A wavelet regression approach for dependence calibration in conditional copula model

Cheikh Tidiane Seck [1] *, *Département de Mathématiques, Université Alioune Diop, Bambey, Senegal*
Aba Diop, Département de Mathématiques, Université Alioune Diop, Bambey, Senegal

Abstract.
In presence of covariates, dependence modeling can be done using conditional copula. In case where the copula function belongs to a given parametric family, an important question is to determine the relationship between the copula dependence parameter and the covariates. In this paper, we propose a wavelet regression approach to estimate this relation often described by the so-called *calibration function*. We discuss asymptotic minimax properties of the linear and non-linear wavelet regression estimators and show their performance via a simulation study. An application to meteorological data reveals that the temperature influences the dependence structure between the maximum and the minimum relative humidity variables, when it takes high values.

Keywords : conditional copula ; calibration function ; copula parameter ; wavelet regression.

1 Introduction

Currently, copulas are widely used for modeling dependence structures of random variables. They have been applied in various domains such as : finance, insurance, survival analysis, meteorology, etc. and a large class of parametric copula models, describing different types of dependence, have been experimented in practice. However, when the dependence structure of a given vector of random variables is influenced by the values of another measured covariate, a conditional copula model seems more convenient to be employed. In this paper, we are interested in estimating non-parametrically the functional relationship between the dependence parameter, say θ, of a given parametric copula family $\mathfrak{C}_\theta$ and some real covariate X. This relationship is often described as

$$\theta(X) = g^{-1}(\eta(X)), \qquad \text{i.e.} \qquad g(\theta(X)) = \eta(X), \tag{1}$$

where g^{-1} is a known inverse link function ensuring that the functional parameter $\theta(\cdot)$ takes its values on the correct range, and η is the so-called *calibration function* which adjusts the level of dependence on the covariate values.

[1]* Corresponding author : Cheikh Tidiane Seck, e-mail: cheikhtidiane.seck@uadb.edu.sn. BP 30 Bambey Senegal

Since the extension of Sklar's theorem [11] for conditional distributions, bringing more flexibility to copulas, dependence modeling via conditional copula has gained an increasing interest among researchers. For instance, dealing with the Clayton copula family, [4] proposed a parametric approach, where the parameter θ is a simple linear function in the covariate X, and utilized the maximum likelihood method to estimate the coefficient in the linear relation. Assuming known marginal distributions, [2] provided a nonparametric approach based on local polynomial techniques to estimate the calibration function η within a local likelihood framework. This approach has been extended by [1] in the context of unspecifed marginal distributions. In the same spirit, [20] proposed a penalized estimation approach that allows parsimonious and enhanced interpretation of the dependence structure of interest. As well, [17] employed cubic splines in a Bayesian context for estimating the calibration function η. And recentlty, [8] proposed a testing methodology for various parametric forms of the calibration function η by letting the marginal distributions unspecified.

In this paper, we propose a wavelet-based estimation approach of the calibration function η, by assuming general margins. Indeed, wavelet series allow parsimonious expansion of various functions. Because of their good localization properties, wavelet bases adapt well to local features of many kinds of functions, including inhomogeneous and discontinuous ones. The approximation properties of wavelet bases are discussed at length in [14]. For more details on the wavelet theory, we refer to [13, 15, 16, 19] and references therein.

The methodology of this paper is based on a regression framework. Precisely, we deal with a fixed design wavelet regression model, where the response variable is defined by the quantity $Z := \eta(X) = g(\theta(X))$, and the predictor is the covariate X, with support $[a, b] \subset \mathbb{R}, a \leq b$. To make our approach possible, we will first partitionne the support $[a, b]$ in finite number of bins and construct, for each bin, an empirical value representing the functional copula parameter $\theta(\cdot)$ in that bin. Then, since the link function g is known, we will apply wavelet denoising techniques on the series of corresponding values for η (obtained as $\eta(\cdot) = g(\theta(\cdot))$), which will be seen as a theoretical signal corrupted by an additive noise.

The paper is organized as follows. Section 2 describes the methodology. After recalling the wavelet expansion on the interval [0,1], we present the wavelet shrinkage method along with Mallat's (1989) pyramidal algorithm. We also discuss in this section asymptotic minimax properties of the linear and nonlinear wavelet shrinkage estimators. In Section 3, we make a simulation study to show the performance of our approach. Section 4 treats an application to a real dataset, while Section 5 concludes the paper.

2 Methodology

2.1 Wavelet expansion on the interval

Without loss of generality we may take $[a, b] = [0, 1]$ and assume that the function η to be estimated belongs to $L^2([0, 1])$, the space of all measurable and square integrable functions defined on $[0,1]$. Thus, we will use wavelet bases on the interval $[0,1]$ to define its expansion. Let ϕ be a scaling function and ψ its associated mother wavelet. Assume that both ϕ and ψ are compactly supported, and define for all integers $j \in \mathbb{Z}$, $k \in \mathbb{Z}$,

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k) \qquad \text{and} \qquad \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k).$$

Let $j_0 \in \mathbb{N}$ be a fixed positive integer, in [5] it is constructed an orthonormal wavelet basis for the space $L^2([0,1])$, with exactly 2^j basis functions at each resolution level $j \geq j_0$. Precisely, the family $\{\phi_{j_0,k} : k = 0, \ldots, 2^{j_0} - 1\} \bigcup \{\psi_{j,k} : j \geq j_0, k = 0, \ldots, 2^j - 1\}$ forms an orthonormal basis for $L^2([0,1])$. Thus, our calibration function η can be decomposed as follows :

$$\eta(x) = \sum_{k=0}^{2^{j_0}-1} \alpha_{j_0,k}\phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty}\sum_{k=0}^{2^j-1} \beta_{j,k}\phi_{j,k}(x), \quad x \in [0,1], \tag{2}$$

where $\alpha_{j_0,k} = \langle \phi_{j_0,k}, \eta \rangle = \int_0^1 \phi_{j_0,k}(x)\eta(x)dx$ and $\beta_{j,k} = \langle \phi_{j_0,k}, \eta \rangle = \int_0^1 \psi_{j,k}(x)\eta(x)dx$ are respectively called the scale coefficients and detail coefficients ; the parameter j_0 is a fixed resolution level.

Note that the orhonormal bases proposed in [5] are boundary adapted. That is, the corresponding wavelet transform automatically handles the boundary effects. There also exist other methods for correcting the boundary bias of wavelet estimators such as : periodization, symmetrization and zero padding.

2.2 A wavelet shrinkage method

Assume that we have n independent and identically distributed observations (Y_{1i}, Y_{2i}, X_i), $i = 1, \ldots, n$ of a random triple (Y_1, Y_2, X). X represents a continuous covariate, with support $[a,b]$, $-\infty < a \leq b < \infty$. (Y_1, Y_2) is a couple of continuous random variables, with marginal distributions F_1 and F_2, the dependence of which is influenced by the covariate X. Let H_X, F_{1X} and F_{2X} denote respectively the joint and conditional marginal distributions given X of the pair (Y_1, Y_2). By Sklar's Theorem (conditional version), for any x in the support of X there is a unique copula C_x such that

$$H_X(y_1, y_2/x) = C_x(F_{1X}(y_1/x), F_{2X}(y_2/x)), \quad y_1, y_2 \in \mathbb{R}. \tag{3}$$

In this paper, we suppose that C_x belongs to a given parametric copula family and its form does not change with x so that we can write $C_x(\cdot,\cdot) \equiv C(\cdot,\cdot,\theta(x))$, where $\theta(x)$ is the dependence parameter satisfying model (1). In contrast, if the conditional margins F_{1X} and F_{2X} are known, we deal with real observations $U_{1i} = F_{1X}(Y_{1i}/X_i)$ and $U_{2i} = F_{2X}(Y_{2i}/X_i)$ from the model

$$(U_{1i}, U_{2i})/X_i \sim C(u_1, u_2; \theta(X_i)), \tag{4}$$

with $\theta(X_i) = g^{-1}(\eta(X_i))$, for $i = 1, \ldots, n$. If F_{1X} and F_{2X} are unknown, we replace U_{1i} and U_{2i} by pseudo-observations as, for example, in [1] :

$$\hat{U}_{1i} = \hat{F}_{1X}(Y_{1i}|X_i); \qquad \text{and} \qquad \hat{U}_{2i} = \hat{F}_{2X}(Y_{2i}|X_i), \quad i = 1, \ldots, n, \tag{5}$$

where for $j = 1, 2$

$$\hat{F}_{jX}(y|x) = \sum_{i=1}^{n} w_{n,i}(x, h)\mathbb{I}(Y_{ji} \le y),$$

with

$$w_{n,i}(x, h) = \frac{K_h(X_i - x)}{\sum_{k=1}^{n} K_h(X_k - x)}, \qquad \text{and} \qquad K_h(\cdot) = \frac{1}{h} K\left(\frac{\cdot}{h}\right).$$

h denotes a bandwidth for controlling the smoothness and $K(\cdot)$ is a symmetric kernel function.

Our aim is to determine the relationship between the copula parameter $\theta(\cdot)$ and the covariate X ; that is to estimate the calibration function η. The definition of the quantity $Z := g(\theta(X)) = \eta(X)$ suggests us using a regression framework. But, we do not have direct observations of the random variable Z, which depends upon the functional copula parameter $\theta(\cdot)$. In order to use the regression setting, we will construct, for the random variable Z, a series of observations based on a suitable finite grid-points in the support $[a, b]$ of the covariate X.

Let $\Delta > 0$ be an arbitrary real number. Let $x_l, l = 1, \ldots, m$ (with m a positive integer) be a finite grid of points defined in such a way that the support $[a, b]$ is partitionned by all the intervals (bins) I_l centered at each point x_l, with radius Δ, i.e.

$$I_l = \{x \in [a, b] : |x - x_l| \le \Delta\}.$$

For each $l = 1, \ldots, m$ define also the bloc of pairewise observations

$$B_l = \{(Y_{1i}, Y_{2i}) : 1 \le i \le n,\ X_i \in I_l\}.$$

To obtain observations for the random quantity Z, which depends on the unknown functional parameter $\theta(\cdot)$, we must approximate the function $\theta(\cdot)$ over each bin I_l centered at x_l. To this end, we first choose Δ small enough in such a way that function $\theta(\cdot)$ is invariant and equal to a real constant θ_l within each bin I_l ; i.e. $\theta(x) = \theta(x_l) =: \theta_l$, for all $x \in I_l$. Then, we estimate each constant parameter θ_l by only using the pairewise observations (Y_{1i}, Y_{2i}) in the corresponding bloc B_l of size n_l. This procedure yields a series of empirical values, say $\hat{\theta}_l, l = 1, \ldots, m$ approximating the functional copula parameter $\theta(\cdot)$, respectively over the different bins $I_l, l = 1, \ldots, m$.

To estimate each parameter $\theta_l, l = 1, \ldots, m$, we make use of Kendall's tau inversion method. We first compute the empirical Kendall's tau version with the observations of the bloc B_l, and then invert the theoretical Kendall's tau formulae (for the considered copula family) to obtain an estimate $\hat{\theta}_l$ for each constant parameter θ_l. Thus, applying the link function g, we get a series of empirical values $Z_l = g(\hat{\theta}_l), l = 1, \ldots, m$, which may be considered as random observations for the response variable $Z = g(\theta(X)) = \eta(X)$.

It is clear that for each l, $\hat{\theta}_l$ is an unbiased estimator of θ_l, because the empirical kendall's tau, $\hat{\tau}_l = 2/C_{n_l}^2 \sum_{i<j} \Delta_{ij} - 1$, with $\Delta_{ij} = \mathbb{I}(Y_{1i} < Y_{1j}, Y_{2i} < Y_{2j}) + \mathbb{I}(Y_{1i} > Y_{1j}, Y_{2i} > Y_{2j})$, obtained for the bloc B_l is an unbiased estimator of the theoretical version for the population from which the bloc B_l is drawn. Since g is known, this implies that $Z_l = g(\hat{\theta}_l)$ is an unbiased estimator of $g(\theta_l)$, i.e.

$$\mathbb{E}Z_l = \mathbb{E}g(\hat{\theta}_l) = g(\theta_l) = g(\theta(x_l)) = \eta(x_l).$$

This allows us to consider the following nonparametric model

$$Z_l = \eta(x_l) + \sigma(x_l)\varepsilon_l, \quad l = 1, \ldots, m, \tag{6}$$

where ε_l is a white noise with zero-mean and unit variance, $\sigma^2(x_l) = \mathrm{Var}(Z_l)$, and $\eta(\cdot)$ represents the *calibration function* that we want to recover here nonparametrically.

To this end, we propose a wavelet shrinkage approach. That is, we will consider the function $\eta(\cdot)$ as a theoretical signal, of which, noisy observations are given by a realisation of the series $Z_l = g(\hat{\theta}_l), l = 1, \ldots, m$; and we denoise these observations by using wavelet transforms. For fixed design models, it is usually assumed, without loss of generality, that the sample (grid) points x_l are within the unit interval $[0, 1]$ and are equidistant ; i.e. $x_l = \frac{l}{m}, l = 1, \ldots, m$. Furthermore, the number of sample points x_l should be a power of two ; i.e. $m = 2^J,\ J \in \mathbb{N}^*$. These assumptions allow to perform both the Discrete Wavelet Transform (DWT) and its inverse (IDWT) using Mallat's (1989) pyramidal algorithm.

Let W be the orthogonal transform matrix associated with a given wavelet basis. Then applying this algorithm yields an approximation $\hat{\eta}$ of the calibration function η after the following steps :

1. Initialize the vector of wavelet coefficients to a sequence of realizations $z = (z_1, \ldots, z_m)$ of $(Z_1, \ldots, Z_m)$;

2. Apply the forward DWT to obtain a vector of wavelet coefficients : $\omega = Wz$;

3. Apply a thresholding function $\delta(\cdot)$ to obtain the estimated coefficients : $\hat{\omega} = \delta(\omega)$;

4. Apply the inverse IDWT to obtain an approximation of the function η over the grid : $\hat{\eta} = W^T\hat{\omega}$,

where W^T represents the transpose of W, and $\hat{\eta} = (\hat{\eta}_1, \ldots, \hat{\eta}_m)$ is a vector of m components approximating the function η over the grid-points, i.e. $\hat{\eta}_l = \widehat{\eta(x_l)}, l = 1, \ldots, m$.

2.3 Asymptotic properties

In this section we discuss asymptotic minimax properties for wavelet shrinkage estimators. We present both the linear and the nonlinear shrinkage rules.

Linear wavelet estimator

Consider the pairewise observations $(x_l, Z_l), l = 1, \ldots, m = 2^J$, with J a given positive integer, from model (6). Natural estimators for the scale coefficients $\alpha_{j_0,k}$ and detail coefficients $\beta_{j,k}$ can be respectively defined as :

$$\widehat{\alpha}_{j_0,k} = \frac{1}{m}\sum_{l=1}^{m} Z_l \phi_{j_0,k}(x_l) \quad \text{and} \quad \widehat{\beta}_{j,k} = \frac{1}{m}\sum_{l=1}^{m} Z_l \psi_{j,k}(x_l).$$

Given a resolution level $j_m = j_0$, the linear shrinkage rule "kills" all the detail coefficients $\{\beta_{j,k} : j \geq j_m, k = 0, \ldots, 2^j - 1\}$ by posing :

$$\widehat{\beta}_{j,k} = \beta_{j,k}\mathbb{I}(j < j_m),$$

where $\mathbb{I}(\cdot)$ designs the indicator function. This leads to the linear wavelet estimator

$$\widehat{\eta}_{j_m}(x) = \sum_{k=0}^{2^{j_m}-1} \widehat{\alpha}_{j_m,k}\phi_{j_m,k}(x),$$

which corresponds to the estimation of the projection of function η onto the sub-space V_{j_m}, element of a multiresolution analysis $(V_j)_{j\in\mathbb{Z}}$ generated by the father wavelet ϕ.

The optimality, in the minimax sense, of the linear wavelet estimator $\widehat{\eta}_{j_m}$ is often investigated over Besov function classes $B^s_{r,q}$ and for L_p-risks. For more details see, e.g., [6], [14]. Under certain regularity conditions including $2^{j_m} \simeq m^{\frac{1}{2s+1}}$, the linear estimator $\widehat{\eta}_{j_m}$ attains the optimal rate of convergence : $m^{\frac{-s}{2s+1}}$. The next theorem gives an upper bound for the L_2-risk over Besov balls $B^s_{r,q}(M)$ of radius $M > 0$. The following hypotheses are needed, with $K(x,y) = \sum_{k\in\mathbb{Z}} \phi(x-k)\phi(y-k)$ the projection kernel of the sub-space V_0.

(H.1) The father wavelet ϕ is bounded, compactly supported and admits $N+1$ derivatives, with N a positive integer.

(H.2) There exists an integrable function F, with $\int |x|^{N+1}F(x)dx < \infty$ such that : $|K(x,y)| \leq F(x-y)$.

(H.3) The kernel K satisfies : for all $y \in \mathbb{R}$, $\int_{-\infty}^{\infty} K(x,y)dx = \delta_{0k},\ \forall k = 0, 1, \ldots, N.$

Theorem 2.1 *[Kerkyacharian and Picard (1992)]*
Assume that hypotheses (H.1-2-3) are satisfied. Let $0 < s < N+1, 2 \leq r \leq \infty, 1 \leq q \leq \infty$. *If the resolution level* j_m *satisfies :* $2^{j_m} = m^{\frac{1}{2s+1}}$, *then*

$$\sup_{\eta\in B^s_{r,q}(M)} \mathbb{E}\|\widehat{\eta}_{j_m} - \eta\|_2^2 \leq Cm^{\frac{-2s}{2s+1}}. \tag{7}$$

Proof. It is the same as that of Theorem 10.2 in [14].

Remark. As for $r \geq p = 2$, the linear minimax risk is optimal, i.e. of order $O(m^{\frac{-2s}{2s+1}})$ (see, Corollary 10.3 in [14]), Theorem 2.1 implies that the linear estimator $\widehat{\eta}_{j_m}$ is optimal whenever the resolution level $j_m = \frac{\log_2(m)}{2s+1}$.

A major drawback of the linear shrinkage rule is that the optimal rate of convergence depends on the reguarity s of the function η which is unknown in practice. The procedure is thus not appropriate, when the function η is spatially inhomogemeous with lower regularity. For such functions one usually relies on nonlinear shrinkage rules.

Nonlinear wavelet estimator

Let j_0, j_m be two integers with $j_m > j_0$; j_0 designs the coarest resolution level. There are two popular ways to define the nonlinear (or thresholding) wavelet estimator : hard-thresholding and soft-thresholding rules. Given a positive threshold $t > 0$, the nonlinear wavalet estimator is generally defined as

$$\widehat{\eta}^*_m(x) = \sum_{k=0}^{2^{j_0}-1} \widehat{\alpha}_{j_0,k}\varphi_{j_0,k}(x) + \sum_{j=j_0}^{j_m-1}\sum_{k=0}^{2^j-1} \gamma^*(\widehat{\beta}_{jk}, t)\psi_{j,k}(x)$$

where $\gamma^*(\cdot, t)$ is a thresholding function, i.e.

$$\gamma^*(x,t) = \begin{cases} sgn(x)(|x|-t)_+ & \text{for soft-thresholding,} \\ x\mathbb{I}(|x|>t) & \text{for hard-thresholding} \end{cases}$$

where $sgn(x)$ designs the sign of x and $x_+ = \max(x, 0)$.
The optimality, in the minimax sense, of the nonlinear wavelet estimator $\widehat{\eta}^*_m$ has been also investigated over Besov function classes and for L_p-risks. The additional hypothesis, compared to the linear case, is the compactness of the function supports. It is proved (see, e.g. [6], [14]) that nonlinear wavelet estimators are near optimal (up to a logarithmic factor) under certain conditions. The next Theorem gives an upper bound for the L_2-risk in a special case (see, Proposition 10.3 in [14]). The following hypotheses are needed :

(G.1): $p = 2, 1 \leq r = q < 2, s > \frac{1}{r}$;

(G.2): $2^{j_0} \simeq m^{\frac{1}{2s+1}}$;

(G.3): $2^{j_m} \geq m^{\alpha/(s-\frac{1}{r}+\frac{1}{2})}$ avec $\alpha = \frac{s}{2s+1}$;

(G.4): $\gamma^*(\widehat{\beta}_{j,k}, t) = \widehat{\beta}_{j,k}\mathbb{I}(|\widehat{\beta}_{j,k}| > t)$ and $t = c\sqrt{\frac{\log(m)}{m}}$.

Theorem 2.2 *[Härdle et al. (1998)]*
Suppose that the father wavelet ϕ and its associated mother wavelet ψ, are both bounded and compactly supported. Let N be a positive integer such that the derivative $\phi^{(N+1)}$ exists and is bounded. If Hypotheses $(G.1-2-3-4)$ are satisfied and $s < N+1$, then there exists a positive constant C, large enough, such that

$$\sup_{\eta \in B^s_{r,r}(M,L)} \mathbb{E}\|\widehat{\eta}^*_m - \eta\|_2^2 \leq C(\log(m))^{\gamma} m^{\frac{-2s}{2s+1}}, \qquad \gamma = 1 - \frac{r}{2}, \tag{8}$$

where $B^s_{r,r}(M,L) = \{f : \|f\|_{srr} \leq M \text{ and } supp(f) \subset [-L, L],\ L > 0\}$.

Proof. It is the same as that of Proposition 10.3 in [14]. A sketch is given in Appendix.

Remark. Whenever hypothesis (G.1) holds, we have by Theorem 10.3 in [14], that the lower bound for the L_2-risk over the Besov class $B^s_{r,r}(M, L)$ is equal to $m^{\frac{-2s}{2s+1}}$. Combining this with statement (8) entails the near optimality of the nonlinear estimator $\widehat{\eta}^*_m$.

3 Simulation study

In this section we evaluate the finite sample performance of our wavelet shrinkage estimators. To this end, we borrow from [2], two specific forms of the calibration function η : a linear specification $\eta(x) = 0.8x - 2$ and a quadratic specification $\eta(x) = 2 - 0.3(x - 2/3)^2$, and compute the integrated square error of the wavelet estimator. We deal with a uniformly distributed covariate X on $[0, 1]$. Data are generated from three parametric copula families : Clayton, Gumbel and Frank according to the following scheme. For each copula family :

- Generate $n \geq 1$ values $x_i, i = 1, \ldots, n$ for the covariate X, from the uniform law $\mathcal{U}(0, 1)$;

- Compute the parameter value corresponding to each x_i : $\theta(x_i) = g^{-1}(\eta(x_i))$, where the link function g^{-1} is given by : $g^{-1}(t) = \exp(t)$ for Clayton, $g^{-1}(t) = \exp(t) + 1$ for Gumbel, and $g^{-1}(t) = t$ for Frank copulas ;

- For each $i = 1, \ldots, n$, generate a pair of values (u_i, v_i) from the considered copula with parameter $\theta(x_i)$;

- Create observations for the couple (Y_1, Y_2) by using Gaussian conditional marginals : $y_{1i} = \Phi^{-1}(u_i)$ and $y_{2i} = \Phi^{-1}(v_i)$ $i = 1, \ldots, n$ where Φ is the standard Gaussian distribution.

The wavelet shrinkage estimator is computed using Mallat's algorithm described in subsection 2.2, where a soft-thresholding is applied with a universal threshold $t = \sigma\sqrt{2 \log m}$. The least asymmetric Daubechies' wavelet with 10 vanishing moments was used. We estimate the noise σ by taking the standard deviation of the observed series $z_l, l = 1, \ldots, m$. All computations are done using the package "WaveThresh" available in R. The integrated square error, approximated by

$$ISE = \frac{1}{m}\sum_{l=1}^{m}(\widehat{\eta(x_l)} - \eta(x_l))^2,$$

is reported after 1000 replications. TABLE 1 shows the results for the linear specification $\eta(x) = 0.8x - 2$, whereas TABLE 2 presents the results for the quadratic one $\eta(x) = 2 - 0.3(x - 2/3)^2$.

Note that since the number of sample points m must be a power of two in the wavelet context, when m varies we multiply the corresponding sample size n by two in order to keep the same scale of proportionality. The results show that our wavelet shrinkage estimator has a good performance, when both m and n increase and the ratio $\frac{m}{n}$ tends to a constant. We also observe that the speed of convergence is faster in the quadratic specification than in the linear specification. Further, One can also see that the perfomance is better in the Clayton and Gumbel cases than in the Frank case. This can be due to the fact that the approximation of the Kendall's tau inversion for Frank copula may not be precise because of the Debye function.

(m,n)	$(8,250)$	$(16,500)$	$(32,1000)$	$(64,2000)$	$(128,4000)$	$(256,8000)$
Clayton	0.2017	0.1288	0.0911	0.0726	0.0670	0.0587
Gumbel	0.1848	0.1292	0.0926	0.0757	0.0632	0.0620
Frank	0.2511	0.2474	0.1277	0.1211	0.1186	0.1173

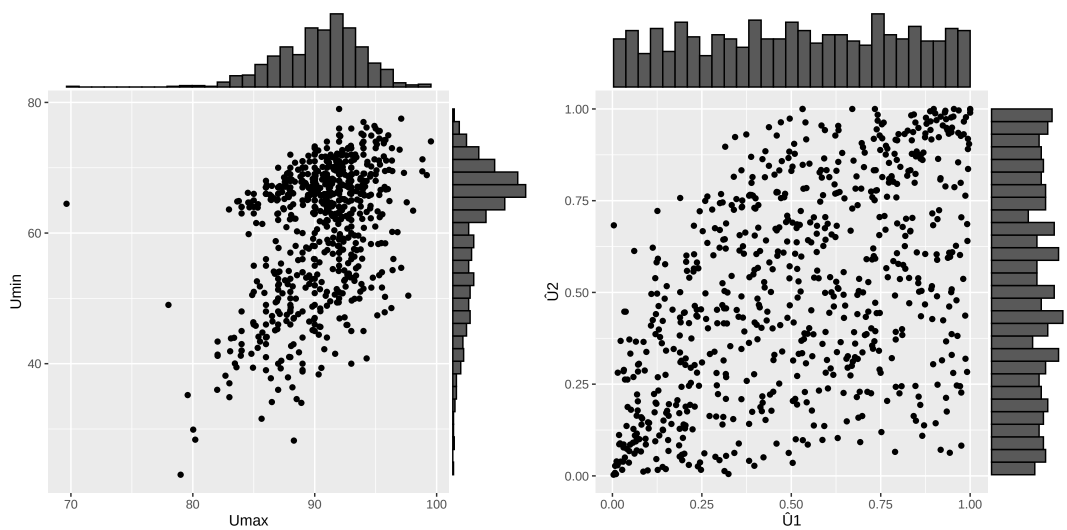
Table 1: *Integrated square error of the wavelet estimator in case of linear specification :* $\eta(x) = 0.8x - 2$.

(m,n)	$(8,250)$	$(16,500)$	$(32,1000)$	$(64,2000)$	$(128,4000)$	$(256,8000)$
Clayton	0.0189	0.0107	0.0066	0.0044	0.0030	0.0027
Gumbel	0.1309	0.0950	0.0773	0.0740	0.0652	0.0666
Frank	0.1410	0.1340	0.1247	0.1238	0.1180	0.1188

Table 2: *Integrated square error of the wavelet estimator in case of quadratic specification :* $\eta(x) = 2 - 0.3(x - 2/3)^2$.

4 Application

In this section we apply our results to meteorological data provided by ANACIM (National Agency for Civil Aviation and Meteorology of Senegal) during the period 1960-2019. The extracted data concern $n = 708$ monthly observations of the variables : maximum relative humidity ($Umax$), minimum relative humidity ($Umin$), expressed in percentage, and maximum temperature ($Tmax$) expressed in celcius degree (°C). We are interested in the influence of the temperature ($Tmax$) on the dependence between the maximum and the minimum relative humidity variables. More precisely, we study the effect of temperature on the strength of the dependence between maximum and minimum relative humidity. Figure 1 shows scatterplots and histograms of the two variables $Umax$ and $Umin$ and their transformed conditional marginals $(\hat{U}_1, \hat{U}_2)$ estimated non parametrically by using formulas defined in (5), with bandwidth $h = 0.1$ chosen arbitrarily. One can see that there is an upper tail dependence between the two humidity variables. Now the question is : Does the temperature have an effect on this dependency ?



(a) (b)

Figure 1: Scatterplots and Histograms of (a) the humidity variables $Umax$, $Umin$ and (b) their transform conditional marginals.

In order to apply the methodology described in Subsection 2.2, where we restrict the values of the covariate in $[0, 1]$, we transform our covariable $Tmax$ into a uniform scale in [0,1] by setting :

$$X = \frac{Tmax - \min(Tmax)}{\max(Tmax) - \min(Tmax)}.$$

We can now apply this methodology. We select $m = 2^{j_m}$ points in the interval $[0, 1]$, with $j_m = 4, 5, 6$, and binwidth $\Delta = 1/2m$. We obtain the results shown in Figure 2, where we represent the Kendall'tau as a function of the temperature (which represents here the covariate X) and for $j_m = 6$. We obtain the sames curves for $j_m = 4$ and $j_m = 5$.

We can observe a same behavior of the Kendall's tau parameter τ for both copulas Clayton and Gumbel. The temperature does not influence the dpendency between the two humidity variables if it takes a scaled value less than 0.7 corresponding to 25°C. But from this value, the temperature produces some fluctuations on the Kendall's tau behavior ; meaning that the temperature influences the dependency between the two humidity variables.

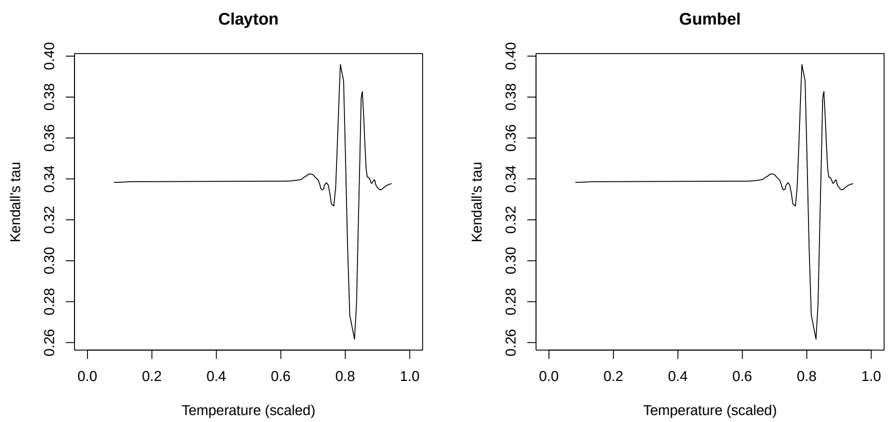


Figure 2: Wavalet estimation of Kendall's tau τ for Clayton and Gumbel copulas : $j_m = 6$.

5 Conclusion

We have applied a wavelet-based regression approach to estimate the relationship between the copula dependence parameter and some covariate in the framework of a conditional copula model. This approach presents some advantages such as the fast computation of the wavelet estimators and their near optimality over a wide classe of regular functions and for a large range of L_p-risks. Our results have been applied for a dataset to analyse the influence of the temperature on the dependence structure between maximum and minimum relative humidity variables. We find that, relatively to this dataset, the temperature influences the dependence structure of these two humidity variables, only when it takes larger values greater than for example 25°C. Using a general likelihood ratio test to assess the significativity of this influence could be interesting issue for the future. As well, it might be interesting to compare our approach to that of [2], who utilized a local linear estimation approach for the calibration function η.

References

[1] F. Abegaz, I. Gijbels, N. Veraverbeke, Semiparametric estimation of conditional copulas. *Journal of Multivariate Analysis*,110 (2012), 43–73.

[2] E. F. Acar, V. R. Craiu, F.Yao, Dependence calibration in conditional copulas : a non-parametric approach, *Biometrics*, 67 (2011), 445–453.

[3] E.F. Acar, R.V. Craiu, F. Yao, Statistical testing of covariate effects in conditional copula models, Electron. J. Stat. 7 (2013) 2822–2850.

[4] M. Craiu (2009), Paramatric estimation of copulas.*U.P.B. Sci. Bull., Series A,* Vol. 71, Iss. 3, 2009.

[5] A. Cohen, I. Daubechies, P. Vial (1993), Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, Elsevier.

[6] D.L. Donoho, I. Johnstone, G. Kerkyacharian, D. Picard (1996), Density estimation by wavelet thresholding. *The Annals of Statistics*, Vol. 24, No. 2, 508-539.

[7] J. Fan, C. Zhang, and J. Zhang, (2001). Generalized likelihood ratio statistics and wilks phenomenon. *Annals of Statistics*, 29(1):153–193. MR1833962

[8] I.Gijbels, N.Veraverbeke, and M.Omelka, Conditional copulas, association measures and their applications, *Comput. Statist. Data Anal.*,vol.55, no. 5, pp. 1919–1932, 2011.

[9] I. Gijbels, M. Omelka, M. Pešta, N. Veraverbeke, Score tests for covariate effects in conditional copulas, *Journal of Multivariate Analysis*, 159 (2017), 111–133.

[10] Nelsen, R. B. (2006), *An introduction to copulas.* Springer, New York. Second Edition.

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

[12] N. Veraverbecke, M. Omelko, I. Gijbels, Estimation of a conditional copula and association measures. *Scandinavian J. Statist.* 38, (2011) 766–780.

[13] Daubechies, I. (1992). *Ten Lecture On Wavelets*. In : CBMS-NSF Regional Conference series in Applied Mathematics, vol.61. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

[14] Härdle, W. Kerkyacharian, G. Picard, D. Tsybakov, A. (1998). Wavelets, Approximation and Statistical Application. Springer-Verlag, New-York.

[15] Mallat, S. (1989). *A theory for multiresolution signal decomposition : The Wavelet Representation*. IEE. Transformation on Pattern Analysis and Machine Intelligence 11, 674-693.

[16] Meyer, Y. (1992). *Wavelets and Operators*. In : Cambridge Studies in Advenced Mathematics, vol.37. Cambridge University Press, Cambridge.

[17] A. Sabeti, Bayesian Inference for Bivariate Conditional Copula Models with Continuous or Mixed Outcomes. Ph.D. Thesis, 2013. Department of Statistical Sciences, University of Toronto.

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

[19] Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. Institute of Statistics and Decision Science.

[20] J. Zou (2015), Nonparametric Methods for Interpretable Copula Calibration and Sparse Functional Classification. Ph. D. Thesis, 2015. Graduate Department of Statistical Science, University of Toronto