Brazilian Journal of Probability and Statistics
2020, Vol. 34, No. 3, 439–463
https://doi.org/10.1214/19-BJPS449

Copula estimation through wavelets

Francyelle L. Medina[a], Pedro A. Morettin[b] and Clélia M. C. Toloi[b]

[a]*Federal University of Pernambuco*
[b]*University of São Paulo*

Abstract. Recently some nonparametric estimation procedures have been proposed using kernels and wavelets to estimate the copula function. In this context, knowing that a copula function can be expanded in a wavelet basis, we propose a new nonparametric copula estimation procedure through wavelets for independent data and times series under an α-mixing condition. The main feature of this estimator is that we make no assumptions on the data distribution and there is no need to use ARMA–GARCH modelling before estimating the copula. Convergence rates for the estimator were computed, showing the estimator consistency. Some simulation studies are presented, as well as analysis of real data sets.

1 Introduction

In applications of insurance and risk management, copulas have been extensively studied as an important tool to describe the dependence structure between random variables and stochastic processes. These functions were introduced by Sklar (1959) and most of the literature focuses on parametric families of copulas like Gaussian, Student t, Frank, Clayton, etc. For further discussion about mathematical properties and definitions, see Nelsen (2005).

Several methods have been used for copula estimation. In the parametric approach, it is necessary to select a copula family and then estimate the parameters, usually by maximum likelihood. For time series data, the usual procedure is to fit ARMA–GARCH models and then estimate some parametric copula by considering the standardized residuals (for details, see Patton (2012)).

Nonparametric estimation methods have been widely used. For independent data, Genest, Massiello and Tribouley (2009) proposed a methodology based on the wavelet decomposition of the copula density, called rank-based estimator. Another reference is Autin, Le Pennec and Tribouley (2010), that used a nonlinear procedure based on thresholding methods.

Fermanian and Scaillet (2003) proposed copula estimators based on kernels. The procedure involves estimation of densities, distribution functions, quantiles and finally estimating the copula function, using the Sklar theorem. Morettin et al. (2010) present a new wavelet estimator, smoothing the empirical copula. They presented some simulation studies to assess the estimator performance and showed that the estimator outperformed the kernel-based estimator. But no proof of consistency was given. Morettin et al. (2011) used the same approach of Fermanian and Scaillet (2003) and derived statistical properties of the estimator.

In this work, we propose a new copula estimator through wavelets, for independent case and time series data. It is shown, under regularity conditions, that the estimator is consistent.

This paper is organized as follows. In Section 2, we propose the new estimation method of copulas through wavelets, for the case of independent data and for time series data. We present two theorems that show the consistency of the estimator for both cases. In Section 3, we perform some simulation studies and in Section 4 we apply the proposed techniques to some real data sets. In Section 5, we conclude with remarks about the applicability and advantages of the wavelet approach.

Key words and phrases. Copula, nonparametric estimation, wavelets, α-mixing processes.
Received July 2017; accepted May 2019.

2 Wavelet estimators

Since the copula function $C(u,v) \in L^2([0,1]^2)$, considering an appropriated wavelet basis, it can be expanded as

$$C(u,v) = \sum_{\mathbf{k}} c_{l,\mathbf{k}} \Phi_{l,\mathbf{k}}(u,v) + \sum_{j \geq l} \sum_{\mathbf{k} \in \mathbb{Z}^2} \sum_{\mu = h,v,d} d^{\mu}_{j,\mathbf{k}} \Psi^{\mu}_{j,\mathbf{k}}(u,v), \tag{2.1}$$

where

$$c_{l,\mathbf{k}} = \int_{[0,1]^2} C(u,v) \Phi_{l,\mathbf{k}}(u,v)\, du\, dv,$$
$$d^{\mu}_{j,\mathbf{k}} = \int_{[0,1]^2} C(u,v) \Psi^{\mu}_{j,\mathbf{k}}(u,v)\, du\, dv. \tag{2.2}$$

For details on copulas, see Nelsen (2005), and for details on wavelets and wavelets expansions, for the bivariate case, see Vidakovic (1999) and Morettin (2014).

Therefore, to estimate the copula function given by (2.1), it is only necessary to estimate the wavelet coefficients given by (2.2).

In this section, we propose and discuss copula estimation techniques for i.i.d. case and time series data.

It is known that the space $L^2([0,1]^2)$ can also be generated by the father wavelets $\{\Phi_{l,\mathbf{k}}(x,y), \mathbf{k} = (k_1,k_2)\}_{\mathbf{k}}$, hence instead of (2.1) we may consider

$$C_l(u,v) = \sum_{\mathbf{k}} c_{l,\mathbf{k}} \Phi_{l,\mathbf{k}}(u,v), \tag{2.3}$$

with

$$c_{l,\mathbf{k}} = \int_{[0,1]^2} C(u,v) \Phi_{l,\mathbf{k}}(u,v)\, du\, dv$$
$$= \int_{[0,1]^2} \left[\int_s^1 \int_r^1 \Phi_{l,\mathbf{k}}(u,v)\, du\, dv \right] c(r,s)\, dr\, ds, \tag{2.4}$$

where l is an arbitrary resolution level and $c(r,s)$ is the copula density.

Considering $r = F(x)$ and $s = G(y)$, it is easy to see that

$$c_{l,\mathbf{k}} = \mathbb{E}_{h(x,y)} \left[\int_{G(Y)}^1 \int_{F(X)}^1 \Phi_{l,\mathbf{k}}(u,v)\, du\, dv \right]. \tag{2.5}$$

2.1 Estimation for i.i.d. case

In order to develop the estimation procedure, let (X_i, Y_i), $i = 1, \dots, n$, be a random sample from a distribution function $H(\cdot,\cdot)$, where the marginal distribution functions $F(\cdot)$ and $G(\cdot)$ are unknown. Let F_n and G_n be their empirical counterparts respectively, that is, $F_n(X_i) = \frac{1}{n}\sum_{k=1}^n \mathbb{I}\{X_k \leq X_i\}$ and $G_n(Y_i) = \frac{1}{n}\sum_{k=1}^n \mathbb{I}\{Y_k \leq Y_i\}$, where $\mathbb{I}\{\mathbf{x} \in \mathcal{B}\}$ denotes the indicator function, that is, $\mathbb{I}\{\mathbf{x} \in \mathcal{B}\} = 1$ if $\mathbf{x} \in \mathcal{B}$ and $\mathbb{I}\{\mathbf{x} \in \mathcal{B}\} = 0$ otherwise.

From (2.5), the proposed estimator for $c_{l,\mathbf{k}}$ is given by

$$\tilde{c}_{l,\mathbf{k}} = \frac{1}{n} \sum_{i=1}^n \left[\int_{G_n(Y_i)}^1 \int_{F_n(X_i)}^1 \Phi_{l,\mathbf{k}}(u,v)\, du\, dv \right],$$

and the estimator for $C(u,v)$ is defined by

$$\tilde{C}_l(u,v) = \sum_{\mathbf{k}} \tilde{c}_{l,\mathbf{k}} \Phi_{l,\mathbf{k}}(u;v).$$

In order to show the performance of the proposed wavelet estimator, we carry out numerical studies, in which we will use the Mean Integrated Squared Error (MISE), defined by

$$\begin{aligned}
&\mathrm{MISE}\big(\tilde{C}_l(u,v), C(u,v)\big) \\
&\quad = \mathbb{E}_{h(x,y)} \big\| \tilde{C}_l(u,v) - C(u,v) \big\|_2^2 \\
&\quad = \mathbb{E}_{h(x,y)} \bigg[\int_0^1 \int_0^1 \big(\tilde{C}_l(u,v) - C(u,v)\big)^2 \, du \, dv \bigg].
\end{aligned} \tag{2.6}$$

To derive some properties of the wavelet estimator, suppose that the following assumptions hold:

(A1) C belongs $L_2([0,1]^2)$ and to the ball of radius $M > 0$ in the Besov space $\mathfrak{B}_2^{s,q}$.

(A2) For every integer $h \in \mathbb{Z}$, the joint distribution $J((X_t; Y_t); (X_{t+h}; Y_{t+h}))$ exists and there is a positive constant $M > 0$ such that, for every bounded zero-mean random variable $H(X_t; Y_t)$ we have

$$\mathbb{E}\big[\big|H(X_t; Y_t).H(X_{t+h}; Y_{t+h})\big|\big] \le M \mathbb{E}\big[\big|H(X_t; Y_t)\big|\big] \mathbb{E}\big[\big|H(X_{t+h}; Y_{t+h})\big|\big].$$

(A3) A bivariate process $\{(X_t, Y_t), t \in \mathbb{Z}\}$ is*α-mixing* and the coefficients $\alpha(p)$ are such that, for $r > 2$,

$$\sum_{p=N}^{\infty} [\alpha(p)]^{1-\frac{2}{r}} = O(N^{-1}).$$

(A4) $\{X_t, t \in \mathbb{Z}\}$ and $\{Y_t, t \in \mathbb{Z}\}$ are both α-mixing processes.

Then, we have the following theorem, that shows the estimator consistency for the independent case.

Theorem 2.1. *Under the assumption* (A1), *given a sample of size n from a bivariate distribution* $H(\cdot,\cdot)$, *with an unknown copula function* $C(\cdot,\cdot)$, *choose* l^*, *such that*

$$2^{l^*} \le n^{\frac{1}{2(s+4)}} < 2^{l^*+1}.$$

Let $\tilde{C}_{l^*}(\cdot,\cdot)$ *be the estimator of* $C(\cdot,\cdot)$ *up to resolution level* l^*.

Then, there exists a constant $K > 0$ *such that*

$$\sup_{C \in \mathfrak{B}^q_{s,2}(M)} n^{\frac{s+2}{s+4}} \mathrm{MISE}\big(\tilde{C}_{l^*}(u,v), C(u,v)\big) \le K.$$

Proof. See the Appendix. □

2.2 Estimation for time series case

Considering the proposed estimator for the time series case, the objective is to use some dependence structure and to assume that the processes are α-mixing.

Let $\{\mathbf{V}_t = (X_t, Y_t), t \in \mathbb{Z}\}$ be a two-dimensional stationary stochastic process, for all $t \in \mathbb{Z}$, and suppose that we have observations $\{\mathbf{V}_t, t = 1, \dots, n\}$. Thus, the estimator $\tilde{c}_{l,\mathbf{k}}$ is defined by

$$\tilde{c}_{l,\mathbf{k}} = \frac{1}{n} \sum_{t=1}^{n} \bigg[\int_{G_n(Y_n)}^1 \int_{F_n(X_n)}^1 \Phi_{l,\mathbf{k}}(u,v) \, du \, dv \bigg].$$

It follows that the copula estimator is given by

$$\tilde{C}_l(u,v) = \sum_{\mathbf{k}} \tilde{c}_{l,\mathbf{k}} \Phi_{l,\mathbf{k}}(u,v).$$

Then, we have the following result, showing the consistency of the wavelet estimator for time series data.

Theorem 2.2. *Under the assumptions* (A1)–(A4), *let n be the size of a sample from the process* $\{\mathbf{V}_t, t \in \mathbb{Z}\}$. *Choose l^*, such that*

$$2^{l^*} \le n^{\frac{1}{2(s+2)}} < 2^{l^*+1}.$$

Let $\tilde{C}_{l^*}(\cdot;\cdot)$ the estimator of $C(u;v)$. Then, for a constant $K > 0$, we have

$$\mathrm{MISE}\big(\tilde{C}_{l^*}(u;v), C(u;v)\big) \le Kn^{-1}.$$

Proof. See the Appendix. □

Considering the proposed estimators based on wavelets, either the i.i.d. case or the time series data case, the idea is to start from an adequate resolution level J, which depends on the sample length n.

The procedure for estimating the copula function through the proposed method is as follows:

(1) As suggested by Genest, Massiello and Tribouley (2009), compute the index J for which $2^J \le \sqrt{n} < 2^{J+1}$.
(2) Denote each element of the sample matrix $\mathbf{A}_{p\times p}$ by (a_{p_1,p_2}), where $p_1, p_2 \in \{1,\dots,p\}$. The matrix $\mathbf{B}$ is obtained by symmetrizing $\mathbf{A}$, where

$$\mathbf{B} = \begin{pmatrix} {}_*\mathbf{A}_* & {}_*\mathbf{A} & {}_*\mathbf{A}_* \\ \mathbf{A}_* & \mathbf{A} & \mathbf{A}_* \\ {}_*\mathbf{A}_* & {}_*\mathbf{A} & {}_*\mathbf{A}_* \end{pmatrix},$$

in which ${}_*\mathbf{A}_* = (a_{p+1-p_1,p+1-p_2})$, $\mathbf{A}_* = (a_{p_1,p+1-p_2})$ and ${}_*\mathbf{A} = (a_{p+1-p_1,p_2})$.
(3) Apply the Fast Wavelet Transform on $\mathbf{B}$ and extract the element in the second row and second column of the transform.
(4) Compute the estimated scaling coefficients by the 2D wavelet inverse transform algorithm,

$$\tilde{C}_l(u,v) = \sum_{\mathbf{k}\in\mathbb{Z}^2} \tilde{\alpha}_{l,\mathbf{k}}\Phi_{j,\mathbf{k}}(u,v), \quad (u,v) \in (0,1)^2.$$

(5) Select l^* and construct the estimated copula $C^* = \tilde{C}_{l^*}$.

3 Simulation studies

In this section, we present the performance of the wavelet estimators, proposed in Section 2, via simulation studies. The procedure was implemented with the Matlab (2013) software and the wavelet toolbox package (see Misiti et al. (1996)). The steps taken are as follows:

(1) draw a sample (X_i, Y_i), for $i = 1,\dots,n$;
(2) compute the empirical copula function on the grid $(\frac{i}{n}, \frac{j}{n})$, for which

$$\mathbb{C}_n\left(\frac{i}{n}; \frac{j}{n}\right) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{I}\{X_k \le X_{(i)}; Y_k \le Y_{(j)}\};$$

(3) compute the copula estimator $\hat{C}$;
(4) compute the true copula $C(\frac{i}{n}; \frac{j}{n})$ on the grid;

(5) repeat the steps (1)–(3) “m” times and compute the Bias and mean squared errors (MSE), defined as

$$\text{Bias} = \frac{1}{m}\sum_{k=1}^{m}(\hat{C}_k - C), \qquad \text{MSE} = \frac{1}{m}\sum_{k=1}^{m}(\hat{C}_k - C)^2.$$

3.1 I.i.d. case

For the i.i.d. case, we consider the random vector (X, Y), where

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N_2\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}; \begin{bmatrix} \sigma_x^2 & \gamma_{x,y} \\ \gamma_{x,y} & \sigma_y^2 \end{bmatrix}\right),$$

in which, $\gamma_{x,y}$ is the covariance function between the random variables X and Y.

The simulation study was performed with independent and dependent components. In both cases, we generate 5000 samples of size $n = 1024$. The results are shown in Table 1 and Table 2. All values are expressed as multiples of 10^{-4}.

3.1.1 *Independent components.* We generated samples from (X, Y), where $\mu_x = 1.33$, $\mu_y = 4$, $\sigma_x^2 = 0.8$, $\sigma_y^2 = 2.86$ and $\gamma_{x,y} = 0$.

Looking at Table 1, for levels $l = 4$ and $l = 5$, we see that the estimators have good performance in terms of Bias and MSE. The results may be considered satisfactory for independent components.

Table 1 *Mean, Bias and MSE of the estimator—Wavelet Daubechies D2—i.i.d. case with independent components*

	Copulas						
$\times 10^{-4}$	$C(0.01; 0.01)$	$C(0.05; 0.05)$	$C(0.25; 0.25)$	$C(0.50; 0.50)$	$C(0.75; 0.75)$	$C(0.95; 0.95)$	$C(0.99; 0.99)$
True	1.00	25.00	625.00	2500.00	5625.00	9025.00	9801.00
Wavelet estimator D2 (5000 samples)							
$l = 1$							
Mean	1.90561	24.92	625.60	2501.69	5625.93	9027.76	9851.28
Bias	0.90561	−0.07571	0.60058	1.69088	0.93426	2.76271	50.2832
MSE	0.00622	0.16637	3.6186	6.13468	3.50512	0.15379	2.52971
$l = 2$							
Mean	1.25984	24.50	625.74	2501.65	5625.80	9028.50	9814.29
Bias	0.25984	−0.49238	0.73695	1.64967	0.79662	3.50903	13.2999
MSE	0.00071	0.02078	0.34866	0.59704	0.33520	0.02173	0.01788
$l = 3$							
Mean	0.86686	24.50	625.72	2501.67	5625.84	9028.50	9087.87
Bias	−0.13313	−0.49926	0.71719	1.67639	0.83611	3.50460	6.87077
MSE	0.00099	0.02047	0.34133	0.58830	0.32950	0.02113	0.00531
$l = 4$							
Mean	0.98855	24.50	625.70	2501.69	5625.82	9028.50	9805.67
Bias	−0.01144	−0.50075	0.70948	1.69390	0.81623	3.50504	4.67279
MSE	0.001062	0.02078	0.33915	0.58574	0.32783	0.02260	0.00327
$l = 5$							
Mean	0.97892	24.50	625.75	2501.62	5625.89	9028.50	9805.67
Bias	−0.02107	0.50252	0.75044	1.62857	0.89153	3.50348	4.67398
MSE	0.00095	0.02120	0.34447	0.59008	0.33201	0.02246	0.00318

Table 2 *Mean, Bias and MSE of the estimator—Wavelet Daubechies D2—i.i.d. case with dependent components*

	Copulas						
$\times 10^{-4}$	$C(0.01; 0.01)$	$C(0.05; 0.05)$	$C(0.25; 0.25)$	$C(0.50; 0.50)$	$C(0.75; 0.75)$	$C(0.95; 0.95)$	$C(0.99; 0.99)$
True	26.90	197.20	1509.80	3739.90	6508.80	9197.20	9826.90
	Wavelet estimator D2 (5000 samples)						
	$l = 1$						
Mean	28.02	246.54	1693.19	3974.67	6689.60	9247.27	9839.80
Bias	5.547882	24.68301	92.17984	117.39426	90.38550	25.04703	6.43748
MSE	0.025848	0.33976	3.78879	6.04942	3.66830	0.34989	0.03637
	$l = 2$						
Mean	28.24	246.79	1693.09	3974.68	6689.39	9247.31	9840.41
Bias	5.65745	24.80838	92.12631	117.39958	90.27855	25.07040	6.74231
MSE	0.02690	0.34450	3.78089	6.04487	3.65641	0.35090	0.03949
	$l = 3$						
Mean	29.29	247.85	1693.24	3974.80	6689.53	9246.78	9840.04
Bias	6.18244	25.34008	92.20058	117.45853	90.34739	24.80288	6.55926
MSE	0.03151	0.36085	3.78883	6.05331	3.66433	0.33809	0.02950
	$l = 4$						
Mean c	28.46279	253.11	1694.23	3975.69	6690.42	9247.86	9835.70
Bias	5.76884	27.96876	92.69756	117.90462	90.79356	25.34236	4.39136
MSE	0.01991	0.42927	3.83516	6.10562	3.70585	0.35227	0.01050
	$l = 5$						
Mean	27.77	235.93	1699.63	3979.61	6694.94	9241.76	9865.37
Bias	5.42284	19.37992	95.39718	119.8644	93.05555	22.29249	19.22409
MSE	0.01437	0.21517	4.06349	6.31467	3.89012	0.26519	0.14911

3.1.2 *Dependent components*. Considering dependent components, we generated a sample from (X, Y), where $\mu_x = 3.05$, $\mu_y = 6.44$, $\sigma_x^2 = 1.13$, $\sigma_y^2 = 3.98$ and $\gamma_{x,y} = 1.49$. The results are shown in Table 2.

Comparing the results in Tables 1 and 2, we observe that the values are different in terms of Bias and MSE for all resolution levels. Also, the values are higher than of dependent components, but by considering that the values are expressed as multiples of 10^{-4}, the results for both cases are satisfactory.

3.2 Time series data

We consider the copula estimator for the VAR(1) process:

$$\mathbf{V}_t = \mathbf{A} + \mathbf{B}\mathbf{V}_{t-1} + \epsilon_t, \tag{3.1}$$

where $\mathbf{V}_t = (X_t; Y_t)^\top$, $\epsilon_t \sim N(0; \mathbf{\Sigma})$ and $\mathbf{A} = (1; 1)^\top$. The matrices $\mathbf{B}$ and $\mathbf{\Sigma}$ are defined taking into account the type of components. For both, we generate 5000 samples of size $n = 1024$. All values are expressed as multiples of 10^{-4}.

3.2.1 *Independent components*. For this case, let

$$\mathbf{B} = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.75 \end{bmatrix} \quad \text{and} \quad \mathbf{\Sigma} = \begin{bmatrix} 0.75 & 0 \\ 0 & 1.25 \end{bmatrix}.$$

The results are shown in Table 3. These values show that at levels $l = 4$ and $l = 5$, the proposed estimators have good performance, compared to other nonparametric estimators.

Table 3 *Mean, Bias and MSE of the estimator—Wavelet Daubechies D2—case with independent components*

	Copulas						
$\times 10^{-4}$	$C(0.01; 0.01)$	$C(0.05; 0.05)$	$C(0.25; 0.25)$	$C(0.50; 0.50)$	$C(0.75; 0.75)$	$C(0.95; 0.95)$	$C(0.99; 0.99)$
True	1.00	25.00	625.00	2500.00	5625.00	9025.00	9801.00
	Wavelet estimator D2 (5000 samples)						
	$l = 1$						
Mean	1.79043	25.69	630.85	2502.78	5624.26	9027.92	9851.29
Bias	0.79043	0.69320	5.85242	2.78267	−0.73069	2.92745	50.29317
MSE	0.00066	0.01840	0.42841	0.81192	0.42911	0.01671	0.25307
	$l = 2$						
Mean	1.19	24.62	630.91	2502.62	5624.21	9028.70	9814.31
Bias	0.18936	−0.37974	5.91207	2.62732	−0.78241	3.70217	13.31907
MSE	0.00076	0.02233	0.41422	0.78837	0.41536	0.02264	0.017952
	$l = 3$						
Mean	0.87626	24.51	630.89	2502.63	5624.32	9028.68	9807.86
Bias	−0.12373	−0.48726	5.89316	2.62973	−0.67675	3.68293	6.86647
MSE	0.00105	0.02195	0.40864	0.77857	0.40861	0.02208	0.00529
	$l = 4$						
Mean	0.98333	24.50	630.88	2502.66	5624.36	9028.58	9805.67
Bias	−0.01667	−0.49575	5.88007	2.66045	−0.63316	3.58812	4.67924
MSE	0.00114	0.02220	0.40662	0.77657	0.40668	0.02339	0.00329
	$l = 5$						
Mean	0.98539	24.50	630.92	2502.61	5624.35	9028.60	9805.69
Bias	−0.01460	−0.49638	5.91641	2.61533	−0.64523	3.59959	4.68638
MSE	0.00105	0.02250	0.41132	0.78310	0.410305	0.02330	0.00322

For further details, see Fermanian and Scaillet (2003), Morettin et al. (2010) and Morettin et al. (2011).

3.2.2 *Dependent components.* In this case, we considered samples, in which

$$\mathbf{B} = \begin{bmatrix} 0.25 & 0.2 \\ 0.2 & 0.75 \end{bmatrix} \quad \text{and} \quad \mathbf{\Sigma} = \begin{bmatrix} 0.75 & 0.5 \\ 0.5 & 1.25 \end{bmatrix}.$$

We observe that the values in Table 4 are similar for all resolution levels. The results in terms of Bias are higher than those presented by Morettin et al. (2010), but are similar to those of Fermanian and Scaillet (2003). The difference can be due to the use of scaling functions only in the wavelet expansion.

Up to this point, we have used the estimator based on the expansion (2.3). Now, we will consider the estimation procedure based on the expansion (2.1), given by

$$\hat{C}(u; v) = \sum_{\mathbf{k}} \hat{c}_{l,\mathbf{k}} \Phi_{l,\mathbf{k}}(u; v) + \sum_{j \geq l}^{J} \sum_{\mathbf{k} \in \mathbb{Z}^2} \sum_{\mu = h,v,d} \hat{d}^{\mu}_{j,\mathbf{k}} \Psi^{\mu}_{j,\mathbf{k}}(u; v), \tag{3.2}$$

and then use a threshold for the wavelet coefficients $\hat{d}^{\mu}_{j,\mathbf{k}}$. Usually, we may use hard or soft thresholds, defined by

$$\delta^{H}_{\lambda}(x) = \begin{cases} 0, & \text{if } |x| \leq \lambda, \\ x, & \text{if } |x| > \lambda, \end{cases}$$

Table 4 *Mean, Bias and MSE of the estimator—Wavelet Daubechies D2—case with dependent components*

	Copulas						
$\times 10^{-4}$	$C(0.01; 0.01)$	$C(0.05; 0.05)$	$C(0.25; 0.25)$	$C(0.50; 0.50)$	$C(0.75; 0.75)$	$C(0.95; 0.95)$	$C(0.99; 0.99)$
True	26.90	197.20	1509.80	3739.90	6508.80	9197.20	9826.90
	Wavelet estimator D2 (5000 samples)						
	$l = 1$						
Mean	20.52574	172.051	1376.47	3586.99	6465.42	9181.81	9855.91
Bias	6.39937	−25.1177	−132.371	−152.891	−43.4135	−15.3651	28.9833
MSE	0.01052	0.19315	2.20303	2.91158	0.54573	0.07141	0.08477
	$l = 2$						
Mean	17.3965	165.603	1374.81	3585.62	6465.04	9185.07	9821.37
Bias	−9.52861	−31.5727	−134.02	−154.26	−43.7941	−12.0981	−5.55452
MSE	0.02029	0.21037	2.22979	2.93473	0.54438	0.08432	0.00473
	$l = 3$						
Mean	17.7891	164.393	1374.55	3585.35	6465.20	9184.22	9817.49
Bias	−9.13599	−32.7832	−134.28	−154.53	−43.6367	−12.9511	−9.43256
MSE	0.01867	0.20959	2.23056	2.93672	0.55730	0.08182	0.014811
	$l = 4$						
Mean	18.3893	164.139	1374.60	3585.32	6465.16	9184.03	9815.85
Bias	−8.53580	−33.0367	−134.233	−154.568	−43.6722	−13.1460	−11.0753
MSE	0.022259	0.208256	2.242146	2.950065	0.558231	0.080473	0.022071
	$l = 5$						
Mean	18.2113	164.098	1374.62	3585.38	6465.17	9187.01	9815.71
Bias	−8.71379	−33.0777	−134.221	−154.503	−43.6685	−13.1644	−11.2071
MSE	0.021615	0.207158	2.240509	2.946415	0.560300	0.080146	0.021664

and

$$\delta_\lambda^S(x) = \begin{cases} 0, & \text{if } |x| \leq \lambda, \\ \sin(x)\big(|x| - \lambda\big), & \text{if } |x| > \lambda, \end{cases}$$

respectively. For more details, see Vidakovic (1999).

Thus, the final estimator is given by

$$\hat{C}(u; v) = \sum_{\mathbf{k}} \hat{c}_{l,\mathbf{k}} \Phi_{l,\mathbf{k}}(u; v) + \sum_{j \geq l}^{J} \sum_{\mathbf{k} \in \mathbb{Z}^2} \sum_{\mu = h,v,d} \delta_\lambda^{H,S}(\hat{d}_{j,\mathbf{k}}^{\mu}) \Psi_{j,\mathbf{k}}^{\mu}(u; v).$$

In this research, we choose as the threshold the high quantile proposed by Morettin et al. (2010), in which

$$\delta_Q(x) = \begin{cases} 0, & \text{if } x \leq Q_p(x), \\ x, & \text{if } x > Q_p(x), \end{cases}$$

where $Q_p(x)$ is the p-quantile of x. We take $p = 0.9$ in what follows.

We generated 5000 samples of size $n = 1024$ of the VAR(1) model given by (3.1), with dependent components. The results are in Table 5.

Comparing the values of Table 4 and Table 5, we can note that there are not many changes in terms of Bias and MSE for the copula estimation on the borders, but the estimations for other quantiles are lower in terms of the Bias when the threshold method is used.

Table 5 *Mean, Bias and MSE of the estimator—Wavelet Daubechie D2, case of dependent components with quantile threshold* ($p = 0.9$)

	Copulas						
$\times 10^{-4}$	$C(0.01; 0.01)$	$C(0.05; 0.05)$	$C(0.25; 0.25)$	$C(0.50; 0.50)$	$C(0.75; 0.75)$	$C(0.95; 0.95)$	$C(0.99; 0.99)$
True	26.90	192.20	1509.80	3739.90	6508.80	9197.20	9826.90
	Wavelet estimator D2—threshold (5000 samples)						
	$l = 1$						
Mean	20.52574	184.39444	1482.05	3656.44	6518.90	9181.81	9822.95
Bias	−6.39936	−12.78198	−26.78521	−83.44168	10.062352	−15.36508	−3.97346
MSE	0.01052	0.10116	0.47938	1.22934	0.37617	0.07141	0.00726
	$l = 2$						
Mean	17.3965	185.1919	1483.57	3656.22	6517.98	9185.07	9821.37
Bias	−9.52861	−11.98454	−25.27142	−83.66274	9.15075	−12.09881	−5.55453
MSE	0.02029	0.11018	0.50373	1.23807	0.361494	0.08432	0.00473
	$l = 3$						
Mean	17.78912	185.72434	1483.41	3656.32	6518.07	9184.22	9817.49
Bias	−9.13599	−11.45208	−25.42999	−83.57007	9.23171	−12.95116	−9.43256
MSE	0.01867	0.12077	0.50173	1.24342	0.36941	0.08182	0.01481
	$l = 4$						
Mean	18.38931	185.65573	1483.51	3656.32	6518.10	9184.03	9815.84
Bias	−8.53580	−11.52069	−25.32560	−83.56696	9.26386	−13.14606	−11.07530
MSE	0.02225	0.11923	0.50913	1.25208	0.37141	0.08047	0.02207
	$l = 5$						
Mean	18.21132	185.73713	1483.50	3656.31	6518.09	9184.01	9815.71
Bias	−8.71379	−11.43929	−25.33263	−83.57427	9.26093	−13.16446	−11.20713
MSE	0.02161	0.12264	0.50845	1.25159	0.37387	0.08014	0.02166

Figures 1 and 2 show the graphical representations of the estimators for different resolution levels, without and with the threshold for dependent components.

3.3 Additional simulations

To evaluate the results of the proposed methodology, we consider an additional simulation study, as presented by Autin, Le Pennec and Tribouley (2010). Consider the empirical loss functions, given by

$$\begin{aligned}\mathrm{Error}(\hat{C}_{l^*}, C_\theta) &= \frac{1}{N^2}\|\hat{C}_{l^*} - C_\theta\|_2^2 \\ &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left\{\hat{C}_{l^*}\left(\frac{i}{N}, \frac{j}{N}\right) - C_\theta\left(\frac{i}{N}, \frac{j}{N}\right)\right\}^2,\end{aligned}$$

and

$$\mathrm{RE}(\hat{C}_{l^*}, C_\theta) = \mathrm{Error}(\hat{C}_{l^*}, C_\theta) \times \left[\frac{1}{N^2}\|C_\theta\|_2^2\right]^{-1},$$

where RE is the *relative error*, $\hat{C}_{l^*}$ is the estimated copula function on the grid $(\frac{i}{N}, \frac{j}{N})$, $i, j = 1, \dots, N$ and C_θ is the parametric copula, with fixed θ.

The procedure is as follows:

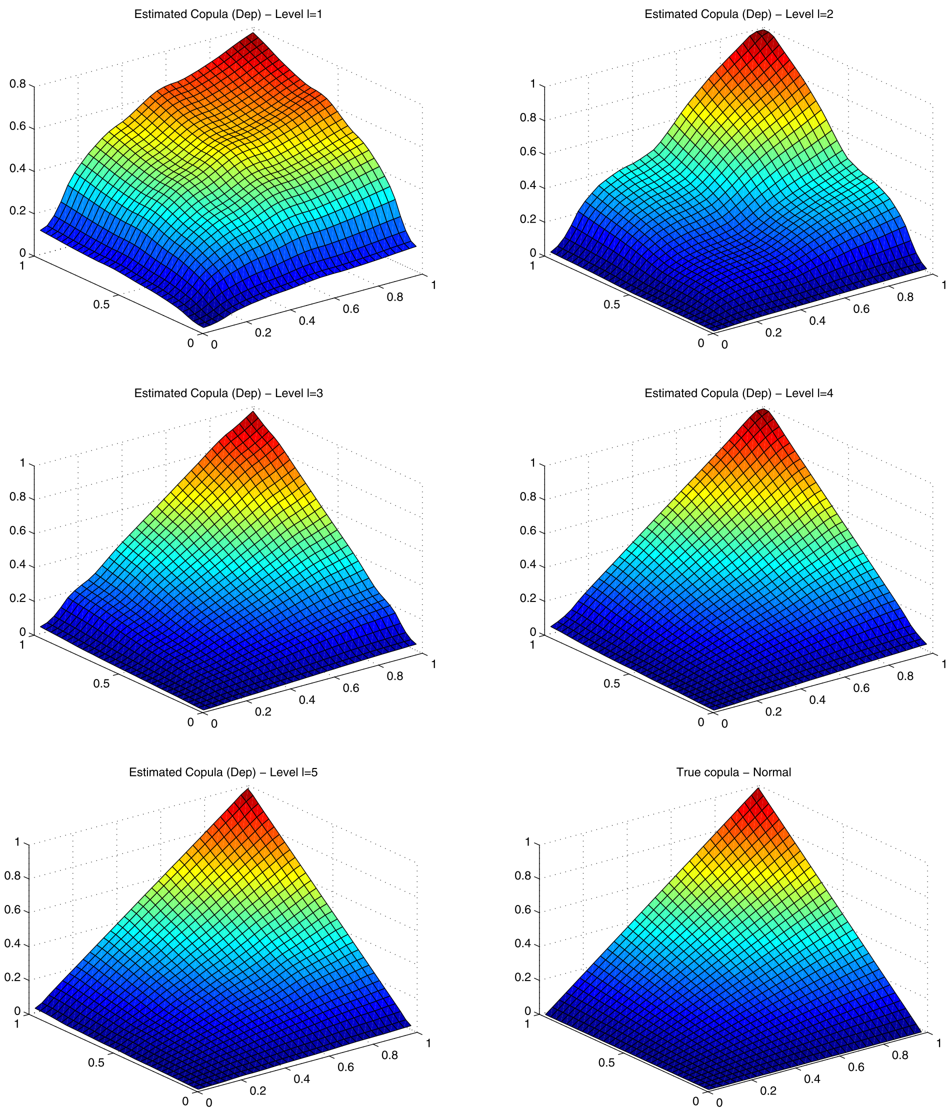

Figure 1 *Graphical representations of the estimated copula at different levels—dependent components without thresholds. Last graphic is for the Normal copula.*

(1) draw a sample of size n from a parametric copula C_θ;
(2) compute a copula estimator using wavelets;
(3) consider $N = n$ and compute the RE between the parametric copula and the estimated copula on the grid;
(4) repeat the steps (1)–(3) r times;
(5) compute the mean and the standard deviation (SD) of the r replications of RE.

We considered $r = 5000$ replicates for samples of size $n = 256$ and $n = 1024$, from Normal, Student-t, Gumbel, Clayton and Frank copula, with fixed parameters. For the normal copula, the correlation coefficient is 0.5. For the Student-t copula the correlation coefficient is 0.5 and degree of freedom is 5. For the Gumbel, Clayton and Frank copula, we take the parameter of dependence as 5. The idea is to study the estimation performance for different sample sizes.

From Table 6 and Table 7, we notice that the SD of the RE decreases when sample size increases. This behavior is the same for all resolution levels.

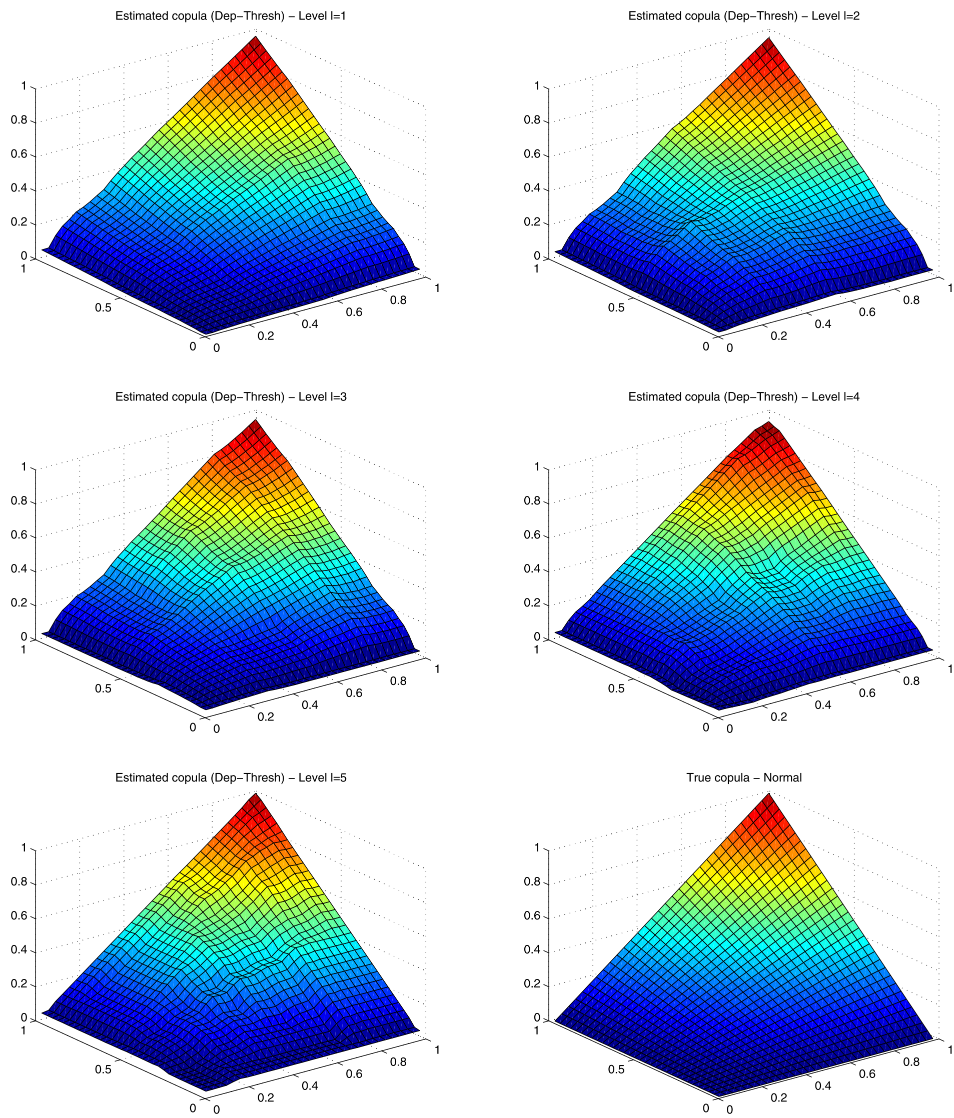

Figure 2 *Graphical representations of the estimated copula at different levels—dependent components with quantile threshold* ($p = 0.9$). *Last graphic is for the Normal copula.*

4 Applications

In this section, we apply the proposed estimation procedure for two pairs of series.

4.1 Ibovespa—IPC

Ibovespa is an index of about 50 stocks that are traded on the BM&FBOVESPA (São Paulo Stock Exchange). IPC (Índice de Precios y Cotizaciones) is an index of 35 stocks that are traded on the Mexican Stock Exchange.

We consider daily returns recorded from September 4th, 1995 to December 30th, 2004 with 1981 observations. The correlation coefficient is moderate, equal to 0.5516. Figure 3 shows the scatter plot of returns of the Ibovespa and IPC.

To verify if the proposed estimator is appropriate, we calculated the error and the relative error between the proposed wavelet estimator (for different wavelets D2, D4, D6 and D8) and five estimated parametric copulas. The values are reported in Table 8, where we see that the

Table 6 *Mean and standard deviation(SD) of RE for the wavelet estimator and some parametric copulas, $n = 256$*

$n = 256$	Wavelets							
	Daubechie D2		Daubechie D4		Daubechie D6		Daubechie D8	
Copulas	Mean	SD	Mean	SD	Mean	SD	Mean	SD
				$l = 1$				
Normal	43.3428	36.5873	43.5456	36.4782	42.7880	36.5374	44.5405	36.5577
Student-t	43.2367	36.2312	43.3472	36.1408	42.6834	36.2340	44.3287	36.2324
Gumbel	39.5915	35.7352	38.9729	35.6866	38.1589	35.7829	40.0677	35.7841
Clayton	41.3565	38.6787	40.8990	38.5900	39.9253	38.7219	42.1046	38.7085
Frank	42.9201	37.1642	42.7577	37.1098	42.0791	37.1950	43.8146	37.1927
				$l = 2$				
Normal	44.0426	36.5550	44.4046	36.5556	44.3168	36.5514	44.3604	36.5497
Student-t	43.9318	36.2537	44.2702	36.2531	44.1876	36.2506	44.2372	36.2585
Gumbel	39.0460	35.7829	39.3049	35.7858	39.1993	35.7865	39.2538	35.7889
Clayton	40.8397	38.7152	41.1343	38.7115	41.0108	38.7124	41.0713	38.7162
Frank	43.2385	37.1989	43.5473	37.2062	43.4611	37.2030	43.5114	37.2035
				$l = 3$				
Normal	44.9479	36.5558	45.0958	36.5598	45.0678	36.5578	45.0677	36.5588
Student-t	44.8298	36.2573	44.9745	36.2613	44.9484	36.2593	44.9481	36.2594
Gumbel	39.6670	35.7895	39.7959	35.7938	39.7575	35.7917	39.7527	35.7917
Clayton	41.4846	38.7158	41.6213	38.7195	41.5805	38.7163	41.5768	38.7167
Frank	44.0507	37.2066	44.1870	37.2115	44.1612	37.2084	44.1612	37.2086
				$l = 4$				
Normal	45.4800	36.5596	45.5347	36.5612	45.5368	36.5602	45.5394	36.5604
Student-t	45.3584	36.2601	45.4129	36.2606	45.4170	36.2600	45.4176	36.2603
Gumbel	40.0904	35.7915	40.1378	35.7923	40.1366	35.7925	40.1381	35.7912
Clayton	41.9263	38.7154	41.9747	38.7160	41.9738	38.7156	41.9753	38.7152
Frank	44.5467	37.2091	44.5975	37.2109	44.5998	37.2098	44.6017	37.2101
				$l = 5$				
Normal	45.7792	36.5626	45.7935	36.5624	45.7967	36.5622	45.7998	36.5623
Student-t	45.6582	36.2612	45.6730	36.2614	45.6764	36.2616	45.6790	36.2615
Gumbel	40.3360	35.7940	40.3472	35.7935	40.3498	35.7935	40.3520	35.7938
Clayton	42.1822	38.7166	42.1934	38.7161	42.1961	38.7163	42.1986	38.7163
Frank	44.8271	37.2120	44.8409	37.2120	44.8440	37.2120	44.8466	37.2120

values of the copula estimate with the wavelet D2 and those using the Student-t copula, are similar.

Figure 4 shows the estimated copula and respective contour plots for this case. To evaluate how the data are associated, we propose an empirical estimation of tail dependence using the estimated copulas, as presented by Caillault and Guégan (2005). Figure 5 shows the estimated empirical tail dependence measures for the Ibovespa and IPC series. These are important tools to describe the properties of the copulas with respect to their tail behavior. These quantities are also useful for computing risk measures.

4.2 Net profit—Sales margin

We consider now annual rates of the sales performance of 1018 companies in Brazil, 2006 according to Exame magazine. This data set was used by Latif and Morettin (2010), who suggested to analyze the normalized ranks of Net profit (US$) and Sales margin (%), denoted

Table 7 *Mean and standard deviation(SD) of RE for the wavelet estimator and some parametric copula,* $n = 1024$

$n = 1024$	Wavelets							
	Daubechie D2		Daubechie D4		Daubechie D6		Daubechie D8	
Copulas	Mean	SD	Mean	SD	Mean	SD	Mean	SD
				$l = 1$				
Normal	11.0341	8.6375	11.0549	8.6368	11.0471	8.6367	11.0788	8.6373
Student-t	11.2959	9.1673	11.3151	9.1663	11.3085	9.1672	11.3383	9.1667
Gumbel	9.8137	8.8952	9.8232	8.8950	9.8139	8.8954	9.8470	8.8953
Clayton	10.4294	9.5182	10.4388	9.5182	10.4289	9.5188	10.4655	9.5188
Frank	0.9561	9.2673	10.9718	9.2675	10.9654	9.2676	10.9945	9.2677
				$l = 2$				
Normal	11.1450	8.6372	11.1601	8.6373	11.1607	8.6374	11.1624	8.6372
Student-t	11.4071	9.1674	11.4216	9.1674	11.4224	9.1674	11.4241	9.1673
Gumbel	9.8933	8.8956	9.9056	8.8957	9.9052	8.8957	9.9071	8.8957
Clayton	10.5110	9.5190	10.5240	9.5188	10.5236	9.5186	10.5255	9.5188
Frank	11.0574	9.2677	11.0706	9.2674	11.0713	9.2676	11.0731	9.2677
				$l = 3$				
Normal	11.2099	8.6375	11.2169	8.6376	11.2175	8.6376	11.2179	8.6376
Student-t	11.4719	9.1674	11.4788	9.1674	11.4794	9.1674	11.4798	9.1674
Gumbel	9.9452	8.8959	9.9511	8.8959	9.9513	8.8959	9.9516	8.8958
Clayton	10.5653	9.5189	10.5714	9.5189	10.5717	9.5189	10.5720	9.5189
Frank	11.1177	9.2675	11.1241	9.2675	11.1247	9.2675	11.1251	9.2676
				$l = 4$				
Normal	11.2441	8.6376	11.2472	8.6377	11.2478	8.6377	11.2480	8.6377
Student-t	11.5061	9.1675	11.5092	9.1675	11.5098	9.1675	11.5101	9.1675
Gumbel	9.9731	8.8959	9.9758	8.8959	9.9761	8.8959	9.9764	8.8959
Clayton	10.5945	9.5189	10.5972	9.5189	10.5977	9.5189	10.5979	9.5189
Frank	11.1497	9.2676	11.1525	9.2676	11.1531	9.2676	11.1533	9.2676
				$l = 5$				
Normal	11.5250	8.6377	11.2640	8.6377	11.2643	8.6377	11.2644	8.6377
Student-t	11.5250	9.1675	11.5261	9.1675	11.5263	9.1675	11.5265	9.1675
Gumbel	9.9886	8.8959	9.9895	8.8959	9.9897	8.8959	9.9898	8.8959
Clayton	10.6107	9.5189	10.6116	9.5189	10.6118	9.5189	10.6120	9.5189
Frank	11.1672	9.2676	11.1683	9.2676	11.1685	9.2676	11.1687	9.2676

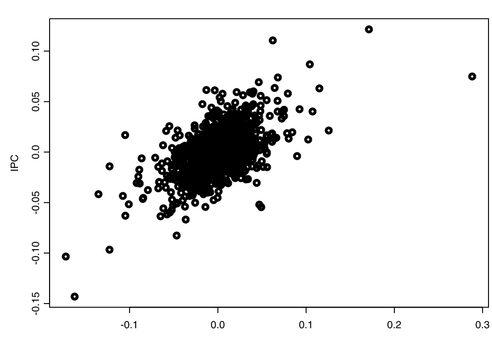

Figure 3 *Scatter plot for the returns of Ibovespa and IPC series.*

Table 8 *Error and RE of copula estimators, for the Ibovespa and IPC series*

Parametric		Wavelet D2		Wavelet D4		Wavelet D6		Wavelet D8	
Copula	Estimative	Error	RE	Error	RE	Error	RE	Error	RE
Normal	0.4356	0.3418	2.6173	0.3418	2.6175	0.3418	2.6176	0.3418	2.6177
Student-t	0.4439 (7.45)*	**0.2933**	**2.2401**	0.2934	2.2403	0.2934	2.2403	0.2934	2.2405
Gumbel	1.3559	1.1778	9.0066	1.1778	9.0069	1.1778	9.0070	1.1778	9.0071
Clayton	0.6529	0.8760	7.0072	0.8760	7.0075	0.8760	7.0076	0.8760	7.0077
Frank	2.8910	0.3743	2.8398	0.3743	2.8400	0.3743	2.8401	0.3744	2.8402

*The value inside of $(\cdot)$ represents the estimated degrees of freedom.

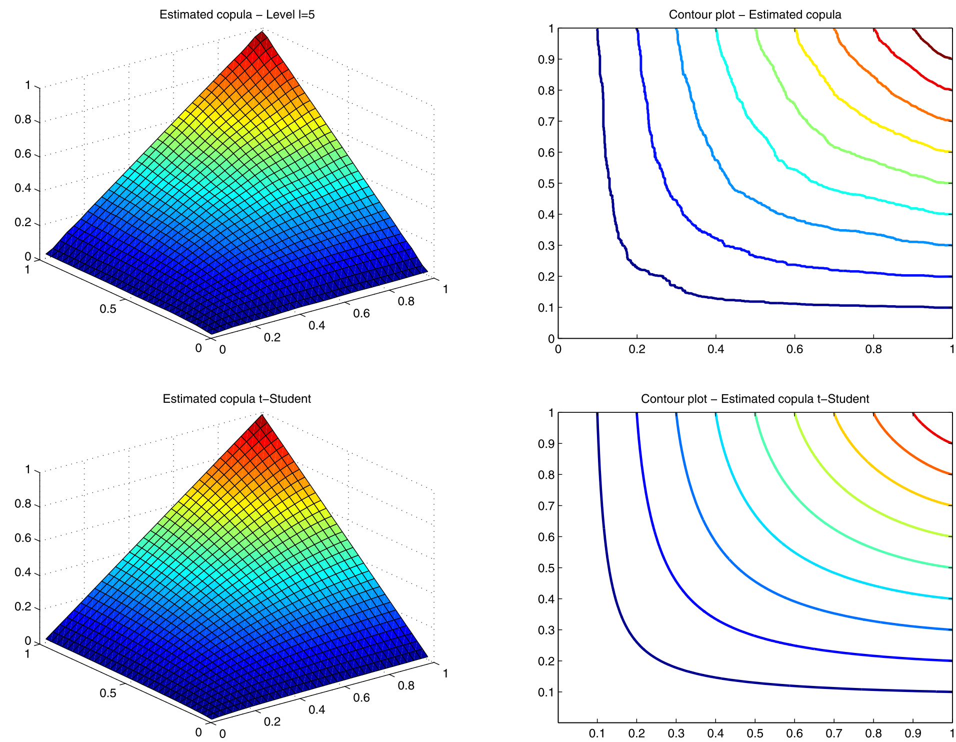

Figure 4 *Graphical representations of the estimated copulas (wavelets in the first line and Student-t in the second line) and contour plots for the Ibovespa and IPC series.*

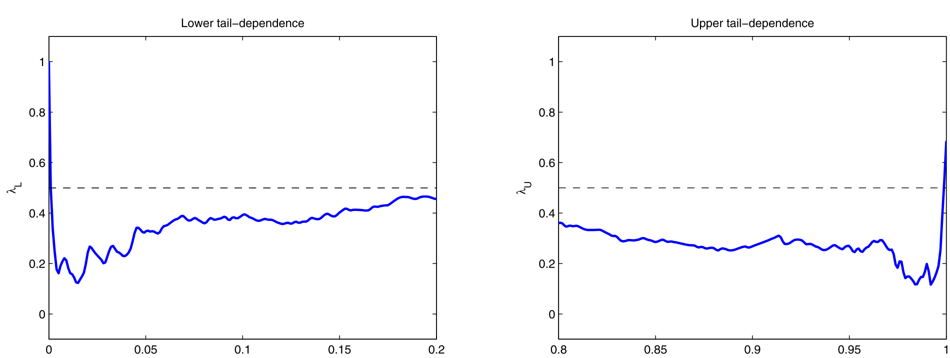

Figure 5 *Graphical representations of the estimated tail dependence for the Ibovespa and IPC series.*

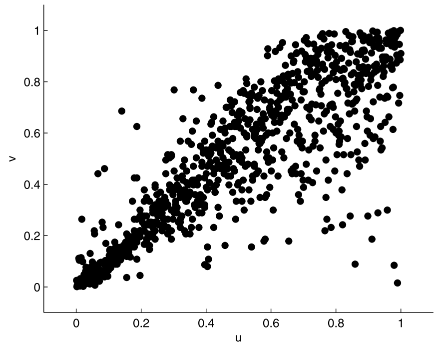

Figure 6 *Scatter plot of normalized ranks for the Net profit and Sales margin series.*

Table 9 *Error and RE of copula estimators, for the Net profit and Sales margin series*

Parametric		Wavelet D2		Wavelet D4		Wavelet D6		Wavelet D8	
Copula	Estimative	Error	RE	Error	RE	Error	RE	Error	RE
Student-t	0.85 (3.35)*	1.05	6.76	1.05	6.76	1.05	6.76	1.05	6.77
Gumbel	2.47	2.85	18.64	2.85	18.64	2.85	18.64	2.85	18.65
Clayton	3.57	**0.56**	**3.73**	0.56	3.73	0.56	3.74	0.56	3.74
Frank	10.39	0.99	6.33	0.99	6.33	0.99	6.33	0.99	6.33

*The value inside of $(\cdot)$ represents the estimated degrees of freedom.

by u and v, respectively, in order to find the dependence structure. The correlation coefficient between u and v is 0.8455. See Figure 6.

In Table 9, we present the error and the relative error between the proposed wavelet estimator (for different wavelets D2, D4, D6 and D8) and five estimated parametric copulas. The closest values are for the estimated copula with Wavelet D2 and the estimated Clayton copula.

Figure 7 shows the graphical representations of the estimated wavelet, the estimated Clayton copula and contour plots. Figure 8 shows the estimated empirical tail dependence measures.

5 Conclusions

In this paper, we proposed a new procedure for the estimation of a copula function by direct expansion on wavelet bases. An advantage of the wavelet approach is that it can be used directly with the original series, without estimating densities, distribution functions or assumptions about the data distribution.

Although the idea here was to use these estimators with time series data, they can also be applied to random samples (i.i.d. data). The aim was to propose a methodology with better results in terms of Bias and of MSE, compared with the results obtained through kernels. We have established consistency of the estimators for i.i.d. and time series data. We reported some simulation studies to assess the performance of the proposed estimators, and the findings show that they perform equally or outperform previous proposals. We also applied the proposed estimation procedure to real data sets.

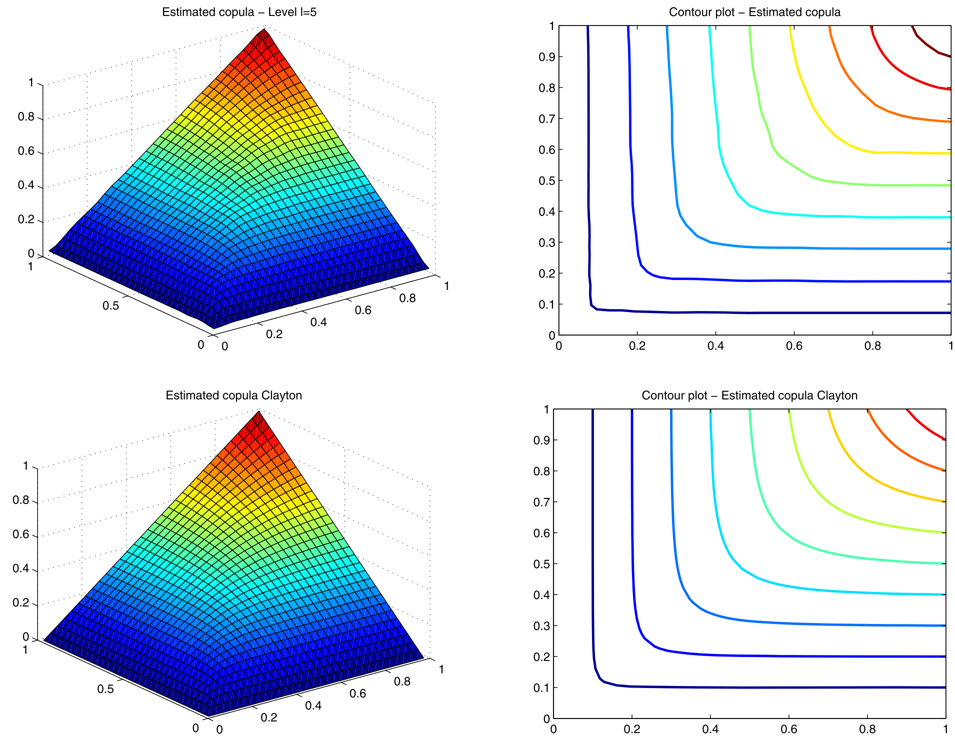

Figure 7 *Graphical representations of the estimated copulas (wavelets in the first line and Clayton in the second line) and contour plots for the Net profit and Sales margin series.*

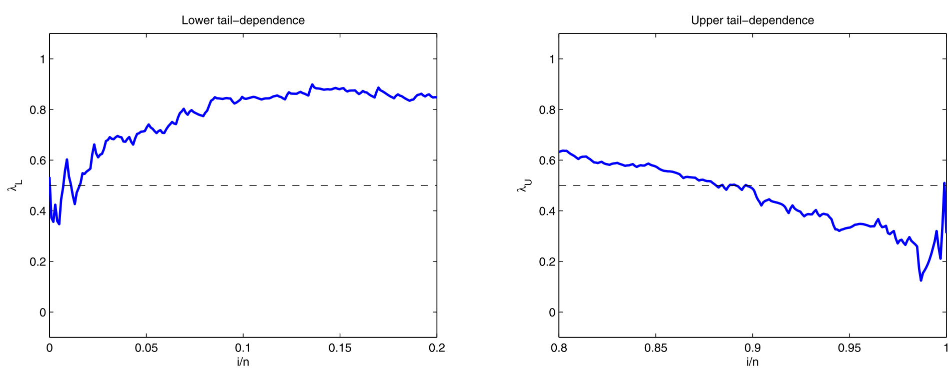

Figure 8 *Graphical representations of the estimated tail dependence for the Net profit and Sales margin series.*

Appendix

Proof of Theorem 2.1

Let

$$\{\Phi_{l,\mathbf{k}}(x; y), \mathbf{k} = (k_1, k_2)\}_{\mathbf{k}} \cup \{\Psi^{\mu}_{j,\mathbf{k}}(x; y), \mathbf{k} = (k_1, k_2), \mu = h, v, d\}_{j \geq l, \mathbf{k}}$$

be an orthonormal basis of $L^2([0, 1]^2)$. We have that

$$\operatorname{MISE}(\tilde{C}_l(u; v), C(u; v)) \leq \mathbb{E}_{h(x,y)} \|\tilde{C}_l(u; v) - C_l(u; v)\|_2^2 + \Bigg\| \sum_{j \geq l} \sum_{\mathbf{k} \in \mathbb{Z}^2} \sum_{\mu = h,v,d} d^{\mu}_{j,\mathbf{k}} \Psi^{\mu}_{j,\mathbf{k}}(u; v) \Bigg\|_2^2. \quad \text{(A.1)}$$

If $C \in \mathfrak{B}_p^{s,q}$, with $s > 0$, $1 \le p, q < \infty$, then

$$\|C\|_{s,p,q} = \|c_{l.}\|_p + \left(\sum_{j\ge l}(2^{j(s+\frac{2}{p}+1)}\|d_{j.}\|_p)^q\right)^{\frac{1}{q}}.$$

Under assumption (A1), using the Hölder inequality, with $\frac{1}{q}+\frac{1}{q'}=1$ and $p=2$, for the second term in (A.1),

$$\begin{aligned}\left\|\sum_{j\ge l}\sum_{\mathbf{k},\mu} d^{\mu}_{j,\mathbf{k}}\Psi^{\mu}_{j,\mathbf{k}}(u;v)\right\|_2 &\le \left(\sum_{j\ge l}\left(2^{j(s+2)}\left\|\sum_{\mathbf{k},\mu} d^{\mu}_{j,\mathbf{k}}\Psi^{\mu}_{j,\mathbf{k}}(u;v)\right\|_2\right)^q\right)^{1/q} \\ &\quad\times\left(\sum_{j\ge l}(2^{-j(s+2)})^{q'}\right)^{1/q'} \\ &\le \|C\|_{s,2,q}\left(\sum_{j\ge l}2^{-j(s+2)q'}\right)^{1/q'} \\ &\le M2^{-2l(s+2)}. \end{aligned}\tag{A.2}$$

Note that, as in Genest, Massiello and Tribouley (2009), for the first term in (A.1)

$$\begin{aligned}\mathbb{E}_{h(x,y)}\|\tilde{C}_l(u;v)-C_l(u;v)\|_2^2 &\le 2\mathbb{E}_{h(x,y)}\|\tilde{C}_l(u;v)-\hat{C}_l(u;v)\|_2^2 \\ &\quad+2\mathbb{E}_{h(x,y)}\|\hat{C}_l(u;v)-C_l(u;v)\|_2^2,\end{aligned}\tag{A.3}$$

where

$$\hat{c}_{l,\mathbf{k}} = \frac{1}{n}\sum_{i=1}^n\left[\int_{G(Y_i)}^1\int_{F(X_i)}^1 \Phi_{l,\mathbf{k}}(u;v)\,du\,dv\right],$$

with

$$\hat{C}_l(u;v) = \sum_{\mathbf{k}}\hat{c}_{l,\mathbf{k}}\Phi_{l,\mathbf{k}}(u;v).$$

Now, we want to find upper bounds for (A.3), separately.
For the first term in (A.3),

$$\begin{aligned}&\mathbb{E}_{h(x,y)}\|\tilde{C}_l(u;v)-\hat{C}_l(u;v)\|_2^2 \\ &\quad=\sum_{\mathbf{k}}\mathbb{E}_{h(x,y)}[(\tilde{c}_{l,\mathbf{k}}-\hat{c}_{l,\mathbf{k}})^2] \\ &\quad=\sum_{\mathbf{k}}\mathbb{E}_{h(x,y)}\left(\frac{1}{n}\sum_{i=1}^n\int_{G_n(Y_i)}^1\int_{F_n(X_i)}^1\Phi_{l,\mathbf{k}}(u;v)\,du\,dv\right. \\ &\qquad\left.-\frac{1}{n}\sum_{i=1}^n\int_{G(Y_i)}^1\int_{F(X_i)}^1\Phi_{l,\mathbf{k}}(u;v)\,du\,dv\right)^2 \\ &\quad\le\sum_{\mathbf{k}}\mathbb{E}_{h(x,y)}(\|\Phi_{l,\mathbf{k}}(u;v)\|_\infty)^2 \\ &\qquad\times\left\{\frac{1}{n}\sum_{i=1}^n[(1-G_n(Y_i))(1-F_n(X_i))\right. \\ &\qquad-(1-G(Y_i))(1-F(X_i))+(1-F_n(X_i))(1-G(Y_i))\end{aligned}$$

$$
\begin{aligned}
&- (1 - F_n(X_i))(1 - G(Y_i))]\Big\}^2 \\
&\leq \sum_{\mathbf{k}} \mathbb{E}_{h(x,y)} (\|\Phi_{l,\mathbf{k}}(u;v)\|_\infty)^2 \\
&\quad \times \left\{ \frac{1}{n} \sum_{i=1}^n [|G(Y_i) - G_n(Y_i)| + |F(X_i) - F_n(X_i)|] \right\}^2.
\end{aligned}
$$

Let

$$
\Delta(X_i) = F(X_i) - F_n(X_i) \quad \text{and} \quad \Delta(Y_i) = G(Y_i) - G_n(Y_i),
$$

for fixed $\epsilon > 0$, $\Delta(Y_i) = \Delta(Y_i)(\mathbb{I}\{|\Delta(Y_i)| > \epsilon\} + \mathbb{I}\{|\Delta(Y_i)| \leq \epsilon\})$.

And for

$$
\sum_{i=1}^n |\Delta(X_i)| \leq n\epsilon + n(\mathbb{I}\{|\Delta(X_i)| > \epsilon\}),
$$

$$
\begin{aligned}
\sum_{\mathbf{k}} \mathbb{E}_{h(x,y)} [(\tilde{c}_{l,\mathbf{k}} - \hat{c}_{l,\mathbf{k}})^2] &\leq \sum_{\mathbf{k}} \mathbb{E}_{h(x,y)} (\|\Phi_{l,\mathbf{k}}(u;v)\|_\infty)^2 \\
&\quad \times \left\{ \frac{1}{n} \sum_{i=1}^n [|\Delta(X_i)| + |\Delta(Y_i)|] \right\}^2 \\
&\leq \sum_{\mathbf{k}} (\|\Phi_{l,\mathbf{k}}(u;v)\|_\infty)^2 \mathbb{E}_{h(x,y)} \{4\epsilon^2 \\
&\quad + (3 + 4\epsilon)[\mathbb{I}\{|\Delta(X_i)| > \epsilon\} + \mathbb{I}\{|\Delta(Y_i)| > \epsilon\}]\}.
\end{aligned}
$$

But we have that

$$
\begin{aligned}
\|\Phi_{l,\mathbf{k}}(u;v)\|_\infty &= \|2^l \Phi(2^l u - k_1; 2^l v - k_2)\|_\infty \\
&= 2^l \sup_{(2^l u - k_1; 2^l v - k_2)} |\Phi(2^l u - k_1; 2^l v - k_2)| \\
&= 2^l \|\Phi(w;z)\|_\infty, \quad \text{with } w = 2^l u - k_1 \text{ and } z = 2^l v - k_2,
\end{aligned}
$$

for $k_1 = 0, \dots, 2^l - 1$, $k_2 = 0, \dots, 2^l - 1$.

Since $\sum_{\mathbf{k}} = 2^{2l}$, applying the Dvoretzky–Kiefer–Wolfowitz inequality (see Dvoretzky, Kiefer and Wolfowitz (1956)), provided that $\epsilon = \{\frac{\delta \log(n)}{2n}\}^{\frac{1}{2}}$, where δ could be as large as desired, one finds

$$
\begin{aligned}
&\mathbb{E}_{h(x,y)} \|\tilde{C}_l(u;v) - \hat{C}_l(u;v)\|_2^2 \\
&\quad \leq 2^{4l} \left\{ k_1 \frac{\log(n)}{n} + k_2 n^{-\delta} + k_3 (\log(n))^{\frac{1}{2}} n^{-(\delta + \frac{1}{2})} \right\},
\end{aligned} \tag{A.4}
$$

where k_1, k_2 and k_3 depend on $\|\Phi(w;z)\|_\infty$ and δ.

For the second term in (A.3)

$$
\begin{aligned}
&\mathbb{E}_{h(x,y)} \|\hat{C}_l(u;v) - C_l(u;v)\|_2^2 \\
&= \sum_{\mathbf{k}} \mathbb{E}_{h(x,y)} \left[\left(\frac{1}{n} \sum_{i=1}^n \int_{G(Y_i)}^1 \int_{F(X_i)}^1 \Phi_{l,\mathbf{k}}(u;v)\, du\, dv \right. \right. \\
&\quad \left. \left. - \frac{1}{n} \sum_{i=1}^n E_{h(x_i,y_i)} \left(\int_{G(Y_i)}^1 \int_{F(X_i)}^1 \Phi_{l,\mathbf{k}}(u;v)\, du\, dv \right) \right)^2 \right].
\end{aligned}
$$

Let

$$W_i = \int_{G(Y_i)}^{1}\int_{F(X_i)}^{1} \Phi_{l,\mathbf{k}}(u;v)\,du\,dv - \mathbb{E}_{h(x_i,y_i)}\left(\int_{G(Y_i)}^{1}\int_{F(X_i)}^{1} \Phi_{l,\mathbf{k}}(u;v)\,du\,dv\right)$$
$$= \int_{G(Y_i)}^{1}\int_{F(X_i)}^{1} \Phi_{l,\mathbf{k}}(u;v)\,du\,dv - c_{l,\mathbf{k}},$$

where W_i are i.i.d. random variables, with $\mathbb{E}_{h(x_i,y_i)}(W_i) = 0$. Then, applying the Rosenthal's inequality (see Rosenthal (1970)), we have that

$$\begin{aligned}
&\mathbb{E}_{h(x,y)}\|\hat{C}(u;v) - \mathbb{E}_{h(x,y)}(\hat{C}(u;v))\|_2^2 \\
&= \sum_{\mathbf{k}} \mathbb{E}_{h(x,y)}\left[\left(\frac{1}{n}\sum_{i=1}^{n} W_i\right)^2\right] \\
&\leq \sum_{\mathbf{k}} \frac{1}{n^2}\mathbb{E}_{h(x,y)}\left[\left(\sum_{i=1}^{n} W_i\right)^2\right] \\
&\leq \sum_{\mathbf{k}} K\left[\sum_{i=1}^{n}\mathbb{E}_{h(x,y)}(W_i^2) + \left(\sum_{i=1}^{n}\mathbb{E}_{h(x,y)}(W_i^2)\right)\right] \\
&\leq \sum_{\mathbf{k}} \frac{1}{n^2} K[n2^{2l} + (n2^{2l})] \\
&\leq \sum_{\mathbf{k}} \frac{1}{n} K' 2^{2l} = K'\frac{2^{4l}}{n}.
\end{aligned} \tag{A.5}$$

Using (A.2), (A.4) and (A.5), we have that

$$\mathrm{MISE}(\tilde{C}_l(u;v), C(u;v)) \leq 2^{4l+1}\left[k_1\frac{\log(n)}{n} + k_2 n^{-\delta} + k_3(\log(n))^{\frac{1}{2}} n^{-(\delta+\frac{1}{2})}\right] + K'\frac{2^{4l+1}}{n} + M2^{-2l(s+2)}.$$

But the expression $K'\frac{2^{4l+1}}{n} + M2^{-2l(s+2)}$ has a minimum when the two terms are balanced. For more details about this procedure, see Härdle et al. (1998). In this case, $\mathrm{MISE}(\tilde{C}_l(u;v), C(u;v))$ has a minimum when l^* is such that $2^{l^*} \leq n^{\frac{1}{2(s+4)}} < 2^{l^*+1}$.

Then,

$$\sup_{C\in\mathfrak{B}^q_{s,2}(M)} \mathrm{MISE}(\tilde{C}_{l^*}(u;v), C(u;v)) \leq K n^{\frac{-(s+2)}{s+4}}$$
$$\Rightarrow \sup_{C\in\mathfrak{B}^q_{s,2}(M)} n^{\frac{s+2}{s+4}}\mathrm{MISE}(\tilde{C}_{l^*}(u;v), C(u;v)) \leq K.$$

This completes the proof of Theorem 2.1.

Proof of Theorem 2.2

In the same way as in the i.i.d. case, to study the performance of $\tilde{c}_{l,\mathbf{k}}$ under dependence structure, we have that

$$\mathrm{MISE}(\tilde{C}_l(u;v), C(u;v)) \leq \mathbb{E}_{h(x,y)}\|\tilde{C}_l(u;v) - C_l(u;v)\|_2^2 + \left\|\sum_{j\geq l}\sum_{\mathbf{k}\in\mathbb{Z}^2}\sum_{\mu=h,v,d} d^{\mu}_{j,\mathbf{k}}\Psi^{\mu}_{j,\mathbf{k}}(u;v)\right\|_2^2.$$

Under the assumption (A1),

$$\Big\| \sum_{j \geq l} \sum_{\mathbf{k},\mu} d^{\mu}_{j,\mathbf{k}} \Psi^{\mu}_{j,\mathbf{k}}(u; v) \Big\|^2_2 \leq M 2^{-2l(s+2)}. \tag{A.6}$$

Then, it is only necessary to study

$$\begin{aligned} \mathbb{E}_{h(x,y)} \|\tilde{C}_l(u; v) - C_l(u; v)\|^2_2 &\leq 2\mathbb{E}_{h(x,y)} \|\tilde{C}_l(u; v) - \hat{C}_l(u; v)\|^2_2 \\ &\quad + 2\mathbb{E}_{h(x,y)} \|\hat{C}_l(u; v) - C_l(u; v)\|^2_2, \end{aligned} \tag{A.7}$$

where

$$\hat{c}_{l,\mathbf{k}} = \frac{1}{n} \sum_{t=1}^{n} \Big[\int_{G(Y_t)}^{1} \int_{F(X_t)}^{1} \Phi_{l,\mathbf{k}}(u; v)\, du\, dv \Big].$$

We have, for the first term in (A.7)

$$\begin{aligned} &\mathbb{E}_{h(x,y)} \|\tilde{C}_l(u; v) - \hat{C}_l(u; v)\|^2_2 \\ &\quad = \sum_{\mathbf{k}} \mathbb{E}_{h(x,y)} [(\tilde{c}_{l,\mathbf{k}} - \hat{c}_{l,\mathbf{k}})^2] \\ &\quad \leq \sum_{\mathbf{k}} \mathbb{E}_{h(x,y)} (\|\Phi_{l,\mathbf{k}}(u; v)\|_\infty)^2 \\ &\qquad \times \Big\{ \frac{1}{n} \sum_{t=1}^{n} [|G(Y_t) - G_n(Y_t)| + |F(X_t) - F_n(X_t)|] \Big\}^2 \\ &\quad = \sum_{\mathbf{k}} \mathbb{E}_{h(x,y)} (\|\Phi_{l,\mathbf{k}}(u; v)\|_\infty)^2 \Big\{ \frac{1}{n} \sum_{t=1}^{n} [|\Delta(X_t)| + |\Delta(Y_t)|] \Big\}^2, \end{aligned}$$

where $\Delta(Y_i) = \Delta(Y_i)(\mathbb{I}\{|\Delta(Y_i)| > \epsilon\} + \mathbb{I}\{|\Delta(Y_i)| \leq \epsilon\})$. For fixed $\epsilon > 0$, which $\epsilon = \frac{1}{n}$, we have that

$$\begin{aligned} &\sum_{\mathbf{k}} \mathbb{E}_{h(x,y)} [(\tilde{c}_{l,\mathbf{k}} - \hat{c}_{l,\mathbf{k}})^2] \\ &\quad \leq 2^{2l} (\|\Phi(w; z)\|_\infty)^2 \\ &\qquad \times \sum_{\mathbf{k}} \Big\{ k_1 \frac{1}{n^2} + \Big(k_2 + k_3 \frac{1}{n}\Big) [P\{|\Delta(X_i)| > \epsilon\} + P\{|\Delta(Y_i)| > \epsilon\}] \Big\}. \end{aligned}$$

Proceeding as in Yu (1993), let $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ be a stationary sequence of random variables, with the same distribution function $F(x)$. If $F(x)$ is continuous and

$$\sum_{t=1}^{\infty} \frac{1}{n^2} \operatorname{Cov}\{X_t, S_{n-1}\} < \infty, \tag{A.8}$$

when $t = 1, \ldots, n$, $S_n = \sum_{t=1}^{n} X_t$ and $n \to \infty$, we have that

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{q.c} 0. \tag{A.9}$$

For a stationary sequence, the condition (A.8) can be replaced by

$$\frac{1}{n} \sum_{t=1}^{n} \operatorname{Cov}\{X_t, X_n\} \longrightarrow 0.$$

Under assumption (A4), and knowing that the mixing condition implies ergodicidity, by the equation (A.9), we have that

$$\sum_{\mathbf{k}} \mathbb{E}_{h(x,y)}\big[(\tilde{c}_{l,\mathbf{k}} - \hat{c}_{l,\mathbf{k}})^2\big] = o(1).$$

For the second term in (A.7), we have that

$$\mathbb{E}_{h(x,y)}\|\hat{C}_l(u;v) - C_l(u;v)\|_2^2 \le \mathbb{E}_{h(x,y)}\|\hat{C}(u;v) - \mathbb{E}_{h(x,y)}(\hat{C}(u;v))\|_2^2 + \|\mathbb{E}_{h(x,y)}(\hat{C}(u;v)) - C(u;v)\|_2^2. \tag{A.10}$$

For a stationary sequence,

$$\|\mathbb{E}_{h(x,y)}(\hat{C}(u;v)) - C(u;v)\|_2^2 = 0.$$

Then, for the first term in (A.10)

$$\begin{aligned}
&\mathbb{E}_{h(x,y)}\|\hat{C}(u;v) - \mathbb{E}_{h(x,y)}(\hat{C}(u;v))\|_2^2 \\
&= \sum_{\mathbf{k}} \mathbb{E}_{h(x,y)}\big[(\hat{c}_{l,\mathbf{k}} - \mathbb{E}_{h(x,y)}(\hat{c}_{l,\mathbf{k}}))^2\big] \\
&= \sum_{\mathbf{k}} \mathbb{E}_{h(x,y)}\left[\left(\frac{1}{n}\sum_{t=1}^{n} W_t\right)^2\right] \\
&= \frac{1}{n^2}\sum_{\mathbf{k}} \mathbb{E}_{h(x,y)}\left[\sum_{t=1}^{n} W_t^2 + \sum_{t=1}^{n}\sum_{\substack{h=1\\ t\neq h}}^{n} W_t W_h\right] \\
&= \frac{1}{n^2}\sum_{\mathbf{k}}\left[\sum_{t=1}^{n} \mathbb{E}_{h(x,y)}(W_t^2) + 2\sum_{h=1}^{n-1}(n-h)\mathbb{E}_{h(x,y)}(W_n W_{n-h})\right],
\end{aligned} \tag{A.11}$$

where

$$W_t = \int_{G(Y_t)}^{1}\int_{F(X_t)}^{1} \Phi_{l,\mathbf{k}}(u;v)\,du\,dv - \mathbb{E}_{h(x,y)}\left(\int_{G(Y_t)}^{1}\int_{F(X_t)}^{1} \Phi_{l,\mathbf{k}}(u;v)\,du\,dv\right).$$

For the first term in (A.11),

$$\begin{aligned}
&\frac{1}{n^2}\sum_{\mathbf{k}}\left[\sum_{t=1}^{n} \mathbb{E}_{h(x,y)}(W_t^2)\right] \\
&\le \frac{1}{n^2}\sum_{\mathbf{k}}\sum_{t=1}^{n} \mathbb{E}_{h(x,y)}\left[\left(\int_{G(Y_t)}^{1}\int_{F(X_t)}^{1} \Phi_{l,\mathbf{k}}(u;v)\,du\,dv\right)^2\right] \\
&\le \frac{1}{n^2}\sum_{\mathbf{k}}\sum_{t=1}^{n} \mathbb{E}_{h(x,y)}\left[\left(2^{-l}\int_{2^l G(Y_t)-k_2}^{2^l-k_2}\int_{2^l F(X_t)-k_1}^{2^l-k_1} \phi(r)\phi(s)\,dr\,ds\right)^2\right] \\
&\le \frac{2^{-2l}}{n^2}\sum_{\mathbf{k}}\left[\sum_{t=1}^{n} \mathbb{E}_{h(x,y)}(M_1^2)\right] = \frac{2^{-2l}}{n}\sum_{\mathbf{k}} M_1^2,
\end{aligned}$$

where $2^l u - k_1 = r$, $2^l v - k_2 = s$, and

$$M_1 = \sup_{l,\mathbf{k}}\left\|\int_{2^l G(Y_t)-k_2}^{2^l-k_2}\int_{2^l F(X_t)-k_1}^{2^l-k_1} \phi(r)\phi(s)\,dr\,ds\right\|_\infty.$$

For the second term in (A.11)

$$\begin{aligned}\frac{2}{n^2}&\sum_{\mathbf{k}}\left[\sum_{h=1}^{n-1}(n-h)\mathbb{E}_{h(x,y)}(W_n W_{n-h})\right]\\&=\frac{2}{n^2}\sum_{\mathbf{k}}\left[\sum_{h=1}^{\gamma}(n-h)\mathbb{E}_{h(x,y)}(W_n W_{n-h})\right.\\&\quad\left.+\sum_{h=\gamma+1}^{n-1}(n-h)\mathbb{E}_{h(x,y)}(W_n W_{n-h})\right].\end{aligned}\tag{A.12}$$

For the first term in (A.12), under the assumption (A2), we get

$$\begin{aligned}&\frac{2}{n^2}\sum_{\mathbf{k}}\left[\sum_{h=1}^{\gamma}(n-h)\mathbb{E}_{h(x,y)}(W_n W_{n-h})\right]\\&\le\frac{2M}{n}\sum_{\mathbf{k}}\left[\sum_{h=1}^{\gamma}\left(1-\frac{h}{n}\right)\big(\mathbb{E}_{h(x,y)}|W_n|\big)\big(\mathbb{E}_{h(x,y)}|W_{n-h}|\big)\right].\end{aligned}$$

But, for all $t=1,\dots,n$,

$$\begin{aligned}\mathbb{E}_{h(x,y)}\big(|W_t|\big)&=\mathbb{E}_{h(x,y)}\left|\int_{G(Y_t)}^{1}\int_{F(X_t)}^{1}\Phi_{l,\mathbf{k}}(u;v)\,du\,dv\right.\\&\quad\left.-\mathbb{E}_{h(x,y)}\left(\int_{G(Y_t)}^{1}\int_{F(X_t)}^{1}\Phi_{l,\mathbf{k}}(u;v)\,du\,dv\right)\right|\\&=2^{-l}\mathbb{E}_{h(x,y)}\left|\int_{2^lG(Y_t)-k_2}^{2^l-k_2}\int_{2^lF(X_t)-k_1}^{2^l-k_1}\phi(r)\phi(s)\,dr\,ds\right.\\&\quad\left.-\mathbb{E}_{h(x,y)}\left(\int_{2^lG(Y_t)-k_2}^{2^l-k_2}\int_{2^lF(X_t)-k_1}^{2^l-k_1}\phi(r)\phi(s)\,dr\,ds\right)\right|,\end{aligned}$$

where

$$\int_{2^lG(Y_t)-k_2}^{2^l-k_2}\int_{2^lF(X_t)-k_1}^{2^l-k_1}\phi(r)\phi(s)\,dr\,ds\le\int_{[a',a]^2}\phi(r)\phi(s)\,dr\,ds.$$

Consider that $\phi(r)\phi(s)$ is nonnulll at $[a',a]^2$. So, we have that

$$\left|\int_{[a',a]^2}\phi(r)\phi(s)\,dr\,ds-\mathbb{E}_{h(x,y)}\left(\int_{[a',a]^2}\phi(r)\phi(s)\,dr\,ds\right)\right|\le 1,$$

and we have that $\mathbb{E}_{h(x,y)}(|W_t|)\le 2^{-l}$ uniformly in $(x;y)$, for all $t=1,\dots,n$.

Then, since $\sum_{h=1}^{\gamma}h<\infty$, we have that

$$\begin{aligned}&\frac{2M}{n}\sum_{\mathbf{k}}\left[\sum_{h=1}^{\gamma}\left(1-\frac{h}{n}\right)\big(\mathbb{E}_{h(x,y)}|W_n|\big)\big(\mathbb{E}_{h(x,y)}|W_{n-h}|\big)\right]\\&\quad\le\frac{2M}{n}\sum_{\mathbf{k}}\left[\sum_{h=1}^{\gamma}\left(1-\frac{h}{n}\right)2^{-2l}\right]\\&\quad\le\frac{2M}{n}\sum_{h=1}^{\gamma}\left(1-\frac{h}{n}\right)=\frac{2M\gamma}{n}-\frac{2M}{n^2}\sum_{h=1}^{\gamma}h.\end{aligned}$$

For the second term in (A.12)

$$\frac{2}{n^2}\sum_{\mathbf{k}}\left[\sum_{h=\gamma+1}^{n-1}(n-h)\mathbb{E}_{h(x,y)}(W_nW_{n-h})\right]$$

$$\leq \frac{2}{n^2}\sum_{\mathbf{k}}\left|\sum_{h=\gamma+1}^{n-1}(n-h)\mathbb{E}_{h(x,y)}(W_nW_{n-h})\right|$$

$$\leq \frac{2}{n}\sum_{\mathbf{k}}\left[\sum_{h=\gamma+1}^{n-1}\left|\mathbb{E}_{h(x,y)}(W_nW_{n-h})\right|\right],$$

For all $t=1,\ldots,n$, since $\mathbb{E}_{h(x,y)}(W_t)=0$, we have that $|\mathbb{E}_{h(x,y)}(W_nW_{n-h})|=|\operatorname{Cov}(W_n, W_{n-h})|$.

Moreover, using the Davydov inequality presented by Davydov (1968) and Rio (1993), we obtain

$$\frac{2}{n}\sum_{\mathbf{k}}\left[\sum_{h=\gamma+1}^{n-1}\left(1-\frac{h}{n}\right)\left|\mathbb{E}_{h(x,y)}(W_nW_{n-h})\right|\right]$$

$$\leq \frac{2}{n}\sum_{\mathbf{k}}\left[\sum_{h=\gamma+1}^{n}\left(1-\frac{h}{n}\right)2\frac{r}{r-2}(2\alpha(h))^{1-\frac{2}{r}}\right.$$

$$\left.\times\left[\mathbb{E}_{h(x,y)}\left(|W_n|^{,r}\right)\right]^{\frac{1}{r}}\left[\mathbb{E}_{h(x,y)}\left(|W_{n-h}|^{r}\right)\right]^{\frac{1}{r}}\right].$$

For all $t=1,\ldots,n$ and $(x;y)$, $|W_t|\leq 2^{-l}$ and $|W_t|^r\leq(2^{-l})^r$, so we have that $\mathbb{E}_{h(x,y)}(|W_t|^r)\leq 2^{-lr}$.

Then

$$\frac{2}{n}\sum_{\mathbf{k}}\left[\sum_{h=\gamma+1}^{n}\left(1-\frac{h}{n}\right)2\frac{r}{r-2}(2\alpha(h))^{1-\frac{2}{r}}\right.$$

$$\left.\times\left[\mathbb{E}_{h(x,y)}\left(|W_n|^{r}\right)\right]^{\frac{1}{r}}\left[\mathbb{E}_{h(x,y)}\left(|W_{n-h}|^{r}\right)\right]^{\frac{1}{r}}\right]$$

$$\leq \frac{1}{n}2^{2l}2^{3-\frac{1}{r}}\frac{r}{r-2}\sum_{h=\gamma+1}^{n}\left(1-\frac{h}{n}\right)(\alpha(h))^{1-\frac{2}{r}}2^{-2l}$$

$$\leq \frac{1}{n}K_r\sum_{h=\gamma+1}^{n}(\alpha(h))^{1-\frac{2}{r}},$$

where $K_r=2^{3-\frac{1}{r}}\frac{r}{r-2}$.

Under the assumption (A3), $\sum_{h=\gamma+1}^{n}(\alpha(h))^{1-\frac{2}{r}}=O(\gamma^{-1})$, and we have that

$$\frac{1}{n}K_r\sum_{h=\gamma+1}^{n}(\alpha(h))^{1-\frac{2}{r}}=\frac{1}{n}K_rO(\gamma^{-1}).$$

With the results of (A.12) we can conclude for the first term in (A.10) that

$$\mathbb{E}_{h(x,y)}\left\|\hat{C}_l(u;v)-\mathbb{E}_{h(x,y)}\left(\hat{C}_l(u;v)\right)\right\|_2^2$$

$$\leq \frac{(M_1)^2}{n}+\frac{2M\gamma}{n}-\frac{2M}{n^2}\sum_{h=1}^{\gamma}h+\frac{1}{n}K_rO(\gamma^{-1}).$$

And then, by (A.7) and (A.10),

$$\mathrm{MISE}(\tilde{C}_l(u; v), C(u; v)) \le \frac{(M_1)^2}{n} + \frac{2M\gamma}{n} - \frac{2M}{n^2} \sum_{h=1}^{\gamma} h \frac{1}{n} K_r O(\gamma^{-1})$$
$$\le K\left[\frac{2}{n} + M 2^{-2l(s+2)}\right].$$

As in i.i.d. case, the expression presents two antagonistic terms that must be balanced for which the expression has a minimum value.

So, $\mathrm{MISE}(\tilde{C}_l(u; v), C(u; v))$ has a minimum when l^* is such that $2^{l^*} \le n^{\frac{1}{2(s+2)}} < 2^{l^*+1}$. Then

$$\mathrm{MISE}(\tilde{C}_{l^*}(u; v), C(u; v)) \le K n^{-1},$$

which completes the proof of Theorem 2.2.

Acknowledgment

This work was supported by CNPq and Fapesp (2013/00506-1) grants.

References

Autin, F., Le Pennec, E. and Tribouley, K. (2010). Thresholding methods to estimate copula density. *Journal of Multivariate Analysis* **101**, 200–222. MR2557629 https://doi.org/10.1016/j.jmva.2009.07.009

Caillault, C. and Guégan, D. (2005). Empirical estimation of tail dependence using copulas: Application to Asian markets. *Quantitative Finance* **5**, 489–501. MR2241324 https://doi.org/10.1080/14697680500147853

Davydov, Y. A. (1968). Convergence of distributions generated by stationary stochastic processes. *Theory of Probability and Its Applications* **13**, 691–696. MR0243586

Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics* **27**, 642–669. MR0083864 https://doi.org/10.1214/aoms/1177728174

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

Genest, C., Massiello, E. and Tribouley, K. (2009). Estimating copula densities through wavelets. *Insurance: Mathematics and Economics* **44**, 170–181. MR2517883 https://doi.org/10.1016/j.insmatheco.2008.07.006

Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelets, Approximation, and Statistical Applications*. *Lecture Notes in Statistics* **129**. New York: Springer. MR1618204 https://doi.org/10.1007/978-1-4612-2222-4

Latif, S. A. and Morettin, P. A. (2010). Introdução a Cópulas e Aplicações na Avaliação do Desempenho de Empresas. *Revista Brasileira de Estatística* **71**, 121–151.

Matlab (2013). MATLAB version 8.1.0.604 (R2013a). Natick, Massachusetts: The Mathworks, Inc.

Misiti, M., Misiti, Y., Oppenheim, G. and Poggi, J. (1996). *Wavelet Toolbox User's Guide*. The MathWorks, Inc.

Morettin, P. A. (2014). *Ondas e Ondaletas*, 2a ed. São Paulo: EDUSP.

Morettin, P. A., Toloi, C. M. C., Chiann, C. and de Miranda, J. C. S. (2010). Wavelet smoothed empirical copula estimators. *Revista Brasileira de Finanças* **8**, 263–281.

Morettin, P. A., Toloi, C. M. C., Chiann, C. and de Miranda, J. C. S. (2011). Wavelet estimation of copulas for time series. *Journal of Times Series Econometrics* **3**, 1–29. MR2928656 https://doi.org/10.2202/1941-1928.1033

Nelsen, R. (2005). *An Introduction to Copulas*, 2nd ed. New York: Springer. MR1653203 https://doi.org/10.1007/978-1-4757-3076-0

Patton, A. (2012). A review of copula models for economic time series. *Journal of Multivariate Analysis* **110**, 4–18. MR2927506 https://doi.org/10.1016/j.jmva.2012.02.021

Rio, E. (1993). Covariance inequalities for strongly mixing processes. *Annales de L'Institut Henri Poincaré* **29**, 587–597. MR1251142

Rosenthal, H. P. (1970). On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. *Israel Journal of Mathematics* **8**, 273–303. MR0271721 https://doi.org/10.1007/BF02771562

Sklar, A. (1959). Fonctions de répartition à *n* dimensions et leurs marges. *Publications de I'Institut de Statistique de L'Université de Paris* **8**, 229–231. MR0125600

Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. New York: Wiley. MR1681904 https://doi.org/10.1002/9780470317020

Yu, H. (1993). A Glivenko–Cantelli lemma and weak convergence for empirical processes of associated sequences. *Probability Theory and Related Fields* **95**, 357–370. MR1213196 https://doi.org/10.1007/BF01192169

F. L. Medina
Department of Statistics
Federal University of Pernambuco
Recife 50740-560
Brazil
E-mail: francy@de.ufpe.br

P. A. Morettin
C. M. C. Toloi
Department of Statistics
University of São Paulo
São Paulo 05508-090
Brazil
E-mail: pam@ime.usp.br
clelia@ime.usp.br