Independence or … copulas! Capturing the dependence among large losses using *extreme-value copulas*

THIAGO DUTRA DE ARAUJO
Universidade de São Paulo
JOÃO VINÍCIUS DE FRANÇA CARVALHO
Universidade de São Paulo

Abstract

There are events that, when they occur, generate several losses simultaneously, leading an insurance company that operates in different lines of business to make multiple payments of indemnity. Thus, the premise of independence among risks, typical in the pricing of these events, must be constantly questioned. One possibility of dependence modeling among losses is the use of copulas, a technique that has been extensively used since the 1990s. However, the empirical literature about this is limited, and articles using real-world microdata are even rarer. The purpose of this article is to estimate the dependence structure among large losses from a single event that generated multiple claims in different lines of business using *extreme-value copulas* (EVC) applying to real-world insurance data and comparing to other families of copulas. The study is divided into two parts: (i) adjustment of dependence on simulated data, whose dependence structure is known, and (ii) capture of the dependence structure in real-world dataset of 13,734 losses belonging to different lines of business, incurred by a single event, whose dependence is unknown. Overall, the non-parametric method performed better than the parametric ones, producing more consistent estimates. In several real situations, EVC were more adequate to capture the dependence among large losses than elliptical or Archimedean copulas, for motor insurance and for large risks. Still, there are important differences in the dependence structures when the sum of the losses is evaluated, even if each part is not a large loss. Thus, the regulator must be aware of this fact when dimensioning the minimum capital requirement.

Keywords: extreme-value copulas; dependence among lines of business; tail dependence; large losses; multiple claims

1. Introduction

In the History of Brazil's independence process, there was a famous speech attributed to D. Pedro I: "*Independence or death*." However, if D. Pedro I was a risk manager in an insurance company, probably the alternative to independence among events wouldn´t be death. At that time, a (rudimentary) notion of dependence among events was already present in the Brazilian Commercial Code in the year 1850 (Lopes, Peris, Chan, & Borelli, 2015). Three points established in that Code highlight this aspect: (i) life insurance for free people was prohibited from being offered together with maritime insurance (at the time, slaves were understood as property); (ii) it was possible to contract different insurance policies for the same objects, under the same or different conditions, and; (iii) it was allowed to contract insurance policies for the ship (hull), its freight (transport) and farms (physical property) in the same policy, provided that the value of each object was clearly determined[i]. In all these points, there was a natural mechanism of dependence among risks, policies or lines of business.

With that, it is noticed – since that time – a special attention to the occurrence of extreme events: events (the so-called '*sea fortunes*') that have low probability of occurrence and whose amount to be indemnified to the insured is a very high value – the severity of the claim is a matter of concern in several articles of that Commercial Code. The occurrence of this type of event is very harmful to the insurance company to the point of even leading to bankruptcy.

To reduce exposure to the bankruptcy risk, the insurance industry has sought to improve its pricing models for claims incurred for this type of event. *Solvency II*[ii] presents the most recent guidelines on insurance risk management, establishing the minimum required capital for insurance companies so that there is a 99.5% probability of guaranteeing the operational continuity for the next 12 months. Another good reason for this work is the increase in the frequency and severity of claims incurred by natural disasters; such phenomena cost billions of dollars to insurance companies (Dietz & Walker, 2019; Bokusheva, 2018).

As an example, a study[iii] showed that typhoon Hagibis, which occurred in Japan in 2019, generated US$ 15 billion in financial losses, of which US$ 9 billion was insured. For sake of comparison, in that same year, the total premiums collected by insurance companies in Brazil were around US$ 28 billion[iv]; that is, if it had occurred in Brazil, this single event would correspond to 32% of the total premiums collected in the year. Another relevant example is that of the Covid-19 pandemic, which occurred in 2020: preliminary reports[v] showed that this event generated claims in several insurance lines (*health insurance*, *life insurance*, among others). If the same insurance company operates in several lines, an unexpected increase in claims can lead to bankruptcy, showing the importance of calculating good measures of dependence among risks in different lines of business.

When analyzing the occurrence of claims generated by extreme events, it is generally not convenient to consider that they are independent (Chukwudum, 2019). Thus, in order to price this type of event, two aspects are usually considered: the *severity* and *dependence* of the incurred claims (Kley, Klüppelberg, & Paterlini, 2020; Sweeting & Fotiou, 2013). There are several ways to model the dependence among large losses: simulations, stress testing, use of copulas, among others (Shen, 2019). In particular, the use of copulas has become relevant in modeling dependence among events in the pricing of claims (Bücher, Irresberger, & Weiss, 2017). In the case of extreme events, the use and choice of copula is an important part of this modeling process: several copulas sufficiently satisfactorily capture the dependence among intermediate values, but do not do so for extreme values (Furman, Kuznetsov, Su, & Zitikis, 2016).

Therefore, this article seeks to analyze both the choice of copula and the method of estimating parameters that are most appropriate to the dependence among large losses incurred by the same generating event. With this, we expect to reduce the scarcity of publications that approach this theme empirically, making use of real-world microdata, since much of the related international literature discusses theoretical methodologies, however little applied.

2. Literature review and related studies

Measuring the dependence among two variables is not exclusive, nor did it originate in Actuarial Science: since the end of the 19th century, Sir Francis Galton and Karl Pearson had already introduced the concept of correlation between variables, which is very relevant in the study of regressions (Frees & Valdez, 1998). For years, the linear correlation coefficient has been used as a measure of dependence among variables, being applied, for example, in calculating risks linked to asset portfolios, measuring the dependence among the returns of these assets (Szego, 2002).

However, from the second half of the twentieth century, some statisticians began to study other measures of dependence. In 1951, Maurice Fréchet proposed the following problem: "given the marginal distributions of a set of random variables, what can be said about the joint distribution of these variables?" (Fréchet, 1951). In 1959, Abe Sklar sent a letter to Fréchet containing some results on this problem and, in it, introduced the word **copula** in the statistical universe, based on the linguistic use of this word (term that links the subject of a sentence to the subject complement): a function that links a joint distribution of variables to their marginal distributions (Sklar, 1959).

In the late 1970s and early 1980s, other problems motivated the deepening of the study of dependence among variables, such as, for example, the construction of survival tables for joint lives, under the hypothesis that individuals' lifetimes aren't independent of each other (Clayton, 1978). However, it was in the 1990s that the use of copulas was widespread, largely to measure dependences among risks incurred in finance and insurance (Embrechts, 2009).

Also, in this decade, new quantitative risk management methodologies were developed, due to the emergence of new products and the implementation of new regulatory guidelines in the insurance market. In these methodologies, limitations were already noted in the use of the linear correlation coefficient to measure dependencies. Embrechts (2009) cites two of them: cases in which a low correlation had extremely dependence in extreme values and a false security generated by the linear correlation coefficient, allowing companies to be exposed to risks greater than expected.

At the end of the 1990s, the first applications of copulas in insurance pricing emerged. At this time, were also described some copula families and their use as an integral part of risk measurement (Frees & Valdez, 1998). Among the copulas described, one that became quite relevant was the *Gaussian copula*, as it was well suited to several sets of variables. However, this excessive confidence of managers in modeling risk dependence using only Gaussian copulas brought numerous financial setbacks to several companies, allowing other copula families to be developed (Embrechts, 2009; Nelsen, 2006).

Thus, several studies have sought to identify which copula is the most appropriate to measure dependencies in risk management. For example, Su & Furman (2017) proposed a method of choosing from *copulas for multiple risk factors* claiming that, in several cases of financial risk management, the choice of copula is based on the ease of its analytical treatment instead of observing all the characteristics of the problem in question.

In *Life insurance*, there are many applications with the use of copulas in modeling the dependence among mortalities of various populations (Giussani & Bonetti, 2019; Chen, MacMinn, & Sun, 2017; Zhu, Tan, & Wang, 2017). According to the authors, there is a need for improvements in the modeling of this type of dependence due to reforms of pension systems, creation of new types of pension funds and implementation of new products in life insurance. In *Non-Life insurance*, recent applications used copulas to price cybersecurity insurance, measuring the dependence among time and speed of infection on a network (Xu & Hua, 2019). Other studies have valued the use of copulas especially in cases where the frequency and severity of claims have different dependence structure within the same dataset (Hua, 2017).

From the 2000s, several works began to use copulas to model the dependence among a particular type of events: extreme events (Szego, 2002). As these events are located on the tails of the distributions, this modeling was known in the literature as *tail dependence*. Since then, several studies have sought to analyze tail dependence measures, such as the upper tail dependence parameter (Li, 2018) and the maximum tail dependence index (Sun, Yang, & Zitikis, 2020).

Strong dependencies on the upper tail of claims severity distributions show that excessive financial losses tend to occur jointly, and not independently (Brechmann, Czado, & Paterlini, 2014). As examples, we can mention: (i) the attack on the World Trade Center, in 2001, which generated large financial losses from insurance companies due to underwriting risks and the subprime crisis in 2008, in which insurance companies suffered large losses in backed investments in collateralized mortgages and debentures (Eling & Toplek, 2009), and; (ii) the dependence among extreme events that involved two lines of business in the Nigeria insurance market: auto insurance and fire insurance. In this case, this dependence occurred when vehicles carrying flammable products collided with other vehicles, causing explosions and multiple accidents (Chukwudum, 2019).

To model dependence on the tails of distribution, there is a copula family called *extreme-value copulas*, widely used by rating agencies and regulators, as it reveals major differences in the assessment of the insurer's risk of deficit and the ruin probabilities, compared to *elliptical copulas* (Eling & Toplek, 2009; Kelliher et al., 2020). In addition, there are also applications using this copula family in modeling dependence among returns from different financial time series and asset classes (Bormann & Schienle, 2018; Ruenzi, Ungeheuer, & Weigert, 2020).

The theoretical framework of extreme-value copulas is still under development from (at least) two perspectives: (i) the comparison between these and the *Archimedean copulas*, and; (ii) the study of parametric and non-parametric forms that estimate the dependence function of these copulas. Pappadà, Perrone, Durante, & Salvadori (2016) used copulas that are both Archimedean and extreme-value copula to compare, based on their dependence functions, which family is the most appropriate to measure the dependence among structural risks arising from natural disasters. Kamnitui, Genest, Jaworski, & Trutschnig (2019) analyzed different methodologies that estimated the dependence function of these copulas, given a certain dependence level.

In Brazil, there are few studies that mention the use of copulas and, among them, even rarer that model tail dependence. Regarding the use of copulas, Melo (2008) used a copula family to model the dependence among two insurance lines of business (residential and business), known as comprehensive insurance, whose premium calculation is strongly impacted if assumed the hypothesis of independence among that losses. Tanaka & Carvalho (2019) used the elliptical and Archimedean copula families to re-estimate the dependence structure among the Brazilian insurance lines of business, calculating solvency capital based in copulas, which were compared with the results obtained using SUSEP's (Brazilian insurance regulatory agency) methodology, set out in Annex III of CNSP Resolution 321/15[vi]. The results showed that the SUSEP methodology measures, in *personal insurances*, an insufficient level of solvency capital in times of severe crises, exposing insurance companies to the risk of bankruptcy.

Thus, this study will explore two gaps: (i) the use of extreme-value copulas as a possibility of choosing the most ranked copula for real-world microdata from an insurance company, and; (ii) different methodologies for estimating copula parameters and measures of association. First, these gaps will be explored in controlled experiments, with simulated data. Then, we will use real-world microdata to verify the conformity of the adjustment in relation to the simulated data.

3. Methodology

Initially, we present general concepts related to copulas. Next, we describe some copulas families: elliptical, Archimedean and extreme-value copulas. Finally, we list some measures of association and the estimation methods used in this study. All mathematical formulations in this section follow closely Nelsen (2006).

Let X and Y a pair of random variables, with distribution functions $F_X(x) = \mathbb{P}[X \leq x]$ and $F_Y(y) = \mathbb{P}[Y \leq y]$, respectively, and $F(x, y) = \mathbb{P}[X \leq x, Y \leq y]$ the joint probability distribution. Denoting $u = F_X(x)$ and $v = F_Y(y)$, a **copula** is a function $C: [0,1]^2 \rightarrow [0,1]$ with the following property

$$F(x, y) = C(u, v), \tag{1}$$

that is, C is a function that associates the joint distribution of random variables to their distributions. This result – known as *Sklar´s theorem* – is central to the study of copulas. In addition, Sklar demonstrated that if F_X and F_Y are continuous, then C is unique.

As a corollary follows that, being $F_X^{-1}(u) = \sup\{x | F_X(x) \leq u\}$, then

$$C(u, v) = F\big(F_X^{-1}(u), F_Y^{-1}(v)\big). \tag{2}$$

In the next subsections, we describe the formulas of some copulas families.

3.1. Elliptical copulas and Archimedean copulas

Elliptical copulas are copulas associated with elliptical distributions, such as Normal and t-Student distributions (both symmetrical). The two most relevant of elliptical copulas are Gaussian and t-Student, whose formulas are shown in Table 1.

Table 1 – Bivariate elliptical copulas formula

Copula	Formula
Gaussian	$C_\theta^{Ga}(u, v) = \Phi_R^2\big(\Phi^{-1}(u), \Phi^{-1}(v)\big)$
t-Student	$C_{n,\theta}^t(u, v) = t_{n,R}^2\big(t_n^{-1}(u), t_n^{-1}(v)\big)$

Source: own elaboration.

In the Gaussian copula, Φ^{-1} is the inverse of the standard Normal distribution function, Φ^2 is the standard Normal distribution for the bivariate case and R is the correlation matrix given by $r_{11} = r_{22} = 1$ and $r_{12} = r_{21} = \theta$, where θ is the linear correlation coefficient. In the t-Student copula, t_n^{-1} is the inverse of the standard t-Student distribution function and t_n^2 is the standard t-Student distribution for the bivariate case, both with n degrees of freedom. R is the same correlation matrix defined before.

It is relevant to note that neither of these two copulas has an explicit functional form. In addition, similarly to the multivariate Normal distribution, the Gaussian copula is symmetrical and light tailed, failing to capture the tail dependence. Therefore, it isn´t suitable for dependence modeling among losses incurred by extreme events.

Archimedean copulas – widely used in actuarial science and finance (Cossette, Marceau, Mtalai, & Veilleux, 2018) – allow a wide variety of dependence structures to be modeled, in addition to have an explicit functional form. Some Archimedean copulas are listed in Table 2.

Table 2 – Bivariate Archimedean copulas explicit formula

Copula	Explicit formula
Clayton	$C_\theta(u,v) = \max\left(\left[u^{-\theta} + v^{-\theta} - 1\right]^{-\frac{1}{\theta}}, 0\right)$
Frank	$C_\theta(u,v) = -\frac{1}{\theta} \cdot \ln\left(1 + \frac{\left(e^{-u\theta} - 1\right)\left(e^{-v\theta} - 1\right)}{e^{-\theta} - 1}\right)$
Joe	$C_\theta(u,v) = 1 - \left[(1-u)^\theta + (1-v)^\theta - (1-u)^\theta \cdot (1-v)^\theta\right]^{\frac{1}{\theta}}$

Source: own elaboration.

3.2. Extreme-value copulas

A copula C_* is an extreme-value copula if, and only if, there exist a copula C such that

$$C_*(u,v) = \lim_{n\to\infty} C^n\left(u^{\frac{1}{n}}, v^{\frac{1}{n}}\right). \tag{3}$$

Several authors have presented procedures for obtaining extreme-value copulas. Pickands (1981) demonstrated that if C is an extreme-value copula, then C can be written as

$$C(u,v) = \exp\left\{ln(u \cdot v) \cdot A\left(\frac{ln(v)}{ln(u \cdot v)}\right)\right\}, \tag{4}$$

for an appropriate choice of the function $A(.)$, called *dependence function* of the extreme-value copula C. This function must be convex and satisfy the following conditions:

- $A(0) = A(1) = 1;$
- $\max\{t, 1-t\} \leq A(t) \leq 1;$

Table 3 presents a list of extreme-value copulas.

Table 3 – Extreme-value copulas dependence function and explicit formula

Copula	$A(t)$	Explicit formula
Tawn	$\theta t^2 - \theta t + 1$	$C_\theta(u,v) = u \cdot v \cdot exp\left\{\theta \cdot \frac{ln(u) \cdot ln(v)}{ln(u) + ln(v)}\right\}$
Galambos	$1 - \left(t^{-\theta} + (1-t)^{-\theta}\right)^{-\frac{1}{\theta}}$	$C_\theta(u,v) = u \cdot v \cdot \exp\left\{\left[(-log(u))^{-\theta} + (-log(v))^{-\theta}\right]^{-\frac{1}{\theta}}\right\}$
Gumbel-Hougaard	$\left(t^\theta + (1-t)^\theta\right)^{\frac{1}{\theta}}$	$C_\theta(u,v) = \exp\left\{-\left[(-ln(u))^\theta + (-ln(v))^\theta\right]^{\frac{1}{\theta}}\right\}$
Husler-Reiss	$A_{HR}(t)$	$C_\theta(u,v) = C_{HR,\theta}(u,v)$

Source: own elaboration.

For the Husler-Reiss copula:

$$A_{HR}(t) = t \cdot \Phi\left(\theta^{-1} + \frac{\theta}{2} \cdot \ln\left(\frac{t}{1-t}\right)\right) + (1-t) \cdot \Phi\left(\theta^{-1} - \frac{\theta}{2} \cdot \ln\left(\frac{t}{1-t}\right)\right), \tag{5}$$

and

$$C_{HR,\theta}(u,v) = \exp\left\{-ln(u) \cdot \Phi\left(\theta^{-1} + \frac{\theta}{2} \cdot ln\left(\frac{ln(u)}{ln(v)}\right)\right) + ln(v) \cdot \Phi\left(\theta^{-1} + \frac{\theta}{2} \cdot ln\left(\frac{ln(v)}{ln(u)}\right)\right)\right\}. \tag{6}$$

3.3. Measures of association

The linear correlation coefficient has several limitations when measuring dependences (Embrechts, 2009). Thus, we present coefficients that seek to overcome such limitations.

Before presenting them, it is necessary to introduce the concept of *concordance*. Two distinct pairs (x_1, y_1) and (x_2, y_2) are concordant if $x_1 < x_2$ and $y_1 < y_2$ or if $x_1 > x_2$ and $y_1 > y_2$, that is, if the smallest value of x is associated to the smallest value of y (or, similarly, the largest value of x is associated to the largest value of y). Note that this implication is equivalent to write $(x_1 - x_2)(y_1 - y_2) > 0$. If $(x_1 - x_2)(y_1 - y_2) < 0$, the pairs are discordant.

Given this concept, consider two random variables X and Y, whose distributions F_X and F_Y are continuous, and variables X' and Y', independent copies of X and Y, respectively. The sample version of *Kendall´s* τ is given by

$$\tau = \tau_{X,Y} = \mathbb{P}[(X - X')(Y - Y') > 0] - \mathbb{P}[(X - X')(Y - Y') < 0]. \tag{7}$$

If X and Y are continuous, it is also possible to write τ as a function of copula C associated to both X and Y:

$$\tau = -1 + 4 \cdot \int_{[0,1]^2} C(u,v)\, dC(u,v). \tag{8}$$

In addition to this, *Spearman´s* ρ is defined as

$$\rho = \rho_{X,Y} = Cor\left(F_X(X), F_Y(Y)\right) = \frac{\mathbb{E}[F_X(X) \cdot F_Y(Y)] - \mathbb{E}[F_X(X)] \cdot \mathbb{E}[F_Y(Y)]}{\sqrt{\mathbb{V}ar[F_X(X)] \cdot \mathbb{V}ar[F_Y(Y)]}}. \tag{9}$$

If X and Y are continuous, also is possible to write ρ as a function of copula C associated to X and Y:

$$\rho = -3 + 12 \cdot \int_{[0,1]^2} C(u,v)\, dudv \tag{10}$$

Table 4 shows the relationship between τ, ρ and the copulas parameters for some of the copulas presented. The fields marked with * indicate that, although it is not possible to write this relationship analytically, it is possible to estimate these coefficients, as will be detailed in subsection 3.4.

Table 4 – Relationship between τ, ρ and the copulas' parameters

Copula	$\boldsymbol{\tau}$	$\boldsymbol{\rho}$
Gaussian	$\tau(\theta) = \frac{2}{\pi} \cdot arcsen(\theta)$	$\rho(\theta) = \frac{6}{\pi} \cdot arcsen\left(\frac{\theta}{2}\right)$
Clayton	$\tau(\theta) = \frac{\theta}{\theta + 2}$	*
Frank	$\tau(\theta) = 1 - \frac{4}{\theta} + \frac{4}{\theta^2} \int_0^\theta \frac{t}{e^t - 1} dt$	$\rho(\theta) = 1 - \frac{12}{\theta^3} \int_0^\theta \frac{\theta t - 2t^2}{e^t - 1} dt$
Joe	$\tau(\theta) = 1 + \frac{4}{\theta} \int_0^1 \frac{\ln\left(1 - t^\theta\right) \cdot \left(1 - t^\theta\right)}{t^{\theta - 1}} dt$	*

Source: own elaboration.

For extreme-values copulas, it is possible to write τ and ρ as a function of dependence function $A(.)$, using the equations (4), (6) and (8):

$$\tau = \int_0^1 \frac{t \cdot (1-t)}{A(t)} dA'(t) \quad and \quad \rho = -3 + 12 \cdot \int_0^1 \frac{1}{\left(1 + A(t)\right)^2} dt. \tag{11}$$

According to Li (2018), another important measure of association is the tail dependence parameters, used to measure the dependence on the upper and lower tails (that is, when both distributions assume very large or very small values). The *upper tail dependence parameter* λ_U and the *lower tail dependence parameter* λ_L are given by (if the limits exist):

$$\lambda_U = \lim_{k \to 1^-} \frac{1 - 2k + C(k,k)}{1-k} \qquad and \qquad \lambda_L = \lim_{k \to 0^+} \frac{C(k,k)}{k}. \tag{12}$$

It is interesting to note that the quantity $C(k,k)$ indicates the proportion of information that is in the lower left corner of the square $[0,1]^2$, limited superiorly by $F_X(x) = k$ and $F_Y(y) = k$. Therefore, if these variables are perfectly correlated, $C(k,k) = k$ (the maximum value of the copula) and, in this case, $\lambda_U = \lambda_L = 1$. As the values of these tail dependence parameters belong to the range [0,1], the value 0 indicates complete absence of dependence and the value 1, complete dependence. Table 5 shows these parameters for elliptical and Archimedean copulas.

Table 5 – Tail dependence parameters for elliptical and Archimedean copulas

Copula	$\boldsymbol{\lambda_L}$	$\boldsymbol{\lambda_U}$
Gaussian	0	0
t-Student	$2t_n\left(-\sqrt{\frac{(n+1)(1-\rho)}{1+\rho}}\right)$	$2t_n\left(-\sqrt{\frac{(n+1)(1-\rho)}{1+\rho}}\right)$
Clayton	0	$2-2^{\frac{1}{\theta}}$
Frank	0	0
Joe	0	$2-2^{\frac{1}{\theta}}$

Source: own elaboration.

Next, we discuss some methods for estimating copula parameters and measures of association presented.

3.4. Estimation methods

There are several methods that estimate copula parameters, as well as measures of association: some parametric, others non-parametric. In this study, we present three of them: the *method-of-moments* (MME), the *method of maximum pseudo-likelihood* (MPLE) – both parametric – and the *Capéraà-Fougères-Genest method* (MCFG), a non-parametric way of estimating the dependence function $A(.)$ for extreme-value copulas.

By the MME, the copula parameters are estimated from the inverse of the function $\tau(.)$, described in Table 4. To do so, it is necessary estimate τ initially. Thus, consider two random variables X and Y and n observed pairs (x_1, y_1), (x_2, y_2), ..., (x_n, y_n), with $x_i \in X$ and $y_i \in Y$, $1 \leq i \leq n$. The estimator $\hat{\tau}_{X,Y}$ is given by

$$\hat{\tau}_{X,Y} = \frac{2}{n(n-1)} \cdot \sum_{1 \leq i < j \leq n} \{\mathbb{I}[(x_i - x_j)(y_i - y_j) > 0] - \mathbb{I}[(x_i - x_j)(y_i - y_j) < 0]\}, \quad (13)$$

where $\mathbb{I}(t)$ is the indicator function: $\mathbb{I}(t) = 1$, if t is true, or $\mathbb{I}(t) = 0$, otherwise.

In the MPLE, from the n observed pairs (x_1, y_1), (x_2, y_2), ... and (x_n, y_n) is created a sample of pseudo-observations $U_{i,n} = \left(\frac{r_{iX}}{n+1}, \frac{r_{iY}}{n+1}\right)$, where r_{iX} and r_{iY} indicate, respectively, the *ranking* (the position) of x_i and y_i in the increasing sequences formed by $\{x_1, x_2, \ldots, x_n\}$ and $\{y_1, y_2, \ldots, y_n\}$, to estimate copula parameters using the maximum log-likelihood. This means that, for uniparametric copulas, the estimator $\hat{\theta}$ is the value of θ that maximizes the function

$$f(\theta) = \sum_{i=1}^{n} \log[c_\theta(U_{i,n})]. \quad (14)$$

Copulas whose parameters can be estimated via MME can also be estimated via MPLE. However, the MPLE is more efficient than the MME, especially in the case of finite samples or multiparametric copulas (Hofert, Kojadinovic, Mächler, & Yan, 2018).

Finally, the MCFG estimates the extreme-value copulas parameters from the dependence function $A(t)$ in Table 3. For this, consider the function

$$\zeta_{i,n}(t) = \min_{j\in\{X,Y\}} \frac{-\log(U_{ij,n})}{t}\ , 1 \le i \le n \tag{15}$$

where $U_{ij,n}$ is the first ($j = 1$) or second ($j = 2$) coordinate from $U_{i,n}$. The initial estimator $\hat{A}_{CFG}(t)$ was corrected by Gudendorf & Segers (2012), in order to improve its goodness of fit to small samples, which is quite frequent in extreme events. The corrected estimator is given by:

$$\widehat{A'}_{CFG}(t) = \exp\left[\frac{1}{n}\sum_{i=1}^{n}\left(\zeta_{i,n}(1) - \zeta_{i,n}(t)\right)\right] \tag{16}$$

One way to evaluate the copula model selection by fitting a copula family to a data set is the criteria namely the *cross-validation copula information criterion* (CIC), according to which, the higher the CIC, most suitable is the model (Grønneberg & Hjort, 2014). This criterion is equivalent (except for a multiplicative constant) to the number given by

$$\widehat{xv}_n = \frac{1}{n}\cdot\sum_{i=1}^{n}\log\left[c_{\theta_{n,-i}}\left(F_{n,1,-i}(x_i), F_{n,2,-i}(y_i)\right)\right] \tag{17}$$

where $\theta_{n,-i}$ is the estimator to the copula parameter and

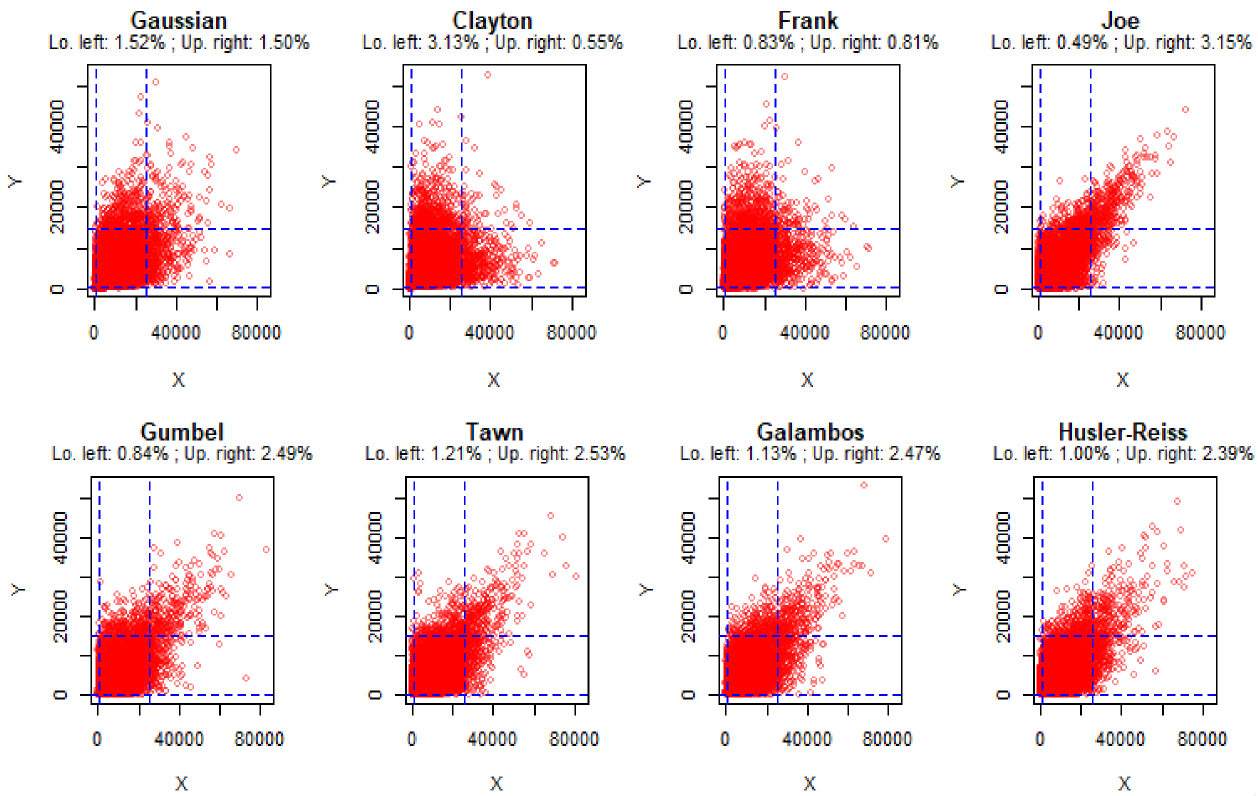
$$F_{n,j,-i}(x) = \begin{cases} \frac{1}{n}\cdot\sum_{k=1,k\neq i}^{n}\mathbb{I}\left(x_{kj} \le x\right), \text{if } x \ge \min_{k\in\{1,\dots,n\}-\{i\}} x_{kj} \\ \frac{1}{n}, \qquad\qquad \text{otherwise} \end{cases} \tag{18}$$

4. Results

4.1. Evaluating the effectiveness of the methods: copula estimation using simulated data

When defining a dependence structure, it is not enough to choose only the value of $\hat{\tau}_{X,Y}$. Figure 1 shows that, for a given value of $\hat{\tau}_{X,Y}$ (in the panels in Figure 1, we fixed $\hat{\tau}_{X,Y} = 0.4$), there are very different dependence structures, resulting from the copula family used – according to Kamnitui et al. (2019).

Figure 1 – 10,000 random simulations from each copula family, all of them with $\hat{\tau}_{X,Y} = 0.4$



Source: own elaboration.

In the subtitles, the percentages shown below the copula name represent the frequency of pairs in the *lower left* and *upper right* rectangles of the joint distribution (regions known as lower and upper tail, respectively). These regions are highlighted by the horizontal and vertical dashed lines corresponding to the quantile indices of 5% and 95% of each of the marginal's distributions.

Figure 1 shows the symmetry of the lower and upper tails in Gaussian copula; the capture of the dependence on the lower tail, by the Clayton copula; and the capture of the dependence on the upper tail, by extreme-value and Joe copulas. This reinforces that the choice of copula family is the first step in copula dependence modeling, as this choice can substantially affect the underlying dependence structure.

To evaluate the effectiveness of the methods in adequately capturing the real dependence structure, fixed and known dependence structures were imposed from simulated data (before using such methods in real-world microdata, in which the dependence structure is typically unknown). For this, three scenarios were defined: the first (called 'weak') has 10,000 pairs (X, Y) simulated from a Gaussian copula with $\hat{\tau}_{X,Y} = 0.1$; the second ('moderate') has 10,000 pairs simulated from a Joe copula, with $\hat{\tau}_{X,Y} = 0.4$; and the third ('strong'), has 100 pairs simulated from a Husler-Reiss copula, with $\hat{\tau}_{X,Y} = 0.7$. In the three scenarios, were adopted $X \sim Exp(8500^{-1})$ and $Y \sim Exp(5000^{-1})$, due to the fact that Exponencial is often used in severity claim modeling for typical insurance lines of business.

Table 6 presents the CIC for the three scenarios based on the copula families presented, showing that the best ranked family for these scenarios are, indeed, Gaussian, Joe and Husler-Reiss, respectively.

Table 6 – CIC for copulas families, using MME, in the three scenarios

Scenario	Gaussian	t-Student	Clayton	Frank	Joe	Gumbel-Hougaard	Tawn	Galambos	Hulser-Reiss
Weak	**118.537**	**118.535**	**70.839**	**109.322**	**9.217**	**75.841**	**51.404**	**75.254**	**70.826**
Moderate	**2048.75**	**2168.11**	**-32.16**	**1925.08**	**2953.65**	**2736.46**	**2726.26**	**2714.55**	**2673.84**
Strong	**69.2798**	**69.9732**	**26.681**	**61.8349**	**67.1129**	**74.7194**	**53.6635**	**75.1437**	**76.0469**

Source: own elaboration.

As much as the model correctly indicated each copula family for all scenarios, one can see that, in the 'weak' scenario, the CIC values of the Gaussian and t-Student copulas are practically the same. However, the t-Student copula for this CIC has 1373 degrees of freedom (high degrees of freedom indicate proximity to a Gaussian). In the 'moderate' scenario, the negative CIC value of the Clayton copula indicates how far this copula is from being the most suitable to capture this dependence: Figure 1 showed that the Clayton copula has a large concentration of data in the lower left rectangle of the distribution, opposite to Joe copula, who has a large concentration of data in the upper right rectangle.

Once the copula family has been chosen, it is necessary to estimate its parameter. Table 7 presents the estimators using the MME and MPLE methods.

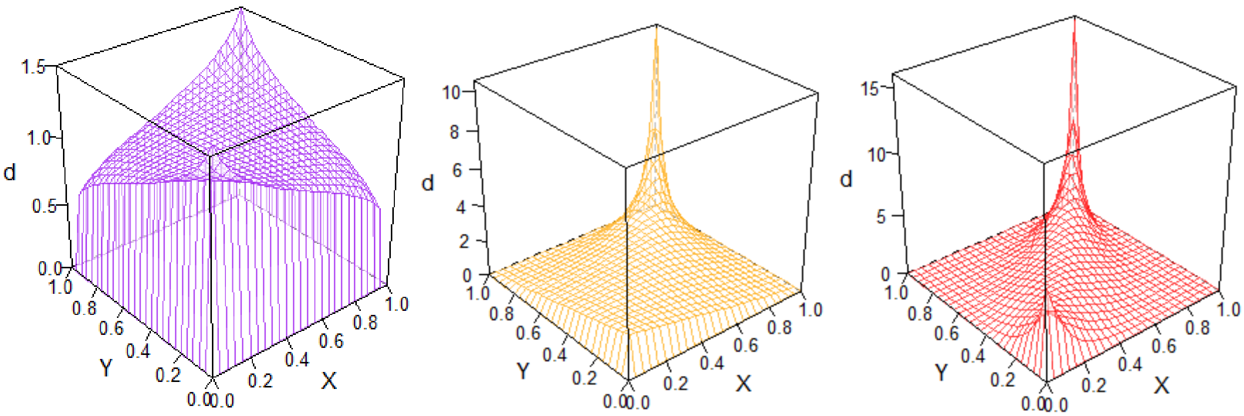
Table 7 – Estimator for the copula parameter in the three scenarios

Scenario	**Method**	$\hat{\theta}$	**Standard deviation**	**Log-likelihood**
Weak	MPLE	0.1539	0.010	119.5354
	MME	0.1547	636.400	119.5323
Moderate	MPLE	2.2358	0.031	2,954.5157
	MME	2.2249	-	2,954.3992
Strong	MPLE	3.2456	0.472	76.6722
	MME	3.0319	49,373.000	76.3523

Source: own elaboration.

Note that the estimator generated by the MPLE method has a higher log-likelihood and lower standard deviation than the one generated by the MME method (it is not possible to calculate the standard deviation using the MME for Joe copula), showing that the first estimator is more adequate than the second, corroborating the results of previous studies (Hofert et al., 2018). Figure 2 shows the joint probability density of X and Y, using the Gaussian, Joe and Husler-Reiss copulas, with the respective parameters $\theta_{Gaussian} = 0.1539$, $\theta_{Joe} = 2.2358$ and $\theta_{H-Reiss} = 3.2456$.

Figure 2 – Joint probability density of (X, Y) with Gaussian (left), Joe (center) and Husler-Reiss (right) copulas

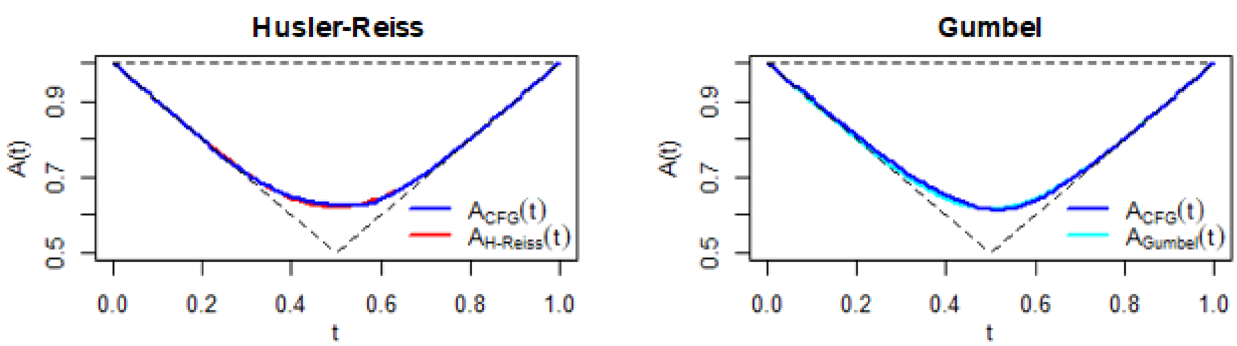


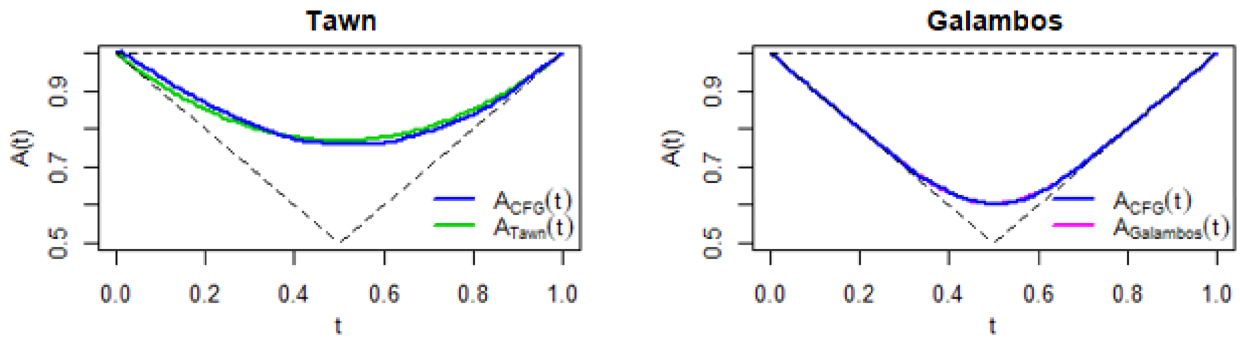
Source: own elaboration.

After successfully capturing the real underlying dependence structure in all three scenarios, we analyze the use of MCFG to obtain an estimator $\hat{A}_{CFG}$ for the *dependence function* for extreme-value copulas. For this, we defined 4 data sets, each with 150 pairs (X, Y) simulated from each of the Gumbel-Hougaard, Tawn, Galambos and Husler-Reiss copulas, with $\hat{\tau}_{X,Y} = 0.7$ (except for Tawn copula: due to restrictions in Equation (11), we chose $\hat{\tau}_{X,Y} = 0.4$). The Exponential distributions of X and Y were the same as in the 'weak', 'moderate' and 'strong' scenarios. Then, we constructed graphs to compare the estimator $\hat{A}_{CFG}$ and the corresponding dependence function, shown in Table 3.

The dashed lines in the graphs in Figure 3 represent the restrictions of the dependence function A(t): $\max\{t, 1-t\} \leq A(t) \leq 1$. According to Gudendorf & Segers (2012) and Hofert et al. (2018), the proximity between the graph and the horizontal dashed line indicates a complete absence of dependence between X and Y. In contrast, the proximity of the graph with the sloping dashed lines indicates complete dependence (comotonicity) among these two variables.

Figure 3 – Comparative graphs between $\hat{A}_{CFG}$ and the respective dependence function





Source: own elaboration.

From Figure 3, we can see that, for the Husler-Reiss, Gumbel-Hougaard and Galambos copulas, there is very little difference between the estimator $\hat{A}_{CFG}$ and the dependence function presented in Table 3. In particular, it is noted that these three copulas captured a high dependence on the lower and upper tails (described parametrically by t < 0.2 and t > 0.8, respectively) on the joint distribution of (X, Y). This fact agrees with what was evidenced by Brechmann et al. (2014): losses resulting from extreme events tend to occur together, and not independently. In addition, the Tawn copula is more sensitive to the difference between $\hat{A}_{CFG}$ and A_{Tawn}, with the majority of the domain of the joint distribution (that is, for t > 0.4) the estimator $\hat{A}_{CFG}$ capturing a higher dependence than that captured by the dependence function A_{Tawn}.

Table 8 shows the values of the upper tail dependence parameter, calculated from Table 4 and Equation (12), considering both the MCFG method and the dependence function (DF) in Table 3. It is shown that the Husler-Reiss, Gumbel-Hougaard and Galambos copulas capture the high dependence on the upper tail (with values close to 0.8) and Tawn copula is more sensitive to the difference between these two methods: the relative differences between the values obtained by the MCFG and DF are 4.57% (Tawn), 0.53% (Gumbel), 0.43% (Galambos) and 0.31% (Husler-Reiss).

Table 8 – Upper tail dependence parameter calculated from MCFG and DF methods

Copula	Method	λ_U	Copula	Method	λ_U
Husler-Reiss	MCFG	0.77526	Tawn	MCFG	0.48376
	DF	0.77290		DF	0.46260
Gumbel	MCFG	0.77307	Galambos	MCFG	0.79658
	DF	0.76900		DF	0.79320

Source: own elaboration.

Finally, we verified the sensitivity of the MCFG method, in relation to DF, for datasets with different sizes – it is quite common in extreme events modeling to have few data available. Thus, the upper tail dependence parameter was calculated for simulated samples with n=100, 50 and 30 pairs and were calculated the relative variations between the values of the coefficients obtained by the MCFG method, in relation to DF. These percentages were organized in Table 9.

Table 9 – Relative difference between upper tail dependence parameters, by the two methods

Copula	$n = 150$	$n = 100$	$n = 50$	$n = 30$
Husler-Reiss	0.31%	-1.02%	-0.12%	-0.04%
Gumbel-Hougaard	0.53%	0.04%	-0.35%	0.91%
Tawn	4.57%	-10.49%	-3.98%	-6.98%
Galambos	0.43%	0.68%	0.77%	0.28%

Source: own elaboration.

The results suggest that, the smaller the sample set is, the more MCFG captures, in Galambos copula, an increasing dependence than that captured by DF, since the differences are always positive and growing (except for n = 30). In Husler-Reiss and Gumbel-Hougaard copulas, however, there are small variations between MCFG and DF. This does not happen for Tawn copula, which presents diffuse results: sometimes MCFG captures a higher dependence than that captured by DF, sometimes lower.

4.2. Capturing the dependence structure using real-world data

The real-world microdata used contain all the insurance policies of an insurance company whose claims incurred by a single event generated payments of indemnities in (at least) two different insurance lines of business in which the company operates, in the period from Jan/2007 to May/2012. All monetary values (in Brazilian Real R$ currency) were brought to the constant currency of May/2012, to make them comparable. The Brazilian insurance lines of business that had claims under these conditions were 0531 (Automobile – Hull), 0553 (Automobile – Optional Civil Liability for Vehicles), 0351 (General Civil Liability), 0378 (Professional Civil Liability), 0118 (Business Comprehensive) and 0520 (Automobile – Passenger Personal Accidents).

Representing each simultaneous claim by an ordered pair (x, y), with x and y indicating the indemnities related to lines of business X and Y, respectively, Table 10 presents the descriptive statistics of these policies, relative to the total indemnified $(x + y)$ generated by these pairs.

Table 10 – Descriptive statistics of microdata

Pair of lines (X, Y)	Number of Observations	Mean of $X + Y$ (R$)	Median of $X + Y$ (R$)	St. deviation of $X + Y$ (R$)	Maximum of $X + Y$ (R$)
(0531,0553)	13676	13,297.03	7,440.30	25,987.45	1,131,667.96
(0351,0378)	25	50,958.47	41,968.50	49,888.17	165,004.24
(0118,0351)	17	108,390.74	15,231.74	293,507.80	1,230,452.83
(0520,0531)	16	48,384.23	38,738.98	35,489.41	128,636.08

Source: own elaboration.

The large number of observations of the pair (0531,0553) is justified by the characteristics of the policies that cover these lines jointly: it is a massified, standardized compound insurance with a long history. In addition, it is noted that, in the four types of pairs of lines listed in Table 10, the mean of the total indemnities per policy is greater than the median. This, together with the high standard deviations, high maximum indemnity amount (per policy) and low number of observations in 3 of the 4 pairs, evidence the presence of extreme events in this company's portfolio. It is relevant to note that in pair (0531,0553) the maximum amount indemnified per policy is higher than 85 times the mean of the payment indemnities (per policy) for the period. In the pair (0118,0351), the maximum amount indemnified per policy is higher than 11 times the mean of the payment indemnities (per policy) for the period.

From these pairs, we analyzed 10 scenarios. The first 7 being relative to the pair (0531,0553), considering: 1) all data of this pair; 2) data related to the 90% quantile index in each of lines 0531 and 0553; 3) same as previous one, with a 95% quantile index; 4) same as scenario 2, with a 99.5% quantile index; 5) data related to the 90% quantile index in the sum of lines 0531 and 0553; 6) same as previous one, with a 95% quantile index; and 7) same as scenario 5, with a 99.5% quantile index.

The other scenarios are: 8) all data in the pair (0351,0378); 9) all data in the pair (0118,0351); and 10) all data from the pair (0520,0531). According to Sweeting & Fotiou (2013) and Sun et al. (2020), an extreme event can generate multiple claims that are not as severe as in an isolated line of business. However, as the insurance company pays the total amount claimed, when adding the indemnity from another line originating from the same event, it may occur that the total amount indemnified exceeds a certain limit, justifying the construction of scenarios 5 to 7, which can generate results quite different from scenarios 2 to 4. Table 11 shows the best ranked copula family (according to the CIC) and the estimated parameter (using the MPLE) for these scenarios.

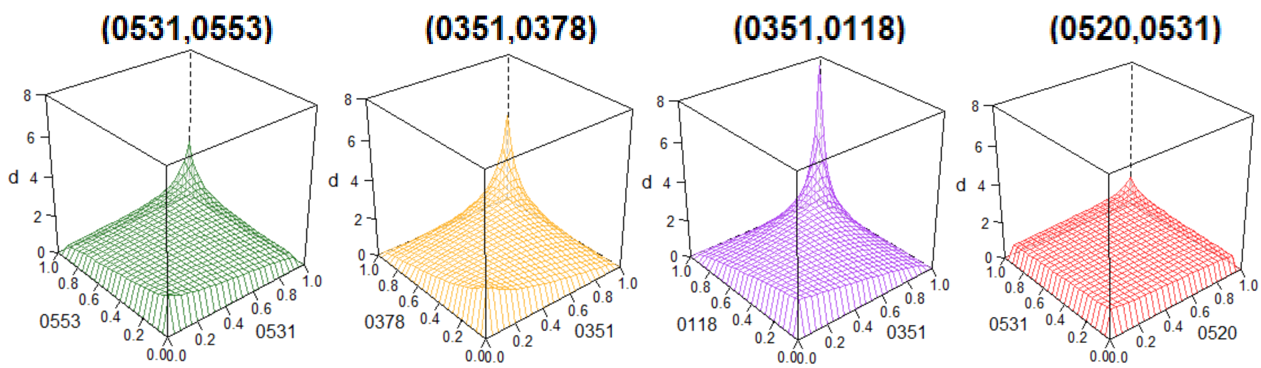
Table 11 – Best ranked copula family and estimated parameter, in each of the scenarios

Scenario	Copula	$\hat{\theta}$	Scenario	Copula	$\hat{\theta}$
1	Gumbel	1.23595±0.008	6	Frank	-2.64214±0.217
2	Husler-Reiss	0.87226±0.067	7	t-Student	-0.37102±0.130
3	Gaussian	0.29163±0.068	8	Galambos	0.75320±0.339
4	Tawn	0.88252±0.313	9	Joe	1.84144±0.795
5	t-Student	-0.39272±0.024	10	Joe	1.07774±0.398

Source: own elaboration.

The scenarios in which all data from the pair were used (scenarios 1, 8, 9 and 10) had their dependences captured by extreme-value or Joe copulas – precisely those that, by Figure 1, capture the high dependence on the upper tail. The joint probability densities for the copulas of these four scenarios (Figure 4) show that if an insurance company wishes to offer policies with multiple coverages, it cannot ignore the fact that large losses occur jointly: there is invariably high density in the high quantiles of both lines.

Figure 4 – Joint probability density, for each of the 4 pairs

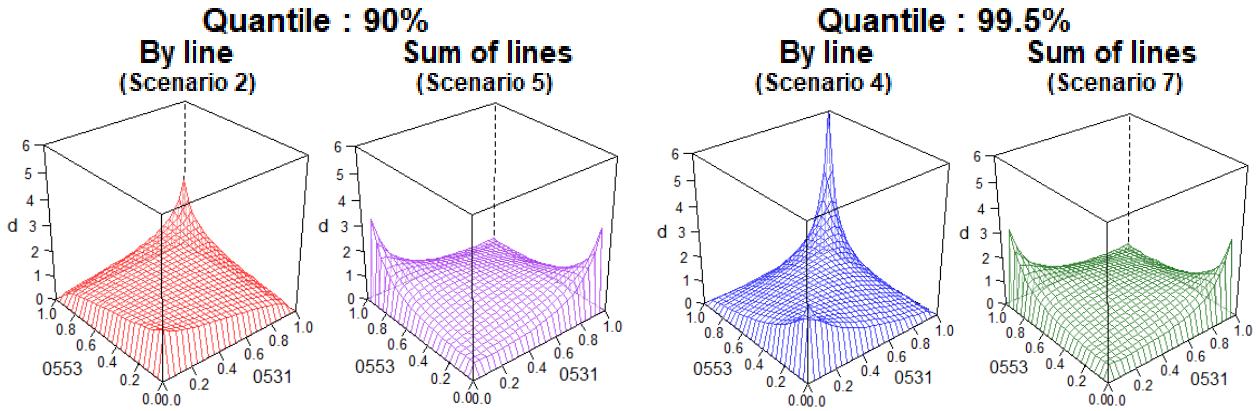


Source: own elaboration.

After restricting the pair (0531,0553) to 90%, 95% and 99.5% quantiles, extreme-value copulas adequately captured the dependence in two scenarios: 2 (90% of each line) and 4 (99.5% of each line). However, when considering the same quantiles, for the sum of the losses (scenarios 5 and 7, respectively), the best ranked copula family was the t-Student, with 4.82 degrees of freedom – in both scenarios.

Figure 5 shows the joint probability densities corresponding to these four scenarios. Note that *extreme-value copulas* are better for capturing dependence in cases which only events that generate large losses in both lines, but do not do so in cases which an event generates a high indemnity in one of the lines of business, but not on the other. When these cases are also included, the dependence structure is modified in order to capture this dependence as well. In scenarios 2 and 4, extreme-value copulas were replaced by a t-Student copula.

Figure 5 – Comparison between joint probability densities, by scenario



Source: own elaboration.

In order to highlight the argument of the previous paragraph, the quantile of 99.5% in line 0531 corresponds to R$137,725.95. The same quantile in line 0553 corresponds to R$55,449.86 and, considering the sum of the losses, this quantile corresponds to R$145,728.54. One of the policies generated claims whose indemnity were R$485,518.77 and R$19,401.52 in these respective lines of business, totaling R$504,920.29 to be indemnified by the insurance company, resulting from a single event. As the amount in line 0553 is less than R$55,449.86, then this policy is not included in scenario 4, but is included in scenario 7, illustrating the difference between the dependence structures in scenarios 2 and 4, in relation to 5 and 7.

Table 12 shows the upper tail dependence parameters for each of the 10 scenarios. The highest value of upper tail dependence parameter occurred in the pair (0118,0351), showing that, in comprehensive business insurance, extreme events tend to severely affect not only the company itself, but also third parties.

Table 12 – Upper tail dependence parameter, for each scenario

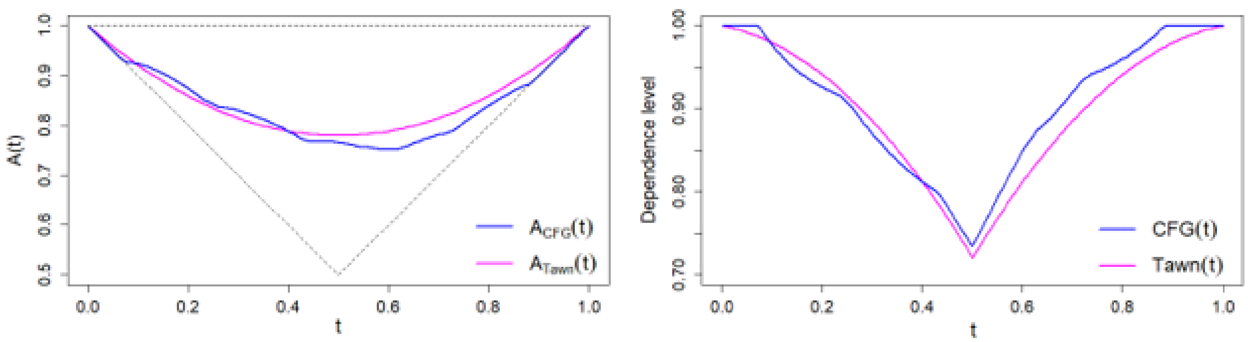
Scenario	**1 (DF)**	**1 (MCFG)**	**2 (DF)**	**2 (MCFG)**	**3**	**4 (DF)**	**4 (MCFG)**
$\boldsymbol{\lambda_U}$	0.24789	0.24491	0.25161	0.24163	0	0.44126	0.46916
Scenario	**5**	**6**	**7**	**8 (DF)**	**8 (MCFG)**	**9**	**10**
$\boldsymbol{\lambda_U}$	0.01127	0	0.01250	0.39841	0.39184	0.54295	0.09754

Source: own elaboration.

The pairs (0531,0553) and (0351,0378) showed little difference between the DF and MCFG method. In addition, if the insurance company considers only the values of λ_U of the scenarios in which the sum of the losses is considered to dimension its solvency capital (and not each line of business individually), it may be exposed to the risk of insolvency if it occurs an extreme event. This argument is evidenced by calculating the variation between the upper tail dependence parameters by comparing the scenarios in which each line is considered individually and its respective scenario, considering the sum of the losses: in scenario 2 (DF), the parameter obtained was 2133% higher than that obtained in scenario 5 and, in scenario 4 (MCFG), the parameter obtained was 3655% higher than that obtained in scenario 7. These high variations strongly reflect the increase in the level of exposure to the insolvency risk of this insurance company, due to some extreme event.

Analyzing the sensitivity of the MCFG method and DF, it is noted that, in scenario 4, the MCFG method indicated an even higher dependence on the upper tail than that indicated by the DF. Both graphs in Figure 6 compare the two methods, with the graph on the right using, for this, the dependence level (where 0 indicates independence and 1, complete dependence). In addition to the high dependence across the domain, note that, for t > 0.4, the MCFG method captures a higher dependence than the DF.

Figure 6 – Comparison between $\hat{A}_{CFG}$ estimator and the dependence function of Tawn copula



Source: own elaboration.

5. Conclusion

There are events that, when they occur, generate several losses simultaneously, in different insurance lines of business. For insurance companies that operates in different lines – and, therefore, make multiple payments of indemnity – it is not recommended to assume that these claims are independent. If the event occurred is an extreme event, whose claims have high severity, an unexpected increase in the indemnified amounts may even lead the insurance company to bankruptcy.

One of the ways to capture the dependence structure among large losses is through copulas. The choice of copula and the method of estimating the parameters is an essential part of capturing this dependence (Su & Furman, 2017). Thus, this article aimed to analyze the choice of copula among the *elliptical*, *Archimedean* and *extreme-value* (EVC) families of copulas, using different methods of estimation (MME, MPLE and MCFG) of the parameters. Once defined as capturing the dependence structure, we estimated the upper tail dependence using real-world microdata from an insurance company.

As a result, the MPLE proved to be more appropriate than the MME method and, in several cases for EVC, the MCFG method captured the upper tail dependence better than the FG. In relation to microdata, EVC proved to be more adequate to capture the dependence among large losses in different scenarios, in relation to the other copulas. The dependence among incurred claims in different lines of business corroborates the fact that SUSEP groups different lines in the so-called 'business classes'. In CNSP Resolution nº360/17, SUSEP predicts and quantifies the dependence that exists among *business classes* in calculating the minimum capital required for the insurance company's solvency, but does not do so for lines of the same class of business. The results of this study showed that this dependence exists and is relevant for the calculation of risk capital, especially for insurance companies that operate in different lines belonging to the same class of business, which may be exposed to greater risks than they believe they are. Still, the dependence on the upper tail obtained by the pair (0118,0351) – the only one in this study whose lines belong to different classes – is 35.7% higher than that calculated by SUSEP, indicating that the insurance company's solvency capital does not correctly predict extremes events, which is dangerous.

In future work, we suggest further applications to other sets of microdata or data referring to other insurance lines of business, such as personal insurance, in which it may be interesting to capture the dependence among the Personal Accidents, Serious Illnesses or Loss of Income lines. Another suggestion is to use a *mixture copulas* or *empirical copulas* to capture different dependence structures within the same data set.

6. References

Bokusheva, R. (2018). Using copulas for rating weather index insurance contracts. *Journal of Applied Statistics*, *45*(13), 2328–2356. https://doi.org/10.1080/02664763.2017.1420146

Bormann, C., & Schienle, M. (2018). Detecting structural differences in tail dependence of financial time series. *Journal of Business and Economic Statistics*, *15*. https://doi.org/10.1080/07350015.2018.1506343

Brechmann, E., Czado, C., & Paterlini, S. (2014). Flexible dependence modeling of operational risk losses and its impact on total capital requirements. *Journal of Banking and Finance*, *40*(1), 271–285. https://doi.org/10.1016/j.jbankfin.2013.11.040

Bücher, A., Irresberger, F., & Weiss, G. N. F. (2017). Testing asymmetry in dependence with copula-coskewness. *North American Actuarial Journal*, *21*(2), 267–280. https://doi.org/10.1080/10920277.2017.1282876

Chen, H., MacMinn, R. D., & Sun, T. (2017). Mortality dependence and longevity bond pricing: a dynamic factor copula mortality model with the GAS structure. *Journal of Risk and Insurance*, *84*, 393–415. https://doi.org/10.1111/jori.12214

Chukwudum, Q. C. (2019). Extreme value theory and copulas: reinsurance in the presence of dependent risks. *Applied Mathematical Sciences*, *13*(2), 67–86. https://doi.org/10.12988/ams.2019.811177

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, *65*(1), 141–151. https://doi.org/10.1093/biomet/65.1.141

Cossette, H., Marceau, E., Mtalai, I., & Veilleux, D. (2018). Dependent risk models with archimedean copulas: a computational strategy based on common mixtures and applications. *Insurance: Mathematics and Economics*, *78*, 53–71. https://doi.org/10.1016/j.insmatheco.2017.11.002

Dietz, S., & Walker, O. (2019). Ambiguity and Insurance: Capital Requirements and Premiums. *Journal of Risk and Insurance*, *86*(1), 213–235. https://doi.org/10.1111/jori.12208

Eling, M., & Toplek, D. (2009). Modeling and management of nonlinear dependencies-copulas in dynamic financial analysis. *Journal of Risk and Insurance*, *76*(3), 651–681. https://doi.org/10.1111/j.1539-6975.2009.01318.x

Embrechts, P. (2009). Copulas: A personal view. *Journal of Risk and Insurance*, *76*(3), 639–650. https://doi.org/10.1111/j.1539-6975.2009.01310.x

Fréchet, M. (1951). Sur les tableaux dont les marges et des bornes sont données. *Revue de l'Institut International de Statistique*, *28*(1/2), 10. https://doi.org/10.2307/1401846

Frees, E. W., & Valdez, E. A. (1998). Understanding relationships using copulas. *North American Actuarial Journal*, *2*(1), 1–25. Retrieved from http://mx.nthu.edu.tw/~jtyang/Teaching/Risk_management/Papers/Quant_Methods/Understanding Relationships Using Copulas.pdf

Furman, E., Kuznetsov, A., Su, J., & Zitikis, R. (2016). Tail dependence of the Gaussian copula revisited. *Insurance: Mathematics and Economics*, *69*, 97–103. https://doi.org/10.1016/j.insmatheco.2016.04.009

Giussani, A., & Bonetti, M. (2019). Marshall–Olkin frailty survival models for bivariate right-censored failure time data. *Journal of Applied Statistics*, *46*(16), 2945–2961. https://doi.org/10.1080/02664763.2019.1624694

Grønneberg, S., & Hjort, N. L. (2014). The Copula Information Criteria. *Scandinavian Journal of Statistics*, *41*(2), 436–459. https://doi.org/10.1111/sjos.12042

Gudendorf, G., & Segers, J. (2012). Nonparametric estimation of multivariate extreme-value copulas. *Journal of Statistical Planning and Inference*, *142*(12), 3073–3085. https://doi.org/10.1016/j.jspi.2012.05.007

Hofert, M., Kojadinovic, I., Mächler, M., & Yan, J. (2018). *Elements of copula modeling with R*. https://doi.org/10.1007/978-3-319-89635-9

Hua, L. (2017). On a bivariate copula with both upper and lower full-range tail dependence. *Insurance: Mathematics and Economics*, *73*, 94–104. https://doi.org/10.1016/j.insmatheco.2017.01.003

Kamnitui, N., Genest, C., Jaworski, P., & Trutschnig, W. (2019). On the size of the class of bivariate extreme-value copulas with a fixed value of Spearman's rho or Kendall's tau. *Journal of Mathematical Analysis and Applications*, *472*(1), 920–936. https://doi.org/10.1016/j.jmaa.2018.11.057

Kelliher, P. O. J., Acharyya, M., Couper, A., Maguire, E., Nicholas, P., Pang, N., … Teggin, P. (2020). Operational risk dependencies. *British Actuarial Journal*, *25*(May), 1–21. https://doi.org/10.1017/S1357321720000033

Kley, O., Klüppelberg, C., & Paterlini, S. (2020). Modelling extremal dependence for operational risk by a bipartite graph. *Journal of Banking and Finance*, 105855. https://doi.org/10.1016/j.jbankfin.2020.105855

Li, H. (2018). Operator tail dependence of copulas. *Methodology and Computing in Applied Probability*, *20*(3), 1013–1027. https://doi.org/10.1007/s11009-017-9592-1

Lopes, F., Peris, R. W., Chan, B. L., & Borelli, E. (2015). Evolução do mercado segurador e crescimento econômico no Brasil. *Revista Eletrônica Do Departamento de Ciências Contábeis & Departamento de Atuária e Métodos Quantitativos (REDECA)*, *2*(2), 21–36. Retrieved from https://revistas.pucsp.br/index.php/redeca/article/view/28561

Melo, E. F. L. de. (2008). Uma aplicação de cópulas de Lévy na agregação de processos multivariados de ruína. *Revista Brasileira de Risco e Seguro*, *4*(7), 47–64. Retrieved from http://www.rbrs.com.br/arquivos/RBRS7-3 Eduardo Fraga.pdf

Nelsen, R. B. (2006). *Introduction to copulas* (2nd ed.). Springer International Publishing.

Pappadà, R., Perrone, E., Durante, F., & Salvadori, G. (2016). Spin-off extreme value and archimedean copulas for estimating the bivariate structural risk. *Stochastic Environmental Research and Risk Assessment*, *30*(1), 327–342. https://doi.org/10.1007/s00477-015-1103-8

Pickands, J. (1981). Multivariate extreme value distributions. *Bulletin of the International Statistical Institute*, (49), 859–878.

Ruenzi, S., Ungeheuer, M., & Weigert, F. (2020). Joint extreme events in equity returns and liquidity and their cross-sectional pricing implications. *Journal of Banking and Finance*, *115*. https://doi.org/10.1016/j.jbankfin.2020.105809

Shen, A. (2019). Bread and butter: the interdependence of risks. *The Actuary Magazine*, *16*(3), 23–24. Retrieved from https://theactuarymagazine.org/wp-content/uploads/2019/06/act-2019-vol16-iss3.pdf

Sklar, A. (1959). Fonctions de repartition a n dimensions et leurs marges. *Pulications de Institut de Statistique de l'Univeristé de Paris*, (8), 229–231. Retrieved from https://mathscinet.ams.org/mathscinet-getitem?mr=125600

Su, J., & Furman, E. (2017). Multiple risk factor dependence structures: Copulas and related properties. *Insurance: Mathematics and Economics*, *74*, 109–121. https://doi.org/10.1016/j.insmatheco.2017.03.003

Sun, N., Yang, C., & Zitikis, R. (2020). A statistical methodology for assessing the maximal strength of tail dependence. *ASTIN Bulletin*, *50*(3), 799–825. https://doi.org/10.1017/asb.2020.21

Sweeting, P., & Fotiou, F. (2013). Calculating and communicating tail association and the risk of extreme loss. *British Actuarial Journal*, *18*(1), 13–72. https://doi.org/10.1017/s1357321712000347

Szego, G. (2002). Measures of risk. *Journal of Banking and Finance*, *26*(1), 1253–1272. https://doi.org/10.1016/S1548-5315(11)70388-6

Tanaka, A. D., & Carvalho, J. V. de F. (2019). Estimação da estrutura de dependências entre classes de seguro por meio de cópulas. *Revista Brasileira de Risco e Seguro*, *15*(25), 1–34. Retrieved from http://www.rbrs.com.br/arquivos/rbrs_25_1.pdf

Xu, M., & Hua, L. (2019). Cybersecurity insurance: modeling and pricing. *North American Actuarial Journal*, *23*(2), 220–249. https://doi.org/10.1080/10920277.2019.1566076

Zhu, W., Tan, K. S., & Wang, C. W. (2017). Modeling Multicountry Longevity Risk With Mortality Dependence: A Lévy Subordinated Hierarchical Archimedean Copulas Approach. *Journal of Risk and Insurance*, *84*, 477–493. https://doi.org/10.1111/jori.12198

---

[i] Articles 686, 687 and 689, available in http://www.planalto.gov.br/ccivil_03/Leis/LIM/LIM556compilado.htm.
[ii] Available in https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32009L0138&from=EN
[iii] Available in http://thoughtleadership.aon.com/Documents/20200122-if-natcat2020.pdf
[iv] According to SUSEP, the total premium was R$ 111 billion, converted into US dollars at a price equal to R$ 3.944.
[v] https://practiceguides.chambers.com/practice-guides/insurance-reinsurance-2020/brazil/trends-and-developments.
[vi] Available in https://www2.susep.gov.br/safe/scripts/bnweb/bnmapi.exe?router=upload/14294.