Journal of Time Series Econometrics

Volume 3, *Issue* 3 | 2011 | *Article* 4

Wavelet Estimation of Copulas for Time Series

Pedro A. Morettin, *University of São Paulo*
Clelia M.C. Toloi, *University of São Paulo*
Chang Chiann, *University of São Paulo*
José C.S. de Miranda, *University of São Paulo*

Recommended Citation:
Morettin, Pedro A.; Toloi, Clelia M.C.; Chiann, Chang; and de Miranda, José C.S. (2011) "Wavelet Estimation of Copulas for Time Series," *Journal of Time Series Econometrics*: Vol. 3: Iss. 3, Article 4.
DOI: 10.2202/1941-1928.1033
Available at: http://www.bepress.com/jtse/vol3/iss3/art4

Wavelet Estimation of Copulas for Time Series

Pedro A. Morettin, Clelia M.C. Toloi, Chang Chiann, and José C.S. de Miranda

Abstract

In this paper, we consider estimating copulas for time series, under mixing conditions, using wavelet expansions. The proposed estimators are based on estimators of densities and distribution functions. Some statistical properties of the estimators are derived and their performance assessed via simulations. Empirical applications to real data are also given.

KEYWORDS: copula, density, time series, wavelet, wavelet estimators

Author Notes: This work was partially supported by FAPESP grants 03/10105 and 08/51097-6. We are grateful to the Editor and two anonymous referees for comments that improved the manuscript. The fourth author thanks NSSJC.

1 Introduction

The establishment of dependence between two or more random variables is an important goal, and copulas can be useful in order to accomplish this task. In many situations the marginals are known (or can be estimated) and the joint distribution is unknown, or may be difficult to estimate. A copula is a function that links marginals to their joint distribution.

There have been many applications of copulas to finance. One important problem in financial econometrics is to model asset returns with the purpose of computing measures of risk, such as the value-at-risk (VaR). One common assumption that is often used is that the returns are Gaussian, but it is well known that they have heavy tails and present high kurtosis. Moreover, to compute these risk measures it is necessary to compute the covariance matrix of a portfolio containing a large number of assets, so modeling a vector of returns is a real challenge. These difficulties led to the use of copulas in finance. Some recent references are Embrechts et al. (1997), Bouyé et al. (2000), Cherubini and Luciano (2001), Embrechts et al. (2003), Patton (2001a,b), Fermanian and Scaillet (2003), Fermanian and Wegkamp (2004) and Fermanian et al. (2004).

Concerning the estimation of copulas, several approaches have been used: parametric methods (maximum likelihood estimates and method of moments), non-parametric methods (empirical copulas and kernel estimation) and semi-parametric methods. Also simulation techniques have had an important role, especially to investigate properties of an estimator. For details see Genest et al. (1995), Deheuvels (1979, 1981) and Shih and Louis (1995).

Most of the work done on the estimation of copulas refers to independent samples of a vector of random variables. Therefore some care should be taken to apply these procedures to time series data. See Morettin et al. (2010) for some considerations on this issue. Fermanian and Scaillet (2003) consider the case of time series and propose estimators of copulas based on kernels. For the case of estimators using wavelets and i.i.d. data, see Genest et al. (2009), Autin et al. (2010) and Gayraud and Tribouley (2010).

In this paper we propose to estimate copulas for time series using wavelets. The basic idea is to consider a wavelet expansion of a function of interest and set some coefficients in this series equal zero. This can be done in basically two ways: the first is to consider a finite number of terms in this expansion, starting from some scale, leading to a linear estimator or smoother. The second is to use some threshold, for example keep only those coefficients with absolute value larger than the threshold, leading to a nonlinear estimator.

To our knowledge only the paper of Fermanian and Scaillet (2003) considers the problem of estimating copulas for time series nonparametrically, in their case, using kernels. Our approach uses wavelets, basically following the same route as in their paper, first estimating densities, then distribution functions and quantiles and finally estimating the copula. A more direct approach, using smoothed empirical copula estimators, is given by Morettin et al. (2010). Basically, this approach is also used by Genest et al. (2009) for the i.i.d. case using ranks and wavelets to estimate the copula density.

One purpose of this paper is to show that wavelets are viable tools to use in conjunction with the problem of copula estimation, since it compares favorably with kernels estimators. This will be seen through some simulations. We also present empirical applications. It will also be seen that the nonparametric approach avoids the usual procedure of first fitting univariate or bivariate GARCH-type models and then fitting some parametric copula to the residual series. Finally, wavelets are known to be suitable for the analysis of functions, in our case density functions, belonging to some function spaces, like Sobolev, Hölder and Besov. We will be interested in deriving some properties of the estimators such as their covariance structure and consistency. Wavelets and kernel based estimators are generally close competitors and it is worthwhile to compare them in some situations, and this is one of our purposes.

The plan of the article is as follows. In Section 2 we introduce the background on copulas, wavelets and density estimation. In Section 3 we discuss the cases of i.i.d data and time series data. Section 4 introduces the proposed wavelet estimators. In Section 5 we derive some statistical properties of the estimators and in Section 6 we present some simulations. Empirical applications are given in Section 7 and the paper ends in Section 8 with some concluding remarks. Proofs of the theorems are deferred to the Appendix.

2 Background

In this section we give the necessary concepts on copulas, wavelets and density estimation.

2.1 Copulas

The concept of copula was introduced by Sklar (1959) and since then the theory and applications have developed in a quick pace.

Definition. A $d-$*dimensional copula* is a function C from $[0,1]^d$ to the interval $[0,1]$, with the properties:

(i) C is *grounded*: for every $\mathbf{u} = (u_1, \ldots, u_d) \in [0,1]^d$, $C(\mathbf{u}) = 0$ if at least one coordinate u_i is equal to 0, $i = 1, \ldots, d$;

(ii) C is *d-increasing*: for every $\mathbf{u}$ and $\mathbf{v}$ in $[0,1]^d$, with $\mathbf{u} \leq \mathbf{v}$, the C-volume $V_C([\mathbf{u}, \mathbf{v}])$ of the box $[\mathbf{u}, \mathbf{v}]$ is non-negative;

(iii) $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$, for all $u_i \in [0,1]$, $i = 1, \ldots, d$.

See Nelsen (2006) for the definition of C-volume and further details on copulas. The following important theorem links the definition of copula with an n-dimensional distribution function and its marginal distributions. Denote by Ran$F$ the range of the d.f. F.

Theorem (Sklar). *Let F be a d-dimensional distribution function with margins $F_1, \ldots, F_d$. Then there exists a d-copula C such that for all $\mathbf{x} = (x_1, \ldots, x_d) \in [-\infty, \infty]^d$, we have*

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)). \tag{1}$$

Conversely, if C is a d-copula and $F_1, \ldots, F_d$ are distribution functions, the function F defined by (1) is a d-dimensional distribution function with margins $F_1, \ldots, F_d$. Moreover, if the margins are all continuous, then C is unique. Otherwise, C is uniquely determined on Ran$F_1 \times \ldots \times$ RanF_d.

Therefore, given Sklar's theorem, it is easy to construct the copula associated to a cumulative distribution, namely

$$C(u_1, \ldots, u_d) = F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)), \tag{2}$$

where $F_i^{-1}(u_i) = \inf\{x_i \mid F_i(x_i) \geq u_i\}$, $i = 1, \ldots, d$, is the quasi-inverse of F_i.

Observe that copulas are multivariate distribution functions with uniform one-dimensional margins.

2.2 Wavelets

Wavelet expansions of functions of interest is now a well established subject in Statistics and other areas. For details see Daubechies (1992) and Meyer (1993). From a *mother wavelet* ψ and a *father wavelet* ϕ (or scaling function) bases of $L_2(\mathbb{R})$ can be obtained using the dilated and shifted versions of ψ and ϕ, namely, $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$ and $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$, where j and k are integers. An orthonormal basis for $L_2(\mathbb{R})$ is generated by taking $\phi_{l,k}(x)$ and $\psi_{j,k}(x)$, with $j \geq l$ and $k \in \mathbb{Z}$, for some coarse, or lower, scale l. Hence, for any $f \in L_2(\mathbb{R})$ we may write, uniquely,

$$f(x) = \sum_k \alpha_k \phi_{l,k}(x) + \sum_{j \geq l} \sum_k \beta_{j,k} \psi_{j,k}(x), \tag{3}$$

where the wavelet coefficients are given by

$$\alpha_k = \int f(x) \phi_{l,k}(x) dx, \tag{4}$$

$$\beta_{j,k} = \int f(x) \psi_{j,k}(x) dx. \tag{5}$$

We now introduce a notation that will make further derivations easier. The idea is to use a single sum in place of (3).

Let $a \in \mathbb{Z}, b \in \mathbb{Z},\ a \leq b$. Denote ${}_a\mathbb{Z}$ the set of all integers z with $z \geq a$ and ${}_a\mathbb{Z}_b$ the set of all integers z such that $a \leq z \leq b$. Let $Ze(l) = \mathbb{Z} \cup ({}_l\mathbb{Z} \times \mathbb{Z})$ and $Ze(l)_J = \mathbb{Z} \cup ({}_l\mathbb{Z}_J \times \mathbb{Z})$. Now define $\psi_\eta,\ \eta \in Ze(l)$ by $\psi_\eta = \phi_{l,\eta}$, if $\eta \in \mathbb{Z}$, and $\psi_\eta = \psi_{j,k}$, if $\eta = (j,k) \in {}_l\mathbb{Z} \times \mathbb{Z}$.

Given an unknown function f we can now write its wavelet expansion as

$$f = \sum_{\eta \in Ze(l)} \beta_\eta \psi_\eta, \tag{6}$$

where $\{\psi_\eta, \eta \in Ze(l)\}$ is a compactly supported wavelet basis. From now on we will denote by ℓ the Lebesgue measure. The wavelet coefficients are given by

$$\beta_\eta = \int f \psi_\eta d\ell. \tag{7}$$

Our interest in what follows will be in the case that f is a density, supposed to belong to $L_2(\mathbb{R})$ or $L_2(\mathbb{R}^d)$. In the case of $f(\mathbf{x})$, for $\mathbf{x} = (x_1, \ldots, x_d)'$, wavelet expansions similar to (3) for f will hold, where the wavelets are obtained as products of one-dimensional wavelets. See Vidakovic (1999) for details. We illustrate here two possibilities of wavelet expansions for the case $d = 2$.

One possibility is to consider a basis with a single scale. Define the bivariate scaling function as $\Phi(x_1, x_2) = \phi(x_1)\phi(x_2)$ and the wavelets by $\Psi^h(x_1, x_2) = \phi(x_1)\psi(x_2)$, $\Psi^v(x_1, x_2) = \psi(x_1)\phi(x_2)$ and $\Psi^d(x_1, x_2) = \psi(x_1)\psi(x_2)$. Let $\mathbf{k} = (k_1, k_2)$. Then a wavelet expansion for $f(x_1, x_2)$ is

$$f(x_1, x_2) = \sum_{\mathbf{k}} c_{\mathbf{k}} \Phi_{l,\mathbf{k}}(x_1, x_2) + \sum_{j=l}^{\infty} \sum_{\mathbf{k}} \sum_{\mu=h,v,d} d^{\mu}_{j,\mathbf{k}} \Psi^{\mu}_{j,\mathbf{k}}(x_1, x_2), \tag{8}$$

where $\Phi_{l,\mathbf{k}}(x_1,x_2) = \phi_{l,k_1}(x_1)\phi_{l,k_2}(x_2)$, $\Psi^h_{j,\mathbf{k}}(x_1,x_2) = \phi_{j,k_1}(x_1)\psi_{j,k_2}(x_2)$, $\Psi^v_{j,\mathbf{k}}(x_1,x_2) = \psi_{j,k_1}(x_1)\phi_{j,k_2}(x_2)$ and $\Psi^d_{j,\mathbf{k}}(x_1,x_2) = \psi_{j,k_1}(x_1)\psi_{j,k_2}(x_2)$, with the wavelet coefficients given by

$$c_{\mathbf{k}} = \int\int f(x_1,x_2)\Phi_{l,\mathbf{k}}(x_1,x_2)dx_1dx_2, \tag{9}$$

$$d^{\mu}_{j,\mathbf{k}} = \int\int f(x_1,x_2)\Psi^{\mu}_{j,\mathbf{k}}(x_1,x_2)dx_1dx_2.$$

Another possibility is to build a basis as the tensor product of two one-dimensional bases with combinations of all scales from each dimension. Here, if $\mathbf{j} = (j_1, j_2), \mathbf{k} = (k_1, k_2)$, we have

$$f(x_1,x_2) = \sum_{\mathbf{k}} c_{\mathbf{k}}\phi_{l,k_1}(x_1)\phi_{l,k_2}(x_2) + \sum_{j_1\geq l}\sum_{\mathbf{k}} \alpha_{j_1,\mathbf{k}}\psi_{j_1,k_1}(x_1)\phi_{l,k_2}(x_2) + \tag{10}$$

$$\sum_{j_2\geq l}\sum_{\mathbf{k}} \beta_{j_2,\mathbf{k}}\phi_{l,k_1}(x_1)\psi_{j_2,k_2}(x_2) + \sum_{j_1,j_2\geq l}\sum_{\mathbf{k}} d_{\mathbf{j},\mathbf{k}}\psi_{j_1,k_1}(x_1)\psi_{j_2,k_2}(x_2),$$

and the wavelet coefficients obtained similarly as in the previous case. The two bases imply different tilings of the time-scale plane. In this work, see Section 4, we consider an extension of the expansion (10) to the d-dimensional case. We have used (8) in Morettin et al. (2010); see Section 3.

There are several possible choices for the wavelets to be used: Haar, compactly supported Daubechies wavelets, Shannon, Meyer, Mexican hat or Morlet wavelet. The latter is often used in physical sciences problems. We will not discuss here the issues concerning the particular choice of a wavelet family. Another possibility is to use B-splines, as in Cosma et al. (2007). We have used the Haar wavelet for the simulations and the empirical applications in this paper.

2.3 Density estimation

In order to estimate a copula it will be necessary to estimate a density function first. There is a huge literature now on nonparametric estimation of probability density functions. See for example Silverman (1986). One popular class of estimators is that of projection estimators (Cencov, 1962) that use orthogonal bases (Fourier, for example).

Concerning wavelet estimators of densities, the literature is also large. We may consider linear or nonlinear wavelet estimators, as we mentioned before for the case of copulas. For a general overview on density estimators via wavelets, see Vidakovic (1999, ch.7) and Härdle et al. (1998). For linear wavelet estimators, some references are Antoniadis and Carmona (1991) and Walter (1992).

Nonlinear wavelet estimators, that use thresholding and shrinking rules, were developed in a series of papers by Donoho and co-authors. See for example, Donoho et al. (1995, 1996), Delyon and Juditsky (1996) and Härdle et al (1998).

In this paper we consider linear wavelet estimators for the marginals and joint densities involved in the determination of a copula.

3 Estimation for i.i.d. and time series data

In this section we briefly discuss the main existing trends in copula estimators for the i.i.d. and time series data.

As we mentioned in Section 1, if a copula belongs to a parametric family of copulas, ML methods can be used. These are all well known and will not be discussed further here. The software S+FinMetrics, a module of S-Plus, implements at least one of these procedures. See Zivot and Wang (2006) for further details.

3.1 Estimation for i.i.d. data

Kernel estimators

Suppose that we have data $(X_i, Y_i), i = 1, \dots, n$. Fermanian et al. (2004) proposed to use

$$\hat{F}_n(x, y) = \frac{1}{n} \sum_{i=1}^{n} K_n(x - X_i, y - Y_i), \tag{11}$$

as a smoothed empirical distribution function estimator. Here $K_n(x, y) = K(a_n^{-1}x, a_n^{-1}y)$ and

$$K(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} k(u, v) du dv,$$

for some bivariate kernel function $k : \mathbb{R}^2 \to \mathbb{R}$, $\int\int k(x,y)dxdy = 1$, and a sequence of bandwidths $a_n \downarrow 0$, as $n \to \infty$. It is proved that for small enough bandwidths a_n, under mild conditions,

$$\sqrt{n} \sup_{x,y} |\hat{F}_n(x,y) - F_n(x,y)| \xrightarrow{P} 0.$$

Similar smoothed estimators $\hat{F}_{1n}$ and $\hat{F}_{2n}$ of the marginal distributions can be proposed, using univariate kernels. A kernel smoothed empirical copula estimator is then obtained from (2), namely

$$\hat{C}_n(u,v) = \hat{F}_n(\hat{F}_{1n}^{-1}(u), \hat{F}_{2n}^{-1}(v)), \quad 0 \le u, v \le 1. \tag{12}$$

Fermanian et al. (2004) proved that the smoothed empirical copula process

$$\hat{Z}_n(x,y) = \sqrt{n}(\hat{C}_n - C)(x,y), \ 0 \le x, y \le 1,$$

converges to a Gaussian process in $L_\infty([0,1]^2)$.

Wavelet estimators

Morettin et al. (2010) proposed a wavelet smoothed empirical copula estimator. Since the copula $C(u,v)$ belongs to $L_2([0,1]^2)$, consider its wavelet expansion, see (8),

$$C(u,v) = \sum_{\mathbf{k}} c_{\mathbf{k}} \Phi_{l,\mathbf{k}}(u,v) + \sum_{j=l}^{\infty} \sum_{\mathbf{k}} \sum_{\mu=h,v,d} d_{j,\mathbf{k}}^{\mu} \Psi_{j,\mathbf{k}}^{\mu}(u,v), \tag{13}$$

with the wavelet coefficients given by

$$c_{\mathbf{k}} = \int\int C(u,v)\Phi_{l,\mathbf{k}}(u,v)dudv, \ \ d_{j,\mathbf{k}}^{\mu} = \int\int C(u,v)\Psi_{j,\mathbf{k}}^{\mu}(u,v)dudv. \tag{14}$$

As estimates of the wavelet coefficients take the *empirical wavelet coefficients*,

$$\hat{d}_{j,\mathbf{k}}^{\mu} = \int C_n(u,v)\Psi_{j,\mathbf{k}}^{\mu}(u,v)dudv, \tag{15}$$

and a similar expression for $c_{\mathbf{k}}$, where C_n is the empirical copula function, which is defined similarly to (12), with the kernel estimators of the distribution functions replaced by the empirical distribution function estimators.

Thus the corresponding estimator for $C(u, v)$ is

$$\hat{C}(u, v) = \sum_{\mathbf{k}} \hat{c}_{\mathbf{k}} \Phi_{l,\mathbf{k}}(u, v) + \sum_{j=l}^{\infty} \sum_{\mathbf{k}} \sum_{\mu} \delta(\hat{d}^{\mu}_{j,\mathbf{k}}, \lambda) \Psi^{\mu}_{j,\mathbf{k}}(u, v), \qquad (16)$$

where $\delta(\cdot, \lambda)$ is a threshold. Hard and soft thresholds are often used. See the above mentioned paper for details and applications.

3.2 Estimation for time series data

As remarked in Section 1, most of the results available in the literature of copulas apply to i.i.d. samples $(X_i, Y_i), i = 1, \ldots, T$, from a distribution function F.

In this section we discuss copula estimation techniques in the presence of time series data. One approach often used is to apply directly the methods available for i.i.d. data (mostly using parametric copula models), which may be misleading.

Fitting univariate and multivariate models

This method, used for example by Dias and Embrechts (2009, 2010) and Patton (2006), consists in estimating the copula for the standardized residuals after fitting linear and/or non-linear univariate or multivariate models to the series. For the use of semiparametric models, see Fan and Chen (2004).

Nonparametric estimation

Let $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ be a d-dimensional stochastic process, and suppose we have observations $\{\mathbf{X}_t, t = 1, \ldots, T\}$. In what follows, let $d = 2$ for simplicity. Assume first that the process is strictly stationary, and let $F(\mathbf{x})$ be the distribution function (d.f.) of $\mathbf{X}_t = (X_{1t}, X_{2t})'$ at $\mathbf{x} = (x_1, x_2)'$.

Let $F_j(x_j), j = 1, 2$, denote the marginal d.f.'s and $f_j(x_j)$ the corresponding probability density functions (p.d.f.). By (2), to estimate the copula C we need to estimate the marginal d.f.'s F_1, F_2, the quantiles $F_1^{-1}(u), F_2^{-1}(v)$ and the joint distribution function F.

Fermanian and Scaillet (2003) (FS from here on) use kernel estimates for C. First estimate the marginals by

$$\hat{f}_j(x_j) = (Th_j)^{-1} \sum_{t=1}^{T} k_j \left(\frac{x_j - X_{jt}}{h_j} \right), \qquad (17)$$

for some kernel function k_j and associated bandwidth h_j, $j = 1, 2$. Then estimate $f(\mathbf{x})$ by

$$\hat{f}(\mathbf{x}) = (T|h|)^{-1} \sum_{t=1}^{T} k(\mathbf{x} - \mathbf{X}_t; h), \tag{18}$$

where $k(\mathbf{x}) = \prod_{j=1}^{2} k_j(x_j)$, $h = diag\{h_1, h_2\}$ and $|h| = h_1 h_2$. Next, estimate F_j by

$$\hat{F}_j(x_j) = \int_{-\infty}^{x_j} \hat{f}_j(x) dx, \tag{19}$$

and F by

$$\hat{F}(\mathbf{x}) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \hat{f}(\mathbf{x}) d\mathbf{x}. \tag{20}$$

Finally, estimate the copula C by

$$\hat{C}(\mathbf{u}) = \hat{F}(\hat{\xi}), \tag{21}$$

where $\hat{\xi} = (\hat{\xi}_1, \hat{\xi}_2)'$ and $\hat{\xi}_j$ is the kernel estimator of the quantile of X_{jt} with probability u_j. FS prove some asymptotic results for the various estimators, assuming that the process is strongly mixing and further conditions on F_j, h_j and h.

Doukham et al. (2009) deals with a special form of weakly dependent sequences, which in principle is easier to be fulfilled than the strong mixing.

4 Wavelet estimators

Of course the estimator (16) can be used for time series data. See Morettin et al. (2010) for details. In this section we propose another wavelet estimator, following the same steps as FS.

From here on we will always assume the following:

Assumption 1. *The d-dimensional process $\{\mathbf{X}_t, t \in \mathbb{Z}\}$, is such that for all t, $\mathbf{X}_t = (X_{1t}, \ldots, X_{dt})'$ has a density function $f_t(\mathbf{x})$ and these density functions are $t-$invariant, i.e., for all $t \in \mathbb{Z}$, and all $\mathbf{x} \in \mathbb{R}^d$, $f_t(\mathbf{x}) = f(\mathbf{x})$. Moreover, $f \in L_2(\mathbb{R}^d) \cap L_\infty(\mathbb{R}^d)$ and there exists, for all i, $1 \leq i \leq d$, the marginal density function $f_i(x_i)$ which is also assumed to belong to $L_2(\mathbb{R})$.*

Hence we do not assume stationarity but consider a larger class of processes. Clearly, this implies that the distribution functions of $\mathbf{X}_t$ as well as the marginals are $t-$invariant. We write $F_t(\mathbf{x}) = F(\mathbf{x})$, and, for all i, $F_{i,t}(x_i) = F_i(x_i)$ for all $t \in \mathbb{Z}$.

Suppose we know $\mathbf{X}_1, \ldots, \mathbf{X}_T$. As mentioned above, the marginal probability density functions (p.d.f.'s) and d.f.'s of each component X_{it} at x_i, $i = 1, \ldots, d$, will be written $f_i(x_i)$ and $F_i(x_i)$, respectively. Our purpose will be to estimate F, F_i and C in (2) using wavelet methods.

We first introduce some additional notation to the one introduced in Section 2 for wavelet expansions.

Let $J = (J_1, \ldots, J_d)$ and $l = (l_1, \ldots, l_d)$, where each $J_i = J_i(T), i = 1, \ldots, d$. Write $\eta = (\eta_1, \ldots, \eta_d) \in \mathcal{Z}e(l)_J := \prod_{i=1}^d Ze(l_i)_{J_i}$, $\eta_r = (j_r, k_r)$, or $\eta_r = k_r \in \mathbb{Z}, r = 1, \ldots, d$, and $\psi_\eta = \psi_{\eta_1} \otimes \cdots \otimes \psi_{\eta_d}$. Let $j(\eta) = (j(\eta_1), \ldots, j(\eta_d))$ such that $j(\eta_r) = j_r$, or $j(\eta_r) = l_r$, in case η_r is integer, $r = 1, \ldots, d$. Let $|j(\eta)| = \sum_{i=1}^d j(\eta_i)$. From here on, we will assume, without loss of generality, that $l = (0, ..., 0) \in \mathbb{Z}^d$.

Since each p.d.f. $f_i(x_i)$ can be expanded as in (3), an estimator is given by

$$\hat{f}_{i,J_i}(x_i) = \sum_k \hat{\alpha}_k \phi_{0,k}(x_i) + \sum_{j \geq 0}^{J_i} \sum_k \hat{\beta}_{j,k} \psi_{j,k}(x_i). \tag{22}$$

Then, we can estimate the d.f. $F_i(x_i)$ by

$$\hat{F}_{i,J_i}(x_i) = \int_{-\infty}^{x_i} \hat{f}_{i,J_i}(y) dy. \tag{23}$$

Using the above notation, we can write

$$\hat{f}_{i,J_i}(x_i) = \sum_{\eta_i \in Ze(0)_{J_i}} \hat{\beta}_{\eta_i} \psi_{\eta_i}(x_i). \tag{24}$$

and the estimate for the d.f is written as in (23).
Now, $f(\mathbf{x})$ can be expanded , using the above notation, as

$$f_J(\mathbf{x}) = \sum_{\eta \in \mathcal{Z}e(0)} \beta_\eta \psi_\eta(\mathbf{x}), \tag{25}$$

As an estimator of $f(\mathbf{x})$ we take

$$\hat{f}_J(\mathbf{x}) = \sum_{\eta \in \mathcal{Z}e(0)_J} \hat{\beta}_\eta \psi_\eta(\mathbf{x}), \tag{26}$$

and we will simply write $\hat{f}_J(\mathbf{x}) = \sum_\eta \hat{\beta}_\eta \psi_\eta(\mathbf{x})$ or even $\hat{f}_J(\mathbf{x}) = \sum \hat{\beta}_\eta \psi_\eta(\mathbf{x})$ where we will always remember that the sum is for $\eta \in \mathcal{Z}e(0)_J$. This will be clear from the context.

This is a *linear estimator*. We may consider nonlinear estimators by replacing $\hat{\beta}_\eta$ in (26), for example, by $\delta(\hat{\beta}_\eta, \lambda)$, where $\delta(\cdot, \lambda)$ is a threshold and λ is a threshold parameter which can be specified in a number of ways. But in this paper we will consider only linear estimates.

The empirical wavelet coefficients are given by

$$\hat{\beta}_\eta = \frac{1}{T} \sum_{t=1}^{T} \psi_\eta(\mathbf{X}_t) \tag{27}$$

and if

$$F_T(\mathbf{x}) = \frac{1}{T} \sum_{t=1}^{T} I\{\mathbf{X}_t \leq \mathbf{x}\}$$

is the empirical distribution function, then the estimator (27) can be written as

$$\hat{\beta}_\eta = \int_{\mathbb{R}^d} \psi_\eta(\mathbf{x}) dF_T(\mathbf{x}).$$

It follows that these are unbiased estimators of the corresponding coefficients.

As an estimator of the distribution function F of $\mathbf{X}_t$ at $\mathbf{x}$ we take

$$\hat{F}_J(\mathbf{x}) = \int_{-\infty}^{\mathbf{x}} \hat{f}_J(\mathbf{y}) d\mathbf{y}. \tag{28}$$

Let $\hat{F}_{i,J_i}$ be the wavelet estimator of the i-th marginal distribution F_i, $i = 1, \ldots, d$. By (2), to estimate the copula at some point $\mathbf{u} = (u_1, \ldots, u_d)$, we propose

$$\hat{C}_J(\mathbf{u}) = \hat{F}_J(\hat{\mathbf{x}}), \tag{29}$$

where $\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_d)$ and $\hat{x}_i = \inf\{y_i \in \mathbb{R} : \hat{F}_{i,J_i}(y_i) \geq u_i\}$, $i = 1, \ldots, d$, is the wavelet estimator of the quantile of X_{it} with probability u_i, $i = 1, \ldots, d$.

5 Properties of the estimators

In this section we derive some properties of the wavelet estimators. We first derive results on the covariance structure of the empirical wavelet coefficients estimators and then on estimators of d.f.'s and copulas. These results are presented in Theorems 1 to 5. Remember that the process will always be assumed to satisfy *Assumption 1.*

Let $f_{t,s}(\mathbf{x},\mathbf{y})$ be the joint density of $\mathbf{X}_t$ and $\mathbf{X}_s$ and $f(\mathbf{x})$ the density of $\mathbf{X}_t$, for every t, s. Define, for $t \neq s$,

$$q_{t,s}(\mathbf{x},\mathbf{y}) = f_{t,s}(\mathbf{x},\mathbf{y}) - f(\mathbf{x})f(\mathbf{y}), \tag{30}$$

and assume that $q_{t,s}$ has the wavelet expansion

$$q_{t,s}(\mathbf{x},\mathbf{y}) = \sum_{\mu}\sum_{\rho} \gamma_{\mu,\rho}^{(t,s)} \psi_\mu \otimes \psi_\rho(\mathbf{x},\mathbf{y}). \tag{31}$$

The proofs of the results that follow are given in the Appendix. We now set down some further assumptions and notation. We denote by $\sum_1$ the sum over the η such that $j(\eta) \leq J$ and by $\sum_2$ the sum over the complement. In some instances we will assume that $J_i \to \infty$, for all $i = 1, \ldots, d$, as $T \to \infty$, but in such a way that $J/T \to 0$, as $T \to \infty$. This will be clear from the context. When we write $\sum_{t \neq s}$ we mean that the sum is over s and t in $1, \ldots, T$, with $t \neq s$. Call $\mathcal{M}$ the set of all $\eta \in \mathcal{Z}e(l)_J$, such that the corresponding wavelets are products of mother wavelets only, and $\mathcal{F} = \mathcal{Z}e(l)_J \setminus \mathcal{M}$. We also will denote scales by $\mathbf{s} = (s_1, \ldots, s_d)$ and $|\mathbf{s}| = \sum_{i=1}^d s_i$. Let $\eta(\mathbf{s},\mathbf{x})$ be the only index η, in scale $\mathbf{s}$, such that $\int_{-\infty}^{\mathbf{x}} \psi_\eta d\mathbf{x}$ is different from zero.

Assumption 2: $\sum_{t \neq s} \gamma_{\eta,\xi}^{(t,s)} = o(T^2),$ *as* $T \to \infty,$ *for all* $\eta, \xi.$

Assumption 3: $\sup_\eta \sum_{t \neq s} \gamma_{\eta,\eta}^{(t,s)} = o(T^2), \;\; T \to \infty.$

Assumption 4: $\sum_\eta \sum_{t \neq s} \gamma_{\eta,\eta}^{(t,s)} = o(T^2), \;\; T \to \infty.$

Assumption 5: $\sum_{\eta \in \mathcal{F}} |\beta_\eta| 2^{-|j(\eta)|/2} < \infty.$

Assumption 6: $2^{\sum J_i} = o(T),$ *as* $T \to \infty.$

Assumption 7: $\sum_{\mathbf{s}} \sum_{\mathbf{s}'} 2^{-(|\mathbf{s}|+|\mathbf{s}'|)/2} |\sum_{t \neq s} \gamma_{\eta(\mathbf{s},\mathbf{x}),\xi(\mathbf{s}',\mathbf{y})}^{(t,s)}| = o(T^2),$ *as* $T \to \infty,$ *for all* $\mathbf{x}, \mathbf{y}$, *plus three similar conditions for* $(\mathbf{s}, \xi \in \mathcal{F}),\ (\eta \in \mathcal{F}, \mathbf{s}'), (\eta \in \mathcal{F}, \xi \in \mathcal{F}).$

Assumption 8: $\sum_{\mathbf{s}} \sum_{\mathbf{s}'} 2^{-[((|\mathbf{s}|+|\mathbf{s}'|)+||\mathbf{s}|-|\mathbf{s}'||)/2]} = o(T),$ *as* $T \to \infty$, *plus three similar conditions for* $(\mathbf{s}, \xi \in \mathcal{F}), (\eta \in \mathcal{F}, \mathbf{s}'), (\eta \in \mathcal{F}, \xi \in \mathcal{F}).$

DOI: 10.2202/1941-1928.1033

Assumption 9: *F is Lipschitz and for all i, $1 \leq i \leq d$, F_i is differentiable, with $\inf_{x \in S} F_i^{'}(x) > 0$, for every bounded interval $S \subset \mathbb{R}$.*

Assumptions 2, 3 and 4 are forms of mixing conditions, since they involve dependence between X_t and X_s. Essentially they say that the covariance between X_t and X_s decreases as $|s-t|$ increases. It would be interesting to see the connections of these assumptions and other forms of asymptotic independence, but we will not pursue this here. We observe, however, that these conditions do not require asymptotic independence but rather asymptotic non-correlatedness only. Assumptions 6, 7 and 8 concern the behavior of the scales, as $T \to \infty$. Assumption 5 is related to the behavior of the wavelet coefficients in the expansion of f. See also Bosq (1998), who uses (27) and kernel estimators for densities. The assumptions 2 to 5 are, in fact, weaker than imposing for example that $\sum_{t,s} |\gamma_{\eta,\xi}^{(t,s)}| < \infty$ and $\sum_{\eta} |\beta_{\eta}| < \infty$; and assumptions 2 to 8 are weaker than imposing $\sum_{t,s} |\gamma_{\eta,\xi}^{(t,s)}| < \infty$, $\sum_{\eta} |\beta_{\eta}| < \infty$ and $2^{2\sum J_i} = o(T)$, as $T \to \infty$.

The next theorem gives the consistency of the empirical wavelet coefficients as well as the covariance structure of these coefficients. It also shows that these empirical coefficients are asymptotically non-correlated.

Theorem 1. *(i) $E(\hat{\beta}_{\eta}) = \beta_{\eta}$.*
(ii) $Cov(\hat{\beta}_{\eta}, \hat{\beta}_{\xi}) = \frac{1}{T}(\int \psi_{\eta}(\mathbf{x})\psi_{\xi}(\mathbf{x})f(\mathbf{x})d\mathbf{x} - \beta_{\eta}\beta_{\xi}) +$
$+\frac{1}{T^2}\sum_{t \neq s} \gamma_{\eta,\xi}^{(t,s)}$.
(iii) Under the Assumption 2, the empirical wavelet coefficients are consistent and asymptotically uncorrelated.

The next result is also of interest.

Theorem 2. *Under the Assumptions 4 and 6 we have*

$$E||\hat{f}_J(\mathbf{x}) - f(\mathbf{x})||_2^2 \to 0, \;\; T \to \infty. \tag{32}$$

As a consequence, we have that $E||\hat{f}_J(\mathbf{x}) - f(\mathbf{x})||_2 \to 0$, as $T \to \infty$.

Theorem 1 is used to derive the correlation structure of the wavelet estimator $\hat{F}_J$ of the distribution function $F(\mathbf{x})$. It is also shown that this estimator is consistent and asymptotically non-correlated.

Theorem 3. *(i) The covariance structure of $\hat{F}_J(\mathbf{x})$ is given by:*

$$Cov(\hat{F}_J(\mathbf{x}), \hat{F}_J(\mathbf{y})) = \sum_{\eta} \sum_{\xi} \int_{-\infty}^{\mathbf{x}} \int_{-\infty}^{\mathbf{y}} \psi_\eta \otimes \psi_\xi d\ell^2 \, Cov(\hat{\beta}_\eta, \hat{\beta}_\xi),$$

where the sums are for all scales up to and including scale J.
(ii) Under the Assumptions 5, 7 and 8 the estimators $\hat{F}_J(\cdot)$ are consistent and asymptotically non-correlated, no matter how $J \to \infty$, as $T \to \infty$.

The result that follows establishes the uniform convergence in probability of $\hat{F}_J$ to F.

Theorem 4. *Under the Assumption 3, we have that*

$$\sup_{\mathbf{x}} |\hat{F}_J(\mathbf{x}) - F(\mathbf{x})| \xrightarrow{P} 0, \quad T \to \infty. \tag{33}$$

From Theorem 4 we have the following result on the estimated copula.

Theorem 5. *Under the Assumptions 3 and 9, for all $\mathbf{u} \in (0,1)^d$, we have that*

$$\hat{C}_J(\mathbf{u}) \xrightarrow{P} C(\mathbf{u}), \quad T \to \infty. \tag{34}$$

6 Some simulations

In this section we present simulation examples of the wavelet estimators proposed in the previous section. We use the examples of Fermanian and Scaillet (2003) and present in the tables their results as well as ours.

The choice of $J = (J_1, J_2)$ is implemented by an heuristic approach. We further assume that $J_1 = J_2 = J^*$ and for $J^* = 2, 3, 4, 5$ we calculated biases, MSE, minimum and maximum values of these quantities and associated ranges. Then a value of J^* was chosen looking at an overall performance of the estimator according to these measures. For a data driven rule, see the next section. Using this rule we would find the same value as here.

(1) First, we consider a stationary bivariate autoregressive process of order one:

$$\mathbf{X}_t = \mathbf{A} + \mathbf{B}\mathbf{X}_{t-1} + \epsilon_t, \tag{35}$$

where $\mathbf{X}_t = (X_{1t}, X_{2t})$, with independent components and thus $C(u_1, u_2) = u_1 u_2$, $\epsilon_t \sim N(0, \Sigma)$, $\mathbf{A} = (1, 1)$, $\text{vec}(\mathbf{B}) = (0.25, 0, 0, 0.75)$ and $\text{vec}(\Sigma) = (0.75, 0, 0, 1.25)$.

DOI: 10.2202/1941-1928.1033

The number of Monte Carlo replications is 1000, while the data length is $T = 2^{10} = 1024$.

Table 1 (a) shows the bias, $E\hat{C}_J - C$, and mean squared error (MSE), $E[(\hat{C}_J - C)^2]$, computed for the Haar wavelet using the chosen value $J^* = 4$, as described above. All values (true value of the copula, bias and MSE) are expressed as multiples of 10^{-4}. Table 1 (b) shows the results of FS. However, we remark that FS used series of the same length 1,024, but with 5,000 replications. Figure 1 shows the estimated copula and the corresponding contour plots. The results may be considered satisfactory and we see that FS estimators have generally larger biases and MSE. As remarked by these authors, of particular interest are the pairs $C(.01, .01)$ and $C(.05, .05)$, since they measure dependence of joint extreme losses.

(2) We now turn to the case where the components of $\mathbf{X}_t$ are dependent processes, with $\mathbf{A} = (1, 1)$, $\text{vec}(\mathbf{B}) = (0.25, 0.2, 0.2, 0.75)$ and $\text{vec}(\Sigma) = (0.75, 0.5, 0.5, 1.25)$.

Since X_{1t} and X_{2t} are positively dependent, we have $C(u_1, u_2) > u_1 u_2$. Based on 1000 Monte Carlo replications with the data length $T = 1024$, Haar wavelets, $J^* = 4$ (chosen as explained above), the results are reported in Table 2 (a), and the corresponding results of FS are in Table 2 (b). Figure 2 shows the estimated copula and the corresponding contour plot, using the Haar wavelet. The values are higher than those for the independent case but still very satisfactory and also better for most of the cases than those obtained by Fermanian and Scaillet.

Table 1: Biases and MSE of estimators: independent case

(a) Haar wavelet estimator (1000 replications)

x10^{-4}	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.50,.50)	C(.75,.75)	C(.95,.95)	C(.99,.99)
True	1.00	25.00	625.00	2500.00	5625.00	9025.00	9801.00
Bias	0.00	-0.01	-0.04	-0.12	-0.07	-0.03	-0.18
MSE	0.00	0.00	0.00	0.00	0.01	0.01	0.01

(b) FS estimator with product of two Gaussian kernels (5,000 replications)

x10^{-4}	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.50,.50)	C(.75,.75)	C(.95,.95)	C(.99,.99)
True	1.00	25.00	625.00	2500.00	5625.00	9025.00	9801.00
Bias	-.09	-0.08	0.40	1.12	-0.90	-0.04	4.66
MSE	0.00	0.01	0.25	0.48	0.25	0.01	0.05

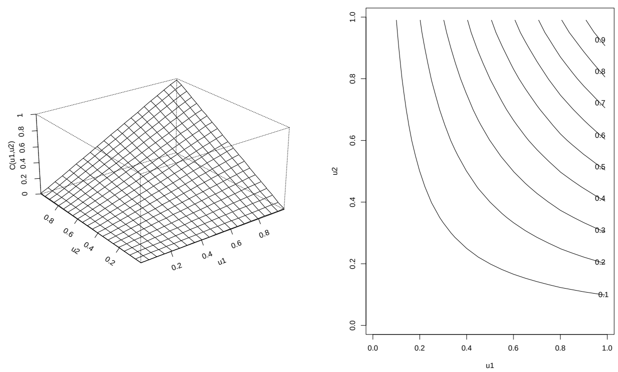

Figure 1: Wavelet estimator for Haar wavelet: independent case.

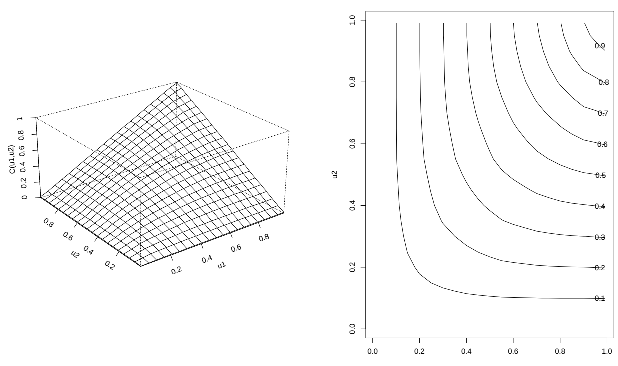

Figure 2: Wavelet estimator for Haar wavelet: dependent case.

Table 2: Biases and MSE of estimators: dependent case

(a) Haar wavelet estimator (1000 replications)

x10^{-4}	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.50,.50)	C(.75,.75)	C(.95,.95)	C(.99,.99)
True	27.08	197.95	1511.74	3747.68	6511.74	9197.95	9827.08
Bias	0.38	-2.36	-20.81	-32.75	-16.27	8.30	15.60
MSE	0.02	0.10	0.45	0.70	0.46	0.10	0.05

(b) FS estimator with product of two Gaussian kernels (5,000 replications)

x10^{-4}	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.50,.50)	C(.75,.75)	C(.95,.95)	C(.99,.99)
True	27.08	197.95	1511.74	3747.68	6511.74	9197.95	9827.08
Bias	-7.47	-34.88	-130.32	-172.28	-130.53	-35.25	-7.65
MSE	0.01	0.18	1.98	3.36	1.99	0.18	0.01

7 Empirical applications

In this section we illustrate the estimation procedure proposed considering two pairs of series. First, we consider daily returns of SP500 and DJIA, as in FS, recorded from 03/01/1994 to 07/07/2000, i.e. 1,700 observations, but considered only $T = 1,024$ to use a fast wavelet transform. Then, we use again an example of FS, considering the pair of stock indices CAC40-DAX35, for the same period as the pair SP500-DJIA and the same number of observations.

We suggest to use a rule of thumb based on Assumption 6. If $T = 2^p$, for an integer $p > 0$, it follows that we must have $\sum_{i=1}^{d} J_i < p$. In our case this results in $J_1 + J_2 < 10$; hence $J^* < 5$. We have chosen $J^* = 4$.

(1) Figure 3 shows the scatterplot of the returns of SP500 and DJIA. There is a high correlation between both series, specifically the contemporaneous correlation coefficient is 0.933. The wavelet estimator of the copula, using the Haar wavelet and $J^* = 4$ is presented in Figure 4.

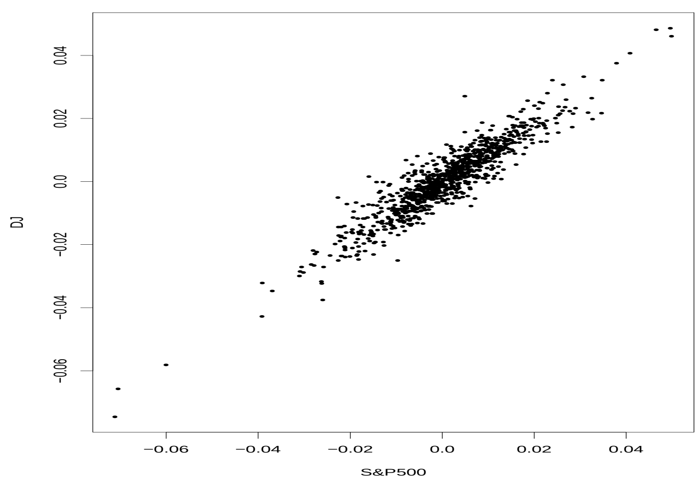

Figure 3: Scatterplot for the returns of SP500 and DJIA.

We see the expected comonotonic behaviour, due to the large dependence (two variables are comonotonic if one is almost surely an increasing function of the other; in this case, the contour curves are formed by two orthogonal lines, parallel to the axes). The Kendall and Spearman coefficients are $\tau = 0.7341$ and $\rho_S = 0.9009$, respectively. The plot is quite similar to the kernel copula estimator of FS.

(2) In Figure 5 we have the scatter plot of the returns of the stock indices CAC40-DAX35, as described above. The contemporaneous correlation coefficient is moderate, 0.67. Figure 6 brings the Haar wavelet estimator of the copula for the two series. The plot suggests a dependence, but not so strong as in the case of SP500-DJ. The Kendall τ is 0.4805 and the Spearman ρ_S is 0.6557.

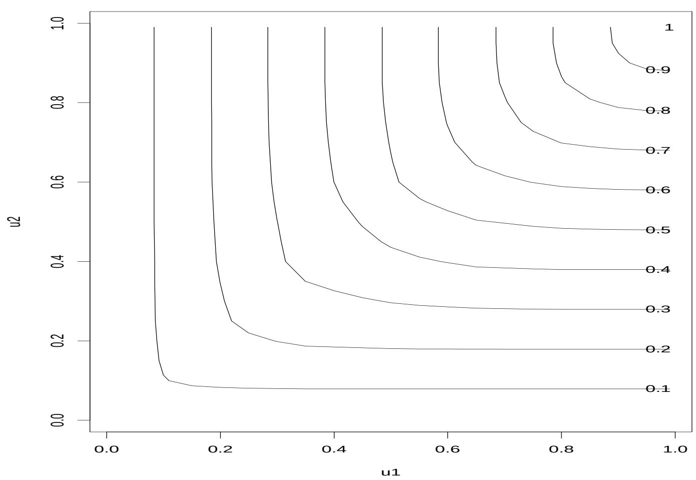

Figure 4: Contour plots of Haar wavelet estimated copula

8 Further remarks

In this work we have developed wavelet estimators of copulas for time series data that are assumed to be generated by a stochastic process. Under appropriate assumptions we have derived some statistical properties of estimators of distribution functions and copulas, including consistency and convergence in probability. We conjecture that the estimators are also asymptotically Gaussian and this will be pursued further.

We obtained the properties for linear wavelet estimators, considering upper scales which depend on the data set size and assuming that the process satisfies certain regularity conditions. It would be of interest to consider also thresholded nonlinear estimators and to derive rates of convergence for the risk of these estimators; this is a plan for future research.

Simulations have shown that the proposed estimators have a good performance compared to other nonparametric, namely kernel, estimators. About

DOI: 10.2202/1941-1928.1033

the processes considered, it was not necessary to assume strict stationarity. Estimators based on empirical copulas may also be considered. See Morettin et al. (2010) for details.

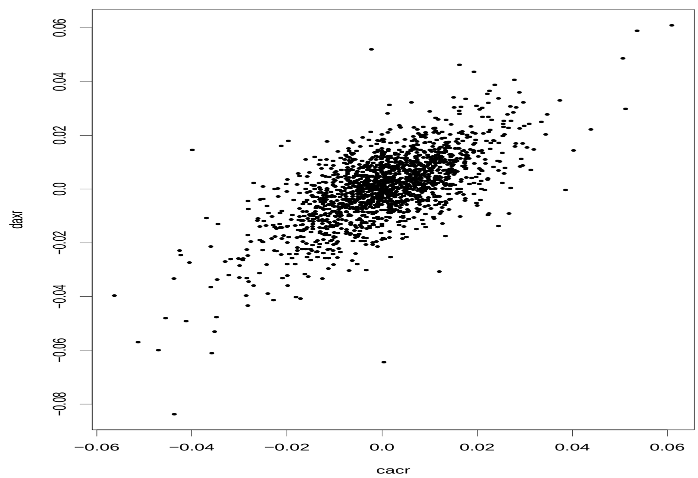

Figure 5: Scatter plot of returns of CAC40 and DAX35 series.

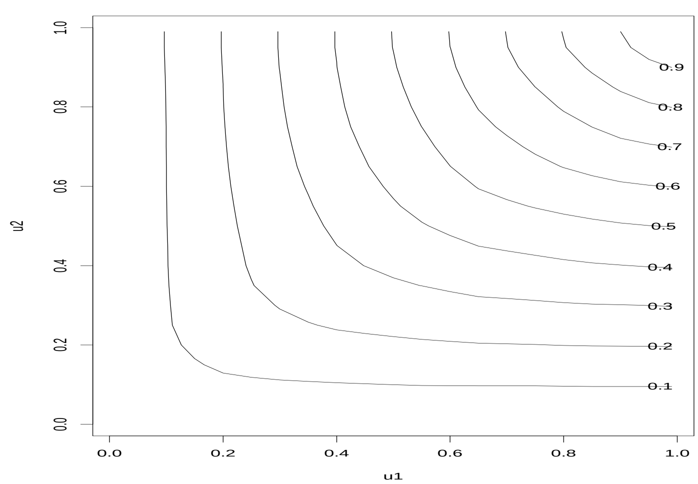

Figure 6: Contour plot of the Haar wavelet estimator for the stock returns of CAC40 and DAX35.

The wavelet estimators for densities can be negative (except for the Haar case). This can also happen to kernel estimators, in case the kernel is negative in some non null set of its domain of definition. To overcome this difficulty, we can proceed in several ways. One is to truncate the estimators to non-negative values and divide the result by the integral. Another is to expand the square root of f, as in Pinheiro and Vidakovic (1997). Walter and Shen (1999) construct non-negative estimators through a construction of non-negative scale functions and using a proper bi-orthogonal system. Building on works of Dechevsky and Penev (1997, 1998) in the univariate and i.i.d. case, Cosma et al. (2007) consider wavelet-based shape-preserving estimators for densities and distribution functions in the multivariate case and for dependent observations. Again, a bi-orthogonal system has to be used. The latter may be useful to derive shape-preserving estimators for copulas. This research is under investigation.

Appendix

Proof of Theorem 1. The first item is immediate, the second is

$$T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}\mathrm{Cov}\{\psi_\eta(\mathbf{X}_t),\psi_\xi(\mathbf{X}_s)\}.$$

Using (30) the covariance becomes

$$T^{-2}\{\sum_{t=s}\left(\int \psi_\eta(\mathbf{x})\psi_\xi(\mathbf{x})f(\mathbf{x})d\mathbf{x}-\beta_\eta\beta_\xi\right)$$
$$+\sum_{t\neq s}\int\int \psi_\eta(\mathbf{x})\psi_\xi(\mathbf{y})q_{t,s}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}\}.$$

By (31) the second term of the previous equation is

$$T^{-2}\sum_{t\neq s}\sum_{\mu}\sum_{\rho}\gamma_{\mu,\rho}^{(t,s)}\int\int \psi_\eta(\mathbf{x})\psi_\xi(\mathbf{y})\psi_\mu(\mathbf{x})\psi_\rho(\mathbf{y})d\mathbf{x}d\mathbf{y}$$
$$=T^{-2}\sum_{t\neq s}\gamma_{\eta,\xi}^{(t,s)},$$

due to the orthonormality of the wavelets. Item (iii) follows from the fact that the integral in the first term of the covariance is bounded by $||\psi_\eta||_\infty||\psi_\xi||_\infty$, since $f(\mathbf{x})$ is a density, and the second term tends to zero by Assumption 2.

Prof of Theorem 2. We can write

$$E\sum_1(\hat{\beta}_\eta-\beta_\eta)^2+\sum_2\beta_\eta^2=\sum_1\mathrm{Var}(\hat{\beta}_\eta)+\sum_2\beta_\eta^2=$$

$$=\sum_1\left(T^{-2}\sum_{t\neq s}\gamma_{\eta,\eta}^{(t,s)}+T^{-1}[\int\psi_\eta^2(\mathbf{x})f(\mathbf{x})d\mathbf{x}-\beta_\eta^2]\right)+\sum_2\beta_\eta^2.$$

Observe that $f\in L_2(\mathbb{R}^n)$ implies that $\sum_\eta\beta_\eta^2<\infty$ and this implies that $\lim_{J\to\infty}\sum_2\beta_\eta^2=0$. On the other hand, $0\leq T^{-1}\sum_1\beta_\eta^2\leq T^{-1}||f||^2\to 0,\ T\to\infty$, and

$$0\leq T^{-1}\sum_1\int\psi_\eta^2(\mathbf{x})f(\mathbf{x})d\mathbf{x}\leq T^{-1}\sum_1||f||_\infty\int\psi_\eta^2(\mathbf{x})d\mathbf{x}=$$

$$=T^{-1}||f||_\infty\sum_1 1\to 0, T\to\infty,$$

whenever we have $\#\{\eta:j(\eta)\leq J\}/T\ \to 0$, as $T\to\infty$, which is equivalent to Assumption 6. Hence the result follows by Assumption 4.

Proof of Theorem 3. (i) is immediate. (ii) Upon substitution of (ii) of Theorem 1 we obtain that the covariance in question is the sum of three terms:

$T^{-2}\sum_{\eta,\xi}\sum_{t\neq s}\gamma_{\eta,\xi}^{(t,s)}\left(\int_{-\infty}^{\mathbf{x}}\int_{-\infty}^{\mathbf{y}}\psi_\eta\otimes\psi_\xi d\ell^2\right)$
$+T^{-1}\sum_{\eta,\xi}\left(\int_{-\infty}^{\mathbf{x}}\int_{-\infty}^{\mathbf{y}}\psi_\eta\otimes\psi_\xi d\ell^2\right)\left(\int\psi_\eta(\mathbf{z})\psi_\xi(\mathbf{z})f(\mathbf{z})d\mathbf{z}-\beta_\eta\beta_\xi\right)=S_1+S_2+S_3$, say.

For S_1 we have that

$$|S_1|=|T^{-2}\sum_\eta\sum_\xi\int_{-\infty}^{\mathbf{x}}\psi_\eta d\ell\int_{-\infty}^{\mathbf{y}}\psi_\xi d\ell\sum_{t\neq s}\gamma_{\eta,\xi}^{(t,s)}|$$

and separating the sums in η and ξ into four sums and summing over scales $\mathbf{s}$ and $\mathbf{s}'$, we have:

$$|S_{11}|=|T^{-2}\sum_{\eta\in\mathcal{M}}\sum_{\xi\in\mathcal{M}}\int_{-\infty}^{\mathbf{x}}\psi_\eta dl\int_{-\infty}^{\mathbf{y}}\psi_\xi dl\sum_{t\neq s}\gamma_{\eta,\xi}^{(t,s)}|$$

$$\leq T^{-2}\sum_{\mathbf{s}}\sum_{\mathbf{s}'}2^{-|\mathbf{s}|/2}2^{-|\mathbf{s}'|/2}M_1^2|\sum_{t\neq s}\gamma_{\eta(\mathbf{s},\mathbf{x}),\xi(\mathbf{s}',\mathbf{y})}^{(t,s)}|,$$

where $M_1 = \max_{\eta:j(\eta)=0} || \int_{-\infty}^{\mathbf{x}} \psi_\eta dl||_\infty$ and $\eta(\mathbf{s},\mathbf{x})$ is the only index η, in scale $\mathbf{s}$, such that $\int_{-\infty}^{\mathbf{x}} \psi_\eta d\mathbf{x}$ is different from zero. The above term converges to zero, as $T \to \infty$, by Assumption 7. Next,

$$|S_{12}| = T^{-2} \sum_{\eta \in \mathcal{M}} \sum_{\xi \in \mathcal{F}} | \int_{-\infty}^{\mathbf{x}} \psi_\eta dl \int_{-\infty}^{\mathbf{y}} \psi_\xi dl \sum_{t \neq s} \gamma_{\eta,\xi}^{(t,s)} |$$

$$\leq T^{-2} | \sum_{\mathbf{s}} \sum_{\xi \in \mathcal{F}} \int_{-\infty}^{\mathbf{x}} \psi_{\eta(\mathbf{s},\mathbf{x})} dl || \int_{-\infty}^{\mathbf{y}} \psi_\xi dl || \sum_{t \neq s} \gamma_{\eta(\mathbf{s},\mathbf{x}),\xi}^{(t,s)} |$$

$$\leq T^{-2} \sum_{\mathbf{s}} M_1 2^{-|\mathbf{s}|/2} \sum_{\xi \in \mathcal{F}} M_1 2^{-|j(\xi)|/2} | \sum_{t \neq s} \gamma_{\eta(\mathbf{s},\mathbf{x}),\xi}^{(t,s)} |$$

$$= T^{-2} M_1^2 \sum_{\mathbf{s}} 2^{-|\mathbf{s}|/2} \sum_{\xi \in \mathcal{F}} 2^{-|j(\xi)|/2} | \sum_{t \neq s} \gamma_{\eta(\mathbf{s},\mathbf{x}),\xi}^{(t,s)} |,$$

and this converges to zero by Assumption 7. Analogously, we have for the third sum

$$|S_{13}| = |T^{-2} \sum_{\eta \in \mathcal{F}} \sum_{\xi \in \mathcal{M}} \int_{-\infty}^{\mathbf{x}} \psi_\eta dl \int_{-\infty}^{\mathbf{y}} \psi_\xi dl \sum_{t \neq s} \gamma_{\eta,\xi}^{(t,s)} |$$

$$\leq T^{-2} M_1^2 \sum_{\mathbf{s}} 2^{-|\mathbf{s}|/2} \sum_{\eta \in \mathcal{F}} 2^{-|j(\eta)|/2} | \sum_{t \neq s} \gamma_{\eta,\xi(\mathbf{s},\mathbf{y})} |$$

and again this converges to zero by Assumption 7. Finally,

$$|S_{14}| = |T^{-2} \sum_{\eta \in \mathcal{F}} \sum_{\xi \in \mathcal{F}} \int_{-\infty}^{\mathbf{x}} \psi_\eta dl \int_{-\infty}^{\mathbf{y}} \psi_\xi dl \sum_{t \neq s} \gamma_{\eta,\xi}^{(t,s)} |$$

$$\leq T^{-2} M_1^2 \sum_{\eta \in \mathcal{F}} \sum_{\xi \in \mathcal{F}} 2^{-(|j(\eta)|+|j(\xi)|)/2} | \sum_{t \neq s} \gamma_{\eta,\xi(\mathbf{s},\mathbf{y})}^{(t,s)} |,$$

which also converges to zero for the same reason. We now turn to S_2.

$$|S_2| = T^{-1} \sum_{\eta} \sum_{\xi} \int_{-\infty}^{\mathbf{x}} \psi_\eta dl \int_{-\infty}^{\mathbf{y}} \psi_\xi dl \int \psi_\eta \psi_\xi f dl.$$

Let $M_\infty = \max_{\eta:j(\eta)=0} \|\psi_\eta\|_\infty$. Then,

$$|S_2| \leq T^{-1} \sum_{\eta} \sum_{\xi} | \int_{-\infty}^{\mathbf{x}} \psi_\eta d\ell || \int_{-\infty}^{\mathbf{y}} \psi_\xi d\ell || \|f\|_\infty \int |\psi_\eta \psi_\xi| dl$$

$$\leq \frac{\|f\|_\infty M_1^2}{T} \sum_\eta \sum_\xi \|\psi_\eta\|_\infty \|\psi_\xi\|_\infty l(\mathrm{supp}\psi_\eta \cap \mathrm{supp}\psi_\xi)$$

$$\leq \frac{\|f\|_\infty M_1^2 M_\infty^2}{T} \sum_\eta \sum_\xi 2^{-(|j(\eta)| \vee |j(\xi)|)} V_0,$$

where $l(A)$ denotes the Lebesgue measure of A and $\mathrm{supp}\psi_\eta$ denotes the support of ψ_η. Notice that $l(\mathrm{supp}\psi_\eta \cap \mathrm{supp}\psi_\xi) \leq 2^{-(|j(\eta)| \vee |j(\xi)|)}.V_0$, where V_0 is the volume of $\mathrm{supp}(\bigotimes_{i=1}^d (\phi_{0,0})_i) = \prod_{i=1}^d [0, a_i]$.

Following a similar argument as for S_1, separating the sum into four sums, we can prove that $|S_2| \leq |S_{21}| + \ldots + |S_{24}| \to 0$, as $t \to \infty$, by Assumption 8.

Finally, for S_3, a similar argument holds, noticing that since $f \in L^2$, there exists η^* such that, for all η, $|\beta_\eta| \leq |\beta_{\eta^*}|$, and separating again S_3 into four sums, it is easy to see that each one converges to zero, as $T \to \infty$.

We now prove that the estimators are asymptotically unbiased, no matter how $J(T) \to \infty$, as $T \to \infty$. Since

$$E(\hat{F}_J(\mathbf{x})) = \int_{-\infty}^{\mathbf{x}} E\left(\sum_1 \hat{\beta}_\eta \psi_\eta(\mathbf{z})\right) d\mathbf{z} = \int_{-\infty}^{\mathbf{x}} \sum_1 \beta_\eta \psi_\eta(\mathbf{z}) d\mathbf{z},$$

we have

$$|F(\mathbf{x}) - E(\hat{F}_J(\mathbf{x}))| = |\int_{-\infty}^{\mathbf{x}} \sum_2 \beta_\eta \psi_\eta(\mathbf{z}) d\mathbf{z}|$$

$$= |\sum_2 \beta_\eta \int_{-\infty}^{\mathbf{x}} \psi_\eta d\ell| \leq \sum_{\mathbf{s}} |\beta_{\eta(\mathbf{s},\mathbf{x})}| \int |\psi_{\eta(\mathbf{s},\mathbf{x})}| d\ell + \sum_{\eta \in \mathcal{F}} |\beta_\eta| \int |\psi_\eta| d\ell$$

$$\leq M_1.|\beta_{\eta^*}| \sum_{\mathbf{s}} 2^{-|\mathbf{s}|/2} + \sum_{\eta \in \mathcal{F}} |\beta_\eta| 2^{-|j(\eta)|/2}$$

and the last term tends to 0, as $J \to \infty$, uniformly in $\mathbf{x}$, using Assumption 5. The sums in $\mathbf{s}$ above are for scales $\mathbf{s}$ in the complement of $\{\mathbf{s} : \mathbf{s} \leq J\}$. Therefore, we have $||F(\mathbf{x}) - E\hat{F}_J(\mathbf{x})||_\infty \to 0$, as $J \to \infty$. Hence the theorem follows.

Proof of Theorem 4. We have that

$$|\hat{F}_J(\mathbf{x}) - F(\mathbf{x})| \leq |\int_{-\infty}^{\mathbf{x}} \sum_1 (\hat{\beta}_\eta - \beta_\eta) \psi_\eta d\ell| + |\sum_2 \beta_\eta \int_{-\infty}^{\mathbf{x}} \psi_\eta d\ell| = S_1 + S_2.$$

The summation in S_2 will be done over all the scales $\mathbf{s}$, as in the proof of Theorem 3: $\sum_{\mathbf{s}} \sum_{\eta:j(\eta)=\mathbf{s}} \beta_\eta \int_{-\infty}^{\mathbf{x}} \psi_\eta d\ell$, but again there is at most one η for each scale such that the integral is not zero. Hence we have that

$$|S_2| = |\sum_{\mathbf{s}} \beta_{\eta(\mathbf{s},\mathbf{x})} \int_{-\infty}^{\mathbf{x}} \psi_\eta d\ell| \leq \sum_{\mathbf{s}} |\beta_{\eta(\mathbf{s},\mathbf{x})}|| \int_{-\infty}^{\mathbf{x}} \psi_\eta d\ell|$$

$$\leq M_1 \sum_{\mathbf{s}} \sup_\eta |\beta_{\eta(\mathbf{s},\mathbf{x})}| 2^{-|\mathbf{s}|/2},$$

for a positive constant M_1, and the sums in $\mathbf{s}$ mean that $\mathbf{s}$ is in the complement of $\mathbf{s} \leq J$. Hence,

$$|S_2| \leq M_1 \sup_\eta |\beta_\eta| \sum_{\mathbf{s}} 2^{-|s|/2} \leq M_1 \sup_\eta |\beta_\eta| \prod_{i=1}^{d} \sum_{s_i \geq 0} 2^{-s_i/2} = M_3 \sup_\eta |\beta_\eta|,$$

and the last term tends to zero as $J \to \infty$. The sums in $\mathbf{s}$, and η in the last expression are for all $\mathbf{s}$ and η such that $\mathbf{s}$ and $j(\eta)$ are not less than or equal to J.

Let us turn to S_1. Again, summing over scale,

$$|S_1| = |\sum_{\mathbf{s} \leq J} \sum_{\eta:j(\eta)=\mathbf{s}} (\hat{\beta}_\eta - \beta_\eta) \int_{-\infty}^{\mathbf{x}} \psi_\eta d\ell|$$

$$\leq \sup_{\eta:j(\eta)\leq J} |\hat{\beta}_\eta - \beta_\eta| \sum_{\mathbf{s}} M_1 2^{-|\mathbf{s}|/2} \leq M_1^* \sup_{\eta:j(\eta)\leq J} |\hat{\beta}_\eta - \beta_\eta|,$$

since the sum converges.

Call $v_\eta = \hat{\beta}_\eta - \beta_\eta$; we have $E(v_\eta) = 0$ and

$$\mathrm{Var}(v_\eta) = T^{-1} \left(\int \psi_\eta^2 f d\ell - \beta_\eta^2 \right) + T^{-2} \sum_{t \neq s} \gamma_{\eta,\eta}^{(t,s)}$$

$$\leq T^{-1} \left(\| f \|_\infty \int \psi_\eta^2 d\ell + \beta_\eta^2 \right) + T^{-2} \sum_{t \neq s} \gamma_{\eta,\eta}^{(t,s)}$$

$$\leq T^{-1} \left(\| f \|_\infty + \sup_\xi \beta_\xi^2 \right) + T^{-2} \sup_\xi |\sum_{t \neq s} \gamma_{\xi,\xi}^{(t,s)}|$$

$$= \frac{M_2}{T} + T^{-2} \sup_\xi |\sum_{t,s} \gamma_{\xi,\xi}^{(t,s)}| = g(T),$$

since $\sup_\xi \beta_\xi^2 < \infty$, given that $f \in L^2$. Hence, for all η, $\mathrm{Var}(v_\eta) \to 0$, as $T \to \infty$, if Assumption 3 holds. Therefore, we can write $\mathrm{Var}(v_\eta) \le g(T)$, such that $g(T) \to 0$, as $T \to \infty$. Let $h = \sup_\eta |v_\eta|$. We have

$$P(h > \lambda\sqrt{g(T)}) = P(\exists\eta : |v_\eta| > \lambda\sqrt{g(T)}) = 1 - P(\forall\eta : |v_\eta| \le \lambda\sqrt{g(T)}).$$

Now, on one hand, $P(|v_\eta| \le \lambda\sqrt{\mathrm{Var}(v_\eta)}) \ge 1 - 1/\lambda^2$, and on the other hand $P(|v_\eta| \le \lambda\sqrt{g(T)}) = 1 - \alpha_\eta$, say. But $1 - \alpha_\eta \ge 1 - 1/\lambda^2$ and so

$$P(h > \lambda\sqrt{g(T)}) \le 1 - (1 - \sum_\eta \alpha_\eta) = \sum_\eta \alpha_\eta \le \sum_\eta 1/\lambda^2,$$

and the last sum is $\#\{\eta : j(\eta) \le J\}/\lambda^2$. Therefore, it is enough to choose λ such that $2^{\sum J_i}/\lambda^2 \to 0$, as $T \to \infty$ and $\lambda\sqrt{g(T)} \to 0$, as $T \to \infty$. For example, $\lambda = 1/(g(T)^{1/4})$ and the scales J_i are such that $2^{\sum J_i}\sqrt{g(T)} \to 0$, as $T \to \infty$.

Proof of Theorem 5. Write $\hat{C}_J(\mathbf{u}) - C(\mathbf{u})$ as $\hat{F}_J(\hat{\mathbf{x}}) - F(\hat{\mathbf{x}}) + F(\hat{\mathbf{x}}) - F(\mathbf{x})$, hence $|\hat{C}_J(\mathbf{u}) - C(\mathbf{u})| \le |\hat{F}_J(\hat{\mathbf{x}}) - F(\hat{\mathbf{x}})| + |F(\hat{\mathbf{x}}) - F(\mathbf{x})|$. Using Theorem 4, it is enough to show that the second term of the r.h.s. converges to zero in probability as $J, T \to \infty$. By Assumption 9,

$$||F(\hat{\mathbf{x}}) - F(\mathbf{x})|| \le \kappa||\hat{\mathbf{x}} - \mathbf{x}|| \le \kappa.d. \max_{1\le i\le d} |\hat{x}_i - x_i|$$

Now, since $\hat{F}_J$ uniformly converges to F in probability, so do the estimated marginals, i.e., for all $i, 1 \le i \le d$, $\hat{F}_{i,J_i} \overset{u.P}{\to} F_i$.

Denote $S_i = (x_i - 1, x_i + 1)$ and $m_i = \inf_{z\in S_i} F'(z)$.

Thus, using the uniform convergence in probability of the marginal distribution estimators, we can write

$$\forall\delta > 0\ \forall\epsilon\ \ 0 < \epsilon < \frac{m_i}{2}\ \exists J_{*,i}\ \forall J_i \ge J_{*,i}\ \forall x \in \mathbb{R}\ P(|\hat{F}_{i,J_i}(x) - F_J(x)| < \epsilon) > 1 - \delta$$

from which we have

$$\forall x_i \in \mathbb{R}\ \forall\delta > 0\ \forall\epsilon\ \ 0 < \epsilon < \frac{m_i}{2}\ \exists J_{*,i}\ \forall J_i \ge J_{*,i} P(|\hat{x}_i - x_i| < \frac{\epsilon}{m_i}) > 1 - \delta.$$

Using Bonferroni's inequality we get

$$\forall(x_1, ..., x_d) \in \mathbb{R}^d\ \forall\delta > 0\ \forall\epsilon\ \ 0 < \epsilon < min_{1\le i\le d}\{\frac{m_i}{2}\}$$

$$\exists J_* = J_{*,1} \vee ... \vee J_{*,d}\ \forall J \geq J_* \quad P(\forall i\ 1 \leq i \leq d \quad |\hat{x}_i - x_i| < \frac{\epsilon}{m_i}) > 1 - d\delta$$

Thus, briefly,

$$P(|F(\hat{\mathbf{x}}) - F(\mathbf{x})| < \kappa \max_{i:1\leq i\leq d}(\frac{\epsilon}{m_i})) > 1 - d\delta,$$

i.e., for all $\mathbf{x} \quad |F(\hat{\mathbf{x}}) - F(\mathbf{x})| \xrightarrow{P} 0$.

References

[1] Antoniadis, A. and Carmona, R. (1991). Multiresolution analysis and wavelets for density estimation. TR, University of California, Irvine.

[2] Autin, F., Lepennec, E. and Tribouley, K. (2010). Thresholding methods to estimate the copula density. *Journal of Multivariate Analysis*, **101**, 200–222.

[3] Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes*. Second Edition. Lecture Notes in Statistics. Springer-Verlag.

[4] Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G. and Roncalli, T. (2000). *Copulas for Finance: A Reading Guide and Some Applications*. Groupe de Recherche Opérationelle, Credit Lyonnais, Paris.

[5] Cencov, N.N. (1962). Evaluation of an unknown distribution density from observations. *Doklady*, **3**, 1559–1562.

[6] Cherubini, U. and Luciano, E. (2001). Value-at-risk trade-off and capital allocation with copulas. *Economic Notes*, **30**, 235–256.

[7] Cosma, A., Scaillet, O. and von Sachs, R. (2007). Multivariate wavelet-based shape-preserving estimation for dependent observations. *Bernoulli*, **13**, 301–329.

[8] Daubechies, I. (1992) *Ten Lectures on Wavelets*. Philadelphia: SIAM.

[9] Dechevsky, L. and Penev, S. (1997). On shape preserving probabilistic wavelet approximators. *Stochastic Analysis and Applications*, **15**, 187–215.

[10] Dechevsky, L. and Penev, S. (1998). On shape preserving wavelet estimators of cumulative distribution functions and densities. *Stochastic Analysis and Applications*, **16**, 423–462.

DOI: 10.2202/1941-1928.1033

[11] Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés- Un test non paramétrique d'independence. *Académie Royale de Belgique-Bulletin de la Classe des Sciences-5e Série*, **65**, 274–292.

[12] Deheuvels,P. (1981) A non parametric test for independence. *Publications de l'Institut de Statistique de l'Université de Paris*, **26**, 29–50

[13] Delyon, B. and Juditsky, A. (1996). On minimax wavelet estimators. *Journal of Applied and Computational Harmonic Analysis*, **3**, 215–228.

[14] Dias, A. and Embrechts, P. (2009). Testing for structural changes in exchange rates dependence beyond linear correlation. *European Journal of Finance*, **15**, 619–637.

[15] Dias, A. and Embrechts, P. (2010). Modeling exchange rate dependence dynamics at different time horizons. *Journal of International Money and Finance*, **29**, 1687–1705.

[16] Donoho, D.L., Johnstone, I., Kerkyacharian, G. and Picard, D. (1995). Wavelet shrinkage: Asymptopia? *Journal of the Royal Statistical Society*, Series B, **57**, 301–369.

[17] Donoho, D.L., Johnstone, I., Kerkyacharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding. *Annals of Statistics*, **24**, 508–539.

[18] Doukham, P., Fermanian, J.-D. and Lang, G. (2009). An empirical central limit theorem with applications to copulas under weak dependence. *Statistical Inference for Stochastic Processes*, **12**, 65–87.

[19] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer.

[20] Embrecths, P., Lindskog, F. and McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. In Handbook of Heavy Tailed Distributions in Finance, ed. S. Rachev, Elsevier, Ch. 8, 329–384.

[21] Fan, Y. and Chen, X. (2004). Estimation of copula-based semiparametric time series models.*Econometric Society 2004 Far Eastern Meetings*, 559. Econometric Society.

[22] Fermanian, J.-D. and Scaillet, O. (2003). Nonparametric estimation of copulas for time series. *Journal of Risk*, **5**, 25–54.

[23] Fermanian, J.-D. and Wegkamp, M. (2004). Time dependent copulas. Preprint.

[24] Fermanian, J.-D., Radulovic, D. and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, **10**, 847–860.

[25] Gayraud, G. and Tribouley, K. (2010). A test of goodness of fit for the copula densities. Forthcoming, *Test*.

[26] Genest, C., Ghoudi, K. and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, **82**, 543–552.

[27] Genest, C., Masiello, E. and Tribouley, K. (2009). Estimating copula densities through wavelets. *Insurance: Mathematics and Economics*, **44**, 170–181.

[28] Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelets, Approximation, and Statistical Applications*. Lecture Notes in Statistics, 129. New York: Springer.

[29] Meyer, Y. (1993). *Wavelets: Algorithms and Applications*. Philadelphia: SIAM.

[30] Morettin, P.A., Toloi, C.M.C., Chiann, C. and de Miranda, J.C.S. (2010). Wavelet smoothed empirical copula estimators. *Brazilian Review of Finance*, **8**, 263–281.

[31] Nelsen, R. (2006). *An Introduction to Copulas*. Lecture Notes in Statistics. Second Edition. New York: Springer.

[32] Patton, A. (2001a). Modelling time-varying exchange rate dependence using the conditional copula. UCSD WP 2001-09.

[33] Patton, A. (2001b). Estimation of copula models for time series of possibly different lengths. UCSD WP 2001-17.

[34] Patton, A.J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Revue*, **47**, 527–556.

[35] Pinheiro, A.S. and Vidakovic, B. (1997). Estimating the square root of a density via compactly supported wavelets. *Computational Statistics and Data Analysis*, **25**, 399–415.

DOI: 10.2202/1941-1928.1033

[36] Shih, J. and Louis, T. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrika*, **51**, 1384-1399.

[37] Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

[38] Sklar, A.(1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229–231.

[39] Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. New York: Wiley.

[40] Walter, G.G. (1992). Approximating of the delta function by wavelets. *Journal of Approximation Theory*, **71**, 329–343.

[41] Walter, G.G. and Shen, X. (1999). Continuous non-negative wavelets and their use in density estimation. *Communications in Statistics-Theory and Methods*, **28**, 1–17.

[42] Zivot, E. and Wang, J. (2006). *Modeling Financial Time Series With S-Plus*. Second Edition. New York: Springer.