Dependence Modeling in Non-Life Insurance: Copula Functions and Capital Adequacy - A Case Study of AXA Insurance

Mekdad Slime, Mohammed El Kamli, Abdelah Ould Khal

Abstract—**The traditional modeling of dependence among branches in non-life insurance has historically rested on the assumption of independence between risks. However, recent studies have brought to light that overlooking the interdependence between risks is not prudent. An illustrative example is found in automobile insurance, where a discernible connection exists between the number of claims in the Automobile Damage branch and those in the Automobile Civil Liability branch.**

This paper aims to delve into the utilization of copula functions for modeling the interrelationships among claims charges across diverse branches within a non-life insurance portfolio. Moreover, it seeks to assess the consequential impacts of these dependencies on the capital requirements of AXA Insurance in the Moroccan market. Our research concentrates on four pivotal branches within AXA Insurance, namely "work accidents", "automobile civil liability", "disability", and "fire, marine, aviation, and transport". Through a thorough analysis encompassing dependency measurements, independence tests, and various graphical representations, we have substantiated the existence of interdependencies among the claim costs associated with these four risk branches. Importantly, our findings underscore that assuming independence leads to a slight underestimation of the capital requirement needed for effective risk accumulation management by the insurer.

Index Terms—**dependence modeling, non-life insurance, financial and insurance mathematics, copulas, risk measures, actuarial science.**

I. Introduction

In the course of our lives, individuals confront a myriad of risks, ranging from accidents and theft to illnesses and natural disasters. To mitigate these uncertainties, people turn to insurance companies, which shoulder these risks on their behalf. Through the collection of premiums or contributions and strategic investments in the financial market, insurance companies play a crucial role in fortifying the economy. While this financial prowess enhances the strength of insurance firms, they must possess the capacity to honor their commitments.

In light of this, a robust legal framework is established to ensure the safeguarding of subscribers' interests. The insurance sector in Morocco, like many other countries, is subject to a comprehensive set of control mechanisms. The Risk-Based Solvency (SBR: Solvabilité Basée sur les Risques) represents the latest Moroccan standard in prudential regulation. The capital requirement outlined in the current Moroccan standard is codified in Article 239 of Dahir No. 1-02-238 dated 25 Rejeb 1423 (October 3, 2002), which promulgates Law No. 17-99 [1].

Enacted through Law No. 64-12, which came into effect on April 14, 2016, the creation of the Supervisory Authority for Insurance and Social Welfare (ACAPS = Autorité de contrôle des assurances et de la prévoyance sociale) marked a significant stride in the ongoing modernization of the Moroccan financial sector. According to its annual statistics [2], in 2021, the premiums issued by insurance and reinsurance companies demonstrated noteworthy figures, reaching 22.9 billion dirhams for life premiums with acceptances and 27.3 billion dirhams for non-life premiums.

Within the "life" branch, the statistics reveal a substantial 32.2% surge in savings premiums (Support Dirhams and Support Units of Account), surpassing 19 billion dirhams, along with a 10.2% increase in death premiums, totaling 3.2 billion dirhams. Similarly, for the "non-life" sector, ACAPS reports an 8.6% upswing in motor insurance premiums, reaching approximately 13 billion dirhams. Furthermore, there is a noticeable uptick in the catastrophic events segment, rising by 9.5% to 521.8 million dirhams. Bodily accidents also saw an increase, reaching nearly 4.77 billion dirhams, reflecting an 8% rise, while technical risks amounted to approximately 276.1 million dirhams, indicating a substantial 59% increase.

These statistics underscore the pivotal role played by insurance companies. Their impact extends beyond the social realm, where they promote well-being and compensate for third-party losses. Additionally, they emerge as crucial contributors to the nation's economic growth.

Up until the early 1990s, insurance professionals predominantly operated under the assumption of risk independence, employing Gaussian models and Pearson correlation in their historical operational methods. Actuaries essentially viewed risk independence as a fundamental methodological prerequisite. Nevertheless, the evolution of financial markets over the past few decades has unequivocally shown that dismissing interdependence among risks is no longer tenable [3], [4], [5], [6], [7], [8]. Practically speaking, the assumption of risk independence is seldom realized. Take, for example, the domain of automobile insurance, where it's evident that the frequency and severity of claims in the Automobile Damage branch are intricately linked to those in the Automobile Civil Liability branch. This realization has spurred a heightened interest in embracing innovative models. In scenarios where

Manuscript received November 11, 2023; revised February 26, 2024.

Mekdad Slime is a PhD Student at Faculty of Sciences, Mohammed V University in Rabat, Laboratory of Mathematical, Statistics and Application, B.P. 1014 - Rabat, Morocco (e-mail: mekdad_slime@um5.ac.ma).

Mohammed El Kamli is a Professor of Mathematics at Faculty of Sciences, Economic, Juridical and Social - Souissi, Mohammed V University in Rabat, Laboratory of Economic Analysis and Modelling (LEAM), B.P. 1014 - Rabat, Morocco (e-mail: m.elkamli@um5r.ac.ma).

Abdellah Ould Khal is a Professor of Statistics at Faculty of Sciences, Mohammed V University in Rabat, Laboratory of Mathematical, Statistics and Application, B.P. 1014 - Rabat, Morocco (e-mail: a.ouldkhal@um5r.ac.ma).

modeling dependence structures becomes imperative, copula functions become invaluable, enabling the construction of non-Gaussian models.

A copula is a mathematical function that creates a connection between marginal distributions to generate a joint distribution. The application of copulas essentially embodies a ”Bottom-Up” approach (refer to paragraph II-A-2), with the primary goal of modeling dependence. The term ”copula,” derived from the Latin word meaning ”link,” was first introduced by Sklar in 1959 [9].

There are several families of copulas [10], [11], and there exist many properties relating to the copulas [12], [13], [14], [15], [16], [17]. Recently, this approach has drawn significant attention across various fields. In 2016, Chunping Li et al. [18] introduced a novel reliability model for two dependent performance characteristics utilizing marginal distribution functions and copula theory. Their emphasis on accounting for dependence was demonstrated through a numerical example using train wheels wear data. In 2018, Yuko Otani et al. [19] analyzed the dependence structure of international equity and bond markets employing a regime-switching copula model. They identified asymmetric dependence in equity markets, highlighting stronger dependency for negative returns compared to positive returns. Applying flexible copulas in three country pairs (UK-US, Japan-US, and Italy-US), their study revealed significant implications from the empirical analysis. In 2019, Shaoqian Huang et al. [20] proposed a method for expanding sample capacity in scenarios with limited or costly samples. Their approach, based on the Distance-Trend Double Effects (DTDE), considers sample component evolution and distances between samples. More recently, Ghizlane Kouaiba et al. [21] investigated the coherence of Value at Risk (VaR) as a risk measure when applied to non-elliptical distributions. They focused on elliptical copulas within a given portfolio, examining whether VaR remains coherent in an elliptical or spherical space.

Copula theory has also attracted considerable attention in various insurance framework applications. For instance, in 2012, Dorothea Diers et al. [22] demonstrated the modeling of dependence structures for non-life insurance risks using the Bernstein copula. The following year, Fang and Madsenb [23] introduced the characteristics of the modified Gaussian pseudo-copula, providing examples of its application in both insurance and finance. In a more recent study in 2018, Hanène Mejdoub and Mounira Ben Arab [24] explored the sensitivity of risk capital estimation to the dependence structure between losses from four non-life business lines of a Tunisian insurance company, utilizing the D-Vine Copula approach. Furthermore, in 2021, Li-Mei Qi et al. [25] employed a copula-stochastic optimization model to investigate revenue insurance premium ratemaking for Jujube in the Aksu region, Xinjiang, China. These studies underscore the diverse and impactful applications of copula theory in the field of insurance.

Our study is organized as follows: In Section II, we revisit the primary definitions of copulas and the dependencies between risks. Section III introduces the data being analyzed. Section IV delves into the interconnections among our four risk branches, employing Pearson's linear correlation coefficient, Kendall's Tau, and Spearman's Rho. Section V explores the application of copula functions in modeling the relationships among claims charges within the four risk branches. Finally, we conclude in the last section.

II. Materials and Methods

A. Dependency types and measures

Definition 2.1: [10]

Let X and Y be two real random variables and let $F_{X;Y}$ be the joint distribution function of X and Y.
We said that X and Y are independent if,

$$(\forall x, y \in \mathbb{R}) : F_{X;Y}(x; y) = F_X(x)F_Y(y).$$

For a significant period, Pearson's correlation coefficient has been the predominant tool for modeling dependence between random variables. Nevertheless, practitioners are now increasingly exploring alternative dependency measures, such as Spearman's Rho and Kendall's Tau. [10], [26].

Definition 2.2: (Pearson's correlation coefficient)

Let X and Y be two random variables with finite variances. The linear correlation coefficient between X and Y is given by,

$$\rho(X; Y) = \frac{Cov(X; Y)}{\sqrt{Var(X)Var(Y)}}.$$

From this definition, we can deduce the following points,

- $\rho(X; Y) = 0 \iff$ The variables X and Y exhibit linear independence.
- $\rho(X; Y) \neq 0 \iff$ The variables X and Y are linearly associated or correlated.

Definition 2.3: (Kendall's Tau)

Let $(X_1; Y_1)$ and $(X_2; Y_2)$ be two continuous random vectors independent and identically distributed. The Kendall's Tau is given by,

$$\tau(X; Y) = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

From this definition, we can deduce the following points,

- Kendall's Tau is symmetric.
- $-1 \leq \tau \leq 1$.
- If X and Y are independents, then $\tau = 0$.

Definition 2.4: (Spearman's Rho)

Let $X = (X_1; X_2; ...; X_n)$ and $Y = (Y_1; Y_2; ...; Y_n)$ be two random vectors and let $R = (R_1; R_2; ...; R_n)$ and $S = (S_1; S_2; ...; S_n)$ the random vectors of the ranks of X and Y respectively. Spearman's Rho is given by,

$$\rho_S(X; Y) = \frac{Cov(R; S)}{\sqrt{Var(R)Var(S)}}.$$

Spearman's Rho, developed by Charles Spearman in 1904, is defined as the correlation between the ranks or orders of observations. [27].
From this definition, we can deduce the following points,

- $-1 \leq \rho_S \leq 1$.
- If X and Y are independents, then $\rho_S = 0$.

Now, we can categorize two approaches: the top-down approach and the bottom-up approach.

1) Top-Down approach: The essence of this approach lies in deriving the marginal distributions F_X and F_Y along with the dependence structure from the joint distribution $F_{X;Y}$ (*See Fig. 1*).

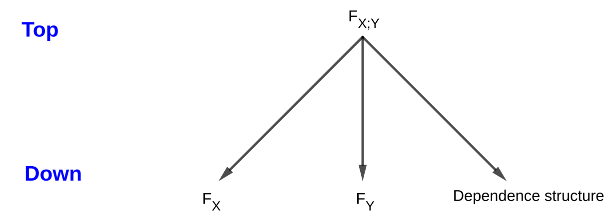

Fig. 1. Top-Down approach.

2) Bottom-Up approach: The essence of this approach is to infer the joint distribution $F_{X;Y}$ from the marginal distributions F_X and F_Y along with the dependence structure. In other words, this approach operates in contrast to the Top-Down approach (*See Fig. 2*).

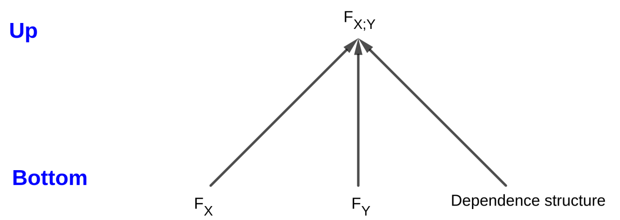

Fig. 2. Bottom-Up approach.

B. Copulas

The copula serves as a function that connects individual marginal distributions to construct the joint distribution, making its application a quintessential "Bottom-Up" approach with the primary objective of modeling dependence relationships.

The following definitions, related to copulas, can be found in [13], [14], [15], [16], [17].

Definition 2.5: Let X and Y be two uniform random variables. We call copula (of dimension 2 or 2-dimensional) any joint cumulative distribution function C whose marginals are the uniform distribution on $[0;1]$. It is defined by,

$$(\forall (x;y) \in [0;1]^2) : C(x;y) = \mathbb{P}(X \leq x; Y \leq y).$$

Definition 2.6: The density c of a bivariate copula function C, if it exists, is defined as follows,

$$(\forall (x;y) \in [0;1]^2) : c(x;y) = \frac{\partial^2 C(x;y)}{\partial x \partial y}.$$

Theorem 2.1: (Sklar's Theorem [9])

Let X and Y be two random variables with joint distribution function $F_{X;Y}$ and F_X, F_Y the marginal distribution functions of X and Y respectively.

There exists a bi-copula C such that, for all element $(x;y)$ in $\mathbb{R}^2$, $F_{X;Y}(x;y) = C(F_X(x); F_Y(y))$.

If F_X and F_Y are continuous, then C is uniquely defined.

Remark 2.1:

1) The function defined above is a joint cumulative distribution function of marginals F_X and F_Y.
2) If F_X and F_Y are continuous, then,
$$C(x;y) = F_{X;Y}(F_X^{-1}(x); F_Y^{-1}(y)).$$
3) By direct calculation, if $F_{X;Y}$ is an absolutely continuous function, then the density f of $F_{X;Y}$ is given, for almost all $(x;y) \in \mathbb{R}^2$, by,
$$f(x;y) = c(F_X(x); F_Y(y)) f_X(x) f_Y(y).$$
Therefore, the density of a random vector $(X;Y)$ can be expressed as a function involving the density of its copula and the cumulative distribution functions of its marginals.

Theorem 2.2: (Invariance theorem)

Let C be the copula function associated with the random vector $(X;Y)$. Let f and g be two strictly increasing functions on $Im(X)$ and $Im(Y)$ respectively.

We can therefore establish the following equality, $C(f(X); g(Y)) = C(X;Y)$.

This study primarily employs three copulas: Gumbel copula, Frank copula, and Clayton copula, as documented in references [10] and [11]. The parameter θ signifies the degree of dependence between the variables.

- A higher θ corresponds to a more pronounced dependence.
- Positive values of θ indicate a positive dependence induced by the copula.

1) Gumbel copula: The Gumbel copula is employed in insurance to analyze the effects of high-intensity events on the correlation between risk levels. It is defined by,

$(\forall \theta > 1)(\forall x; y \in]0;1]) :$
$$C(x,y) = \exp\{-[(-\ln x)^\theta + (-\ln y)^\theta]^{\frac{1}{\theta}}\}.$$

This copula enables the consideration of positive dependencies and the representation of risks characterized by a more pronounced dependence structure in the upper tail.

2) Frank copula: The Frank copula is defined by,

$(\forall \theta \neq 0)(\forall x; y \in [0;1]) :$
$$C(x,y) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta x}-1)(e^{-\theta y}-1)}{(e^{-\theta}-1)}\right).$$

This copula allows the modeling of negative dependencies. However, the latter does not present any dependence on the distribution tails.

3) Clayton copula: The Clayton copula allows for the consideration of positive dependencies and the representation of risks characterized by a more pronounced dependence structure in the lower tail. It is defined by,

$$(\forall \theta > 0)(\forall x; y \in [0;1]) : C(x,y) = (x^{-\theta} + y^{-\theta} - 1)^{\frac{-1}{\theta}}.$$

III. Presentation and Preparation of the Data to be Studied

In the context of this research, the dataset available to us includes information from a leading insurance company in the Moroccan market, specifically "AXA Insurance," covering data for individuals, professionals, and companies. The utilized database contains claims from the years 2005 to 2018, offering a detailed overview via a 13-year occurrence histogram.

Our research primarily centers on evaluating the risk associated with variations in the claims burden within the portfolio, a pivotal factor that can substantially influence the financial stability of a non-life insurance company. To exemplify the dataset, we have considered a randomly selected sample comprising 105 data points. Our sample is limited to four crucial branches within AXA Insurance: work accidents (AT), automobile civil liability (AUTO), disability (INCAPACITE), and fire, marine, aviation, and transport (AUTOFLOT). The selection of these specific branches stems from their significant relevance in comprehending and addressing potential threats to the insurer's solvency. The sample was chosen randomly, ensuring the inclusion of data from all accident years, thereby avoiding bias toward either the most recent or the oldest years. This random approach was adopted to ensure that the observed correlation within the sample accurately mirrors the correlation present across the entire dataset. The functions of the R code are accessible from [28], [29].

Figure 3 provides an excerpt of the dataset, offering a concise yet informative snapshot of the data.

	dev	surv	AT	INCAPACITE	AUTO	AUTOFLOT
1	2005	2005	1145107	129155.4	2594115	1033974.2
2	2006	2005	3910952	381857.5	10079082	1476169.0
3	2007	2005	6609556	502549.5	17899666	1927735.8
4	2008	2005	8153475	604454.6	24255763	2318378.6
5	2009	2005	9845033	691634.1	27614771	2487257.2
6	2010	2005	11398727	772469.4	29009101	2590459.1
7	2011	2005	12152911	847517.3	30271914	2651549.2
8	2012	2005	12695081	943499.7	31219843	2706352.4
9	2013	2005	12989676	1150469.6	31757728	2707235.3
10	2014	2005	13343884	1212045.5	32475805	2720460.8
11	2015	2005	13726681	1260269.1	32741817	2727036.5
12	2016	2005	13972597	1297816.3	32946211	2738159.9
13	2017	2005	14105824	1346467.0	33106509	2745255.0
14	2018	2005	14217341	1374709.6	33195632	2745653.3
15	2006	2006	1272023	169429.4	3049124	206171.0
16	2007	2006	4821913	446623.5	9272037	970241.3
17	2008	2006	6985127	580382.8	16912242	1577865.1
18	2009	2006	9220214	720232.4	22098330	2007095.4
19	2010	2006	11857206	844281.2	25055254	2185749.1
20	2011	2006	13220560	932559.0	27359726	2280567.5

Fig. 3. Data extract.

The data extracted from the claims database comprises claim amounts that have not undergone revaluation. To ensure consistency across diverse claim costs, we opted to standardize all payment amounts to a common reference date: the year 2018. This adjustment leads to an "as if" statistic, rendering the data comparable to the evaluation years. The process involves indexing the claims and applying the insurance conditions from the valuation year to align the dataset. The underlying principle is that a claim of 100,000 MAD in 2005 is not directly comparable to a claim of 100,000 MAD in 2018 due to the influence of economic developments and inflation on claim costs over time. To rectify this, we adjusted the claims using a pertinent index as a reference point, specifically utilizing the consumer price index, accessible at [30].

Let n be the year of listing, I_k be the index of year k, S_k be the value of a claim in year k and S_k^n be the "as if" value of a claim for year k seen in year n. We have then, $S_k^n = S_k \times \dfrac{I_n}{I_k}$.

As an example, we have,
$''as\,if''\ value =$
$$Claim\ cost\ in\ 2005 \times \frac{Consumer\ price\ index\ in\ 2018}{Consumer\ price\ index\ in\ 2005}.$$
The "as if" statistics are derived for each claim base amount, representing what would be paid if the claims had occurred in 2018. All subsequent analyses, including descriptive statistics, are conducted based on these adjusted amounts.

Figure 4 presents a snippet of the "As if" data, providing a succinct yet informative overview of the dataset.

	dev	surv	AT	INCAPACITE	AUTO	AUTOFLOT
1	2005	2005	2101498	237025.7	6595915	1897547.7
2	2006	2005	7177372	700784.3	18497115	2709062.9
3	2007	2005	12129846	922278.0	32849437	3537777.7
4	2008	2005	14963245	1109294.1	44514137	4254684.6
5	2009	2005	18067589	1269285.7	50678583	4564610.4
6	2010	2005	20918925	1417634.5	53237456	4754006.2
7	2011	2005	22303002	1555362.4	55554967	4866118.8
8	2012	2005	23297992	1731509.1	57294606	4966693.6
9	2013	2005	23838633	2111339.9	58281730	4968313.8
10	2014	2005	24488675	2224343.9	59599545	4992585.2
11	2015	2005	25191183	2312843.9	60087730	5004652.9
12	2016	2005	25642487	2381750.3	60462832	5025066.5
13	2017	2005	25886985	2471034.0	60757012	5038087.4
14	2018	2005	26091641	2522864.8	60920569	5038818.5
15	2006	2006	2334415	310936.5	5595748	378364.8
16	2007	2006	8849166	819642.8	17016027	1780585.2
17	2008	2006	12819093	1065117.5	31037318	2895695.4
18	2009	2006	16920922	1321769.3	40554819	3683418.1
19	2010	2006	21760325	1549423.4	45981361	4011283.1
20	2011	2006	24262350	1711430.7	50210525	4185293.7

Fig. 4. "As if" data.

In a particular segment of the insurance industry, a discernible characteristic is identified concerning the net claim amounts: their independence and consistent adherence to a uniform distribution persist from one fiscal year to the subsequent. This observation suggests that the statistical attributes governing the net claims, including their probability distribution and interdependence, maintain stability over time within the specified branch. Such steadfastness in distribution and independence is fundamental for ensuring the reliability of models and analyses in the realms of risk assessment and evaluations of financial stability.

IV. Dependency research

This section aims to elucidate the interrelationships among our four risk branches from a statistical perspective. Pearson's linear correlation coefficient, Kendall's Tau, and Spearman's Rho are three metrics employed to assess the extent of dependence existing among our four variables.
To further validate the presence of dependence, we subjected our data to rigorous statistical tests of independence [31]. These tests served to either affirm or negate the observed dependence. To visually represent these dependencies, we utilized two graphical tools: a dendrogram and a scatter plot. These visualizations effectively illustrated the predominant dependency structures within our data. Finally, a Kendall plot (or K-plot) [32] made it possible to specify and validate the dependence structures.

A. Dependency measures

Tables I, II, and III showcase the results of Pearson's linear correlation, Spearman's Rho, and Kendall's Tau analyses, respectively. These tables offer valuable insights into the intricate relationships among our four risk branches.

TABLE I
PEARSON'S LINEAR CORRELATION.

	AT	INCAPACITE	AUTO	AUTOFLOT
AT	1	0.6301000	0.9067453	0.8125878
INCAPACITE		1	0.5369057	0.2496704
AUTO			1	0.8811172
AUTOFLOT				1

TABLE II
SPEARMAN'S RHO.

	AT	INCAPACITE	AUTO	AUTOFLOT
AT	1	0.6744039	0.9004147	0.8470350
INCAPACITE		1	0.5585009	0.3740411
AUTO			1	0.9067800
AUTOFLOT				1

TABLE III
KENDALL'S TAU.

	AT	INCAPACITE	AUTO	AUTOFLOT
AT	1	0.5054945	0.7542125	0.6681319
INCAPACITE		1	0.3717949	0.2542125
AUTO			1	0.7593407
AUTOFLOT				1

Highlighting the statistical relationships within our data, we observe a strong linear correlation, with a Pearson coefficient of 0.9067453, between the branches AT and AUTO. Conversely, there appears to be no discernible linear correlation between the INCAPACITE and AUTOFLOT branches.

In alignment with the robust Pearson linear correlation coefficient observed between AT and AUTO, we also find substantial values for both Spearman's Rho (0.9004147) and Kendall's Tau (0.7542125), further emphasizing the strong association between these branches. Conversely, when examining the connections between the INCAPACITE and AUTOFLOT branches, we notice relatively weaker values for Spearman's Rho (0.3740411) and Kendall's Tau (0.2542125), indicative of a less pronounced relationship.

To comprehensively assess the significance of these findings, it is imperative to conduct statistical tests of independence. These tests will determine whether the aforementioned coefficients significantly differ from zero, providing critical insights into the underlying dependencies within our data.

B. Statistical tests of correlation

Tables IV, V, and VI present the P-values obtained for correlation tests at a significance level of 5%. These P-values serve as indicators to assess the statistical significance of the observed correlations, aiding in the interpretation of the strength and reliability of the relationships within the dataset.

TABLE IV
P-VALUE OF THE PEARSON CORRELATION TEST.

	AT	INCAPACITE	AUTO	AUTOFLOT
AT		8.019e-13	$<$ 2.2e-16	$<$ 2.2e-16
INCAPACITE			6.639e-09	0.0102
AUTO				$<$ 2.2e-16
AUTOFLOT				

TABLE V
P-VALUE OF THE SPEARMAN CORRELATION TEST.

	AT	INCAPACITE	AUTO	AUTOFLOT
AT		$<$ 2.2e-16	$<$ 2.2e-16	$<$ 2.2e-16
INCAPACITE			8.481e-10	0.0001783
AUTO				$<$ 2.2e-16
AUTOFLOT				

TABLE VI
P-VALUE OF THE KENDALL CORRELATION TEST.

	AT	INCAPACITE	AUTO	AUTOFLOT
AT		1.513e-14	$<$ 2.2e-16	$<$ 2.2e-16
INCAPACITE			1.9e-08	0.0001814
AUTO				$<$ 2.2e-16
AUTOFLOT				

The significance tests were conducted using the *cor.test* function within the R software, providing a robust statistical framework for assessing the significance of correlations in our analysis [28].

In every conducted test, we have firmly rejected the null hypothesis (H_0) of independence, as indicated by P-values consistently falling below the 5% threshold. This compelling evidence leads us to conclude a substantial correlation among the Pearson, Kendall, and Spearman coefficients within the six pairs under scrutiny.

C. Graphical dependency analysis

The statistical study provided information only on the intensity of the positive or negative dependence of the various correlations. The next step involves conducting a graphical analysis to delve deeper into characterizing the dependencies among these four branches of risk.

Scatter plots.

Scatter plots involve graphically representing the cloud of points $(x_i; y_i)$, with $i = 1; 2; ...; n$.
In cases where data points align along a diagonal line, it signifies a scenario characterized by a linear correlation. Conversely, if the points are uniformly scattered across the plane, this suggests a state of independence.

Figure 5 provides a visual representation through a Scatter diagram, offering a comprehensive view of the relationships within the four couples that have been subject to study. This graphical depiction allows for a nuanced examination of the patterns and correlations present among the variables under investigation.

Although the number of observations is limited for these pairs of risk branches studied, a notable dependency structure becomes apparent. The diagram in figure 5 shows a fairly strong concentration of data points, indicating the presence of interdependence among these six pairs. Nevertheless, the scatter plot analysis alone doesn't provide a more in-depth

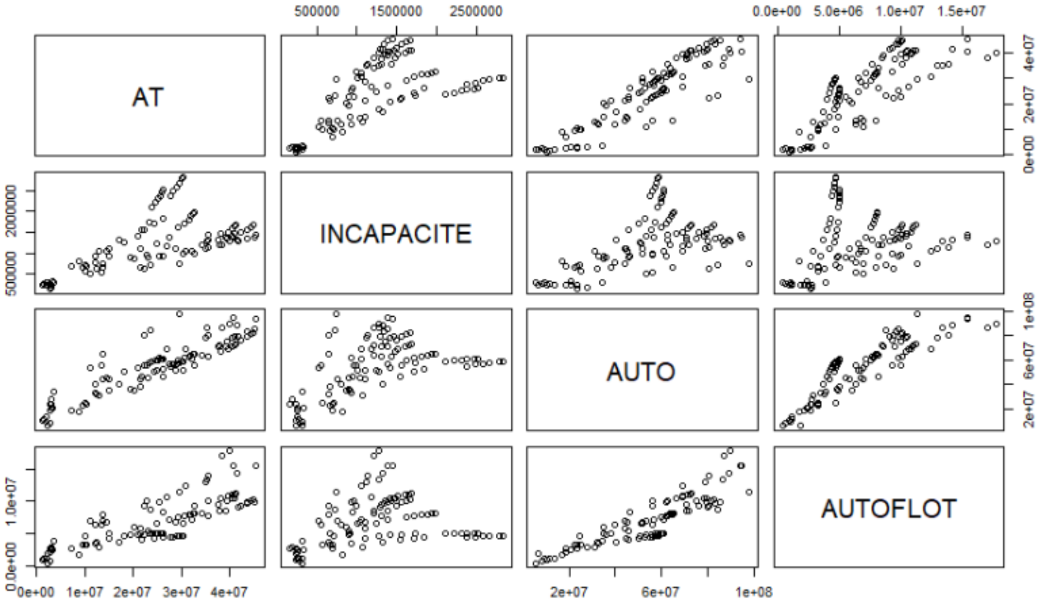

Fig. 5. Scatter diagram for the couples studied.

understanding. It is through the examination of dependograms that we anticipate gaining richer insights into these dependencies.

The dependograms.

The dependogram of the real random variables X and Y is the cloud of points $(F_n(x_i); G_n(y_i))$, for all $i = 1; 2; ...; n$, where F_n and G_n are the empirical distribution functions of X and Y respectively defined by,

$$\forall i \in \{1; 2; ...; n\}: \quad F_n(x_i) = \frac{Rang(x_i)}{n} \in [0; 1],$$

and $\forall i \in \{1; 2; ...; n\}: \quad G_n(y_i) = \dfrac{Rang(y_i)}{n} \in [0; 1].$

According to figure 6, it is evident that, in the context of the AT/AUTO pair, there is a subtle concentration of data points along the upward diagonal, while the upper-left and lower-right corners exhibit an absence of points. This particular distribution implies a mild linear dependence between these branches.

In contrast, when analyzing the other three pairs, we observe strictly positive dependencies, characterized by a distinct alignment of data points along the ascending diagonal. Furthermore, noteworthy tail dependencies exist between specific branches, such as a "top-right" tail dependency between the AUTO/AUTOFLOT branches and the AT/AUTOFLOT branches. This suggests simultaneous occurrences of substantial claims loads between these branches.

The K-Plot:

The graphical depiction presented in figure 7 elucidates the K-plot for the four pairs currently under scrutiny. This visual representation not only offers a comprehensive view of the interactions and interdependencies within the examined couples but also serves as a valuable tool for discerning patterns and trends in the data. The K-plot proves to be instrumental in unraveling the intricacies of relationships, providing a practical utility in analyzing and understanding the dynamics at play.

Upon analyzing the insights gleaned from the presented K-plot, it becomes apparent that a conspicuous positive dependence prevails across all branches. The graphical representation vividly illustrates the robust interconnection and mutual influence characterizing the relationships among the various elements under scrutiny.

D. Discussion and conclusion

The confirmation of dependencies among claim costs associated with the four risk branches under examination has been established through meticulous dependency measurements, independence tests, and various graphical representations. Notably, within the portfolio, a prominent observation emerges: all pairs of branches exhibit significant interdependence. These dependencies, predominantly positive, are highlighted by dependograms that subtly emphasize the upper tails, with no discernible accentuation of dependence observed in the lower tails.

Consequently, with this understanding, we can confidently proceed by narrowing our focus to Archimedean copulas for modeling these dependencies. This deliberate choice allows us to exclude elliptical copulas, specifically designed for symmetrical dependencies, as they are deemed less suitable given the observed asymmetry in the interdependence patterns within the claim costs.

V. Dependency modeling and copula estimation

For each of the four branches, it has been noted that the annual net claims charges demonstrate interdependence and exhibit similar distribution patterns. This observation allows us to propose adjustments to the legislation governing the annual charges associated with these branches. This step is pivotal, as we incorporated these distributions in simulating copulas, subsequently refining them for the portfolio.

The distributions typically employed in non-life insurance to model claims costs encompass the Normal distribution, Log-Normal distribution, Exponential distribution, Gamma distribution, Pareto distribution, and Weibull distribution.

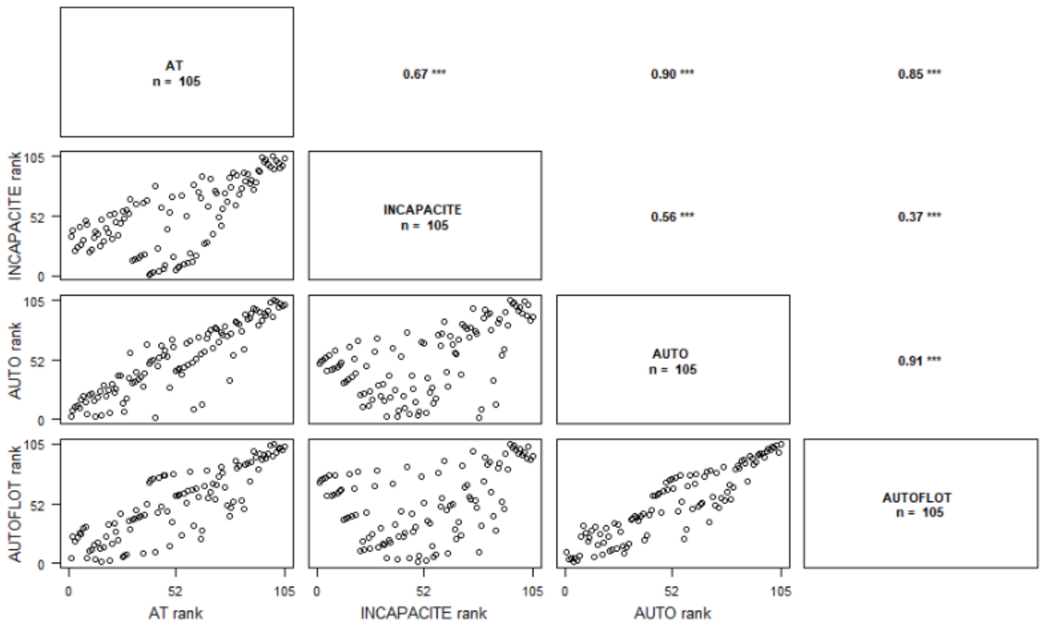

Fig. 6. Dependogram for the couples studied.

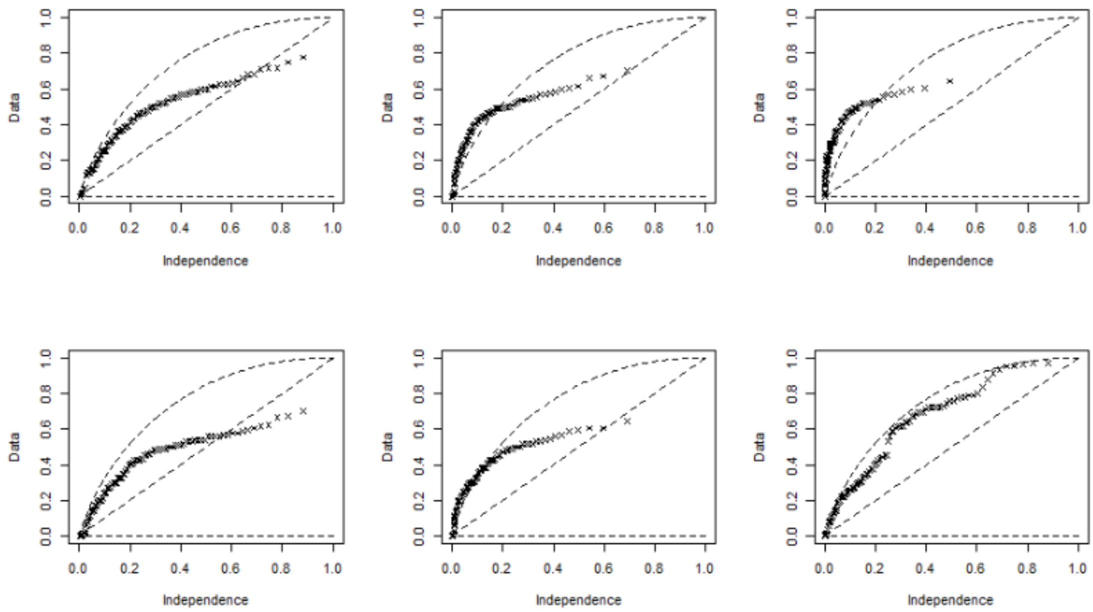

Fig. 7. K-plot for the couples studied.

This parametric exploration entails approximating and estimating the distribution followed by the variable representing claims costs, utilizing well-established probability distributions. We have opted to evaluate the data's compatibility with various distributions commonly used in practical applications, selecting the one that maximizes the Log-likelihood as the preferred choice. The initial estimation involves calculating the densities for each risk branch, as illustrated in figure 8.

The overall shape of the curves permits us to confine our examinations to the following distributions:
$F = \{Normal, Log-Normal, Gamma, Weibull\}$.
Utilizing the maximum likelihood method within the R environment, we have derived the outcomes, as showcased in table VII.

TABLE VII
LogL Value.

Distro tested	AT	INCAPACITE	AUTO	AUTOFLOT
Normale	-1804.681	-1489.07	-1862.983	-1674.072
Log-Normale	-1832.595	-1495.197	-1887.942	-1681.132
Gamma	-	-	-	-
Weibull	-1808.228	-1484.538	-	-1668.712

The Gamma distribution adjustment proves inadequate for meeting the specific requirements of this context, making it unsuitable for application. Shifting our focus to the distribution with the higher Log-likelihood, we discern the appropriate distribution and its associated parameters.

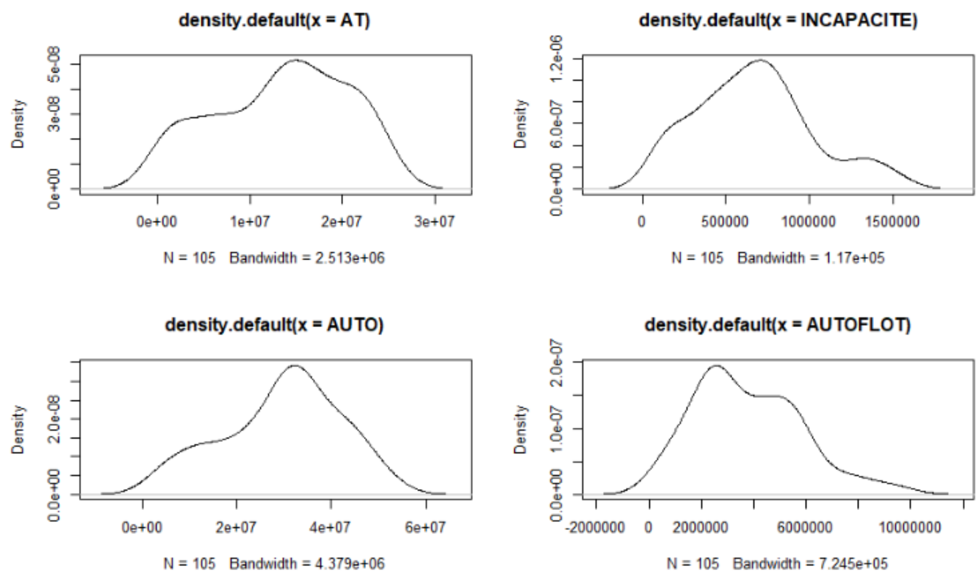

Fig. 8. Estimated density of the four branches.

- The costs associated with the AT branch conform to a Normal distribution with a mean of 13354374 and a variance of 7049728.
- The costs related to the INCAPACITE branch adhere to a Weibull distribution with parameters 1.996363e+00 and 7.547724e+05.
- The costs tied to the AUTO branch are modeled by a Normal distribution with a mean of 29572843 and a variance of 12283369.
- The costs affiliated with the AUTOFLOT branch follow a Weibull distribution with parameters 1.916060e+00 and 4.245075e+06.

Upon juxtaposing the graphs in figure 8 with the estimated densities, a noteworthy disparity becomes evident. Consequently, we conducted a quality assessment of the adjustment using the Kolmogorov-Smirnov test (see table VIII), employing a significance threshold of 5%.

TABLE VIII
TEST RESULT OF KOLMOGOROV-SMIRNOV.

Adjustment	Test statistics	P-value
AT Normale	0.86667	<2.2e-16
INCAPACITE Weibull	0.06808	0.7152
AUTO Normale	0.094121	0.3101
AUTOFLOT Weibull	0.095587	0.2927

The Kolmogorov-Smirnov test is a hypothesis test used here to determine whether a sample follows a given distribution known by its continuous distribution function.
The null hypothesis H_0, which posits that the costs associated with our branches adhere to such a distribution, remains unrefuted in all cases except for the AT branch, where we obtain a P-value significantly below 5%. As a result, the adequacy of this adjustment is deemed insufficient, prompting us to resort to estimating the couples using a semi-parametric method. This choice is driven by our inability to accurately model the marginal distributions, leading us to assume an absence of knowledge regarding their specific forms.

A. Calibration and estimation of the copula

The purpose of this paragraph is to assess the parameters of candidate copulas for modeling the interdependence among the risk branches under consideration. The investigation employs a set of Archimedean copulas, denoted as $C_\theta = \{Clayton; Gumbel; Frank\}$.

To address the limitations associated with the parametric form of the marginals and thereby mitigate modeling risks, we have opted for the non-parametric maximum likelihood method. This approach allows the estimation of parameters independently of the parametric forms of the marginals, a flexibility facilitated by the copulas. It is essential to emphasize that a copula serves to distinguish the structure of dependence between variables induced by the copula from the marginal laws. The estimation process utilizes the $lcopula$ package in the R software, employing the $fitCopula$ function and the mpl method, as detailed in [28].

Method used.

The following outlines the procedure utilized for the calibration and estimation of the copula.

1) Select the copula family to be tested, denoted as C, with $C \in C_\theta$.
2) Estimate the parameter θ of the copula using the maximum likelihood method.
3) Identify the statistic, in this case, a distance metric.
4) Choose the copula that minimizes this distance as the retained copula.

Estimation of the parameters of the tested copulas.

The parameters obtained via the maximum likelihood method are detailed in table IX, showcasing the outcomes of our rigorous statistical analysis.

B. Fit testing

In the subsequent analysis, we fitted a copula to each pair of branches exhibiting significant dependence and juxtaposed the overall claim distributions under both dependent conditions (simulated using the fitted copula) and independent scenarios.

TABLE IX
PARAMETERS ESTIMATED FROM THE MAXIMUM LIKELIHOOD METHOD.

Copula	AT/INCAPACITE	AT/AUTO	AT/AUTOFLOT	INCAPACITE/AUTO
Clayton	2.02	3.815	3.027	1.358
Gumbel	1.518	2.933	2.34	1.361
Frank	5.001	13.57	9.395	3.875

Copula	INCAPACITE/AUTOFLOT	AUTO/AUTOFLOT
Clayton	0.9745	4.528
Gumbel	1.192	3.263
Frank	2.424	12.36

TABLE X
CRAMÉR-VON MISES STATISTIC VALUE.

Copula	AT/INCAPACITE	AT/AUTO	AT/AUTOFLOT	INCAPACITE/AUTO
Clayton	0.071544	0.08002	0.077774	non-finie
Gumbel	0.32849	0.080816	0.099357	0.32033
Frank	0.1653	0.02415	0.05133	0.19681

Copula	INCAPACITE/AUTOFLOT	AUTO/AUTOFLOT
Clayton	non-finie	0.074506
Gumbel	0.18854	0.069484
Frank	0.12342	0.051382

The Cramér-Von Mises statistic is defined by, $\sum_{i=1}^{n}\left[C_{\theta_n}\left(\frac{R_i}{n+1};\frac{S_i}{n+1}\right)-C_n\left(\frac{R_i}{n+1};\frac{S_i}{n+1}\right)\right]^2$, where C_n be the empirical copula and C_{θ_n} be the copula with parameter θ_n, simulated thanks to the function *pcopula* under R. S_n gauges the disparity between the empirical copula and the one subjected to testing and fitting. Consequently, a diminished distance signifies superior adjustment, making the copula associated with the smallest distance the preferred choice. The results are succinctly summarized in table X.

This test can be performed using the *gofCopula* function of the *Copula* package of the R software [28].

Remark 5.1:

1) In the copula evaluations for the AT/INCAPACITE pair, Clayton's copula stands out as the most appropriate choice.
2) Frank's copula demonstrates superior performance in the assessments conducted for the AT/AUTO, AT/AUTOFLOT, INCAPACITE/AUTOFLOT, INCAPACITE/AUTO, and AUTO/AUTOFLOT pairs.

These results are not surprising if we refer to the dependogram. The adequacy of these results was checked with the *BiCopSelect* function of the *Copula* package under the R software [28]. This function makes it possible, for given ranks, to return the best copula in the sense of the Akaike information criterion (AIC). The identical outcomes for the copulas AT/INCAPACITE and AT/AUTO were achieved using this integrated function within R. For the other copulas, the function favors the Gumbel copula for the couple AT/AUTOFLOT, Clayton's copula for the couple INCAPACITE/AUTO, Clayton's copula for the couple INCAPACITE/AUTOFLOT, and Gumbel's copula for the couple AUTO/AUTOFLOT.

C. Context and modeling assumptions

Actuaries face a fundamental challenge in determining the minimum equity value objectively. To achieve this, it is crucial to construct a robust mathematical model known as the risk model. This model involves depicting the company's financial status or capital at the conclusion of a period based on the probability of ruin accepted by the insurer.

Model description.

In this model, we denote:

- S: the total claims burden of the portfolio over the period.
- FP: the capital or reserve available at the start of the period.
- P: the pure premium for the period associated with S.
- π: the risk premium for the period associated with S.
- B: the technical benefit.
- RF: the financial result or the capital at the end of the period.

Model assumptions.

We adopt the following assumptions:

- The time horizon under consideration is one year, corresponding to the calendar year.
- The sole random variable in the model is the annual claims burden, and it is considered gross of reinsurance.
- The insurance company commences operations at the onset of the period with an initial reserve or equity amounting to FP monetary units (where FP $>$ 0), assumed to be equivalent to the solvency margin.
- The combined cost of claims associated with two risk branches is, in reality, the total of the claim costs associated with each of these two branches.

We have then, $P=\mathbb{E}(S)$; $\pi=P+loads=(1+7\%)P$ and $B=\pi-S$.

The annual result called X is given by,

$$\begin{aligned} X &= technical\,profit + financial\,result \\ &= \pi - S + (2\%\times FP + 3\%\times\pi - 1.5\times S) \\ &= 0.02\times FP + 1.03\times\pi - 1.015\times S. \end{aligned}$$

It is important to note that the initial equity level at the start of the period has minimal effect on the variable X. X is a stochastic variable with an undisclosed distribution, and its behavior is approximated empirically through the conducted simulations.

D. Comparison of the distributions obtained by considering the dependence and independence

To contrast the distributions obtained when considering the dependence between branches with those assumed to be independent, we assessed each scenario through:

- Empirical features of the induced distributions.
- Outcomes of the capital requirement assessment model, computed using risk measures such as VaR (Value at Risk) and TVaR (Tail Value at Risk).

Tables XI and XII present the empirical characteristics of the distributions obtained with the copulas selected in section V-B, along with comparisons to the scenario where branches are treated as independent.

TABLE XI
MEAN AND STANDARD DEVIATION OF AGGREGATE CLAIMS DISTRIBUTIONS (DEPENDENT CASE).

Dependent case	Mean (in MAD)	Standard deviation
AT/INCAPACITE	7021957	8106885
AT/AUTO	21327521	12705221
AT/AUTOFLOT	8536537	7066780
INCAPACITE/AUTO	15015863	16641610
INCAPACITE/AUTOFLOT	2202613	2122780
AUTO/AUTOFLOT	16654465	15567569

TABLE XII
MEAN AND STANDARD DEVIATION OF AGGREGATE CLAIMS DISTRIBUTIONS (INDEPENDENT CASE).

Independent case	Mean (in MAD)	Standard deviation
AT/INCAPACITE	7038903	8123682
AT/AUTO	21437890	12829351
AT/AUTOFLOT	8541502	7076901
INCAPACITE/AUTO	15065758	16787444
INCAPACITE/AUTOFLOT	2229621	2142270
AUTO/AUTOFLOT	16662892	15596804

It's worth noting that considering dependence through an Archimedean copula between the costs of two branches in determining the distribution of their aggregate loss load has the following effects:

- Decreases the mean of the aggregate claims charge.
- Diminishes the standard deviation of the distribution.

Table XIII provides a comprehensive depiction of the outcomes derived from the capital requirement assessment model. These results are meticulously calculated employing two prominent methodologies: Value at Risk (VaR) and Tail Value at Risk (TVaR). Through this table, we delve into a detailed analysis of the implications of these methodologies on our capital requirements, offering insights into the potential risks associated with our operations.

Discussion and conclusion.

Let us delve into a specific example, focusing on the AT/INCAPACITE pair within a scenario that accounts for dependency. In this case, the one-year horizon Value at Risk (VaR) is computed as 4700226 MAD with a 99% confidence level. This implies a 99% probability that the portfolio will not incur a loss exceeding 4700226 MAD over the next year. Conversely, there exists only a 1% chance that the loss will surpass this threshold. In contrast, under the independent scenario, the VaR amounts to 3058069 MAD.

These findings indicate that integrating dependency modeling via copula theory results in a marginal increase in the capital requirement. This observation remains consistent across two pivotal risk metrics (VaR and TVaR) and various confidence thresholds (99% and 99.9%).

VI. CONCLUSION

This paper addresses a practical scenario involving the consideration of dependencies among risks within a non-life insurance portfolio. Utilizing copula theory, we investigated the impact of dependencies between claims costs across different insurance branches on the insurer's capital requirement.

TABLE XIII
COMPARISON OF THE VALUES (IN MAD) OF RISK MEASURES OF THE AGGREGATE LOAD DISTRIBUTIONS IN THE DEPENDENT AND INDEPENDENT CASE FOR THE COUPLES STUDIED.

COUPLE (Copula)	$VaR_{99\%}$	$VaR_{99.9\%}$	$TVaR_{99\%}$	$TVaR_{99.9\%}$
AT/INCAPACITE (Clayton's copula)				
Dependent case	4700226	11139383	7465478	13258919
Independent case	3058069	8177050	5355384	10028852
Difference (in %)	53.7%	36.2%	39.4%	32.2%
AT/AUTO (Frank's copula)				
Dependent case	-1440575	9248083	3254814	12517139
Independent case	-9963186	1362474	-5135628	5025902
Difference (in %)	85.5%	578.77%	163.37%	149.05%
AT/AUTOFLOT (Frank's copula)				
Dependent case	10284096	16545389	13081034	18599346
Independent case	5753096	12298494	8548691	14499514
Difference (in %)	78.76%	34.53%	53.02%	28.28%
INCAPACITE/AUTO (Frank's copula)				
Dependent case	-80879.02	9014534.00	4084426	12486668
Independent case	-973469.5	8307161.9	3130036	11891462
Difference (in %)	91.7%	8.52%	30.49%	5.01%
INCAPACITE/AUTOFLOT (Frank's copula)				
Dependent case	10411388	13839275	11867860	14952250
Independent case	10007899	12989761	11413690	14306610
Difference (in %)	4.03%	6.54%	3.98%	4.51%
AUTO/AUTOFLOT (Frank's copula)				
Dependent case	6174742	15766206	10448113	19301208
Independent case	757697.2	11162957.3	5217391	14343313
Difference (in %)	714.94%	41.24%	100.26%	34.57%

Our findings revealed that when the assumption of independence is not validated, the insurer tends to marginally underestimate the capital requirement for risk aggregation, particularly in cases where risks exhibit positive dependence. This discrepancy underscores the need to reevaluate the significance of incorporating dependencies among risk branches within the studied portfolio.

REFERENCES

[1] Bulletin officiel, $N^o 5054-2$ Ramadan 1423 (7-11-2002).

[2] Autorité de contrôle des assurances et de la prévoyance sociale. Code des assurances. https://www.acaps.ma/en.

[3] Li, D. X., "On default correlation: A copula function approach," *The Journal of Fixed Income*, vol. 9, no.4, pp. 43-54, 2000. http://dx.doi.org/10.2139/ssrn.187289.

[4] Embrechts, P., McNeil, A., Straumann, D., "Correlation and dependence in risk management: properties and pitfalls. Risk managemen," *Cambridge University Press*, pp. 176-223, 2002.

[5] Frey, R., McNeil, A. J., "Dependent Defaults in Models of Portfolio Credit Risk," *Journal of Risk*, vol. 6, no. 1, pp. 59-92, 2003.

[6] Embrechts, P., Lindskog, F., Mcneil, A., "Modelling Dependence with Copulas and Applications to Risk Management," *Handbook of Heavy Tailed Distributions in Finance*, North-Holland, vol. 1, pp. 329-384, 2003.

[7] Rosenberg, J. V., Schuermann, T., "A general approach to integrated risk management with skewed, fat-tailed risks," *Journal of Financial Economics*, vol. 79, no. 3, pp. 569-614, 2006. DOI: 10.1016/j.jfineco.2005.03.001.

[8] Siburg, K. F., Stoimenov, P., Weiß, G. N., "Forecasting portfolio-Value-at-Risk with nonparametric lower tail dependence estimates," *Journal of Banking & Finance*, vol. 54, pp. 129-140, 2015. DOI: 10.1016/j.jbankfin.2015.01.012.

[9] Sklar, M., "Fonctions de répartition à n dimensions et leurs marges," *Annales de l'ISUP*, vol. 8, no. 3, pp. 229-231, 1959.

[10] Joe, H., "Dependence Modeling with Copulas," *CRC press*, 2014.

[11] Nelsen, R. B., "An introduction to copulas," *Springer science & business media*, 2007. DOI: 10.1007/0-387-28678-0.

[12] Cherubini, U., Luciano, E., Vecchiato, W., "Copula methods in finance," *John Wiley & Sons*, 2004. DOI:10.1002/9781118673331.

[13] Jaworski, P., Durante, F., Hardle, W. K., Rychlik, T., "Copula theory and its applications," *Springer*, 2010. DOI: 10.1007/978-3-642-12465-5.

[14] Trivedi, P. K., Zimmer, D. M., "Copula modeling: an introduction for practitioners," *Foundations and Trends® in Econometrics*, vol. 1, no. 1, pp. 1-111, 2007. DOI: 10.1561/0800000005.

[15] Cherubini, U., Mulinacci, S., Gobbi, F., Romagnoli, S., "Dynamic copula methods in finance," *John Wiley & Sons*, 2011.

[16] Durante, F., Sempi, C. "Principles of copula theory," *CRC press*, 2015. DOI: 10.1201/b18674.

[17] Fabien, F., Perrin, T. T., Neptune-La Defense, T., "Copula: a new vision for economic capital and application to a four line of business company," *Astin Conference*, 2003. http://www.actuaries.org/astin/colloquia/berlin/faivre.pdf

[18] Li, C., Hao, H., "A copula-based degradation modeling and reliability assessment," *Engineering Letters*, vol. 24, no. 3, pp. 295-300, 2016.

[19] Otani, Y., Imai, J., "An Empirical Analysis of the Dependence Structure of International Equity and Bond Markets Using Regime-switching Copula Model," *IAENG International Journal of Applied Mathematics*, vol. 48, no. 2, pp. 191-205, 2018.

[20] Huang, S., Zhou, Q., Wang, H., "Sample Expansion Approach and Simulation Analysis for Low-Capacity Samples Based on Distance-Trend Double Effects," *IAENG International Journal of Applied Mathematics*, vol. 49, no. 4, pp. 521-531, 2019.

[21] Kouaiba, G., Mentagui, D., "Modeling of Market Risk by Elliptical Copulas Under the Condition of Coherence: Theory and Application," *IAENG International Journal of Applied Mathematics*, vol. 52, no. 1, pp. 116-121, 2022.

[22] Diers, D., Eling, M., Marek, S. D., "Dependence modeling in non-life insurance using the Bernstein copula," *Insurance: Mathematics and Economics*, vol. 50, no. 3, pp. 430-436, 2012.

[23] Fang, Y., Madsen, L., "Modified Gaussian pseudo-copula: Applications in insurance and finance," *Insurance: Mathematics and Economics*, vol. 53, no. 1, pp. 292-301, 2013.

[24] Mejdoub, H., Arab, M. B., "Impact of dependence modeling of non-life insurance risks on capital requirement: D-Vine Copula approach," *Research in International Business and Finance*, vol. 45, pp. 208-218, 2018.

[25] Qi, L. M., Yao, R. Y., Zhang, X. Z., Zhang, Y. J., Wang, X. Y., Tao, J. P., Zhu, J. M., "Research on Revenue Insurance Premium Ratemaking of Jujube Based on Copula-Stochastic Optimization Model," *Journal of Mathematics*, vol. 2021, pp. 1-8, 2021. DOI: 10.1155/2021/9959638.

[26] Zaid, M. A., "Correlation and regression analysis," Turkey: The Statistical, Economic and Social Research and Training Centre for Islamic Countries (SESRIC), 2015. https://www.academia.edu/37839237/Correlation_and_Regression_Analysis_TEXTBOOK_ORGANISATION_OF_ISLAMIC_COOPERATION_STATISTICAL_ECONOMIC_AND_SOCIAL_RESEARCH_AND_TRAINING_CENTRE_FOR_ISLAMIC_COUNTRIES_OIC_ACCREDITATION_CERTIFICATION_PROGRAMME_FOR_OFFICIAL_STATISTICS

[27] Spearman, C., "General intelligence, objectively determined and measured," *The American Journal of Psychology*, vol. 15, no. 2, pp. 201-292, 1904. DOI: 10.2307/1412107.

[28] Hofert, M., Kojadinovic, I., Mächler, M., Yan, J., "Elements of Copula Modeling with R," *Springer International Publishing*, 2018. DOI: 10.1007/978-3-319-89635-9.

[29] Copula: Multivariate Dependence with Copulas. http://cran.r-project.org/web/packages/copula/index.html.

[30] Perspective Monde: Outil pédagogique des grandes tendances mondiales depuis 1945. https://perspective.usherbrooke.ca/bilan/servlet/BMTendanceStatPays?langue=fr&codePays=Mar&codeStat=FP.CPI.TOTL.ZG&codeTheme=2.

[31] Hollander, M., Wolfe, D. A., Chicken, E., "Nonparametric statistical methods," *John Wiley & Sons*, 2013.

[32] Genest, C., Boies, J. C., "Detecting dependence with Kendall plots," *The American Statistician*, vol. 57, no 4, pp. 275-284, 2003. DOI: 10.1198/0003130032431.