{2}$$

Where c represents the copula density. This expression will prove useful later for parameter estimation. (2) allows for tailored   dynamic marginal modeling considering all   characteristics of each F_i, including the mean, standard deviation, skewness, kurtosis and any type of short and long memory serial dependence, plus a search for the best fit for the dependence structure   through a large number of copula families that may be considered. This results in flexible multivariate distributions with any choice of margin distributions. An important example is the family of the *meta*-elliptical distributions (Fang, Fang and Kotz, 2002, 2005) which, unlike the family of elliptical distributions, do not impose any constraints on their margins. This flexibility have motivated applications of copulas in finance.   They include asset allocation, credit scoring, default risk modeling, derivative pricing, and risk management, see Bouyè, Durrleman, Bikeghbali, Riboulet, and Roncalli (2000), Li (2000), Costinot, et al. (2000), McNeil, Frey e Embrechts (2005), Cherubini and Luciano (2002), Hu (2002), Embrechts, Lindskog, and McNeil (2003), Cherubini, Luciano, and Vecchiato (2004), among many others.

Copulas can be employed in probability theory to characterize other types of associations (Nelsen (1999), Joe (1997)). The copula based dependence measures (for example, Kendall's τ, Spearman's ρ_S, and the tail dependence coefficients)   are invariant to any increasing transformation of individual series.

2.2 Copulas with time varying parameters

Most of the above mentioned applications deal with *unconditional copulas*. In practice, conditional distributions with respect to past observations are more powerful when describing the underlying model. This is also true for copulas, which may efficiently fit static and dynamic forms of dependence. Patton (2006) extended Sklar's theorem and   introduced the conditional copula in the bivariate case.

Consider a stationary continuous process $(X_{1,t}, X_{2,t})_{t \in Z}$. Let, $F_t, F_{1,t}$, and $F_{2,t}$ represent their joint and marginal cdfs at time t.   In the case the copula of $(X_{1,t}, X_{2,t})$ is independent of t, the dependence structure of (X_1, X_2) is   given by its constant copula C (see (1)). In the time series context, the Sklar's theorem (Sklar, 1959) may be extended:

$$F_t\left(x_{1,t}, x_{2,t} \mid \mathcal{A}_t\right) = C_t\left(F_{1,t}(x_{1,t} \mid \mathcal{A}_t), F_{2,t}(x_{2,t} \mid \mathcal{A}_t \mid \mathcal{A}_t\right), \tag{3}$$

where C_t is a copula at times t, and

$$\mathcal{A}_t = \sigma\{x_{1,t-1}, x_{2,t-1}, x_{1,t-2}, x_{2,t-2}, \dots\} \quad t = 1, \dots, T,$$

represents the σ-algebra generated by all past joint information up to time t provided by the sample $(x_{1,1}, x_{2,1}), \dots, (x_{1,T}, x_{2,T})$. The same conditioning set for each marginal and for the conditional copula guarantees that we have indeed a copula (not a *pseudo-copula*), and each transformed variable (standard uniform) is independent of the information in the conditioning set of its marginal distribution. Theoretical details may be found in Fermanian and Wegkamp (2004), where the concept of pseudo-copulas was introduced and previous attempts in the direction of modeling time varying dependence structures using copulas were unified. As specified, the common σ-algebra may take into account several lagged values. Therefore there is a sequence of conditional copula functions, which depend on some index t and on the past values $\{(X_{1,t-1}, X_{2,t-1}), \dots\}$ of the random vector.

Previous works on dynamic copula modeling include Rockinger and Jondeau (2001), where a parametric copula conditional to the position of past joint observations in the unit square is combined with previous marginal estimation of GARCH-type models with time varying skewness and kurtosis. Van den Goorbergh, Genest and Werker (2005) studied the behavior of bivariate option prices when the dependence structure of the underlying financial assets follows a dynamic copula model, using rolling windows to estimate the copula parameter. In addition, trends were detected as functions of past volatilities. Semiparametric dynamic copula modeling strategy was proposed by Hafner and Reznikova (2009) where the copula parameter was considered as being a smooth function of time. A similar idea was proposed by Mendes and Melo (2010), where the full dynamic dependence structure existing among assets was assessed by applying local maximum likelihood estimation to copula parameters using the estimated GARCH volatilities as regressors. The regime switching copula of Pelletier (2006), Garcia and Tsafack (2008), and Chollete et al. (2008) allows for two regimes, characterized by different levels of dependence.

Patton (2001, 2006) specified the time variation of the copula parameters by defining an evolution equation based on lagged past observations (forcing variables) and an autoregressive term. This evolution equation may be written as follows. Let C_t be some parametric copula family parameterized by θ_t, and let Θ represent the parameter space. Then

$$\theta_t = \Lambda \left(w + \beta \Lambda^{-1}(\theta_{t-1}) + \alpha \frac{1}{N} \sum_{j=1}^{j=N} | u_{t-j} - v_{t-j} | \right) \qquad (4)$$

where $\Lambda : \mathfrak{R} \to \Theta$ is an strictly increasing function   designed to keep the parameter estimates within their boundaries, and Λ^{-1} represents its inverse.   For example, when θ_t is the linear correlation coefficient r_t or the Kendall's τ, Λ may be the modified logistic transformation

$$\Lambda (x) = \frac{1 - e^{-x}}{1 + e^{-x}}$$

and $\Lambda^{-1} (r) = \log \left(\frac{1+r}{1-r} \right)$. In the case θ_t is the (upper or lower) tail dependence coefficient, the logistic transformation $\Lambda (x) = (1 + e^{-x})^{-1}$ may be used.

A similar evolution equation was used in Dias and Embrechts (2004), Cherubini, Luciano, and Vecchiato (2004), and Fantazzini (2007). In all these papers some member of the large family of GARCH models was used to model the margins and some elliptical copula with time varying parameters was taken to model the dependence structure. Estimates were obtained by maximizing the log-likelihood, which allows for constructing a model selection criterion.

Other references in time-varying copulas modeling in finance include Hafner and Manner (2008) where the parameters vary according to an independent latent random process and are estimated using efficient importance sampling. Giacomini, Härdle, Spokoiny (2009) where adaptive estimation is based on the assumption of local homogeneity for modelling the distribution of returns.

Note that all above cited papers just considered *short memory* on the copula level.

2.3 Long range dependence

Models for *long memory in mean* were first introduced in the *univariate* setting by Granger and Joyeux (1980) and Hosking (1981), following the seminal work of Hurst (1951), and in econometrics by Granger (1980). The important characteristic of an Autoregressive Fractionally Integrated Moving Average (ARFIMA) process is its autocorrelation function (acf) decay rate. In an ARFIMA process, the acf exhibits a hyperbolic decay rate, differently from an ARMA process which presents a geometric decay rate. Long memory in mean has been observed in data from

areas such as meteorology, astronomy, hydrology, and economics, as reported in Beran (1994).

There does not exist a unique definition of long range dependence, and the most commonly used employ the standard autocovariance or autocorrelation functions taking their slow decay to characterize long-memory processes. Let $\{y_t\}_{t=1}^{T}$ be a weakly stationary process and let $\gamma_k = cov(y_t, y_{t+k})$ and $\rho_k = corr(y_t, y_{t+k})$ represent, respectively, its autocovariance and autocorrelation of lag k. We say that $\{y_t\}_{t=1}^{T}$ exhibits long range dependence if $\sum_{k=-\infty}^{\infty} |\gamma_k| = \infty$.

If $\{y_t\}_{t=1}^{T}$ follows an ARFIMA (p, d, q) process, then

$$\phi(B)(1-B)^d y_t = \theta(B)\epsilon_t \qquad (6)$$

where B is the backshift operator, that is, $B^k y_t = y_{t-k}$, $\phi(B) = 1 - \phi_1 B - \cdots \phi_p B^P$ and $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ represent the ordinary autoregressive and moving average components; ϵ_t is a white noise process with zero mean and unit variance. When -0.5 < d < 0.5, the ARFIMA (p, d, q) process is stationary and has autocorrelation function of lag k, as $k \rightarrow \infty$, equal to $\frac{(-d)!}{(d-1)!}k^{2d-1}$. $\{y_t\}$ is covariance stationary for $d < 1/2$ and invertible for $d > -1/2$. The so called anti-persistence case, $d < 0$, describes the behavior of overdifferenced series. Proper long range dependence occurrs when $d > 0$ and if $0 < d < 0.5$ the process presents long-memory behavior.

Persistence and long memory are interesting features found when analyzing long-term patterns of rivers' flows (Hurst, 1951), prices fluctuations, stocks' trading volumes, and so on. It is an intermediate state between two well known situations, namely, the purely random sequence of variables (stationarity, $d = 0$), and the deterministic trend (unit root, $d = 1$).

The ARFIMA framework was naturally extended towards volatility models. The Fractionally Integrated Generalized Autoregressive Conditionally Heteroskedastic (FIGARCH) models were introduced by Baillie, Bollerslev and Mikkelsen (1996) and Bollerslev and Mikkelsen (1996), motivated by the fact that the acf of the squared, log-squared, or the absolute value series of an asset return decays slowly, even when the return series has no serial correlation. Also aiming to model long memory in the second moment, Breidt et al. (1998) introduced the Fractionally Integrated Stochastic Volatility (FISV) model.

Models for heteroskedastic time series with long memory are of great interest in econometrics and finance, and empirical facts about asset returns have motivated the several extensions of GARCH type models (FIGARCH, FIEGARCH, TGARCH, SW-ARCH, LM-ARCH, among many others). Many applied works have detected the presence of long memory in the mean and in the volatility of risky assets, market indexes and exchange rates (for example, Crato (1994), Saqdique and Silvapulle (2001), Lobato and Savin (1998)).

2.4 Long range dependence on the copula level

We now extend the ARFIMA framework towards copula models.

Let C_θ be a parametric copula with $\theta \in \Theta$. Consider a one-to-one transformation $\Lambda : \Re \rightarrow \Theta$. In order to model the dynamic behavior of θ, it is supposed that the copula of $(X_{1,t}, X_{2,t})$ is C_{θ_t}, where $\theta_t = \Lambda(y_t)$ and $y_1, \ldots, y_T$ are realizations of an ARFIMA model.

In other words, to capture the long range time variation in the conditional copula we extend the Patton (2001) evolution equation (4) to:

$$(1-B)^d \theta_t = \Lambda \left(w + \beta \Lambda^{-1}((1-B)^d \theta_{t-1}) + \alpha \frac{1}{N} \sum_{j=1}^{j=N} | u_{t-j} - v_{t-j} | \right) \qquad (7)$$

where $(1-B)^d$ is the fractional difference operator, defined by a binomial series

$$(1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k = 1 - dB - \frac{1}{2} d(1-d)(2-d)B^3 - \ldots \qquad (8)$$

2.5 Estimation of long memory on the level of copulas

Let $(\theta_t, \eta_{1,t}, \eta_{2,t})$ represent the vector of all parameters indexing a bivariate distribution F_t at time t, where θ_t is the copula vector parameter, and $\eta_{1,t}, i = 1,2$ represent the marginal parameters. The conditional joint density may be written as

$$f(x_{1,t}, x_{2,t}; \theta_t, \eta_{1,t}, \eta_{2,t} | \mathcal{A}_t) = c_t (u_t, v_t; \theta_t | \mathcal{A}_t) f_{1,t}(x_{1,t}; \eta_{1,t} | \mathcal{A}_t) f_{2,t}(x_{2,t}; \eta_{2,t} | \mathcal{A}_t) \qquad (9)$$

where, at each time t, c_t is the copula density function, $u_t = F_{1,t}\left(x_{1,t}; \eta_{1,t} \mid \mathcal{A}_t\right)$ and $v_t = F_{2,t}\left(x_{2,t}; \eta_{2,t} \mid \mathcal{A}_t\right)$, and $f_{i,t} = \left(x_{i,t}; \eta_{i,t} \mid \mathcal{A}_t\right)$ is the conditional density of each marginal distribution, $i = 1,2$. We assume that the functional form of the copula remains fixed over the sample, although this could be relaxed.

The fully efficient maximum likelihood estimates of are obtained by maximizing (9). However, parametric estimation of bivariate data under the long-memory copula approach may be accomplished in two steps, the so called IFM method, see Joe (1997). In the first step univariate marginal models are fitted and the estimated cdf $\hat{F}_{i,t}, i = 1,2$, , are used to compute the probability integral transformed data. In the second step the copula parameters are estimated. As pointed out by a referee, under the true distribution, $U_t = F_{1,t}\left(x_{1,t}\right)$ is uniformly distributed on [0,1], but this may not be true under the estimated marginal distribution, $U_t = \widehat{F_{1,t}}\left(X_{1,t}\right)$. When applying the IFM method (Joe, 1997, 2005) it is crucial that the estimated marginal distributions are indistinguishable from the true marginal distributions. Marginal fits should be carefully checked since a poor fit will result in probability integral transforms not being standard uniform or *i.i.d.*. As a consequence, any copula model will be mis-specified. Evaluation whether the transformed series are *i.i.d.* may be accomplished by visual assessment using the autocorrelation function. Diebold et al. (1998) suggest that the autocorrelations should be computed and inspected for the first four moments. Patton (2006) suggests to carry on a formal test by regressing each centered series (up to the fourth power) on their lagged values. The hypothesis that the transformed series are standard uniform may be tested via the Kolmogorov-Smirnov test.

Assuming that the marginal distributions are known or have been accurately estimated, so that a sample $\{(u_t, v_t)\}, t = 1, \dots, T$ is available, and assuming that the copula parameters vary according to some evolution equation similar to (7), the copula parameters θ_t may be obtained by numerically maximizing the pseudo log-likelihood function

$$\mathcal{L}(\theta_t; (u_t, v_t), \dots, (u_T, v_T) \mid \mathcal{A}_t) = \sum_{t=1}^{T} \log c_t\left(u_t, v_v ; \theta_t \mid \mathcal{A}_t\right) \quad (10)$$

which will give (under weak regularity conditions) consistent and efficient estimates whose covariance matrix is given by the inverse of the Fisher information matrix. However, estimation may be numerically difficult and computationally very expensive, and needs to be evaluated with a truncated infinite sum in (8). Some well known algorithms previously applied

to other long-memory models, such as the Durbin-Levinson-Whittle sequences, see Whittle (1963), Shaman (2008), Shaman (2010), and the pioneer work of Sowell (1992), may be considered here.

2.6 Identification and simulation of long range dependence on the level of copulas

We now describe a simulation experiment to illustrate how long memory on the level of copulas may be detected and estimated. We do not implement the maximum likelihood method but instead we provide a very simple and intuitive estimation procedure which could be good starting points when computing the maximum likelihood estimates.

We simulate $\{y_1, y_2, \dots, y_T\}$ from an ARFIMA $(0, d, 0)$ process with $d = 0.35$, and obtain $\{r_1, r_2, \dots, r_T\} = \{\Lambda(y_1), \Lambda(y_2), \dots, \Lambda(y_T)\}$ where Λ is given by (5). The transformed series may be considered a path of linear correlation coefficients. The simulated path $\{r_T\}$ is then used to simulate the copula data $\{(u_1, v_1), (u_2, v_2), \dots, (u_T, v_T)\}$ from a Gaussian copula where, at each time $t, t = 1, 2 \dots, T$, the copula is indexed by r_t. The experiment is repeated 500 times[3].

We are looking for some *proxy*, a series obtained from the copula data which may be used to identify the presence of long memory on the copula level. We inspect the following simulated series: $\{uv\}_{t=1}^{T}$ and $\{u + v\}_{t=1}^{T}$. The Lo's modified rescaled adjusted range test, R/S test (Lo, 1991) for long range dependence was computed for these series and an estimate for d was computed according to the periodogram based Whittle's method (Taqqu et al., 1995).[4] We found presence of long memory for the series {u v} in 65.80 % of the 500 repetitions of the experiment, providing an average d

---

[3] A concern about this simulation experiment is that for each r_t, just one pair of values is generated to represent the copula at time *t*, and this is of course a source of large variability. There is no way to check if data generated reflects the *true* r_t. To overcome this problem we set *T* large, $T = 10000$. We note that the large sample will also increase the power of the R/S test to be applied later on.

[4] The Hurst rescaled range (R/S) analysis was introduced in Mandelbrot and Walis(1969) and used in Mandelbrot (1971) to detect LM in asset prices. We applied the Lo's modified rescaled adjusted range test (R/S test) (Lo, 1991), and computed an approximate maximum likelihood estimator, the Whittle's estimator, implemented in S-Plus based on the algorithm of Haslett and Raftery (1989). All the existing tests and estimators for long range dependence are very tricky to use, heavily dependent on arguments. To get around this problem in the simulation experiments we rejected the null only when it was rejected by the R/S test and the Whittle estimator was statistically significant at the 5\% level.

estimate of 0.123. The tentative *proxy* $\{u + v\}$ was not able to indicate presence of long memory, being successful in only in 4% of the cases. As expected, both marginal series $\{u\}_{t=1}^{T}$ and $\{v\}_{t=1}^{T}$ did not rejected the null of *no* long range dependence in 94.80% and 94.52% of the repetitions. Thus, long range dependence was detected only on the cross-products, and this long term behavior was not carried out to the marginal processes.

Figure1 shows the acfs computed for the $\{y\}_{t=1}^{T}$ ARFIMA series, for the transformed correlation coefficient path $\{r_t\}_{t=1}^{T}$, for the dependence proxies $\{uv\}_{t=1}^{T}$ and $\{u + v\}_{t=1}^{T}$, and for the marginal processes $\{u\}_{t=1}^{T}$ and $\{v\}_{t=1}^{T}$. For this particular simulated data, the R/S statistics were 3.6853 ($d = 0.1253$) and 2.1417, significant at the 1% and 5% levels, respectively for the $\{uv\}$ and the $\{u + v\}$ series.
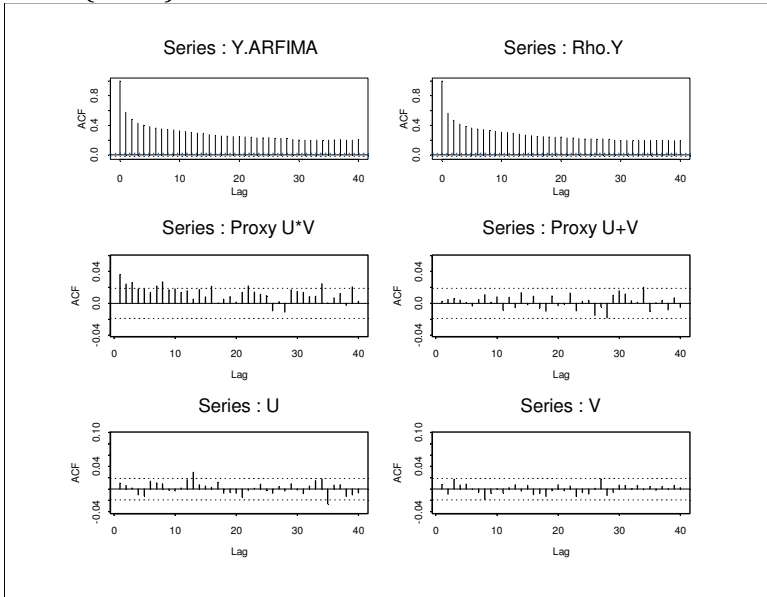

Figure 1: *The acfs of the simulated data: the ARFIMA series, the correlation coefficient path, the proxies, and the marginal processes.*

This simple experiment confirms that long memory may exist only on the dependence level, while the marginal processes possess just short memory or no temporal structure at all. In the case $\{X_{1,t}, X_{2,t}\}$ do have short and/or long memory, they would be implied by their marginal data generating

processes, and not by the copula.   In summary, we have now a tool for detecting long memory in copula parameters for real data[5]

For real data analysis we suggest to estimate the time-varying copula parameters through the specification of an evolution equation such as (7) for each copula parameter (or alternatively for a copula-based measure such as the tail dependence or the Kendall's τ coefficients), and the maximization of the log-likelihood (10). Note that this allows for each copula parameter to be driven by a different ARFIMA process. Thus the degree of persistence may vary among dependence measures.

However, for elliptical copulas for which the linear correlation coefficient is the canonical measure of dependence, we propose to use the cross-product data $\{(u_1, v_1), \dots, (u_T, v_T)\}$ as a proxy for identifying and estimating long range dependence on the copula level using the following *ad hoc* procedure. Fit by maximum likelihood an ARFIMA *(p,d,q)* model (note we are allowing also for short memory) to the proxy data $\left\{ \Lambda^{-1}(u_t, v_t) \right\}_{t=1}^{T}$ and obtain the fitted values $\hat{y}_t$, which are noise free. Then form the path of time-varying correlations $\{\hat{r}_t\}_{t=1}^{T}$ where $\hat{r}_t = \Lambda(\hat{y}_t)$.   The method yields an estimate for *d* and   the   ARFIMA *k-steps* ahead forecasts of the copula parameter. In addition, the   correlation coefficient path may be used as starting values when optimizing (10).

Following a suggestion from a referee, we now carry on a large number of simulations intended to assess the accuracy of the estimation method.   Since empirical works have shown that typically financial returns have a dependence structure best modeled by the *t*-copula (termed by Paul Embrechts as the "desert island copula", see Köck, Schlüter, Weigert, 2008), to mimic a real life situation we simulate from the elliptical *t*-copula. We run the simulation scheme described in (2.6) and apply the estimation method described in (2.5).

Four values for the long range parameter *d* and two values for the degrees of freedom *v* were considered: *d* = 0.05,0.10,0.20,0.25, and *v* = 5 and 10. The number of repetitions of each experiment was 500 and the series length was 3000. Table 1 summarizes the results for *v* = 5 (the results for *v* = 10 are similar). We report the average estimate and the standard error, along

---

[5] We wonder if long range dependence could exist only in some specific dependence parameter and not in the others. For example, just during stressful times, at high quantiles, and therefore in the tail dependence coefficient, but not during low volatility days. Then we would need another exploratory tool.}. Simulations from different long memory processes, and different copula families indexed by real valued parameters, were carried out and confirmed the good performance of the exploratory tool.

with the MSE. The bivariate trajectories were obtained through the conditional copula and, as already commented, this is a key step. As a consequence, the d estimates show a downward bias since some of the long memory property is lost during the copula data generating process.

Table 1: *Average estimate (standard error) from the 500 simulations of models.*

MODELS	$\hat{v}$ $(s.e.; MSE)$	$\hat{d}$ $(s.e.; MSE)$
t-copulas v = 5, d = 0.05	5.022 (0.828;0.678)	0.018 (0.002;0.001)
t-copulas v = 5, d = 0.10	4.992 (0.724;0.526)	0.034(0.007;0.004)
t-copulas v = 5, d = 0.20	5.176 (0.752;0.610)	0.076(0.010;0.015)
t-copulas v = 5, d = 0.25	5.362(0.732;0.692)	0.092(0.009;0.025)

3 - Real data analyzes

We provide two examples using real data from a developed and an emerging market.

Example 1: *Log-returns* on *U. S. market indexes*. A 5-years sample composed by 1827 pairs of daily log-returns on the SP500 and Nasdaq, from June 1, 2000 to June 1, 2007 was obtained from Datastream. Figure 2 shows the evolution through time of the log-returns.
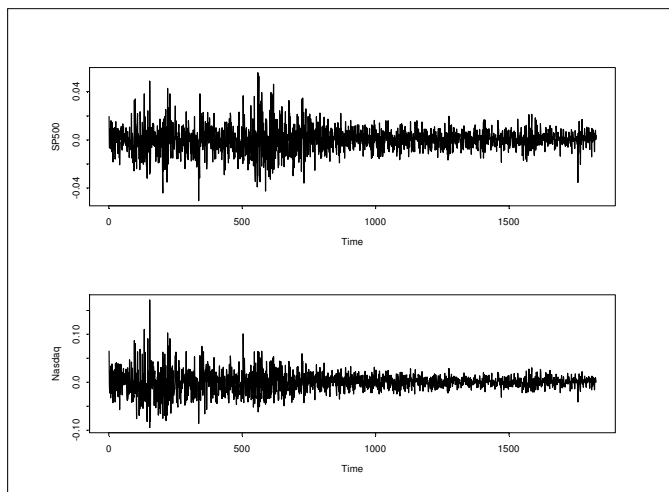

Figure 2: *Time series plot of daily log-returns on the SP500 (top) and Nasdaq (bottom).*

We apply the two-steps estimation method. The KPSS statistics for testing stationarity accepted the null for both returns series at the 1% level. The series show most of the stylized facts found in return series, in particular a wide range of significant autocorrelations in the squares and just a few for the returns. Figure 3 shows the acf computed using the SP500 and Nasdaq log-returns in the upper panel, and their squares in the lower panel. The modified R/S test accepted the null hypothesis of no long-term dependence on the return level, but rejected at the 1% confidence level for squared series. Thus, ARMA (p,q)-FIEGARCH(r,d,s) models were fitted to the log-return series in order to extract all marginal dynamics.
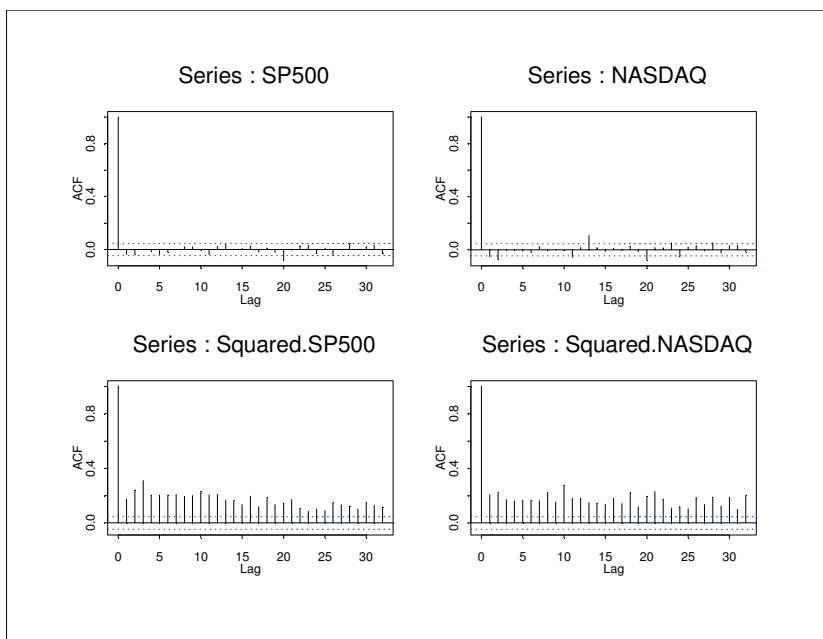

Figure 3: *The acf of SP500 and Nasdaq log-returns (top), and of their squares (bottom).*

For both series the best fit was an ARMA (0,1)-FIEGARCH (2,d,1) plus leverage term, with all estimates being highly statistically significant. The d estimates were, respectively, $d = 0.355$ and $d = 0.616$. The Ljung-Box test applied to the standardized residuals and their squares accepted the null hypotheses of no autocorrelation, indicating that the residuals are free of volatility clusters and temporal dependences.

The series of transformed data $\{u_t\}_{t=1}^{T}$, and $\{v_t\}_{t=1}^{T}$ were obtained from the empirical distribution by ranking the standardized residuals[6]. The modified R/S test applied to the copula proxy $\{u_t v_t\}_{t=1}^{T}$ *rejected* the null hypothesis of *no* long memory. The test statistic was equal to 1.908 (d=0.0859), significant at the 5% level. It is interesting to report that the modified R/S test for long range dependence applied to each series of standardized residuals *accepted* the null hypothesis of no long memory, the test statistics being, respectively, 1.685 and 1.597.

An ARFIMA (0,d,0) was found as the best fit for the $\{\Lambda^{-1}(u_t, v_t)\}_{t=1}^{T}$ (Λ-transformed cross-products) series and the path of estimated correlation coefficients $\{\hat{r}_t\}_{t=1}^{T}$ was derived using the fitted values. Although the d estimate was very small, $d = 0.0109$, it was highly significant ($s.e. = 0.00283$).

Figure 4 shows in the first row the $\{\hat{r}_t\}_{t=1}^{T}$ series, and in the second row its acf. Assuming a t-copula and fixing the estimated correlations sequence, we estimated the (constant) degree of freedom as 11. The log-likelihood value for the time varying t-copula model was 1316.55, and for a Gaussian copula it was 1304.05. For the sake of completeness we also fit a static t-copula to the data, obtaining the global constant maximum likelihood estimates of $(\hat{r} = 0.8733, \hat{v} = 9)$, being the log-likelihood equal to 1305.99.

---

[6] Another possibility for computing the transformed data is to assume a mixture model for the univariate distributions, where the Generalized Pareto distribution would be used for the extreme lower and upper tails and the empirical cdf for the rest of the data.
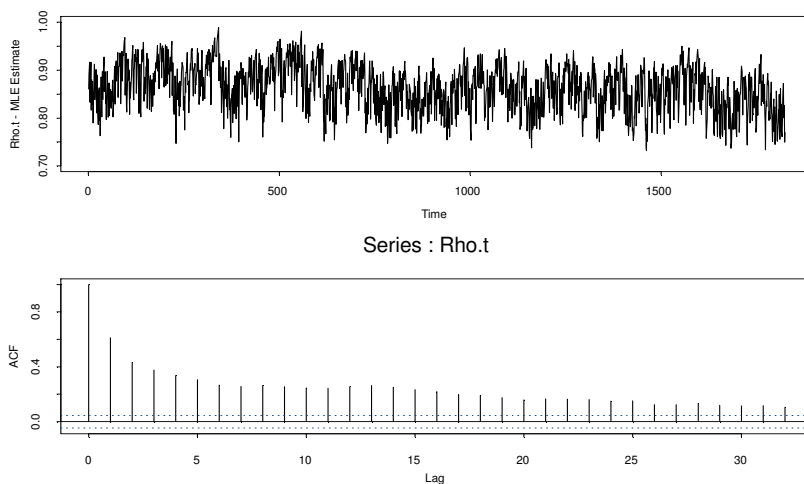
Figure 4: The *estimated path of the correlation coefficients and its acf.*

Usually, the estimation of the time varying dependence and marginal dynamics is not the final objective in a time series application. Each problem will require the forecast of a different functional of the joint conditional distribution. For example, estimation of the one-step ahead conditional Value-at-Risk of a portfolio of assets requires obtaining the one-step ahead copula and marginal predictive distributions. This may be accomplished by simulating the one-step ahead predictive distribution, based on the ARFIMA forecast of the copula parameter and the ARMA-FIEGARCH forecasts of the marginal parameters, to finally compute the desired quantile of some linear combination of the composing variables. The simulations also provide confidence intervals for any quantity of interest.

Example 2: Daily realized volatilities of Brazilian stocks. A sample composed by pairs of daily realized volatilities of two of the most traded stocks in the Brazilian market, Bradesco (BBDC4) and Petrobras (PETR4), from January 2, 2001 to April 30, 2009 (8-years), is obtained from the Bolsa de Valores, Mercadorias e Futuros (Bovespa). The realized volatilities were computed from the high frequency (5-minutes) returns. Figure 5 shows the evolution through time of the $T = 2063$ daily realized volatilities of the Brazilian stocks BBDC4 (top) and PETR4 (bottom).
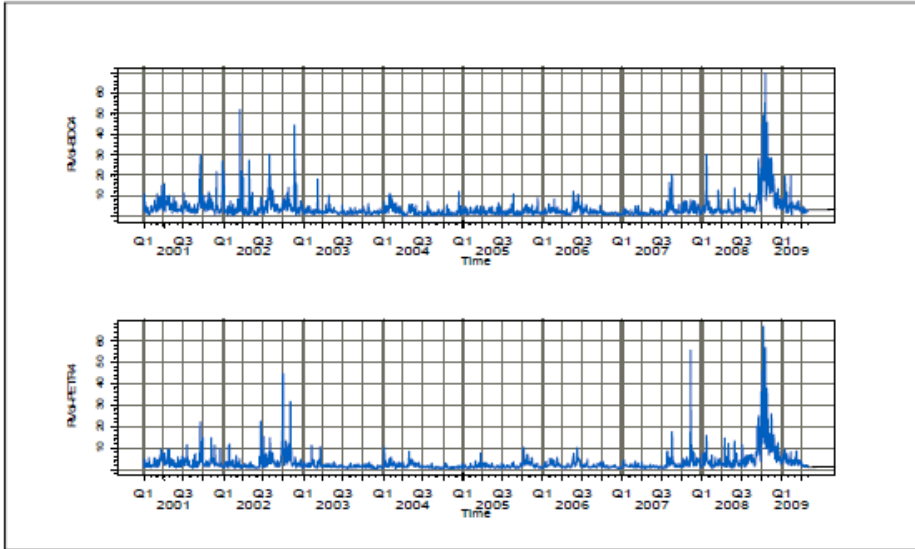
Figure 5: Time series plot of the daily realized volatilities of the Brazilian stocks BBDC4 (top) and PETR4 (bottom).

The modified R/S test strongly rejected the null hypothesis of no long-term dependence on the mean level, thus we fit an ARFIMA (p,d,q) model to the realized volatilities series. For both series the best fit was an ARFIMA (1,d,0) with all estimates being highly statistically significant. The d estimates were, respectively, $d = 0.4418$ and $d = 0.4717$, and the modified R/S test accepted null for both series of residuals. Having extracted the long range dependence of the univariate series of volatilities, we inspect the presence of long memory on the copula level using the residuals from the fits.

The series of transformed data $\{u_t\}_{t=1}^{T}$, and $\{v_t\}_{t=1}^{T}$ were obtained from the empirical distribution by ranking the standardized residuals. The modified R/S test applied to the copula proxies $\{uv\}_{t=1}^{T}$ and $\{u+v\}_{t=1}^{T}$, *strongly rejected* the null hypothesis of no long memory. The test statistics were respectively equal to 2.8334 (d=0.1667) and 3.1089 (d=0.2365), both significant at the 1% level. Figure 6 shows the acf of the proxies. As expected, the modified R/S test for long range dependence applied to each series of standardized residuals accepted the null hypothesis of no long memory.

Again an ARFIMA (0,d,0) was found as the best fit for the Λ-transformed cross-products (d = 0.0931 and t-statistic 5.3887) and the path

of estimated correlation coefficients $\{\hat{r}_t\}_{t=1}^{T}$ were derived using the maximum likelihood fitted values. Figure 7 has in the first row the $\{\hat{r}_t\}_{t=1}^{T}$ series, and in the second row its periodogram, which shows a large concentration of power (variance) in low frequencies.
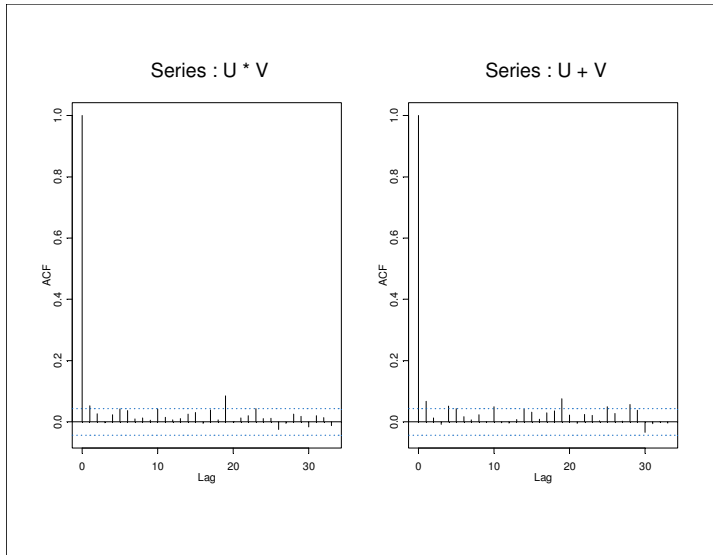

Figure 6: *The autocorrelation function of the proxies $U V$ and $U+V$ from the realized volatilities series.*

Assuming a t-copula and fixing the estimated correlations sequence, we computed the maximum likelihood estimate of the constant degree of freedom as 5. The AIC value for the time varying t-copula model was -466.3. We note that average of the r_t estimates is 0.2772. A static t-copula is also fitted to the data, obtaining the global constant maximum likelihood estimates of $(\hat{r} = 0.3885, \hat{v} = 5)$, being the AIC equal to -445.9.
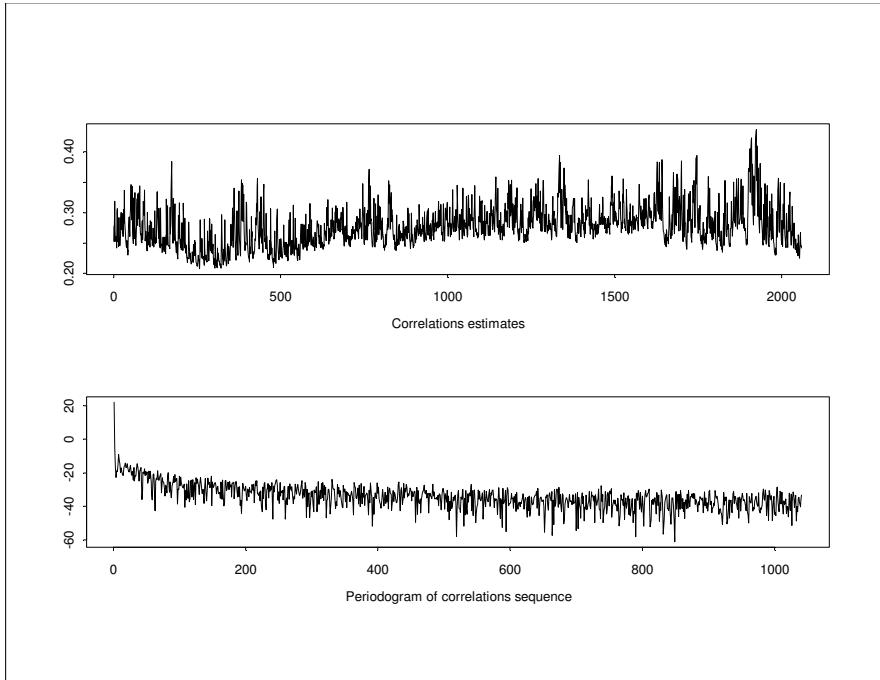
Figure 7: *The estimated path of the correlation coefficients and its periodogram.*

4 – Conclusion

In this paper we have studied long range dependence in the copula parameters sequence.The definition of copulas possessing long memory is quite general and one may assume that each copula parameter is a one-to-one transformation of some non-observable long-memory process. Thus each copula parameter may be driven by a different ARFIMA process and consequently the degree of persistence may vary among dependence measures. In Finance this is certainly important since dependence during high volatility days may be substantially different from that during usual days.

For real data analysis we suggest to carry on the two-steps maximum likelihood estimation method, and in the second step to specify an evolution equation for each copula parameter (or alternatively for a copula-based measure such as the tail dependence or the Kendall's τ coefficients). For dependence structures allowing for types of non-linear dependence, the long

memory may manifest itself in some other measure of association, for example in the tail dependence coefficient.

Since the data cross-products is a simple data transformation containing information about the dependence in the data, we proposed to take the copula data cross-products as proxy for detecting long range dependence on the copula level. The decay rate of the (linear) autocorrelation function of this series was then inspected for long-memory. This seems natural for elliptical copulas for which the linear correlation coefficient is the canonical measure of dependence.

We provided two applications using real data. The examples showed that it is possible for long memory to exist only on the dependence structure level, possessing the marginal processes just short memory or no temporal structure at all. We were initially motivated by problems from the area of finance, but the methodology may be applied to data from any other environment, and we encourage applications in other fields to exploit the potentialities of the new model.

References

Baillie, R.T., Bollerslev, T. and Mikkelsen, H.O. (1996). "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity". Journal of Econometrics, Vol. 74, 3-30.

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

Bollerslev, T. and Mikkelsen, H.O. (1996). "Modeling and pricing long memory in stock market Volatility". Journal of Econometrics, Vol. 73, 151-184.

Box, G.E.P., and Jenkins, G.M. Time Series Analysis: Forecasting and Control, Revised Edition, San Francisco: Holden Day, 1976.

Breidt, F.J., Crato, N. and de Lima, P. (1998). "The detection and estimation of long memory instochastic volatility". Journal of Econometrics, Vol. 83, 325-348.

Cherubini, U., Luciano, E. (2002). "Multivariate Option Pricing with Copulas". ICER Working Papers - Applied Mathematics Series 05-2002, ICER - International Centre for Economic Research.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). Copula Methods in Finance, John Wiley.

Chollete, Heinen, Valdesogo (2009) Modeling International Financial Returns with a Multivariate Regime-switching Copula. Journal of Financial Econometrics, 7(4):437–480.

Costinot, A., Roncalli, T., & Teiletche, J. (2000). Revisiting the dependence between financial markets with copulas.   Working paper.

Crato, N. (1994). Some international evidence regarding the stochastic behavior of stock returns. Applied Financial Economics, 4, 33-39.

Dias, A., Embrechts, P. (2004). Dynamic copula models for multivariate high-frequency data in finance. Discussion paper.

Diebold, Francis X., Gunther, T., and Tay, Anthony S. (1998). "Evaluating Density Forecasts, with Applications to Financial Risk Management", International Economic Review, 39, 863-883.

Engle, R.F. (1982). "Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation".   Econometrica, Vol.   50, 987-1008.

Fang, H-B.,   Fang, K-T., and   Kotz, S.   (2002). "The meta-elliptical distributions with given marginals".   J. Multivar. Analysis, 82, 1-16.

Fang, H., Fang, K-T, Kotz, S. (2005). "Corrigendum to The Meta-Elliptical distribution with given marginal, [J. Multiv. Anal.,   82, 2002, 1-16]". Journal of Multivariate Analysis, 94, 222-223.

Fantazzini, D. (2007).   "Dynamic Copula Modelling for Value at Risk". Frontiers in Finance and Economics, forthcoming. Available at SSRN: http://ssrn.com.

Fermanian, J-D., Wegkamp, M. (2004). "Time dependent copulas´. Working paper.   Available   at http://www.crest.fr/pageperso/lfa/fermanian/fermanian.htm.

Fischer, M., Köck, C., Schlüter, S., and Weigert, F. 2008. "Multivariate copula models at work: outperforming the `Desert Island Copula'?" http://www.statistik.wiso.uni-erlangen.de/forschung/d0079.pdf.

Garcia, R. and Tsafack, G. (2008). "Dependence structure and extreme comovements in international equity and bond markets with portfolio diversification effects".   Working paper, EDHEC Risk Asset. Management Research Centre

Giacomini, E.,   Härdle, W. K., Spokoiny, V.   (2009) Inhomogeneous dependency modelling with time varying copulae. Journal of Business and Economic Statistics, 27:224–234

Granger, C.W.J. (1980) "Long Memory relationships and the aggregation of dynamic models".   Journal of Econometrics, 14, 227-238.

Granger, C.W.J. and Joyeux, R. (1980) "An Introduction to Long Memory Time Series Models and Fractional Differencing". Journal of Time Series Analysis, Vol. 1, 15-29.

Hafner, C. M. and Manner, H. (2008). "Dynamic stochastic copula models: Estimation, inference and applications". METEOR Research Memorandum RM/08/043, Maastricht University.

Hafner, C. M. and Reznikova, O. (2009). "Efficient estimation of a semiparametric dynamic copula model". Manuscript, Institute of Statistics, UCL.

Hosking, J. (1981). ``Fractional Differencing''. Biometrika, Vol. 68}, 165-167.

Ibragimov, R. (2005) "Copula-based characterizations for higher-order markov processes". Forthcoming in Econometric Theory. An extended version is available as Harvard Institute of Economic Research Discussion Paper No. 2094.

Joe, H. (1997). Multivariate Models and Dependence Concepts. Chapman & Hall: London.

Li, D. (2000). On default correlation: A copula function approach. Journal of Fixed Income 9, 43-54.

Lo, A. W. (1991). Long-term memory in stock market prices. Econometrica, 59:1279:1313.

Lobato, I., & Savin, N. E. (1998). "Real and spurious long-memory properties of stock market data". Journal of Business and Economic Statistics, 16, 261-268.

McNeil, A. J., Frey, R. and Embrechts, P. (2005). Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press.

Mendes, B.V.M. and Melo, E.F.L. (2010). "Local Estimation of Dynamic Copula Models". Intern. J. of Theoretical and Applied Finance}, 13, 2, 241-258.

Nelsen, R. B. (2006). An introduction to copula. New York: Springer.

Patton, A. (2001). "Modelling asymmetric exchange rate dependence". Financial Econometric Research Center, WP04-03. Available at http://users.wbs.ac.uk/group/ferc/home

Pelletier, Denis, 2006. "Regime switching for dynamic correlations," Journal of Econometrics, Elsevier, vol. 131(1-2), pages 445-473.

Patton, A. (2006). "Modelling Asymmetric Exchange Rate Dependence", International Economic Review, 47(2), 527-556.

Rockinger, M. and Jondeau, E. (2001). Conditional dependency of financial series: An application of copulas. HEC Paris DP 723.

Shaman, P. "Properties of generalized Levinson–Durbin–Whittle sequences" (2008). Journal of Statistical Planning and Inference, V. 138, I. 9, 2808-2814. Special Issue in Honor of Theodore Wilbur Anderson, Jr. on the Occasion of his 90th Birthday.

Shaman, P. "Generalized Levinson–Durbin sequences, binomial coefficients and autoregressive estimation" (2010). Journal of Multivariate Analysis, V. 101, I. 5, 1263-1273.

Sklar, A. (1959). "Fonctions de repartition an dimensions et leurs marges". Publications de l'Institut de Statistique de l'Université de Paris, 8, 229-231.

Saqdique, S., & Silvapulle, P. (2001). Long-term memory in stock market returns: International evidence. International Journal of Finance and Economics, 6, 59-67.

Sowell, F. B. "Maximum likelihood estimation of stationary univariate fractionally integrated time series models" (1992). Journal of Econometrics, 53, 165-188.

Taqqu, M. S., Teverovsky, V., Willinger, W. (1995). "Estimators for long range dependence: an empirical study". Fractals, 3(4):785-798.

Van den Goorbergh, R. W. J., Genest, C., Werker, J. (2005). "Bivariate option pricing using dynamic copula models". Insurance: Mathematics and Economics, 37 p. 101-114.

Whittle, P. (1963). "On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix". Biometrika, 50, 129–134.