Copula estimation through wavelets

Francyelle L. Medina\textsuperscript{a}, Pedro A. Morettin\textsuperscript{b} and Clélia M. C. Toli\textsuperscript{b}

\textsuperscript{a}Department of Statistics, Federal University of Pernambuco, Recife, Brazil
\textsuperscript{b}Department of Statistics, University of São Paulo, São Paulo, Brazil

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 $\alpha$-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 et al. (2009) proposed a methodology based on the wavelet decomposition of the copula density, called rank-based estimator. Another reference is Autin et al. (2010), that used a nonlinear procedure based on thresholding methods.

\textit{Keywords and phrases.} Copula, Nonparametric estimation, Wavelets, $\alpha$-mixing processes.
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.

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

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

where

$$c_{l,k} = \int_{[0,1]^2} C(u, v) \Phi_{l,k}(u, v) du dv,$$

$$d_{j,k}^{\mu} = \int_{[0,1]^2} C(u, v) \Psi_{j,k}^{\mu}(u, v) du dv. \quad (2.2)$$

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).
Copula estimation

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 $\{\Phi_{l,k}(x,y), k = (k_1, k_2)\}_{k'}$, hence instead of (2.1) we may consider

$$C_l(u,v) = \sum_k c_{l,k} \Phi_{l,k}(u,v), \quad (2.3)$$

with

$$c_{l,k} = \int_{[0,1]^2} C(u,v) \Phi_{l,k}(u,v) dudv$$

$$\quad = \int_{[0,1]^2} \left[ \int_r^1 \int_s^1 \Phi_{l,k}(u,v) dudv \right] c(r,s) drds, \quad (2.4)$$

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

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

2.1 Estimation for i.i.d. case

In order to develop the estimation procedure, let $(X_i, Y_i), i = 1, \ldots, n$, 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, i.e., $F_n(X_i) = \frac{1}{n} \sum_{k=1}^n \mathbb{I}\{X_k \leq X_i\}$ and $G_n(Y_i) = \frac{1}{n} \sum_{k=1}^n \mathbb{I}\{Y_k \leq Y_i\}$, where $\mathbb{I}\{x \in B\}$ denotes the indicator function, i.e., $\mathbb{I}\{x \in B\} = 1$ if $x \in B$ and $\mathbb{I}\{x \in B\} = 0$ otherwise.

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

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

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

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

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
\[ \text{MISE} \left( \hat{C}_l(u, v), C(u, v) \right) = \mathbb{E}_{h(x,y)} \left\| \hat{C}_l(u, v) - C(u, v) \right\|_2^2 = \mathbb{E}_{h(x,y)} \left[ \int_0^1 \int_0^1 \left( \hat{C}_l(u, v) - C(u, v) \right)^2 \, du \, dv \right]. \quad (2.6) \]

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

(A1) \( C \) belongs to the ball of radius \( M > 0 \) in the Besov space \( \mathfrak{B}_{s,q}^2 \).

(A2) For every integer \( h \in \mathbb{Z} \), the joint distribution \( J \left( (X_t; Y_t); (X_{t+h}; Y_{t+h}) \right) \) 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

\[ \mathbb{E} \left[ \left| H(X_t; Y_t) \cdot H(X_{t+h}; Y_{t+h}) \right| \right] \leq M \mathbb{E} \left[ \left| H(X_t; Y_t) \right| \right] \mathbb{E} \left[ \left| H(X_{t+h}; Y_{t+h}) \right| \right]. \]

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

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

(A4) \( \{X_t, t \in \mathbb{Z}\} \) and \( \{Y_t, t \in \mathbb{Z}\} \) are both \( \alpha \)-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 \( H(.,.) \), with an unknown copula function \( C(.,.) \), choose \( l^* \), such that

\[ 2l^* \leq n^{\frac{1}{2(r+1)}} < 2l^* + 1. \]

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

Then, there exists a constant \( K > 0 \) such that

\[ \sup_{C \in \mathfrak{B}_{s,q}^2(M)} n^{-1} \text{MISE} \left( \hat{C}_{l^*}(u, v), C(u, v) \right) \leq K. \]

**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 \( \{ V_t = (X_t, Y_t), t \in \mathbb{Z} \} \) be a two-dimensional stationary stochastic process, for all \( t \in \mathbb{Z} \), and suppose that we have observations \( \{ V_t, t = 1, \ldots, n \} \).

Thus, the estimator \( \hat{c}_{l,k} \) is defined by

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

It follows that the copula estimator is given by

\[
\hat{C}_l(u,v) = \sum_{k} \hat{c}_{l,k} \Phi_{l,k}(u,v).
\]

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 \( \{ V_t, t \in \mathbb{Z} \} \). Choose \( l^* \), such that

\[
2^{l^*} \leq n^{\frac{1}{2l^*+2}} < 2^{l^*+1}.
\]

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

\[
MISE \left( \hat{C}_{l^*}(u;v), C(u;v) \right) \leq Kn^{-1}.
\]

**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 et al. (2009), compute the index \( J \) for which
   \[
   2^J \leq n^{\frac{1}{2J+2}} < 2^{J+1}.
   \]
2. Denote each element of the sample matrix \( A_{p \times p} \) by \( (a_{p_1,p_2}) \), where \( p_1, p_2 \in \{1, \ldots, p\} \). The matrix \( B \) is obtained by symmetrizing \( A \),
where

\[
\mathbf{B} = \left( \begin{array}{ccc}
\mathbf{A}_* & \mathbf{A} & \mathbf{A}_* \\
\mathbf{A}_* & \mathbf{A} & \mathbf{A}_* \\
\mathbf{A}_* & \mathbf{A} & \mathbf{A}_*
\end{array} \right),
\]

in which \( \mathbf{A}_* = \left( a_{p+1-p_1,p+1-p_2} \right) \), \( \mathbf{A}_* = \left( a_{p_1,p+1-p_2} \right) \) and \( \mathbf{A} = \left( a_{p+1-p_1,p_2} \right) \).

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 (R2013a) 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 \(i = 1, ..., n\);
2. compute the empirical copula function on the grid \( \left( \frac{i}{n}, \frac{j}{n} \right) \), for which

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

3. compute the copula estimator \( \hat{C} \);
4. compute the true copula \( C \left( \frac{i}{n}, \frac{j}{n} \right) \) on the grid;
5. repeat the steps (1)-(3) “m” times and compute the Bias and mean squared errors (MSE), defined as

\[
\text{Bias} = \frac{1}{m} \sum_{k=1}^{m} \left( \hat{C}_k - C \right), \quad \text{MSE} = \frac{1}{m} \sum_{k=1}^{m} \left( \hat{C}_k - C \right)^2.
\]
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 5,000 samples of size $n = 1,024$. 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 $\gamma_{x,y} = 0$.

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

<table>
<thead>
<tr>
<th>Copulas</th>
<th>C(.01,.01)</th>
<th>C(.05,.05)</th>
<th>C(.25,.25)</th>
<th>C(.50,.50)</th>
<th>C(.75,.75)</th>
<th>C(.95,.95)</th>
<th>C(.99,.99)</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>1.00</td>
<td>25.00</td>
<td>625.00</td>
<td>2,500.00</td>
<td>5,625.00</td>
<td>9,025.00</td>
<td>9,801.00</td>
</tr>
<tr>
<td>Mean</td>
<td>1.90561</td>
<td>24.92</td>
<td>625.60</td>
<td>2,501.69</td>
<td>5,625.93</td>
<td>9,027.76</td>
<td>9,851.28</td>
</tr>
<tr>
<td>Bias</td>
<td>0.90561</td>
<td>-0.07571</td>
<td>0.60058</td>
<td>1.69088</td>
<td>0.93426</td>
<td>2.76271</td>
<td>50.2832</td>
</tr>
<tr>
<td>MSE</td>
<td>0.000622</td>
<td>0.16637</td>
<td>3.6186</td>
<td>6.13468</td>
<td>3.50512</td>
<td>0.13739</td>
<td>2.52971</td>
</tr>
<tr>
<td>Mean</td>
<td>1.25984</td>
<td>24.50</td>
<td>625.74</td>
<td>2,501.65</td>
<td>5,625.80</td>
<td>9,028.50</td>
<td>9,814.29</td>
</tr>
<tr>
<td>Bias</td>
<td>0.25984</td>
<td>-0.49238</td>
<td>0.73695</td>
<td>1.64967</td>
<td>0.79662</td>
<td>3.50903</td>
<td>33.2999</td>
</tr>
<tr>
<td>MSE</td>
<td>0.00071</td>
<td>0.03078</td>
<td>0.34966</td>
<td>0.59704</td>
<td>0.33520</td>
<td>0.02173</td>
<td>0.01788</td>
</tr>
<tr>
<td>Mean</td>
<td>0.86686</td>
<td>24.50</td>
<td>625.72</td>
<td>2,501.67</td>
<td>5,625.84</td>
<td>9,028.50</td>
<td>9,807.87</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.13313</td>
<td>-0.49926</td>
<td>0.71719</td>
<td>1.67639</td>
<td>0.83611</td>
<td>3.50460</td>
<td>6.8917</td>
</tr>
<tr>
<td>MSE</td>
<td>0.00059</td>
<td>0.02047</td>
<td>0.34133</td>
<td>0.58830</td>
<td>0.32900</td>
<td>0.02113</td>
<td>0.00531</td>
</tr>
<tr>
<td>Mean</td>
<td>0.98855</td>
<td>24.49529</td>
<td>625.70</td>
<td>2,501.69</td>
<td>5,625.8162</td>
<td>9,028.50</td>
<td>9,805.67</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.01144</td>
<td>-0.50075</td>
<td>0.70948</td>
<td>1.65990</td>
<td>0.86123</td>
<td>3.50304</td>
<td>4.6729</td>
</tr>
<tr>
<td>MSE</td>
<td>0.000162</td>
<td>0.02078</td>
<td>0.34915</td>
<td>0.58574</td>
<td>0.32783</td>
<td>0.022260</td>
<td>0.008327</td>
</tr>
<tr>
<td>Mean</td>
<td>0.97892</td>
<td>24.50</td>
<td>625.75</td>
<td>2,501.62</td>
<td>5,625.89</td>
<td>9,028.50</td>
<td>9,805.67</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.02107</td>
<td>0.50232</td>
<td>0.73944</td>
<td>1.63857</td>
<td>0.89133</td>
<td>3.50448</td>
<td>4.67398</td>
</tr>
<tr>
<td>MSE</td>
<td>0.000355</td>
<td>0.021230</td>
<td>0.34447</td>
<td>0.59008</td>
<td>0.33201</td>
<td>0.02246</td>
<td>0.00918</td>
</tr>
</tbody>
</table>

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.

### 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.

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

<table>
<thead>
<tr>
<th>$\times 10^{-4}$</th>
<th>$C(0.01;0.01)$</th>
<th>$C(0.05;0.05)$</th>
<th>$C(0.25;0.25)$</th>
<th>$C(0.50;0.50)$</th>
<th>$C(0.75;0.75)$</th>
<th>$C(0.95;0.95)$</th>
<th>$C(0.99;0.99)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>26.9</td>
<td>197.2</td>
<td>1.509.8</td>
<td>3.739.9</td>
<td>6.568.8</td>
<td>9.197.2</td>
<td>9.826.9</td>
</tr>
<tr>
<td><strong>Wavelet estimator D2 (5,000 samples)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$l=1$ Mean</td>
<td>28.02</td>
<td>246.54</td>
<td>1.693.39</td>
<td>3.971.67</td>
<td>6.689.60</td>
<td>9.247.27</td>
<td>9.839.80</td>
</tr>
<tr>
<td>Bias</td>
<td>5.517882</td>
<td>21.66381</td>
<td>92.17984</td>
<td>117.30426</td>
<td>90.38550</td>
<td>25.01033</td>
<td>6.43748</td>
</tr>
<tr>
<td>MSE</td>
<td>0.025848</td>
<td>0.33976</td>
<td>3.78879</td>
<td>6.04912</td>
<td>3.66830</td>
<td>0.34869</td>
<td>0.06357</td>
</tr>
<tr>
<td>$l=2$ Mean</td>
<td>28.24</td>
<td>246.79</td>
<td>1.693.00</td>
<td>3.974.68</td>
<td>6.689.39</td>
<td>9.247.31</td>
<td>9.840.41</td>
</tr>
<tr>
<td>Bias</td>
<td>6.07745</td>
<td>21.30983</td>
<td>92.12631</td>
<td>117.30036</td>
<td>90.27855</td>
<td>25.07010</td>
<td>6.43231</td>
</tr>
<tr>
<td>MSE</td>
<td>0.026500</td>
<td>0.34150</td>
<td>3.78089</td>
<td>6.04717</td>
<td>3.66941</td>
<td>0.34800</td>
<td>0.06349</td>
</tr>
<tr>
<td>Bias</td>
<td>6.18244</td>
<td>25.34008</td>
<td>92.10068</td>
<td>117.49853</td>
<td>90.31479</td>
<td>24.80288</td>
<td>6.74592</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03151</td>
<td>0.36085</td>
<td>3.78883</td>
<td>6.05331</td>
<td>3.66433</td>
<td>0.33809</td>
<td>0.03949</td>
</tr>
<tr>
<td>$l=4$ Mean</td>
<td>28.46279</td>
<td>253.111</td>
<td>1.694.23</td>
<td>3.975.69</td>
<td>6.690.42</td>
<td>9.247.86</td>
<td>9.835.70</td>
</tr>
<tr>
<td>Bias</td>
<td>5.76884</td>
<td>27.96976</td>
<td>92.69756</td>
<td>117.90162</td>
<td>90.27935</td>
<td>25.34236</td>
<td>6.39336</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03599</td>
<td>0.34292</td>
<td>3.83516</td>
<td>6.10662</td>
<td>3.70685</td>
<td>0.35227</td>
<td>0.07050</td>
</tr>
<tr>
<td>$l=5$ Mean</td>
<td>27.77</td>
<td>235.09</td>
<td>1.696.63</td>
<td>3.979.61</td>
<td>6.694.94</td>
<td>9.241.76</td>
<td>9.865.37</td>
</tr>
<tr>
<td>Bias</td>
<td>5.12284</td>
<td>25.37992</td>
<td>92.87914</td>
<td>119.86441</td>
<td>90.05555</td>
<td>22.22949</td>
<td>6.22489</td>
</tr>
<tr>
<td>MSE</td>
<td>0.01437</td>
<td>0.21517</td>
<td>4.06349</td>
<td>6.31467</td>
<td>3.88012</td>
<td>0.26059</td>
<td>0.14911</td>
</tr>
</tbody>
</table>

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:

$$V_t = A + BV_{t-1} + \epsilon_t,$$

where $V_t = (X_t; Y_t)^\top$, $\epsilon_t \sim N(0; \Sigma)$ and $A = (1; 1)^\top$. The matrices $B$ and $\Sigma$ are defined taking into account the type of components. For both,
we generate 5,000 samples of size \( n = 1,024 \). All values are expressed as multiples of \( 10^{-4} \).

### 3.2.1 Independent components

For this case, let

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

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

<table>
<thead>
<tr>
<th>Copulas ( \times 10^{-4} )</th>
<th>( C(.01; .01) )</th>
<th>( C(.05; .05) )</th>
<th>( C(.25; .25) )</th>
<th>( C(.50; .50) )</th>
<th>( C(.75; .75) )</th>
<th>( C(.95; .95) )</th>
<th>( C(.99; .99) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>1.00</td>
<td>25.00</td>
<td>625.00</td>
<td>2,500.00</td>
<td>5,625.00</td>
<td>9,025.00</td>
<td>9,801.00</td>
</tr>
<tr>
<td>Wavelet estimator D2 (5,000 samples) ( l=1 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.79043</td>
<td>25.69</td>
<td>630.85</td>
<td>2,502.78</td>
<td>5,621.26</td>
<td>9,027.92</td>
<td>9,851.29</td>
</tr>
<tr>
<td>Bias</td>
<td>0.79043</td>
<td>0.69320</td>
<td>5.85242</td>
<td>2.78267</td>
<td>-0.70369</td>
<td>2.92745</td>
<td>50.29317</td>
</tr>
<tr>
<td>MSE</td>
<td>0.00006</td>
<td>0.00040</td>
<td>0.00841</td>
<td>0.00032</td>
<td>0.00012</td>
<td>0.00005</td>
<td>0.00004</td>
</tr>
<tr>
<td>Wavelet estimator D2 (5,000 samples) ( l=2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.19</td>
<td>24.62</td>
<td>630.91</td>
<td>2,502.62</td>
<td>5,621.21</td>
<td>9,028.70</td>
<td>9,814.31</td>
</tr>
<tr>
<td>Bias</td>
<td>0.18056</td>
<td>-0.37974</td>
<td>5.91207</td>
<td>2.62732</td>
<td>-0.78241</td>
<td>3.70217</td>
<td>13.31907</td>
</tr>
<tr>
<td>MSE</td>
<td>0.00066</td>
<td>0.00233</td>
<td>0.11222</td>
<td>0.00837</td>
<td>0.00264</td>
<td>0.00192</td>
<td>0.001702</td>
</tr>
<tr>
<td>Wavelet estimator D2 (5,000 samples) ( l=3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.87626</td>
<td>24.51274</td>
<td>630.89</td>
<td>2,502.63</td>
<td>5,621.34</td>
<td>9,028.89</td>
<td>9,807.38</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.12373</td>
<td>-0.18796</td>
<td>5.89316</td>
<td>2.62933</td>
<td>-0.67675</td>
<td>3.88203</td>
<td>6.88647</td>
</tr>
<tr>
<td>MSE</td>
<td>0.00105</td>
<td>0.00195</td>
<td>0.10864</td>
<td>0.00857</td>
<td>0.00228</td>
<td>0.000529</td>
<td>0.00029</td>
</tr>
<tr>
<td>Wavelet estimator D2 (5,000 samples) ( l=4 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.98333</td>
<td>24.50</td>
<td>630.88</td>
<td>2,502.66</td>
<td>5,624.36</td>
<td>9,028.58</td>
<td>9,805.67</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.01667</td>
<td>-0.19575</td>
<td>5.89007</td>
<td>2.60035</td>
<td>-0.63316</td>
<td>3.89812</td>
<td>4.60924</td>
</tr>
<tr>
<td>MSE</td>
<td>0.00114</td>
<td>0.00225</td>
<td>0.10002</td>
<td>0.10006</td>
<td>0.00239</td>
<td>0.00029</td>
<td>0.00029</td>
</tr>
<tr>
<td>Wavelet estimator D2 (5,000 samples) ( l=5 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.98539</td>
<td>24.50</td>
<td>630.92</td>
<td>2,502.61</td>
<td>5,621.35</td>
<td>9,028.60</td>
<td>9,806.69</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.01400</td>
<td>-0.19638</td>
<td>5.90641</td>
<td>2.61533</td>
<td>-0.61523</td>
<td>3.90959</td>
<td>4.68638</td>
</tr>
<tr>
<td>MSE</td>
<td>0.00105</td>
<td>0.00225</td>
<td>0.11322</td>
<td>0.10305</td>
<td>0.00230</td>
<td>0.00022</td>
<td>0.00022</td>
</tr>
</tbody>
</table>

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. 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

\[
B = \begin{bmatrix} 0.25 & 0.2 \\ 0.2 & 0.75 \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} 0.75 & 0.5 \\ 0.5 & 1.25 \end{bmatrix}.
\]
Table 4  Mean, Bias and MSE of the estimator - Wavelet Daubechies D2 - case with dependent components.

<table>
<thead>
<tr>
<th>Copulas</th>
<th>$C(.01 ; .01)$</th>
<th>$C(.05 ; .05)$</th>
<th>$C(.25 ; .25)$</th>
<th>$C(.50 ; .50)$</th>
<th>$C(.75 ; .75)$</th>
<th>$C(.95 ; .95)$</th>
<th>$C(.99 ; .99)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>26.9</td>
<td>197.2</td>
<td>1.509.8</td>
<td>6.508.8</td>
<td>9.197.2</td>
<td>9.826.9</td>
<td></td>
</tr>
<tr>
<td>Wavelet estimator D2 (5,000 samples)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>0.01052</td>
<td>0.19315</td>
<td>2.20303</td>
<td>2.91158</td>
<td>0.54573</td>
<td>0.07141</td>
<td>0.08477</td>
</tr>
<tr>
<td>MSE</td>
<td>0.02029</td>
<td>0.21037</td>
<td>2.22979</td>
<td>2.93473</td>
<td>0.54438</td>
<td>0.08432</td>
<td>0.00473</td>
</tr>
<tr>
<td>MSE</td>
<td>0.01867</td>
<td>0.20959</td>
<td>2.23056</td>
<td>2.950065</td>
<td>0.55730</td>
<td>0.08182</td>
<td>0.014811</td>
</tr>
<tr>
<td>$l=4$ Mean</td>
<td>18.3893</td>
<td>164.139</td>
<td>1.374.60</td>
<td>3.585.32</td>
<td>6.465.16</td>
<td>9.184.03</td>
<td>9.815.71</td>
</tr>
<tr>
<td>MSE</td>
<td>0.02229</td>
<td>0.20825</td>
<td>2.242146</td>
<td>2.946415</td>
<td>0.558231</td>
<td>0.080473</td>
<td>0.022071</td>
</tr>
<tr>
<td>MSE</td>
<td>0.021615</td>
<td>0.207158</td>
<td>2.240509</td>
<td>2.946415</td>
<td>0.560300</td>
<td>0.080146</td>
<td>0.021664</td>
</tr>
</tbody>
</table>

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

$$
\hat{C}(u; v) = \sum_{k} \hat{c}_{l,k} \Phi_{l,k}(u; v) + \sum_{j \geq 1} \sum_{k \in \mathbb{Z}^2} \sum_{\mu=0}^{J} \hat{d}_{j,k}^{\mu} \Psi_{j,k}^{\mu}(u; v),
$$

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

$$
\delta_{H}^{\lambda}(x) = \begin{cases} 
0, & \text{if } |x| \leq \lambda, \\
 x, & \text{if } |x| > \lambda,
\end{cases}
$$
Figure 1  Graphical representations of the estimated copula at different levels - dependent components without thresholds. Last graphic is for the Normal copula.

and

\[ \delta^S(x) = \begin{cases} 
0, & \text{if } |x| \leq \lambda, \\
\sin(x)(|x| - \lambda), & \text{if } |x| > \lambda,
\end{cases} \]

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

Thus, the final estimator is given by

\[
\hat{C}(u; v) = \sum_{k} \hat{c}_{l,k} \Phi_{l,k}(u; v) + \sum_{j \geq l} \sum_{k \in \mathbb{Z}^2} \sum_{\mu = h,v,d} \delta^{H,S}_{\lambda} \left( \hat{d}_{j,k}^\mu \right) \Psi_{j,k}^\mu(u; v).
\]

In this research, we choose as the threshold the high quantile proposed by Vidakovic.
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 5,000 samples of size $n = 1,024$ of the VAR(1) model given by (3.1), with dependent components. The results are in Table 5.

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

<table>
<thead>
<tr>
<th>Copulas</th>
<th>True</th>
<th>Wavelet estimator D2 - threshold(5,000 samples)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\times 10^{-4}$</td>
<td>C(.01;01)</td>
<td>C(.05;05)</td>
</tr>
<tr>
<td>Mean</td>
<td>25.9</td>
<td>197.2</td>
</tr>
<tr>
<td>MSE</td>
<td>0.01052</td>
<td>0.10116</td>
</tr>
</tbody>
</table>

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.

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

3.3 Additional simulations

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

$$\text{Error} \left( \hat{C}_{l^*}, C_\theta \right) = \frac{1}{N^2} \left\| \hat{C}_{l^*} - C_\theta \right\|_2^2 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\{ \hat{C}_{l^*} \left( \frac{i}{N}, \frac{j}{N} \right) - C_\theta \left( \frac{i}{N}, \frac{j}{N} \right) \right\}^2,$$
and
\[ RE \left( \hat{C}_{l^*}, C_\theta \right) = Error \left( \hat{C}_{l^*}, C_\theta \right) \times \left[ \frac{1}{N^2} \| C_\theta \|_2^2 \right]^{-1}, \]
where \( RE \) is the relative error, \( \hat{C}_{l^*} \) is the estimated copula function on the grid \( \left( \frac{i}{N}, \frac{j}{N} \right), \ i, j = 1, \ldots, N \) and \( C_\theta \) is the parametric copula, with fixed \( \theta \).

The procedure is as follows:

1. draw a sample of size \( n \) from a parametric copula \( C_\theta \);
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 \( r = 5,000 \) replicates for samples of size \( n = 256 \) and \( n = 1,024 \), 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.

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

<table>
<thead>
<tr>
<th>n=256</th>
<th>Wavelets</th>
<th>Copulas</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daubechies D2</td>
<td>Daubechies D4</td>
<td>Daubechies D6</td>
<td>Daubechies D8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1=1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>43.3428</td>
<td>36.5873</td>
<td>43.5466</td>
<td>36.4782</td>
<td>42.7880</td>
<td>36.5374</td>
<td>44.5405</td>
<td>36.5577</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-t</td>
<td>43.2967</td>
<td>36.2372</td>
<td>43.3172</td>
<td>36.1408</td>
<td>42.6340</td>
<td>36.2346</td>
<td>44.3287</td>
<td>36.2234</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gumbel</td>
<td>39.5915</td>
<td>35.7352</td>
<td>38.9290</td>
<td>35.6866</td>
<td>38.1589</td>
<td>35.7829</td>
<td>40.0677</td>
<td>35.7841</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clayton</td>
<td>41.3565</td>
<td>38.6787</td>
<td>40.8900</td>
<td>38.5900</td>
<td>39.9253</td>
<td>38.7219</td>
<td>42.1046</td>
<td>38.7085</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frank</td>
<td>42.9301</td>
<td>37.1642</td>
<td>42.7577</td>
<td>37.1098</td>
<td>42.1950</td>
<td>37.1950</td>
<td>43.8146</td>
<td>37.1927</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1=2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>44.0426</td>
<td>36.5550</td>
<td>44.3046</td>
<td>36.5566</td>
<td>44.3168</td>
<td>36.5514</td>
<td>44.5604</td>
<td>36.5497</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-t</td>
<td>43.9318</td>
<td>36.2537</td>
<td>43.2702</td>
<td>36.2531</td>
<td>43.1876</td>
<td>36.2506</td>
<td>44.2372</td>
<td>36.2588</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clayton</td>
<td>40.8397</td>
<td>38.7152</td>
<td>41.1343</td>
<td>38.7115</td>
<td>41.0108</td>
<td>38.7124</td>
<td>41.0713</td>
<td>38.7162</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frank</td>
<td>43.2385</td>
<td>37.1069</td>
<td>43.5173</td>
<td>37.2062</td>
<td>37.3611</td>
<td>37.2030</td>
<td>43.5114</td>
<td>37.2035</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>44.9479</td>
<td>36.5550</td>
<td>45.0068</td>
<td>36.5598</td>
<td>45.0678</td>
<td>36.5578</td>
<td>45.6077</td>
<td>36.5588</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-t</td>
<td>44.8298</td>
<td>36.2573</td>
<td>44.1975</td>
<td>36.2613</td>
<td>44.9484</td>
<td>36.2530</td>
<td>44.9481</td>
<td>36.2594</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clayton</td>
<td>41.4846</td>
<td>38.7158</td>
<td>41.6213</td>
<td>38.7195</td>
<td>41.5805</td>
<td>38.7163</td>
<td>41.5768</td>
<td>38.7167</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frank</td>
<td>44.0307</td>
<td>37.2066</td>
<td>44.1870</td>
<td>37.2115</td>
<td>44.1612</td>
<td>37.2084</td>
<td>44.1612</td>
<td>37.2086</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>45.4800</td>
<td>36.5596</td>
<td>45.5947</td>
<td>36.5612</td>
<td>45.5908</td>
<td>36.5602</td>
<td>45.5994</td>
<td>36.5604</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-t</td>
<td>45.3584</td>
<td>36.2501</td>
<td>45.4129</td>
<td>36.2606</td>
<td>45.4170</td>
<td>36.2600</td>
<td>45.4176</td>
<td>36.2603</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gumbel</td>
<td>40.0904</td>
<td>35.7915</td>
<td>40.1378</td>
<td>35.7923</td>
<td>40.1366</td>
<td>35.7925</td>
<td>40.1381</td>
<td>35.7912</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clayton</td>
<td>41.9283</td>
<td>38.7154</td>
<td>41.9747</td>
<td>38.7160</td>
<td>41.9738</td>
<td>38.7166</td>
<td>41.9753</td>
<td>38.7152</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frank</td>
<td>45.5467</td>
<td>37.2091</td>
<td>44.5975</td>
<td>37.2109</td>
<td>44.5908</td>
<td>37.2098</td>
<td>44.6017</td>
<td>37.2101</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1=5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>45.7792</td>
<td>36.5626</td>
<td>45.7935</td>
<td>36.5624</td>
<td>45.7907</td>
<td>36.5622</td>
<td>45.7998</td>
<td>36.5623</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-t</td>
<td>45.6582</td>
<td>36.2612</td>
<td>45.6730</td>
<td>36.2614</td>
<td>45.6764</td>
<td>36.2616</td>
<td>45.6790</td>
<td>36.2615</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gumbel</td>
<td>40.3360</td>
<td>35.7940</td>
<td>40.3472</td>
<td>35.7935</td>
<td>40.3488</td>
<td>35.7935</td>
<td>40.3520</td>
<td>35.7938</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clayton</td>
<td>42.1822</td>
<td>38.7166</td>
<td>42.1934</td>
<td>38.7161</td>
<td>42.1901</td>
<td>38.7163</td>
<td>42.1986</td>
<td>38.7163</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frank</td>
<td>44.8271</td>
<td>37.2120</td>
<td>44.8400</td>
<td>37.2120</td>
<td>44.8410</td>
<td>37.2120</td>
<td>44.8466</td>
<td>37.2120</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

wavelets D2, D4, D6 and D8) and five estimated parametric copulas. The values are reported in Table 8, where we see that the 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 be-
### Table 7 Mean and standard deviation (SD) of RE for the wavelet estimator and some parametric copula, n = 1,024.

<table>
<thead>
<tr>
<th>n=1,024</th>
<th>Wavelets</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Daubech D2</td>
<td>Daubech D4</td>
<td>Daubech D6</td>
<td>Daubech D8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copulas</td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Normal</td>
<td>11.0341</td>
<td>8.6375</td>
<td>11.0549</td>
<td>8.6376</td>
<td>11.0788</td>
<td>8.6373</td>
<td></td>
</tr>
<tr>
<td>Frank</td>
<td>0.9561</td>
<td>9.2673</td>
<td>10.9718</td>
<td>9.2676</td>
<td>10.9945</td>
<td>9.2677</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>11.0471</td>
<td>8.6376</td>
<td>11.0471</td>
<td>8.6374</td>
<td>11.0788</td>
<td>8.6373</td>
<td></td>
</tr>
<tr>
<td>Frank</td>
<td>0.9561</td>
<td>9.2676</td>
<td>10.9718</td>
<td>9.2676</td>
<td>10.9945</td>
<td>9.2677</td>
<td></td>
</tr>
</tbody>
</table>

### Table 8 Error and RE of copula estimators, for the Bbovespa and IPC series.

<table>
<thead>
<tr>
<th>Copula</th>
<th>Wavelet D2</th>
<th>Wavelet D4</th>
<th>Wavelet D6</th>
<th>Wavelet D8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.3418</td>
<td>2.6173</td>
<td>0.3418</td>
<td>2.6175</td>
</tr>
<tr>
<td>Student-t</td>
<td>0.3418</td>
<td>2.6175</td>
<td>0.3418</td>
<td>2.6176</td>
</tr>
<tr>
<td>Clayton</td>
<td>0.8760</td>
<td>7.0072</td>
<td>0.8760</td>
<td>7.0075</td>
</tr>
<tr>
<td>Frank</td>
<td>0.3418</td>
<td>2.6177</td>
<td>0.3418</td>
<td>2.6177</td>
</tr>
</tbody>
</table>

*The value inside of [ ] represents the estimated degrees of freedom.*
Figure 3 Scatter plot for the returns of Bovespa and IPC series.

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 Bovespa and IPC series.

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 1,018 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 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.

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

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** 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.

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.

Acknowledgements

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

References


Appendix

Proof of Theorem 2.1

Let

\[ \{ \Phi_{l,k}(x; y), k = (k_1, k_2) \}_{k} \cup \{ \Psi_{j,k}(x; y), k = (k_1, k_2), \mu = h, v, d \}_{j \geq l, k} \]

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

\[
MISE(\tilde{C}_1(u; v), C(u; v)) \leq \mathbb{E}_{h(x,y)} \left\| \tilde{C}_1(u; v) - C_1(u; v) \right\|^2_2 + \left\| \sum_{j \geq l} \sum_{k \in \mathbb{Z}^2} \sum_{\mu = h, v, d} d_{j,k}^\mu \Psi_{j,k}(u; v) \right\|^2_2. \quad (A.1)
\]

If \( C \in \mathbb{R}^{s,q}_p \), with \( s > 0, 1 \leq p, q < \infty \), then

\[
\|C\|_{s,p,q} = \|c_l\|_p + \left( \sum_{j \geq l} \left( 2^{j(s+\frac{2}{p}+1)} \|d_j\|_p \right)^q \right)^{\frac{1}{q}}.
\]

Under assumption (A1), using the Hölder inequality, with \( \frac{1}{q} + \frac{1}{q'} = 1 \) and
For the second term in (A.1),
\[
\left\| \sum_{j \geq l} \sum_{k, \mu} d_{j,k}^\mu \Psi_{j,k}^\mu(u,v) \right\|_2 \leq \left( \sum_{j \geq l} \left( 2^{j(s+2)} \left\| \sum_{k, \mu} d_{j,k}^\mu \Psi_{j,k}^\mu(u,v) \right\|_2 \right)^q \right)^{1/q} \times \\
\left( \sum_{j \geq l} \left( 2^{-j(s+2)} \right)^{q'} \right)^{1/q'} \leq \|C\|_{s,2,q} \left( \sum_{j \geq l} 2^{-j(s+2)} \right)^{1/q'} \leq M 2^{-2(l(s+2))}. \tag{A.2}
\]

Note that, as in Genest et al. (2009), for the first term in (A.1)
\[
\mathbb{E}_{h(x,y)} \left\| \hat{C}_l(u,v) - C_l(u,v) \right\|_2^2 \leq 2 \mathbb{E}_{h(x,y)} \left\| \hat{C}_l(u,v) - \hat{C}_l(u,v) \right\|_2^2 + 2 \mathbb{E}_{h(x,y)} \left\| \hat{C}_l(u,v) - C_l(u,v) \right\|_2^2, \tag{A.3}
\]
where
\[
\hat{C}_{l,k} = \frac{1}{n} \sum_{i=1}^n \int_{G(Y)} \int_{F(X)} \Phi_{l,k}(u,v) du dv,
\]
with
\[
\hat{C}_l(u,v) = \sum_k \hat{C}_{l,k} \Phi_{l,k}(u,v).
\]

Now, we want to find upper bounds for (A.3), separately.
For the first term in (A.3),
\[
\mathbb{E}_{h(x,y)} \left\| \tilde{C}_t(u,v) - \hat{C}_t(u,v) \right\|_2^2 = \sum_k \mathbb{E}_{h(x,y)} \left[ (\tilde{c}_{t,k} - \hat{c}_{t,k})^2 \right] = \sum_k \mathbb{E}_{h(x,y)} \left( \frac{1}{n} \sum_{i=1}^n \int_{G(Y_i)}^1 \int_{F(X_i)^{u,v}}^1 \Phi_{t,k}(u,v) dudv \right)^2 \\
- \frac{1}{n} \sum_{i=1}^n \int_{G(Y_i)}^1 \int_{F(X_i)}^1 \Phi_{t,k}(u,v) dudv \right)^2 \\
\leq \sum_k \mathbb{E}_{h(x,y)} \left( \| \Phi_{t,k}(u,v) \|_\infty \right)^2 \\
\left\{ \frac{1}{n} \sum_{i=1}^n \left[ (1 - G(Y_i))(1 - F(X_i)) \right. \right. \\
- (1 - G(Y_i))(1 - F(X_i)) + (1 - F_n(X_i))(1 - G(Y_i)) \right. \\
- (1 - F_n(X_i))(1 - G(Y_i)) \right\}^2 \\
\leq \sum_k \mathbb{E}_{h(x,y)} \left( \| \Phi_{t,k}(u,v) \|_\infty \right)^2 \times \\
\left\{ \frac{1}{n} \sum_{i=1}^n \left[ G(Y_i) - G_n(Y_i) \right] + \left. |F(X_i) - F_n(X_i)| \right\}^2 \right. \\
\leq \sum_{i=1}^n |\Delta(X_i)| \leq n\epsilon + n \left( \mathbb{I} \{|\Delta(X_i)| > \epsilon \} + \mathbb{I} \{|\Delta(X_i)| \leq \epsilon \} \right).
\]

Let
\[
\Delta(X_i) = F(X_i) - F_n(X_i) \text{ and } \Delta(Y_i) = G(Y_i) - G_n(Y_i),
\]

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

And for
\[
\sum_{i=1}^n |\Delta(X_i)| \leq n\epsilon + n \left( \mathbb{I} \{|\Delta(X_i)| > \epsilon \} \right),
\]

\[
\sum_k \mathbb{E}_{h(x,y)} \left[ (\tilde{c}_{t,k} - \hat{c}_{t,k})^2 \right] \leq \sum_k \mathbb{E}_{h(x,y)} \left( \| \Phi_{t,k}(u,v) \|_\infty \right)^2 \left\{ \frac{1}{n} \sum_{i=1}^n \left[ |\Delta(X_i)| + |\Delta(Y_i)| \right] \right\}^2 \\
\leq \sum_k (\| \Phi_{t,k}(u,v) \|_\infty)^2 \mathbb{E}_{h(x,y)} \left\{ 4\epsilon^2 \\
+ (3 + 4\epsilon) \left( \mathbb{I} \{|\Delta(X_i)| > \epsilon \} + \mathbb{I} \{|\Delta(Y_i)| > \epsilon \} \right) \right\}. 
\]
But we have that

\[
\| \Phi_{l,k}(u; v) \|_{\infty} = \left\| 2^l \Phi(2^l u - k_1; 2^l v - k_2) \right\|_{\infty} \\
= 2^l \sup_{(2^l u - k_1; 2^l v - k_2)} \left| \Phi(2^l u - k_1; 2^l v - k_2) \right| \\
= 2^l \| \Phi (w, z) \|_{\infty}, \text{ with } w = 2^l u - k_1 \text{ and } z = 2^l v - k_2,
\]

for \( k_1 = 0, \ldots, 2^l - 1, \ k_2 = 0, \ldots, 2^l - 1. \)

Since \( \sum_k = 2^{2l} \), applying the Dvoretzky-Kiefer-Wolfowitz inequality (see Dvoretzky et al. (1956)), provided that \( \epsilon = \left\{ \frac{\delta \log(n)}{2n} \right\}^{\frac{1}{2}} \), where \( \delta \) could be as large as desired, one finds

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

where \( k_1, k_2 \) and \( k_3 \) depend on \( \| \Phi (w, z) \|_{\infty} \) and \( \delta \).

For the second term in (A.3)

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

Let

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

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

Using (A.2), (A.4) and (A.5), we have that
\[
MISE \left( \hat{C}_l(u;v), C(u;v) \right) \leq 2^{l+1} \left[ k_1 \log(n) + k_2 \frac{n^{-\delta}}{n} + k_3 \left( \log(n) \right)^{\frac{3}{2}} \frac{n^{-\left( \delta + \frac{1}{2} \right)}}{n} \right] \\
+ K \frac{2^{4l+1}}{n} + M2^{-2l(s+2)}.
\]

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, $MISE \left( \hat{C}_l(u;v), C(u;v) \right)$ has a minimum when $l^*$ is such that $2^{l^*} \leq n \frac{1}{2^{l+1}} < 2^{l^*+1}$.

Then,
\[
\sup_{C \in \mathbb{R}_{s,d}^q(M)} MISE \left( \hat{C}_{l^*}(u;v), C(u;v) \right) \leq Kn \frac{2^{l^*}}{n} \\
\Rightarrow \sup_{C \in \mathbb{R}_{s,d}^q(M)} n^{\frac{l^*+2}{2}} MISE \left( \hat{C}_{l^*}(u;v), C(u;v) \right) \leq K.
\]

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 $\hat{c}_{l,k}$ under dependence structure, we have that
\[
MISE \left( \hat{C}_l(u;v), C(u;v) \right) \leq \mathbb{E}_{h(x,y)} \left\| \hat{C}_l(u;v) - C_l(u;v) \right\|^2_2 \\
+ \left\| \sum_{j \geq l} \sum_{k \in \mathbb{R}_{s,d}^q \mu = h_{x,y}} d_{j,k}^\mu \Psi_{j,k}^{\mu}(u;v) \right\|^2_2.
\]
Under the assumption (A1),
\[
\left\| \sum_{j \geq 1} \sum_{k, \mu} d_{j,k}^{\mu} \Psi_{j,k}(u; v) \right\|_2^2 \leq M 2^{-2(s+2)}. \tag{A.6}
\]

Then, it is only necessary to study
\[
\mathbb{E}_{h(x,y)} \left\| \hat{C}_t(u; v) - C_t(u; v) \right\|_2^2 \leq 2 \mathbb{E}_{h(x,y)} \left\| \hat{C}_t(u; v) - \hat{C}_t(u; v) \right\|_2^2 + 2 \mathbb{E}_{h(x,y)} \left\| \hat{C}_t(u; v) - C_t(u; v) \right\|_2^2, \tag{A.7}
\]
where
\[
\hat{c}_{t,k} = \frac{1}{n} \sum_{i=1}^{n} \left[ \int_{G(Y_i)}^{1} \int_{F(X_t)}^{1} \Phi_{j,k}(u; v) du dv \right].
\]

We have, for the first term in (A.7)
\[
\mathbb{E}_{h(x,y)} \left\| \hat{C}_t(u; v) - C_t(u; v) \right\|_2^2 = \sum_{k} \mathbb{E}_{h(x,y)} \left( (\hat{c}_{t,k} - \hat{c}_{t,k})^2 \right)
\]
\[
\leq \sum_{k} \mathbb{E}_{h(x,y)} \left( ||\Phi_{t,k}(u; v)||_{\infty} \right)^2 \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[ G(Y_i) - G_n(Y_i) \right] + \left[ F(X_t) - F_n(X_t) \right] \right\}.
\]
\[
= \sum_{k} \mathbb{E}_{h(x,y)} \left( ||\Phi_{t,k}(u; v)||_{\infty} \right)^2 \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[ \Delta(X_t) + |\Delta(Y_i)| \right] \right\}^2,
\]
where $\Delta(Y_i) = \Delta(Y_i) \left( I \{ |\Delta(Y_i) > \epsilon \} + I \{ |\Delta(Y_i) \leq \epsilon \} \right)$. For fixed $\epsilon > 0$, which $\epsilon = \frac{1}{n}$, we have that
\[
\sum_{k} \mathbb{E}_{h(x,y)} \left( (\hat{c}_{t,k} - \hat{c}_{t,k})^2 \right) \leq 2 \int_{n}^{2} \left( ||\Phi_{t}(u; z)||_{\infty} \right)^2 \sum_{k} \left\{ k_1 \frac{1}{n^2}
\]
\[
+ \left( k_2 + k_3 \frac{1}{n^2} \right) \left[ P \{ |\Delta(X_t) > \epsilon \} + P \{ |\Delta(Y_i) > \epsilon \} \right] \right\}.
\]

Proceeding as in Yu (1993), let \( \{ X_t, \ t \in \mathbb{Z} \} \) be a stationary sequence of random variables, with the same distribution function \( F(x) \). If \( F(x) \) is continuous and
\[
\sum_{t=1}^{\infty} \frac{1}{n^2} \text{Cov} \{ X_t, S_{n-t} \} < \infty, \tag{A.8}
\]
when \( t = 1, ..., n, \ S_n = \sum_{t=1}^{n} X_t \) and \( n \to \infty \), we have that
\[
\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{\frac{q}{n}} 0. \tag{A.9}
\]
For a stationary sequence, the condition (A.8) can be replaced by

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

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

$$\sum_k \mathbb{E}_{h(x,y)} \left[ (\hat{c}_{t,k} - \hat{c}_{t,k})^2 \right] = o(1).$$

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

$$\mathbb{E}_{h(x,y)} \left\| \hat{C}(u; v) - C_{l}(u; v) \right\|_2^2 \leq \mathbb{E}_{h(x,y)} \left\| \hat{C}(u; v) - \mathbb{E}_{h(x,y)} \left( \hat{C}(u; v) \right) \right\|_2^2 + \left\| \mathbb{E}_{h(x,y)} \left( \hat{C}(u; v) \right) - C(u; v) \right\|_2^2. \quad (A.10)$$

For a stationary sequence,

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

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

$$\mathbb{E}_{h(x,y)} \left\| \hat{C}(u; v) - \mathbb{E}_{h(x,y)} \left( \hat{C}(u; v) \right) \right\|_2^2 = \sum_k \mathbb{E}_{h(x,y)} \left[ (\hat{c}_{t,k} - \mathbb{E}_{h(x,y)}(\hat{c}_{t,k}))^2 \right] = \sum_k \mathbb{E}_{h(x,y)} \left[ \left( \frac{1}{n} \sum_{t=1}^{n} W_t \right)^2 \right] = \frac{1}{n^2} \sum_k \mathbb{E}_{h(x,y)} \left\{ \sum_{t=1}^{n} W_t^2 + \sum_{t=1}^{n} \sum_{h=1}^{n} W_t W_h \right\} - \frac{2}{n^2} \sum_{h=1}^{n-1} (n-h) \mathbb{E}_{h(x,y)} (W_n W_{n-h}), \quad (A.11)$$

where $W_t = \int_{G(Y_t)} \int_{F(X_t)} \Phi_k(u; v) du dv - \mathbb{E}_{h(x,y)} \left( \int_{G(Y_t)} \int_{F(X_t)} \Phi_k(u; v) du dv \right)$.  

imsart-bjps ver. 2014/10/16 file: "Submit BJPS":tex date: April 16, 2019
For the first term in (A.11),
\[
\frac{1}{n^2} \sum_k \left[ \sum_{t=1}^n E_h(x,y) (W^2) \right] \leq \frac{1}{n^2} \sum_k \left[ \sum_{t=1}^n E_h(x,y) \left( \left( \int_{G(Y_t)}^{1} \int_{F(X_t)}^{1} \Phi_{1,k}(u;v) du dv \right) \right)^2 \right]
\]
\[
\leq \frac{1}{n^2} \sum_k \left[ \sum_{t=1}^n E_h(x,y) \left( 2^{-l} \int_{2^l G(Y_t) - k_2}^{2^l - k_2} \int_{2^l F(X_t) - k_1}^{2^l - k_1} \phi(r) \phi(s) dr ds \right)^2 \right]
\]
\[
\leq \frac{2^{-2l}}{n^2} \sum_k \left[ \sum_{t=1}^n E_h(x,y) \left( M^2_k \right) \right] = \frac{2^{-2l}}{n} \sum_k M^2_k,
\]
where \( 2^l u - k_1 = r, 2^l v - k_2 = s \), and
\[
M_1 = \sup_{t,k} \left\| \int_{2^l G(Y_t) - k_2}^{2^l - k_2} \int_{2^l F(X_t) - k_1}^{2^l - k_1} \phi(r) \phi(s) dr ds \right\|_\infty.
\]

For the second term in (A.11)
\[
\frac{2}{n^2} \sum_k \left[ \sum_{h=1}^{n-1} (n-h) E_h(x,y) (W_n W_{n-h}) \right] = \frac{2}{n^2} \sum_k \left[ \sum_{h=1}^n (n-h) E_h(x,y) (W_n W_{n-h}) \right] + \sum_{h=\gamma+1}^{n-1} (n-h) E_h(x,y) (W_n W_{n-h}), \tag{A.12}
\]

For the first term in (A.12), under the assumption (A2), we get
\[
\frac{2}{n^2} \sum_k \left[ \sum_{h=1}^\gamma (n-h) E_h(x,y) (W_n W_{n-h}) \right] \leq \frac{2M}{n} \sum_k \left[ \sum_{h=1}^\gamma \left( 1 - \frac{\gamma}{n} \right) \times \left( E_h(x,y) (W_n) \right) \left( E_h(x,y) (W_{n-h}) \right) \right].
\]

But, for all \( t = 1, \ldots, n \),
\[
E_h(x,y) (|W_t|) = E_h(x,y) \left( \int_{G(Y_t)}^{1} \int_{F(X_t)}^{1} \Phi_{1,k}(u;v) du dv \right)
\]
\[
- E_h(x,y) \left( \int_{G(Y_t)}^{1} \int_{F(X_t)}^{1} \Phi_{1,k}(u;v) du dv \right)
\]
\[
= 2^{-l} E_h(x,y) \left( \int_{2^l G(Y_t) - k_2}^{2^l - k_2} \int_{2^l F(X_t) - k_1}^{2^l - k_1} \phi(r) \phi(s) dr ds \right)
\]
\[
- E_h(x,y) \left( \int_{2^l G(Y_t) - k_2}^{2^l - k_2} \int_{2^l F(X_t) - k_1}^{2^l - k_1} \phi(r) \phi(s) dr ds \right).
\]
where
\[
\int_{2^l G(Y_t) - k_2}^{2^l - k_2} \int_{2^l F(X_t) - k_1}^{2^l - k_1} \phi(r) \phi(s) dr ds \leq \int_{[a', a]^2} \phi(r) \phi(s) dr ds.
\]
Consider that $\phi(r)\phi(s)$ is nonnull at $[a', a]^2$. So, we have that

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

and we have that $\mathbb{E}_{h(x,y)} (|W_t|) \leq 2^{-l}$ uniformly in $(x, y)$, for all $t = 1, ..., n$. Then, since $\sum_{h=1}^{n} h < \infty$, we have that

$$\frac{2M}{n} \sum_{h=1}^{n-1} \left( 1 - \frac{h}{n} \right) \left( \mathbb{E}_{h(x,y)} (|W_n|) \right) \leq \frac{2M}{n} \sum_{h=1}^{n-1} \left( 1 - \frac{h}{n} \right) 2^{-2l} \leq \frac{2M}{n} \sum_{h=1}^{n-1} \left( 1 - \frac{h}{n} \right),$$

For the second term in (A.12)

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

For all $t = 1, ..., n$, since $\mathbb{E}_{h(x,y)} (W_t) = 0$, we have that $\mathbb{E}_{h(x,y)} (W_n W_{n-h}) = |Cov (W_n, W_{n-h})|$. Moreover, using the Davydov inequality presented by Davydov (1968) and Rio (1993), we obtain

$$\frac{2}{n} \sum_{k} \left[ \sum_{h=\gamma+1}^{n-1} \left( 1 - \frac{h}{n} \right) \mathbb{E}_{h(x,y)} (W_n W_{n-h}) \right] \leq \frac{2}{n} \sum_{k} \left[ \sum_{h=\gamma+1}^{n-1} \left( 1 - \frac{h}{n} \right) 2^{-2l} \left( 2\alpha(h) \right)^{1-\frac{r}{2}} \left( W_n |^r \right)^{\frac{r}{2}} \left( W_n |^r \right)^{\frac{r}{2}} \right].$$

For all $t = 1, ..., n$ 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

$$\frac{2}{n} \sum_{k} \left[ \sum_{h=\gamma+1}^{n-1} \left( 1 - \frac{h}{n} \right) 2^{-2l} \left( 2\alpha(h) \right)^{1-\frac{r}{2}} \left( W_n |^r \right)^{\frac{r}{2}} \left( W_n |^r \right)^{\frac{r}{2}} \right] \leq \frac{1}{n} 2^{2l} 3^{\frac{r}{2}} \frac{1}{r - 2} \sum_{h=\gamma+1}^{n} \left( 1 - \frac{h}{n} \right) (\alpha(h))^{1-\frac{r}{2}} 2^{-2l} \leq \frac{1}{n} K_r \sum_{h=\gamma+1}^{n} (\alpha(h))^{1-\frac{r}{2}},$$
where \( K_r = 2^{n-\frac{1}{r} - \frac{r}{2}} \).

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

\[
\frac{1}{n} K_r \sum_{h=\gamma+1}^{n} (\alpha(h))^{1-\frac{\gamma}{r}} = \frac{1}{n} K_r O(\gamma^{-1}).
\]

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

\[
\mathbb{E}_{h(x,y)} \left\| \hat{C}_l(u;v) - \mathbb{E}_{h(x,y)} (\hat{C}_l(u;v)) \right\|_2^2 \leq \frac{(M_1)^2}{n} + \frac{2M \gamma}{n} - \frac{2M}{n^2} \sum_{h=1}^{\gamma} h \frac{1}{n} K_r O(\gamma^{-1}) + \frac{1}{n} K_r O(\gamma^{-1}).
\]

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

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

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

So, \( MISE \left( \hat{C}_l(u;v), C(u;v) \right) \) has a minimum when \( l^* \) is such that

\[
2^{l^*} \leq n^{\frac{1}{2(s+2)}} < 2^{l^*+1}.
\]

Then

\[
MISE \left( \hat{C}_{l^*}(u;v), C(u;v) \right) \leq Kn^{-1},
\]

which completes the proof of Theorem 2.2.