

**Mémoire présenté devant l'ENSAE Paris
pour l'obtention du diplôme de la filière Actuariat
et l'admission à l'Institut des Actuaires
le 06/03/2023**

Par : **Elhadji Magatte DIAGNE**

Titre : **Agrégation du STEC Réserve dans le portefeuille
non vie AXA France : cas des dépendances entre Lobs**

Confidentialité : NON OUI (Durée : 1 an 2 ans)

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

Membres présents du jury de la filière

Entreprise : AXA France

*Nom : **Mehdy LATRECHE***

Signature :

*Membres présents du jury de l'Institut
des Actuaires*

Directeur du mémoire en entreprise :

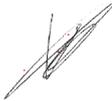
*Nom : **Yanis BENATTIA***

Signature :

**Autorisation de publication et de
mise en ligne sur un site de
diffusion de documents actuariels
(après expiration de l'éventuel délai de
confidentialité)**

Secrétariat:

Signature du responsable entreprise



Bibliothèque:

Signature du candidat



Remerciements

Je remercie Monsieur Mehdy Latreche de m'avoir accueilli au sein de son équipe à la direction risque management d'AXA France IARD et d'avoir supervisé les travaux de ce mémoire.

J'exprime également ma reconnaissance à Monsieur Yanis Benattia, mon tuteur en entreprise qui m'a accompagné tout au long de mon alternance. Je le remercie pour sa disponibilité, son professionnalisme et sa pédagogie. Mes remerciements vont également à l'endroit de Monsieur Damien Primel et tous les autres membres et consultants de l'équipe pour leur disponibilité et leurs conseils.

Je remercie mon responsable pédagogique et enseignant à l'ENSAE, Monsieur Nicolas Baradel qui s'est pleinement investi dans le suivi de mon mémoire.

Enfin, j'exprime ma reconnaissance à mes proches pour leur soutien inconditionnel et à toutes les personnes qui ont contribué de quelque manière que ce soit à l'aboutissement de ce mémoire.

Résumé

L'exposition au risque des compagnies d'assurance et de réassurance et les exigences quantitatives de la réglementation solvabilité 2 (pilier 1) font que ces dernières sont tenues d'avoir une gestion rigoureuse du risque. Pour répondre à ces exigences, AXA France IARD, à travers une approche modèle interne a développé un modèle d'agrégation des risques adapté à son portefeuille pour le calcul du SCR ou STEC (vocabulaire interne).

Dans cette étude, nous nous intéressons à l'estimation des dépendances entre Lobs pour l'agrégation du STEC risque de réserve et la prise en compte de la diversification. Différentes méthodes d'agrégation ont été appliquées pour challenger la méthode initiale qui utilise les paramètres de la FS avec une maille plus fine sous l'hypothèse de corrélation linéaire. Il s'agit principalement des corrélations empiriques basées sur les données du portefeuille avec comme variable les séries de boni mali. Les autres méthodes consistent à une calibration via les copules. Une succession de tests ont également été mis en œuvre pour comparer les méthodes et étudier la stabilité dans le temps.

La méthode choisie est celle des corrélations empiriques calibrées à travers les séries de boni mali. Celles-ci sont ajustées (arrondies et limitées par un seuil inférieur) pour éviter les dépendances négatives et ajouter une marge de prudence. L'étude a montré que le montant du STEC est sensible aux dépendances entre Lobs. La diversification des risques est donc un important levier pour un assureur en solvabilité 2.

Abstract

The risk exposure of insurance and reinsurance companies and the quantitative requirements of Solvency 2 regulations (Pillar 1) mean that the latter are required to have rigorous risk management. To satisfy these requirements, the insurer AXA France IARD, through an internal model approach, has developed a risk aggregation model adapted to its portfolio for the calculation of the SCR or STEC (internal vocabulary).

In this study, we are interested in the estimation of the dependencies between lines of business or Lobs for the aggregation of the STEC reserve risk and the taking into account of diversification. Different aggregation methods have been applied to challenge the initial method that uses the SF parameters with a finer granularity under the linear correlation assumption. These are mainly empirical correlations based on the portfolio data with the boni mali series as variable. The other methods consist of a calibration via the copulas. A series of tests were also implemented to compare the methods and study the stability over time.

The method chosen is that of empirical correlations calibrated through the boni mali series. These are adjusted (rounded and lower bounded) to avoid negative dependencies and add a conservative margin. The study showed that the amount of STEC is sensitive to dependencies between Lobs. Risk diversification is therefore an important leverage for a solvency 2 insurer.

Note de Synthèse

Dans le but de déterminer le besoin en capital (ou STEC dans le vocabulaire d'AXA France), les sociétés d'assurance et de réassurance soumises à la réglementation solvabilité 2 sont tenues d'évaluer les risques liés à leur activité. C'est le cas d'AXA France IARD, qui utilise une approche modèle interne pour la gestion des risques de son portefeuille. Dans le cadre de cette étude, le focus sera mis sur la diversification du risque de réserve en se basant sur de nouvelles méthodes d'agrégation (ou estimation des dépendances) entre les branches ou Lobs du portefeuille.

Définition des concepts

L'agrégation des risques est une étape du processus de mesure des risques au cours de laquelle différents risques sont considérés conjointement afin d'obtenir un profil de risque global (Basel-Committee, 2010).

A travers cette définition découle le concept de diversification qui est lié à la technique d'agrégation des risques. On dira qu'il existe une diversification si la méthode d'agrégation fournit un résultat meilleur que la somme directe des risques marginaux. Une telle diversification s'observe dès lors qu'il existe une hétérogénéité dans un portefeuille.

La diversification est une conséquence directe de la modélisation des dépendances. Elle peut être mesurée par une matrice de corrélation empirique ou plus généralement en utilisant une copule. Il est important de préciser qu'il existe plusieurs formes de dépendance (linéaire / non linéaire, symétrique / asymétrique, de queue, etc.) et que la corrélation n'est qu'une mesure de la dépendance.

Application des méthodes d'agrégation au portefeuille AXA France IARD

Le portefeuille sur lequel porte les travaux comporte 36 Lobs selon la segmentation du modèle interne. Parmi celles-ci, 9 ne disposent pas d'historique de données de sinistres et 2

d'entre elles ont un faible historique. Par rapport aux Lobs de la FS, il existe une correspondance qui permet de faire des choix méthodologiques en cas d'absence de données.

Dans la première méthode (ou méthode initiale), l'agrégation se fait via une matrice de corrélations entre Lobs dont les valeurs sont les paramètres de la FS. On suppose aussi une corrélation de 75% pour les Lobs correspondant à la même Lob FS.

$$STEC_{res} = \sqrt{(STEC_i)_{i=1,\dots,36}^T \times \Sigma \times (STEC_i)_{i=1,\dots,36}} \quad (1)$$

i et j sont les indices des Lobs

$\Sigma = [corr(i, j)]_{i,j \in 1,\dots,36}$ la matrice de corrélations construite

La nouvelle méthode proposée porte sur des corrélations empiriques calculées à partir des séries de taux de Boni Mali par années calendaires. Ces séries sont obtenues en utilisant les triangles de charges de sinistres (paiements sur les sinistres connus + provisions D/D) sur les quels est estimée la volatilité du risque de réserve avec le modèle de Merz & Wuthrich. En effet, en considérant une telle série, les corrélations estimées sont plus susceptibles de correspondre au profil de risque du portefeuille et de refléter la sinistralité. La méthodologie choisie pour la construction de la matrice de corrélation dépend de la structure du portefeuille, la disponibilité des données et de la correspondance par rapport aux Lobs de la FS. Les coefficients empiriques utilisées sont Pearson et Spearman, puis le max des deux sera retenu. C'est cette méthode qui sera privilégiée de part sa simplicité et le fait qu'elle soit applicable sur tout le portefeuille

i et j correspondent à la même Lob FS	Données	Corrélation
X	X	paramètre FS
V	X	75%
V	V	corr(Lob_i , Lob_j)
X	V	corr(Lob FS L , Lob FS K) avec $i \in L$ et $j \in K$

Récapitulatif de la méthodologie de calcul des corrélations

L'agrégation via les copules ont permis d'explorer d'autres formes de dépendances. Les copules estimées sont : la copule de student, les copules archimédiennes Clayton, Frank et Gumbel et la copule multivariée hiérarchique RVine qui est plus générale. Cependant, cette méthode porte sur 25 Lobs en raison du manque de données et la nécessité des distributions marginales pour l'estimation. C'est pourquoi, les résultats obtenus avec ces méthodes serviront de test. En d'autres termes, on identifie la copule qui génère le montant de fonds

propres le plus élevé et on modifie les corrélations empiriques pour être à un niveau plus élevé.

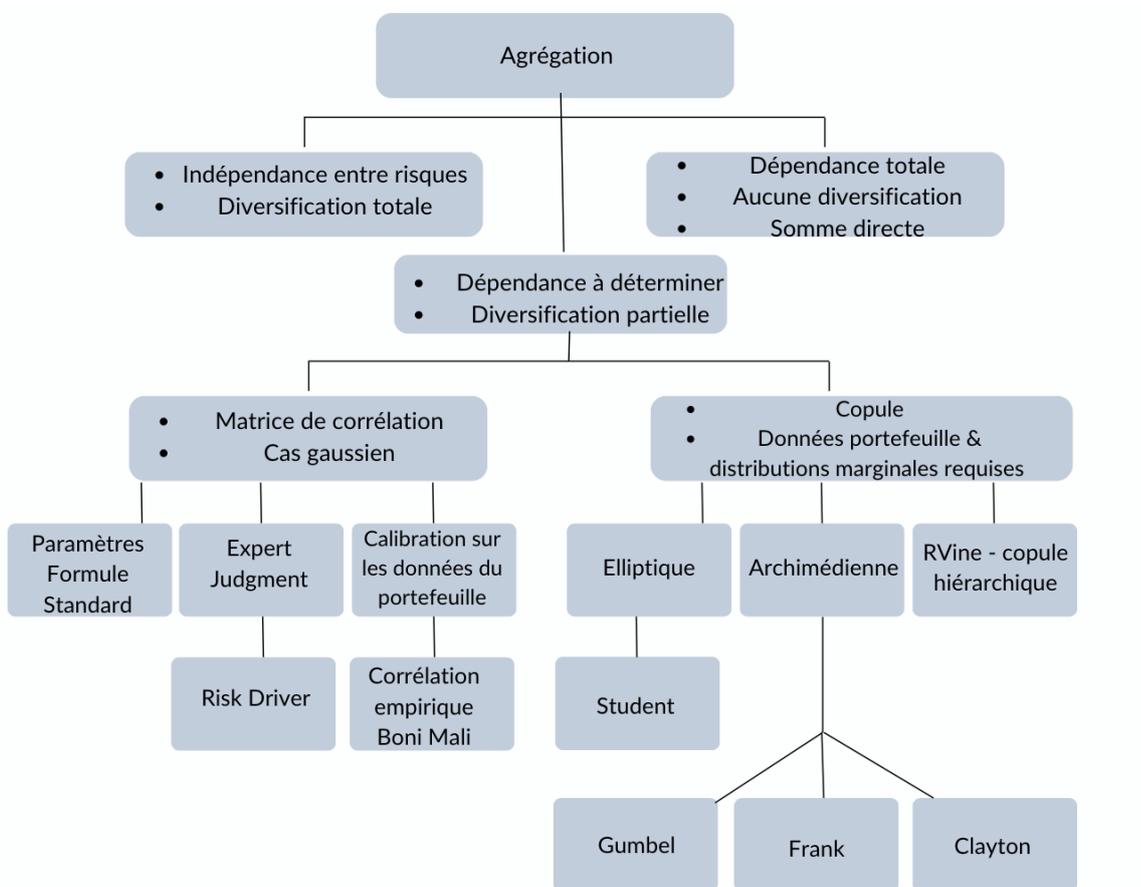


Schéma récapitulatif des méthodes d'agrégation utilisées

Résultats de l'agrégation

Dans cette étude, nous avons défini deux concepts qui nous permettront de déterminer la méthode d'agrégation appropriée :

1. Corrélation équivalente

Elle permet d'obtenir le scénario de dépendance ou niveau équivalent de corrélation entre Lobs fourni par une méthode d'agrégation quelconque. En effet, pour un montant de STEC obtenu par une méthode donnée (matrice de corrélation empirique ou copule), on peut se baser sur la relation suivante pour déterminer la corrélation équivalente ρ correspondante.

$$STEC_{res}^2 = \sum_{i=1}^n x_i^2 + \rho \sum_{i=1}^n x_i S_{-i} \quad \text{avec } S_{-i} = \sum_{j=1, j \neq i}^n x_j \text{ et } x_i = STEC_i$$

Ce paramètre permettra de comparer les différentes méthodes d'agrégation et de réaliser des tests. On remarque qu'il existe une relation quadratique entre la corrélation équivalente et le STEC. Et plus celle-ci est élevée, plus la méthode correspondante est prudente.

2. Paramètres d'arrondi et de seuil inférieur des corrélations empiriques

Ces paramètres sont définis pour ajouter une marge de prudence sur les corrélations empiriques. Arrondir au supérieur les corrélations permet de les rendre stables afin qu'elles puissent être utilisées dans le futur. Le seuil inférieur est fixé pour éviter les dépendances négatives ou trop faibles et de s'assurer que la matrice de corrélation empirique soit définie positive.

Les premiers résultats obtenus portent sur les méthodes empiriques et concernent toutes les Lobs du portefeuille. L'agrégation par les corrélations empiriques obtenues via les taux de boni Mali génère un STEC de l'ordre de 820 millions (corrélation de rang). En termes d'impact, il s'agit d'une baisse de 46% par rapport à la méthode initiale.

Méthode	Standalone	STEC initial	Risk Driver	Correl pearson	Correl rang
STEC (milliards)	2,19	1,53	1,58	0,79	0,82
Impact	-	-	3,1%	-48,2%	-46,3%
Diversification	0,0%	29,9%	27,7%	63,7%	62,3%

Résultats de l'agrégation sur le portefeuille global avec les méthodes empiriques

A première vue, cette méthode est peu prudente pour alimenter les fonds propres et pour palier à cette insuffisance, il convient de trouver la combinaison optimale de paramètres d'arrondi et de seuil inférieur à ajouter.

Pour trouver ces paramètres, nous allons dans un premier temps comparer les résultats d'agrégation de cette approche empirique avec ceux obtenus via les copules (autres formes de dépendances). Cette comparaison se basera sur les corrélations équivalentes puisque les copules sont calibrées sur une partie du portefeuille.

La méthode initiale correspond à un scénario de dépendance de 44,2%. Par conséquent, toute méthode dont la corrélation équivalente est supérieure à ce niveau génère une perte de solvabilité. Pour les corrélations empiriques basées sur les boni mali, on retiendra le max entre le coefficient de pearson et celui des rangs de spearman. Cette méthode correspond à un scénario de -3,3%, alors que pour la copule de Gumbel, on est à 25,5%. En appliquant

les paramètres d'arrondi et de seuil minimal, il est possible de passer de 9,8% (couple (0,0)) à 35% (couple (30,30)).

Methode	Correl Equiv
STEC Standalone	100,0%
STEC initial	44,2%
Max (Correl rang/pearson) param_prud (30,30)	35%
Gumbel	25,5%
Max (Correl rang/pearson) param_prud (0,0)	9,8%
Student	0,2%
Clayton	-1,3%
Frank	-2,4%
RVine	-2,4%
Max (Correl rang/pearson)	-3,3%

Corrélations équivalentes en fonction des méthodes

Au vu de ces résultats, les paramètres optimaux devraient correspondre à un scénario de dépendance au moins équivalent à 25,5%. En se basant sur la matrice ci-dessous on trouve la combinaison (20%,15%).

		Step_Correl						
		0%	5%	10%	15%	20%	25%	30%
Limite_Correl	0%	10%	11%	12%	14%	15%	18%	20%
	5%	13%	14%	14%	16%	18%	19%	20%
	10%	16%	17%	17%	18%	18%	19%	19%
	15%	20%	21%	21%	22%	22%	23%	23%
	20%	24%	25%	25%	26%	26%	27%	27%
	25%	29%	29%	29%	30%	30%	31%	31%
	30%	33%	33%	33%	34%	34%	35%	35%

Les autres étapes consistent à réaliser des tests pour vérifier certains critères :

- Stabilité et Robustesse : il s'agit d'un backtest en considérant les deux années précédentes (2020 et 2019) et une comparaison avec les résultats obtenus via les données comptables.
- Pour tester la précision, des intervalles de confiance ont été construits avec la méthode bootstrap

Les résultats des tests montrent que la combinaison de paramètres choisie permet de vérifier les critères cités. Ainsi, une corrélation minimale de 20% sera retenue et toutes les

autres corrélations seront arrondies à 15% supérieure. La matrice de corrélation empirique qui en découle génère un montant de STEC de 1,28 milliards. Initialement ce montant est de 1,53 milliards, soit une baisse de 16,7%.

STEC agrégé	Impact solva	Diversification	Correl Equiv
1,28	-16,7%	41,6%	28,9%

Les résultats de cette nouvelle méthode sont à l'heure actuelle satisfaisants au regard des hypothèses faites, du gain de diversification observé et du niveau de prudence choisi.

Cependant, la démarche est soumise à des hypothèses plus ou moins fortes et des limites. Il s'agit principalement d'un manque de données. Rappelons que l'estimation des dépendances en grande dimension est une tâche difficile et il faut suffisamment de données pour des estimations précises.

L'étude a montré que le montant du STEC est sensible aux dépendances entre Lobs. La diversification des risques est donc un important levier pour un assureur en solvabilité 2.

Executive Summary

In order to assess the solvency capital requirement (or STEC in AXA France's terminology), insurance and reinsurance companies subject to Solvency 2 regulations are required to measure the risks associated with their business. This is the case for AXA France IARD, which uses an internal model approach to manage the risks in its portfolio. In the context of this study, the focus will be on the diversification of the reserve risk based on new aggregation methods (or dependency estimation) between lines of business or Lobs of the portfolio.

Definition of concepts

Risk aggregation is a step in the risk measurement process in which different risks are considered together to obtain an overall risk profile (Basel-Committee, 2010).

This definition gives rise to the concept of diversification, which is linked to the risk aggregation technique. Diversification exists if the aggregation method provides a better result than the direct sum of the marginal risks. Such diversification is observed whenever there is heterogeneity in a portfolio.

Diversification is a direct consequence of dependency modeling. It can be measured by an empirical correlation matrix or more generally by using a copula. It is important to specify that there are several forms of dependency (linear/non-linear, symmetrical/asymmetrical, tail, etc.) and that correlation is only one specific measure of dependency.

Application of the aggregation methods to the AXA France IARD portfolio

The portfolio studied includes 36 Lobs according to the segmentation of the internal model. Of these, 9 have no historical claims data and 2 of them have little historical data. Compared to the SF (standard formula) Lobs, there is a mapping which allows to make methodological choices in case of absence of data.

In the first method (or initial method), the aggregation is performed via a correlations

matrix between Lobs whose values are the parameters of the SF. We also assume a correlation of 75% for the Lobs corresponding to the same SF Lob.

$$STEC_{res} = \sqrt{(STEC_i)_{i=1,\dots,36}^T \times \Sigma \times (STEC_i)_{i=1,\dots,36}} \quad (2)$$

i and j are Lobs index

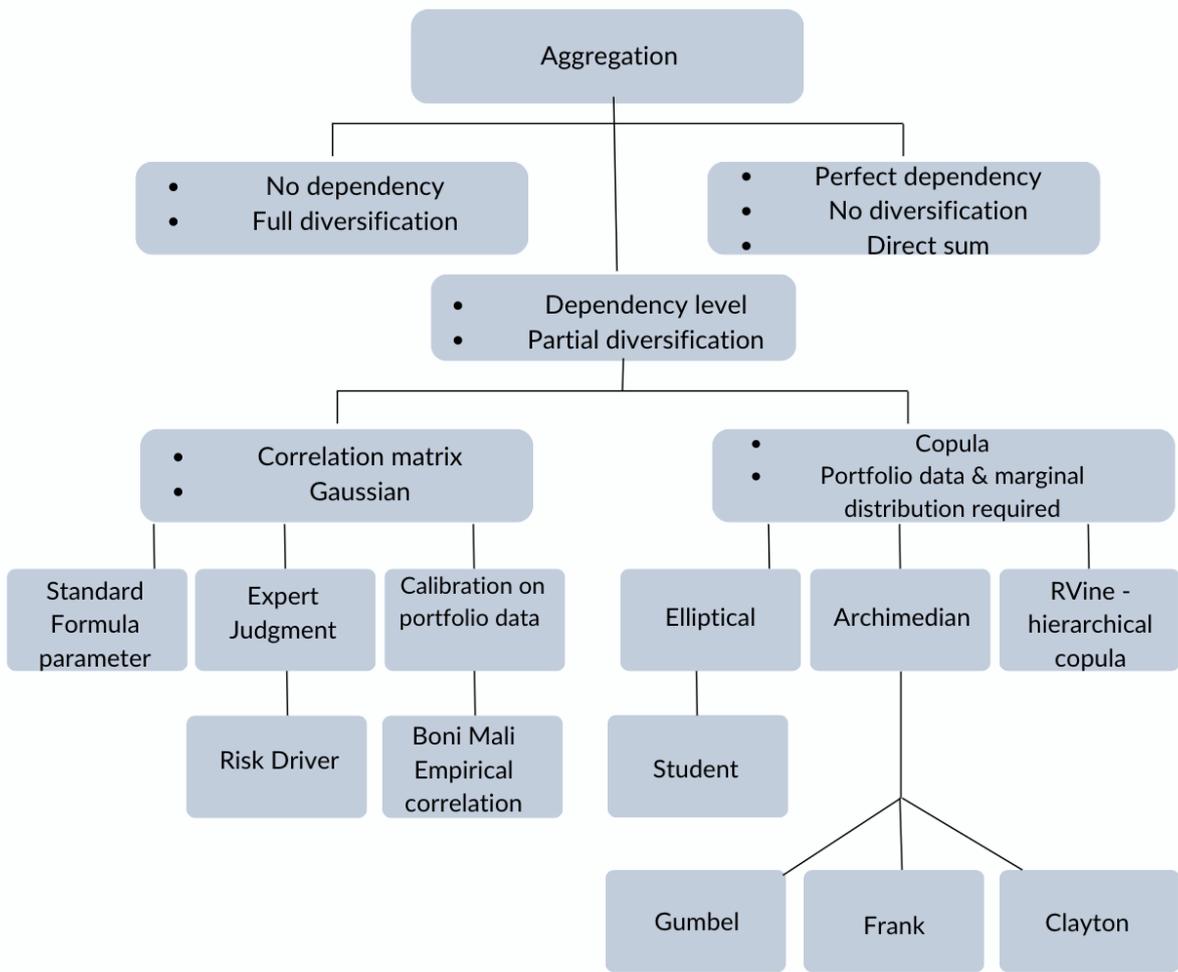
$\Sigma = [corr(i, j)]_{i,j \in 1,\dots,36}$ performed correlation matrix

The new proposed method is based on empirical correlations calculated from the series of Boni Mali rates by calendar year. These series are obtained by using the triangles of incurred claims (payments on known claims + D/D reserves) on which the volatility of the reserve risk is estimated using the Merz Wuthrich model. Indeed, by considering such a series, the estimated correlations are more likely to correspond to the risk profile of the portfolio and to reflect the loss experience. The methodology chosen for the construction of the correlation matrix depends on the structure of the portfolio, the availability of data and the match with the Lobs of the SF. The empirical coefficients used are Pearson and Spearman (rank correlation), then the maximum of these two will be retained. This method will be preferred because of its simplicity and the fact that it is applicable to the whole portfolio.

i and j correspond to the same SF Lob	Data	Correlation
X	X	SF parameter
V	X	75%
V	V	corr(Lob_i , Lob_j)
X	V	corr(Lob SF L , Lob SF K) with $i \in L$ and $j \in K$

Summary of the methodology for calculating correlations

Aggregation via copulas has allowed us to explore other forms of dependencies. The estimated copulas are: the student copula, the archimedean copulas (Clayton, Frank and Gumbel) and the RVine hierarchical multivariate copula which is more general. However, this method is based on 25 Lobs due to the lack of data and the need for marginal distributions for estimation. Therefore, the results obtained with these methods will serve as a test. In other words, the copula that generates the highest amount of STEC is identified and the empirical correlations are modified to be at a higher level.



Summary of the aggregation methods used

Aggregation results

In this study, we have defined two concepts that will allow us to determine the appropriate aggregation method:

1. Single equivalent correlation

It allows us to obtain the dependency scenario or equivalent level of correlation between Lobs provided by any aggregation method. Indeed, for an amount of STEC obtained by a given method (empirical correlation matrix or copula), we can use the following equation to determine the corresponding equivalent correlation ρ .

$$STEC_{res}^2 = \sum_{i=1}^n x_i^2 + \rho \sum_{i=1}^n x_i S_{-i} \quad \text{with } S_{-i} = \sum_{j=1, j \neq i}^n x_j \text{ and } x_i = STEC_i$$

This parameter will allow comparison of the different aggregation methods and for testing purposes. We notice that there is a quadratic relation between the equivalent correlation and the STEC. And the higher this one is, the more conservative the corresponding method is.

2. Parameters of rounding and lower threshold of empirical correlations

These parameters are defined to add a margin of conservativeness to the empirical correlations. Rounding up the correlations makes them stable so that they can be used in the future. The lower threshold is set to avoid negative or too weak dependencies and to ensure that the empirical correlation matrix is positive definite.

The first results obtained concern the empirical methods and cover all the Lobs in the portfolio. Aggregation by the empirical correlations obtained via the Boni Mali rates generates a STEC amount of around 820 millions (rank correlation). In terms of impact, this is a decrease of 46% compared to the initial method.

Method	Standalone	Initial	Risk Driver	Correl pearson	Correl rang
STEC (billions)	2.19	1.53	1.58	0.79	0.82
Impact	-	-	3.1%	-48.2%	-46.3%
Diversification	0.0%	29.9%	27.7%	63.7%	62.3%

Aggregation results on the overall portfolio with empirical methods

At first sight, this method is not conservative enough in feeding the equity and to overcome this shortcoming, it is necessary to find the optimal combination of rounding and lower threshold parameters to add.

To find these parameters, we will first compare the aggregation results of this empirical approach with those obtained via copulas (other forms of dependencies). This comparison will be based on equivalent correlations since the copulas are calibrated on a part of the portfolio.

The initial method corresponds to a dependency scenario of 44.2%. Therefore, any method with an equivalent correlation above this level will generate a solvency loss. For empirical correlations based on boni mali, the max between the pearson coefficient and the spearman rank coefficient is used. This method corresponds to a scenario of -3.3% , whereas for the Gumbel copula, we are at 25.5%. By applying the rounding and minimum threshold parameters, it is possible to go from 9.8% (combination (0,0)) to 35% (combination (30,30)).

Method	Correl Equiv
STEC Standalone	100.0%
STEC initial	44.2%
Max (Correl rang/pearson) param_prud (30,30)	35%
Gumbel	25.5%
Max (Correl rang/pearson) param_prud (0,0)	9.8%
Student	0.2%
Clayton	-1.3%
Frank	-2.4%
RVine	-2.4%
Max (Correl rang/pearson)	-3.3%

Equivalent correlations according to the methods

Based on these results, the optimal parameters should correspond to a dependency scenario at least equivalent to 25.5%. Based on the matrix below, we find the combination (20%, 15%)

		Step_Correl						
		0%	5%	10%	15%	20%	25%	30%
Threshold_Correl	0%	10%	11%	12%	14%	15%	18%	20%
	5%	13%	14%	14%	16%	18%	19%	20%
	10%	16%	17%	17%	18%	18%	19%	19%
	15%	20%	21%	21%	22%	22%	23%	23%
	20%	24%	25%	25%	26%	26%	27%	27%
	25%	29%	29%	29%	30%	30%	31%	31%
	30%	33%	33%	33%	34%	34%	35%	35%

The other steps consist of tests to verify some criteria:

- Stability and Robustness: this is a backtest considering the two previous years (2020 and 2019) and a comparison with the results obtained via the accounting data.
- To test the accuracy, confidence intervals level were calculated with the bootstrap method

The results of the tests show that the chosen combination of parameters allows to verify the mentioned criteria. Thus, a minimum correlation of 20% will be retained and all other correlations will be rounded up to the nearest 15%. The resulting empirical correlation matrix generates a STEC amount of 1.28 billions. Initially this amount is 1.53 billions, a decrease of 16.7%.

Aggregated STEC amount	Solvency impact	Diversification	Equiv Correl
1.28	-16.7%	41.6%	28.9%

The results of this new method are currently satisfactory in terms of the assumptions made, the observed diversification benefit and the chosen level of prudence.

However, the approach is subject to more or less strong assumptions and limitations. This is mainly due to a lack of data. Recall that estimating high-dimensional dependencies is a difficult task and sufficient data are needed for accurate estimates.

The study showed that the amount of STEC is sensitive to the dependencies between Lobs. Risk diversification is therefore an important lever for a solvency 2 insurer.

Table des matières

1	Généralités sur l'agrégation de risques et la diversification	3
1.1	Concepts théoriques d'agrégation de risque	4
1.1.1	La notion de risque	4
1.1.2	Mesure de risque	4
1.1.3	Agrégation et diversification du risque	6
1.1.4	Lien entre dépendance et diversification	8
1.2	Agrégation et diversification en Solvabilité 2	9
1.2.1	Cadre réglementaire Solvabilité 2	9
1.2.2	Diversification en assurance dommage	10
1.2.3	Agrégation en formule standard	13
1.2.4	Cas du modèle interne AXA France P&C	16
2	Modélisation de la dépendance entre risques	21
2.1	Concepts théoriques de dépendance statistique	21
2.1.1	Caractérisation et propriétés	21
2.1.2	La corrélation comme mesure de dépendance	22
2.1.3	Dépendance de queue	26
2.1.4	Précision des corrélations empiriques et intervalle de confiance	26
2.2	La théorie des copules	28
2.2.1	Généralités, définitions et théorème de Sklar	28
2.2.2	Les familles de copule	29
2.2.3	Estimation paramétrique	33
2.2.4	Cas des copules hiérarchiques : RVine Tree Copula	36
3	Application au portefeuille AXA France IARD : méthodes initiales d'agrégation du STEC réserve	41
3.1	Rappels sur la modélisation du risque de réserve	41
3.1.1	Le risque de réserve	42
3.1.2	Charges ultimes, CDR et volatilité des réserves	43

3.1.3	Modélisation des Boni Mali	46
3.2	Dépendance inter Lob basée sur les paramètres de la FS et sur les Risk Driver	47
3.2.1	Caractéristiques des Lobs du portefeuille d'AXA France IARD	48
3.2.2	Méthode d'agrégation initiale : matrice de corrélation avec les paramètres de la FS	52
3.2.3	Méthode d'agrégation basée sur les scénarios Risk Driver	54
4	Nouvelles méthodes de modélisation des dépendances inter Lob et d'agrégation du STEC réserve	57
4.1	Méthodes empiriques basées sur les taux de boni mali	57
4.1.1	Construction de la base de données des taux boni-mali : traitements et calculs	57
4.1.2	Corrélations basées sur les séries de taux de boni mali	61
4.1.3	Construction de la matrice de corrélation inter Lob	62
4.1.4	Agrégation et corrélation équivalente	67
4.2	Calibration des dépendances via différentes familles de copules	69
4.2.1	Données et paramètres des distributions marginales	69
4.2.2	Calibration des copules elliptiques et archimédiennes	71
4.2.3	Calibration de la copule hiérarchique RVine	75
5	Résultats de l'agrégation du STEC réserve et analyse de la diversification	78
5.1	Résultats d'agrégation du STEC suivant les différentes méthodes	78
5.1.1	Agrégation avec une matrice de corrélation empirique	78
5.1.2	Agrégation par les copules	81
5.2	Paramètres de prudence et corrélation équivalente	83
5.2.1	Optimisation des paramètres de prudence (arrondi et seuil inférieur	83
5.2.2	Corrélations équivalentes	85
5.3	Tests et Choix du niveau de dépendance	88
5.3.1	Backtests sur données historiques et données comptables	88
5.3.2	Résultats des tests et choix optimal des paramètres	89
5.4	Résultats de l'agrégation sur le portefeuille global	92
5.4.1	Présentation des résultats	92
5.4.2	Bilan	93

A	Liste des annexes	97
A.1	Algorithme : Calcul de la densité de la copule RVine	97
A.2	Algorithme : Méthode séquentielle de sélection du modèle RVine	98
A.3	Algorithme Bootstrap pour les intervalles de confiance avec R	98
A.4	Corrélations empiriques avec paramètres de prudence optimaux	100
A.5	Intervalle de confiance des corrélations	102
A.6	Coefficient de dépendance de queue à partir de la copule de student	104
A.7	Matrice des arbres de la copule hiérarchique RVine	105
A.8	Matrice des types de copules et paramètres de la copule hiérarchique RVine .	105
	Bibliographie	107

Liste des figures

1.1	Illustration du lien entre corrélation et diversification	9
1.2	Modules et sous-modules de risques du SCR	14
1.3	Avantages et inconvénients : Formule standard vs Modèle interne	18
1.4	Vision globale du calcul du STEC	19
2.1	Plot entre U et V	24
2.2	Exemple d'une structure R-vine de dimension n=7	37
2.3	Inputs de l'algorithme : matrice des constraint set, types et paramètres des copules	39
3.1	Dynamique de vie d'un sinistre	43
3.2	Illustration du CDR	45
4.1	Densité de la copule de Student : exemple sur 3 paires de Lobs	74
5.1	Montant du STEC réserve avec 25 Lobs suivant différentes méthodes	82
5.2	Montant de STEC en milliards en fonction du niveau de corrélation équivalente	86

Liste des tableaux

2.1	Fonctions générateurs et expressions des copules archimédiennes dimension 2	32
2.2	Dépendance de queue et tau de Kendall	32
3.1	Les Lobs du portefeuille P&C AXA France et leurs poids en volumes de primes	50
3.2	Table de correspondance entre Lob AXA et Lob Formule Standard	51
3.3	Corrélations entre Lob Formule Standard	52
3.4	Corrélations Lobs modèle interne basées sur celles des Lobs FS	53
3.5	Liste des 16 risk driver	54
3.6	Matrice qualitative des impacts	55
3.7	Matrice de corrélation suivant l'approche Risk Driver	56
4.1	Traitements de données sur les corrélations de rang	62
4.2	Corrélations inter Lob au sein d'une même Lob FS avec historique de taux de boni mali	64
4.3	Corrélations entre Lob FS incluant les paramètres de la FS	65
4.4	Récapitulatif de la méthodologie de calcul des corrélations	65
4.5	Matrice de Corrélations inter Lob finale	66
4.6	Paramètres des distributions marginales	71
4.7	Estimation des paramètres de la copule de Student	73
4.8	Estimation paramétrique des copules archimédiennes	75
4.9	Estimation de la copule hiérarchique RVine	76
5.1	Montants du STEC réserve par Lob en 2021	79
5.2	Résultats de l'agrégation sur le portefeuille global avec les méthodes de corrélation empiriques	80
5.3	Impact STEC et Diversification sur les 25 Lobs	82
5.4	Montant de STEC en milliards par les corrélations empiriques arrondies et avec seuil minimal	84
5.5	Impact solvabilité des corrélations empiriques arrondies et avec seuil minimal	84

5.6	Corrélations équivalentes en fonction des méthodes	87
5.7	Matrice des scénarios de dépendance de la méthode des corrélations empiriques arrondies et avec seuil minimale	87
5.8	Back test avec les corrélations estimées sur les données historiques	90
5.9	Test avec les corrélations estimées sur les données issues de la comptabilité	90
5.10	Intervalles de confiance des corrélations : exemple avec 4 Lobs	91
5.11	Matrice de corrélation empirique avec paramètres de prudence optimaux	93

Introduction

Les sociétés d'assurance sont exposées à un certain nombre de risques qui peuvent être observés à différents niveaux de leur portefeuille. Ces risques peuvent entraîner d'énormes pertes et par conséquent, menacer la solvabilité de la compagnie. Cette situation justifie l'intérêt d'une démarche prudentielle encadrée par les directives réglementaires Solvabilité II. Dans ce contexte, les entités concernées (sociétés d'assurance et de réassurance) sont tenues de détenir chaque année un montant de fonds propres leur permettant de couvrir le risque de ruine avec une probabilité de 99,5%. Cet indicateur appelé SCR prend en compte toute la structure du portefeuille avec différents niveaux d'agrégation.

De plus en plus, les réflexions autour de solvabilité II sur la prise en compte de l'effet de diversification dans le calcul du SCR suscite l'intérêt de tous les acteurs de l'activité d'assurance notamment les sociétés d'assurance et de réassurance, les régulateurs (ACPR) et les actionnaires. La diversification permet de réduire le risque global qui découle du processus d'agrégation et des hypothèses sur la structure de dépendance entre facteurs de risque. Ce concept de diversification est fondamental en gestion des risques et présente un cadre méthodologique applicable dans le calcul du SCR. La prise en compte des effets de diversification relève de la spécification de l'article 121.5 de la directive Solvabilité II.

Dans la formule standard, le calcul se base sur l'agrégation des SCR unitaires via la matrice de corrélation linéaire qui suppose implicitement que la distribution conjointe des facteurs de risque est gaussienne (famille elliptique). Etant donné que les facteurs de risques peuvent être différents d'un assureur à un autre, cette hypothèse est remise en cause, ce qui justifie l'intérêt du modèle interne. Plusieurs sociétés d'assurance ont adopté cette approche dans le cadre du calcul de leur SCR sous la supervision et le contrôle de l'ACPR. C'est le cas de AXA France qui, dans son modèle interne prend en compte la diversification. Dans le portefeuille non-vie d'AXA France, cette diversification concerne différents niveaux de risques : au sein d'une ligne de business, entre différentes lignes de business ou entre sous modules de risque (réserve, prime, CAT).

Dans le cadre de ce mémoire le travail portera sur la modélisation des dépendances pour l'agrégation du SCR réserve en mettant l'accent sur la diversification entre lignes de business ou Lobs du portefeuille non-vie. Le principal objectif sera d'estimer ces dépendances en se basant sur les données du portefeuille avec une approche statistique (corrélations empiriques et copules). Ces estimations devront également générer un niveau de fonds propres suffisant conformément aux exigences réglementaires. Ce travail permettra de répondre aux questions suivantes :

Quelle est la méthode adoptée dans le modèle interne AXA France IARD pour prendre en compte la diversification dans le calcul du SCR réserve ?

Et comment peut-on améliorer cette méthode en utilisant les techniques avancées d'agrégation en grandes dimensions permettant une meilleure estimation de la structure de dépendance inter Lobs et une prise en compte des exigences réglementaires ?

Pour répondre à ces questions, le présent document sera structuré en trois parties. Dans la première partie, nous présenterons le cadre général de l'étude. Puis nous ferons un rappel sur les approches de modélisation de la dépendance entre risques. Ensuite, la troisième partie portera sur la description de la méthodologie utilisée pour estimer les dépendances inter Lobs. Enfin, la dernière partie présentera les résultats de l'étude en décrivant le choix de la méthode finale et son impact .

Chapitre 1

Généralités sur l'agrégation de risques et la diversification

En présence d'un portefeuille de dimension élevée, la mesure du risque global requiert des méthodologies de plus en plus sophistiquées. Cette problématique est au coeur des travaux en risk management des institutions financières. Ce sujet suscite également l'intérêt des assureurs qui font face à plusieurs facteurs de risque. Par exemple en assurance dommage, le portefeuille de l'assureur peut contenir plusieurs produits avec diverses garanties proposées aux clients qui ont des profils de risque différents dans un environnement économique et financier de plus en plus volatile. Dans ces conditions, et conformément aux exigences quantitatives de la réglementation Solvabilité 2, une mesure optimale du risque global devient une nécessité.

Plus concrètement, si on note par X_i ($i = 1, \dots, d$) les risques marginaux (cas des pertes liées aux frais ou charges de sinistre), le calcul du risque agrégé peut être effectué en utilisant des mesures de risque telles que la variance totale, la VaR ou TVaR qui ont chacune des propriétés mathématiques. Mais l'un des plus grands défis auquel est souvent confronté l'assureur est la détermination de la distribution jointe ou celle de la somme $S = \sum_{i=1}^d X_i$ en présence d'une dépendance entre les risques marginaux.

L'estimation d'une telle distribution multivariée peut se faire dans un cadre elliptique (hypothèse de distribution Gaussienne ou de Student) comme dans le cas de la Formule Standard. Mais la complexité des structures de dépendance est telle que cette hypothèse est plus ou moins forte (Bernard and Vanduffel, 2016). En effet, les distributions elliptiques ne sont pas adaptées lorsque l'on souhaite obtenir de bonnes estimations de probabilités pour des scénarios extrêmes simultanés (McNeil, Frey, and Embrechts, 2015). Typiquement,

il s'agit du problème d'asymétrie ou Skewness qui constitue l'une des principales limites dans l'application des mesures de risques telles que la VaR.

Dans la suite, nous présenterons les différentes approches de mesures de risques en définissant les notions importantes, les techniques d'agrégation et la diversification. Cette présentation se fera d'un point de vue théorique et en lien avec la réglementation Solvabilité 2.

1.1 Concepts théoriques d'agrégation de risque

1.1.1 La notion de risque

Un risque peut être défini comme "hasard", "chance" avec de mauvaises conséquences, pertes ou exposition à une malchance (*Oxford English Dictionary*), le tout étant liés à des événements aléatoires ou de sources d'incertitudes affectant des actions et des prises de décisions. Bien que la plupart ont une connotation négative du risque, il peut aussi présenter des opportunités pour certains investisseurs ou gestionnaires de risque. En effet, une grande partie du secteur financier n'existerait pas sans la présence du risque financier et les opportunités offertes aux investisseurs et autres acteurs du marché.

D'une manière générale, la notion de risque se prête mal à une définition unanime. C'est un concept qui dépend essentiellement du contexte et des stratégies et objectifs recherchés par les parties prenantes. Ainsi, on peut considérer le risque comme " tout événement ou action susceptible de nuire à la capacité d'une entité à atteindre ses objectifs et à exécuter ses stratégies ". En d'autres termes, il correspond à " une probabilité quantifiable de perte ou de rendement inférieur à celui espéré" (McNeil, Frey, and Embrechts, 2015).

1.1.2 Mesure de risque

En finance et en assurance, les mesures de risque les plus fréquemment utilisées sont la VaR et la TVaR. Dans le cadre général, une mesure de risque est une métrique représentable par une fonction $\rho(\cdot)$ qui fait le mapping entre la variable aléatoire X et la valeur de la perte associée à cette variable. $\rho(X)$ peut être une mesure marginale (pour un seul risque) ou agrégée (diversifiée) (Bernard and Vanduffel, 2016).

On supposera que $X_i \in \Omega, (i = 1, \dots, d)$ qui est un ensemble de pertes aléatoires telles que

$X_i, X_j \in \Omega$ implique que $X_i + X_j \in \Omega$ et pour tous réels $a > 0$ et b $aX_i + b \in \Omega$.

Il existe un certain nombre de propriétés ou axiomes que doit satisfaire la métrique ρ pour être considérée comme mesure de risque cohérente (Artzner et al., 1999).

1. **Invariance par translation** : pour tout $X \in \Omega$ et tout nombre réel α , on a

$$\rho(X + \alpha r) = \rho(X) - \alpha$$

Cette axiome assure que $\rho(X + \rho(X)r) = 0$ et permet de vérifier les hypothèses faites sur Ω

2. **Sous-additivité** : pour tout X_1 et $X_2 \in \Omega$, $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.

L'intuition qui ressort de cette axiome est que le risque global d'un portefeuille ne peut être pire que d'additionner les deux risques individuels : c'est le principe de diversification. En général, dans la gestion des risque en finance, cette axiome de sous-additivité fournit une diversification bénéfique même si dans certains contextes elle peut être une limite.

3. **Homogénéité positive**: pour tout $\lambda \geq 0$ et $X \in \Omega$, $\rho(\lambda X) = \lambda \rho(X)$.

Par exemple en doublant le poids d'un portefeuille, le risque est aussi doublé. Cette axiome implique que le risque d'une position est proportionnel à sa taille.

4. **Monotonicité**: Pour tout X et $Y \in \Omega$ avec $X \leq Y$, on a $\rho(Y) \leq \rho(X)$.

Un portefeuille avec des rendements futurs plus élevés comporte moins de risques.

Par rapport aux mesures de risque les plus connues en finance et en assurance, on peut remarquer que la **VaR** ne satisfait pas l'axiome de sous-additivité. Celle-ci n'est donc pas une mesure de risque cohérente selon Artzner et al., 1999. En effet, la VaR ou Value at Risk de niveau α correspond à la fonction quantile dont les propriétés n'assurent pas cette axiome de sous-additivité.

Par ailleurs, la **TVaR** ou **Expected Shortfall** est bien une mesure de risque cohérente car satisfait les 4 axiomes. Elle s'exprime sous forme d'espérance conditionnelle et décrit la

survenance des pertes sous réserve que le seuil soit franchi. Elle donne également une indication plus significative du montant des pertes extrêmes.

Il est important de rappeler que la nécessité qu'une mesure de risque satisfasse la propriété de sous-additivité ou la cohérence d'une manière générale fait l'objet de plusieurs débats et de sujets de recherche en finance. Ainsi, d'autres critères peuvent aussi être explorés. On peut citer le caractère conservateur de la métrique pour justifier le niveau de prudence conformément au cadre réglementaire dans lequel la gestion du risque s'établit. Plus concrètement, il est important de tester la robustesse et la stabilité dans le temps de la métrique en question.

En effet, dans la plupart du temps la détermination de la mesure de risque découle d'une modélisation du risque faite sur les données du portefeuille. La méthodologie appliquée dépend du type de risque (crédit, marché, souscription, etc.) et peut se faire en formule standard ou suivant une approche modèle interne en se basant sur certaines hypothèses (distributions statistiques, structures des données, chocs conjoncturels, etc.). La robustesse de la mesure du risque par rapport à un défaut de spécification du modèle et aux petits changements dans les données est un des tests à réaliser dans un tel contexte. Du point de vue d'un régulateur, la stabilité de la mesure de risque à la variation des hypothèses du modèle est une étape essentielle.

Dans le contexte de Solvabilité II, deux assureurs détenant le même portefeuille devraient obtenir la même VaR pour ce portefeuille. Cependant, dans le cas où le modèle le plus adéquat ne peut pas être identifié avec certitude ou presque, les assureurs peuvent utiliser deux modèles différents et obtenir des résultats de VaR significativement différents (Bernard and Vanduffel, 2016).

1.1.3 Agrégation et diversification du risque

L'agrégation des risques est une étape du processus de mesure des risques au cours de laquelle différents risques sont considérés conjointement afin d'obtenir un profil de risque global (Basel-Committee, 2010). D'une manière générale, il s'agit d'un processus qui consiste à additionner en prenant en compte les dépendances entre risques marginaux.

A travers cette définition découle le concept de diversification qui est lié à la technique d'agrégation des risques. On dira qu'il existe une diversification si la méthode d'agrégation fournit un résultat meilleur que la somme directe des risques marginaux. Une telle diversification s'observe dès lors qu'il existe une hétérogénéité dans un portefeuille ou lorsqu'on est en présence d'une entité qui exerce des activités qui l'exposent à différents marchés et pouvant regrouper plusieurs entités ou zones géographiques. Par exemple en assurance, on peut observer une diversification au sein d'un même type de risque comme dans le cas des contrats d'assurance dommage (entre les branches ou suivant les garanties) ou en rapport avec plusieurs types de risques (crédit, marché et taux, assurance vie et assurance dommage).

En gestion des risques, l'une des applications les plus connues du concept de diversification est le calcul du capital réglementaire. Dans la présente étude, nous nous intéressons à cet aspect, en l'appliquant dans le secteur de l'assurance. Un tel cadre réglementaire dont l'objectif est de promouvoir la prudence et le contrôle de la santé financière des organismes d'assurance limite les gains de diversification. En effet, en période de stress, les exigences en capital doivent être déterminées de sorte à couvrir toutes les pertes potentielles, ce qui tend à réduire les gains de diversification. Ainsi, il s'agit d'une réelle contrainte qui pèse sur la méthode d'agrégation donnant lieu à un niveau de diversification. C'est sur cette base qu'a été défini le niveau de capital réglementaire en Solvabilité 2 relativement à un scénario de choc bicentenaire (plus de détails dans la section 1.2).

Pour illustrer ¹ le concept de diversification on considère un portefeuille comprenant deux risques X_1 et X_2 et ρ la mesure de risque utilisée pour l'agrégation et la détermination du capital réglementaire.

$\rho(X_1)$ et $\rho(X_2)$ représentent respectivement les montants de fonds propres pour les risques X_1 et X_2 et $\rho(X_1 + X_2)$ le montant de fonds propres pour le risque global en supposant que X_1 et X_2 restent inchangés en combinant les deux portefeuilles. Le gain de diversification ou diversification benefit (noté $DB(\rho, X_1, X_2)$) est en général défini par

$$DB(\rho, X_1, X_2) = [\rho(X_1) + \rho(X_2)] - \rho(X_1 + X_2)$$

¹Cette illustration est tirée de Bernard and Vanduffel, 2016 ,page 12

$DB(\rho, X_1, X_2)$ s'interprète comme le gain (ou perte) obtenu en combinant différents portefeuilles. Dans le cas d'une mesure de risque cohérente ou qui satisfait l'axiome de sous-additivité, $DB(\rho, X_1, X_2) \geq 0$.

1.1.4 Lien entre dépendance et diversification

Les mesures de risque et les méthodes d'agrégation utilisent un ensemble de techniques et de concepts mathématiques. La dépendance statistique est l'un des concepts qui apparaît le plus souvent pour étudier et quantifier la diversification. En effet, toute mesure de risques agrégés fait intervenir des dépendances entre ces derniers. Dès lors, une des étapes importantes pour quantifier les effets de diversification est une modélisation adéquate des dépendances entre risques marginaux. Cette étape sera présentée en détails dans le chapitre 2. La forme de dépendance la plus simple est représentée par les corrélations empiriques. Mais nous verrons dans la suite qu'il existe des structures de dépendance plus complexes. Rappelons également que plus le portefeuille à agréger est hétérogène et a une taille importante, plus l'application d'une méthode d'agrégation devient difficile.

Pour illustrer ce lien entre corrélation et effet de diversification on considère la figure 1.1 où ρ représente la corrélation entre les risques 1 et 2.

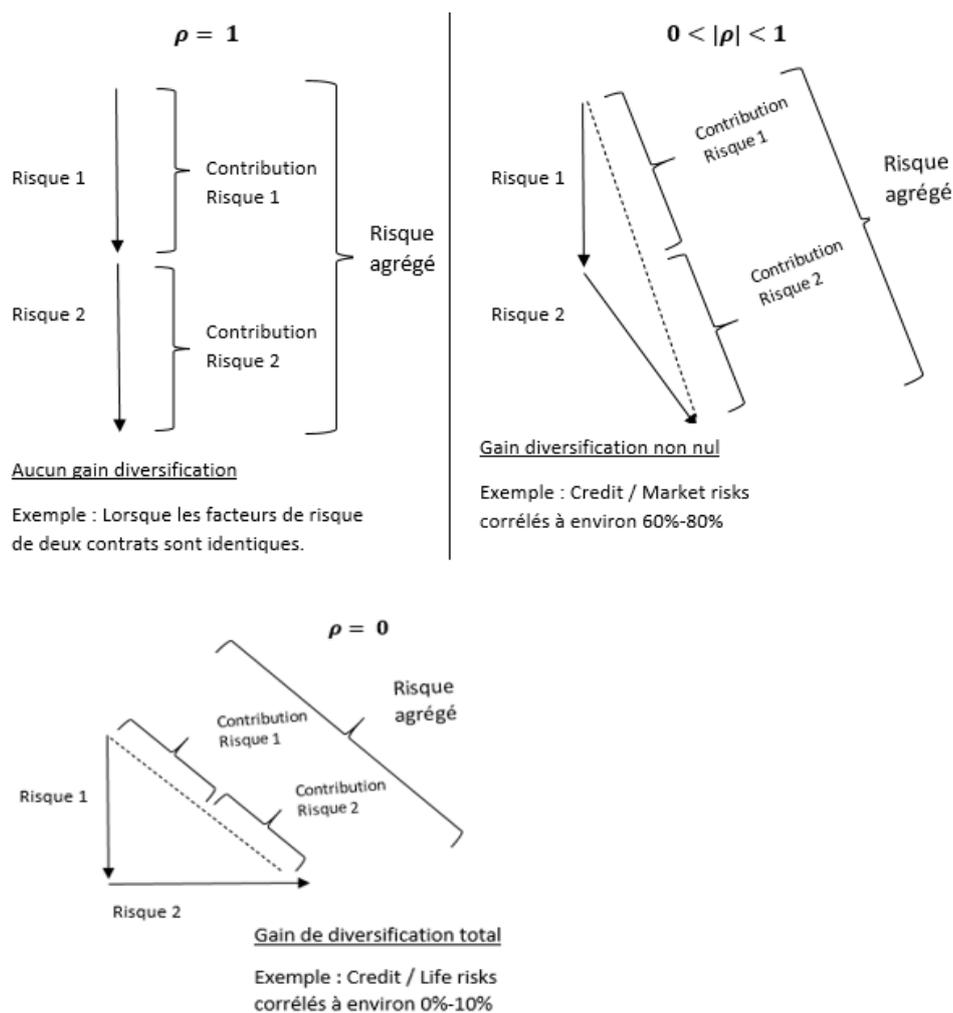


FIGURE 1.1: Illustration du lien entre corrélation et diversification

("Risk diversification benefits in insurance" n.d.) , page 5

1.2 Agrégation et diversification en Solvabilité 2

1.2.1 Cadre réglementaire Solvabilité 2

Entrée en vigueur en janvier 2016, la directive solvabilité 2 fournit un cadre prudentiel pour un nouveau régime harmonisé de solvabilité et de surveillance destiné au secteur des assurances en Europe. Compte tenu de l'environnement économique et financier de plus en plus incertain et les risques émergents, cette directive vise à établir un régime avec des niveaux de protection plus élevés et plus uniformes aux consommateurs favorisant l'égalité de concurrence.

A l'instar des trois piliers de la réglementation bancaire introduite par le régime de Bâle II, la directive solvabilité 2 se divise en 3 piliers.

Le pilier I concerne certains éléments du bilan tels que le niveau des actifs, les provisions techniques et les capitaux d'une entreprise. Il comporte deux séries d'exigences en matière de fonds propres : l'exigence des capitaux de solvabilité (SCR) qui est plus sensible au risque ; et l'exigence des capitaux minimum (MCR) qui est plus faible et plus générique. Le calcul du SCR peut se faire avec une approche standard ou par l'intermédiaire du « modèle interne » qui nécessite une approbation réglementaire.

Dans le pilier II, on retrouve des exigences plus ou moins qualitatives. Il s'agit des normes plus strictes en matière de gestion des risques et de gouvernance. Il donne plus de pouvoirs aux superviseurs pour contester les entreprises en matière de gestion des risques. Les entreprises sont tenues de préparer et de soumettre une ORSA (Own Risk and Solvency Assessment) à leurs superviseurs, en identifiant tous les risques de leur activité ainsi que les capitaux nécessaires pour gérer ces risques.

Le pilier III renforce la transparence pour les superviseurs et le public, par le biais de rapports annuels privés destinés aux superviseurs et de rapports publics sur la solvabilité et la situation financière. Les entreprises doivent fournir des informations plus détaillées sur leurs activités sur une base trimestrielle et annuelle.

Ce dispositif réglementaire est en constante évolution avec des discussions et révisions pour s'adapter aux mouvements du marché et anticiper des chocs éventuels et l'émergence de nouveaux risques.

Si on s'intéresse de près au pilier I, le SCR renvoie aux exigences quantitatives de solvabilité. Il correspond à la Value-at-Risk des fonds propres de base d'une entreprise d'assurance ou de réassurance au niveau de confiance de 99.5% à horizon annuel² ("Calculation of the Solvency Capital Requirement" 2011). La détermination du SCR se fait suivant plusieurs étapes et reflète le risque global du portefeuille de chaque compagnie.

1.2.2 Diversification en assurance dommage

L'accumulation de pertes dans un environnement économique et financier de plus en plus volatile peut entraîner la non-solvabilité des compagnies d'assurance. Une bonne stratégie

²voir article 101 de la directive Européenne https://www.eiopa.europa.eu/rulebook/solvency-ii/article-2324_en

concernant le choix des branches et types de risque à souscrire et la localisation géographique de ces risque est primordiale chez les assureurs. L'une des préoccupations majeures du risk manager en assurance est de s'assurer que le portefeuille de contrats souscrits n'est pas fortement exposé aux pertes potentielles liées à un événement ou séries d'évènements aléatoires (Pestano, 2017).

En assurance dommage, il existe plusieurs facteurs qui peuvent accentuer les risques auxquels sont exposés les assureurs. Ces facteurs peuvent affecter à la fois plusieurs branches. Par exemple, une tempête de grêle peut générer des sinistres dans les branches auto et habitation, un précédent juridique peut entraîner des hausses similaires des coûts de sinistre aussi bien du côté de l'indemnisation des travailleurs que des tiers. Dans une classe de contrats comprenant plusieurs garanties, on peut observer une corrélation en fréquences et/ou en montants entre ces garanties (Partrat, 2003):

- en dommages corporels, pour les garanties frais de soins et incapacité
- en risques industriels, pour les risques directs liés à l'activité et aux pertes d'exploitation.

Concernant les sinistres extrêmes, plusieurs évènements remettent en cause les techniques traditionnelles de gestion de risque en assurance : c'est le cas des Tempêtes Lothar et Martin en 1999 et le sinistre industriel AZF de Toulouse en 2001. Au niveau international, les attentats du World Trade Center (septembre 2001) et la crise financière de 2008 ont eu d'importantes répercussions dans l'environnement économique et financier mondial. A travers cet exemple, on constate que la survenance d'un même évènement peut être à l'origine de sinistres sur diverses branches.

Ces exemples montrent que les facteurs de risques qui affectent les différentes branches d'assurances sont de natures et d'origines diverses. Cette situation met en évidence les structures de dépendance entre les branches d'assurance et justifie leur prise en compte dans la gestion des risques d'entreprise et de marché pour les assureurs.

La diversification (qui peut être vue comme une répartition de risques) est un élément essentiel dans un tel contexte. C'est un moyen pour un assureur de faire en sorte que de mauvais résultats affectant une partie de son portefeuille soient compensés par de bons résultats dans d'autres parties. L'effet de la diversification et la méthode d'agrégation dépendent de la structure du portefeuille. Lorsqu'un assureur souscrit différentes branches

d'activité, son objectif est de diversifier son exposition aux pertes dans le sens où les pertes affectant une catégorie d'activité sont moins susceptibles d'affecter les autres. Par exemple, un assureur peut souscrire des risques d'assurance automobile et d'assurance de biens commerciaux pour le compte de ses assurés dans l'espoir que l'augmentation des demandes d'indemnisation découlant d'un facteur de risque (accident par exemple) puisse être compensée par une réduction des pertes de biens due à des améliorations (progrès, sécurité) dans le secteur industriel ou commercial. Néanmoins, il est possible que les catégories d'activités assurées soient touchées par le même événement, par exemple un incendie (Pestano, 2017).

Plus généralement, la souscription d'un plus grand nombre de contrats d'assurance permet à l'assureur de mieux répartir les risques. En ajoutant l'assurance maritime, l'assurance aviation et l'assurance responsabilité civile générale à sa gamme de produits, l'assureur sera plus diversifié et les mauvais résultats dans un domaine pourront être compensés par de bons résultats dans d'autres branches (Pestano, 2017). Il peut également segmenter ses lignes d'affaires à des échelles plus fines suivant les types de garanties ou en fonction des types d'assurés (entreprises, particuliers, institutions publiques). Cependant, cette segmentation devra tenir compte d'une taille de portefeuille raisonnable.

Par ailleurs, en cas de survenance d'un événement ayant un impact important (catastrophe naturelle), l'assureur ayant souscrit des contrats dans une seule ville ou zone géographique peut être exposé à des pertes immenses sur la globalité du portefeuille. Pour éviter une telle situation, une diversification qui consiste à souscrire des risques dans des zones géographiques différentes peut être une solution. En effet, la proximité géographique des contrats souscrits augmente la probabilité simultanée des pertes. Dans l'article de Pestano, 2017 on trouve un exemple illustratif de ce type de diversification qui est généralement adoptée par les grandes entreprises d'assurance et de réassurance. Il explique que si l'assureur souscrit des contrats en Floride, qui est fortement exposée aux ouragans, il peut aussi vouloir souscrire des contrats en Californie, qui n'est pas exposée aux ouragans mais aux tremblements de terre. L'objectif est que les mauvais résultats d'une saison d'ouragans active en Floride soient compensés par des périodes d'activité sismique réduite en Californie ou vice versa.

1.2.3 Agrégation en formule standard

Depuis l'entrée en vigueur du régime Solvabilité 2, les exigences quantitatives de solvabilité reposent sur les notions de besoin minimum en marge de solvabilité (MCR) et de capital de solvabilité requis (SCR). Les résultats de QIS5 (Quantitative impact study 5) fournissent une démarche plus rigoureuse basée sur la mesure du risque pour le calibrage de la formule standard du capital de solvabilité requis. L'objectif est de réduire le risque que l'assureur ne soit pas en mesure de faire face aux engagements auprès des assurés, de limiter les pertes auxquels les actionnaires peuvent être exposés en cas de ruine et d'alerter au plus tôt les autorités de contrôle. Ce cadre permet également de promouvoir un climat de confiance concernant la stabilité financière du secteur de l'assurance.

Le calcul du SCR peut se faire suivant plusieurs méthodes :

- Formule Standard
- Formule standard et paramètres propres à l'entreprise
- Modèle interne partiel
- Modèle interne complet

Le choix de la méthode par les assureurs dépend de leurs capacités et ressources. C'est le cas du modèle interne complet qui requiert plus de moyens. La méthode adoptée peut aussi dépendre d'autres facteurs tels que la taille et la complexité du portefeuille, le besoin en diversification, l'environnement économique et financier, la concurrence, etc.

En formule standard, le calcul du SCR global se déroule en plusieurs étapes. Il est composé de modules et sous-modules de risques individuels qui sont agrégés comme indiqué au niveau de la figure 1.2 :

La présente étude met l'accent sur le risque de souscription non-vie. Le besoin en capital de ce module est calculé à partir de 3 sous modules correspondants aux facteurs de risque identifiés par l'EIOPA :

- Prime et réserve
- Rachat
- CAT

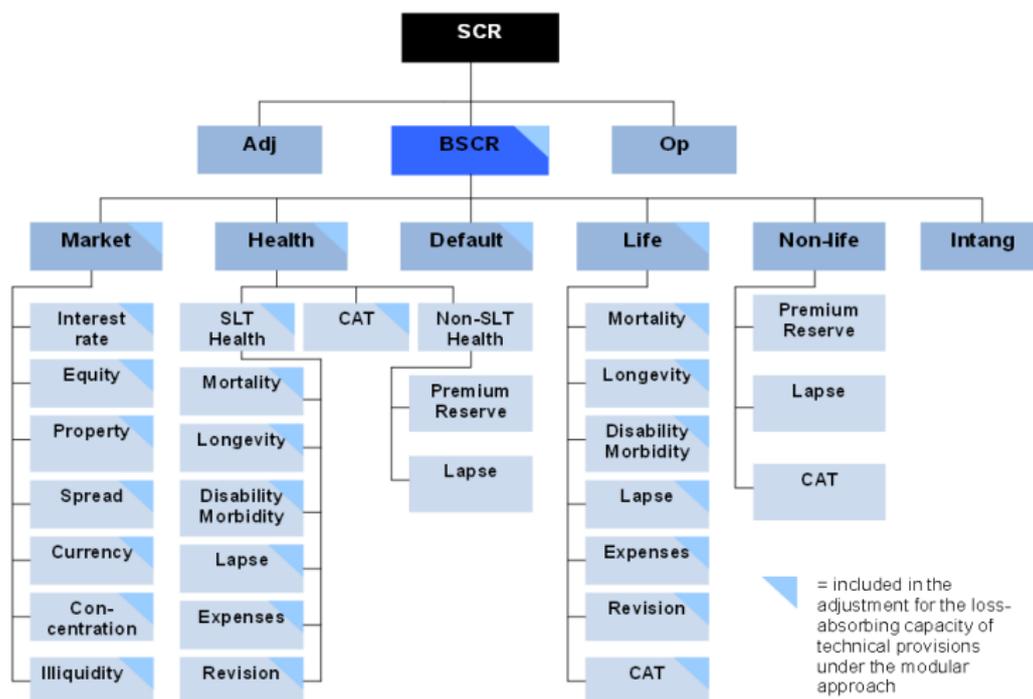


FIGURE 1.2: Modules et sous-modules de risques du SCR

Le risque de souscription non-vie est défini comme :

"Le module reflétant le risque découlant des engagements d'assurance non-vie, compte tenu des périls couverts et des procédés appliqués dans l'exercice de cette activité. Il tient compte de l'incertitude pesant sur les résultats des entreprises d'assurance et de réassurance dans le cadre de leurs engagements d'assurance et de réassurance existants, ainsi que du nouveau portefeuille dont la souscription est attendue dans les douze mois à venir" (Art. 105 (2) de la Directive 2009/138/EC)

Dans ce mémoire, on s'intéressera plus particulièrement au sous-module primes et réserves. Selon la même référence, il s'agit "du risque de perte, ou de changement défavorable de la valeur des engagements d'assurance, résultant de fluctuations affectant la date de survenance, la fréquence et la gravité des événements assurés, ainsi que la date et le montant des règlements de sinistres". Cette définition mentionne plusieurs facteurs qui rendent incertains les résultats des assureurs. Pour faire face à ces déviations, la formule proposée pour le SCR primes et réserves est donnée par :

$$SCR_{primes, reserve} = [Var_{99,5\%}(exp(X)) - 1]V$$

$$V = \sum_{Lob} (V_{reserve}^{Lob} + V_{prime}^{Lob}) = \text{volume total des primes et réserves}$$

Cette quantité représente le volume du portefeuille. Elle est agrégée suivant les différentes branches ou Lob (line of business) qui la composent.

X est une variable aléatoire qui suit une loi normale de moyenne $-\frac{1}{2}\log(\sigma^2 + 1)$ et de variance $\log(\sigma^2 + 1)$.

Il vient alors que $\exp(X)$ suit une distribution lognormale de moyenne 1 et de variance σ . σ représente le facteur de risque "CoV" provenant de la calibration. Il s'agit du standard déviation net combiné, résultant de la combinaison des standard déviation du risque de réserve et du risque de prime.

$$\text{En posant : } \rho(\sigma) = \frac{\exp(N_{0,995}\sqrt{\log(\sigma^2+1)})}{\sqrt{\log(\sigma^2+1)}}$$

Avec $N_{0,995}$: quantile 99.5% quantile loi normale standard.

On peut réécrire la formule du SCR par :

$$SCR_{primes,reserves} = \rho(\sigma)V$$

$$\text{Avec } \rho(\sigma) = [\text{Var}_{99,5\%}(\exp(X)) - 1] \quad .$$

Cette fonction est fixée de telle sorte que, dans l'hypothèse d'une distribution lognormale du risque sous-jacent, une charge de capital-risque correspondant à la Value at risk standard au seuil 99,5% puisse être déterminée. Approximativement, on a $\rho(\sigma) = 3\sigma$.

Le standard deviation net global est déterminé comme suit :

$$\sigma = \sqrt{\frac{1}{V^2} \sum_{i,j} \text{Corr}(LoB_i, LoB_j) \sigma_i \sigma_j V_i V_j}$$

Dans cette formule, les indices i et j désignent les différents LoBs.

En effet, $V_i = V_i^{prime} + V_i^{reserve}$

$\text{Corr}(LoB_i, LoB_j)$: matrice de corrélation entre Lobs. La calibration des corrélations a été faite selon la méthodologie QIS et l'hypothèse de dépendance linéaire.

Les facteurs de volatilités σ_i pour les risques de prime et réserve sont calibrés sur les données de sinistres et primes du portefeuille.

Les limites de l'approche standard

La méthode présentée précédemment sur le calcul du SCR en formule standard, requiert un certain nombre d'hypothèses notamment sur la distribution du facteur de risque. La calibration de la volatilité constitue un processus plus ou moins complexe et dépend des données du portefeuille de l'assureur. Ainsi, la question que l'on se pose est de savoir si cette approche est adaptée à tous les assureurs quelque soit leur profil de risque ? Malgré le fait que toutes les entreprises d'assurances et de réassurances en France sont soumises au cadre réglementaire solvabilité II, et que la plupart partagent le marché français ou européen, appliquer la formule standard peut être moins bénéfique qu'une approche modèle interne dans le cadre du calcul du capital de solvabilité. En effet, le cadre de modélisation de la dépendance linéaire entre risques proposé par la formule standard présente des limites que le modèle interne cherche à corriger (Appert-Raullin, Tann, and Devineau, 2014):

- Problème de la dépendance non-linéaire des événements extrêmes
- Difficulté dans l'interprétation des valeurs de la matrice de corrélation entre lignes d'affaires proposée par la formule standard
- Scénarios des catastrophes naturelles définis dans la formule standard peu adaptés pour un assureur disposant un grand nombre d'entités internationales

1.2.4 Cas du modèle interne AXA France P&C

Intérêts et avantages du modèle interne

De plus en plus, les entreprises d'assurance adoptent une approche modèle interne partiel ou complet pour répondre aux exigences quantitatives sous Solvabilité 2. Le choix d'un tel cadre méthodologique passe par une procédure de candidature auprès des autorités de contrôle qui en donnent l'autorisation. Ces dernières disposent d'un délai de 6 mois à partir de la date de remise du dossier de candidature complet pour statuer sur le caractère adéquat du modèle. Après la validation de la candidature, un dispositif de suivi et contrôle régulier est mis en place par les autorités. En effet, les organismes d'assurance opèrent des changements réguliers pour faire évoluer leur modèle interne compte tenu de l'environnement

économique, la concurrence ou en lien avec d'autres facteurs internes ou externes.

Les graphiques de la figure 2 révèlent que le modèle interne présente un fort intérêt stratégique et organisationnel en raison de la flexibilité qu'il permet. Grâce au modèle interne, les entreprises peuvent aussi observer un gain en fonds propres par rapport à la formule standard. En effet, les entreprises en modèle interne prennent en compte la diversification dans la gestion du risque de leur portefeuille pour mieux estimer le risque global et aboutir à un SCR qui en général peut être moins élevé que celui calculé en formule standard.

Il est également important de rappeler que le modèle interne présente un certain nombre de limites par rapport à la formule standard. Ces limites concernent essentiellement les coûts de développement, de déploiement et d'entretien. Il s'agit des moyens mis à disposition pour la mise en place et le suivi du modèle interne qui doit répondre aux exigences de rigueur méthodologique et fournit des résultats compatibles avec le cadre prudentiel solvabilité 2.

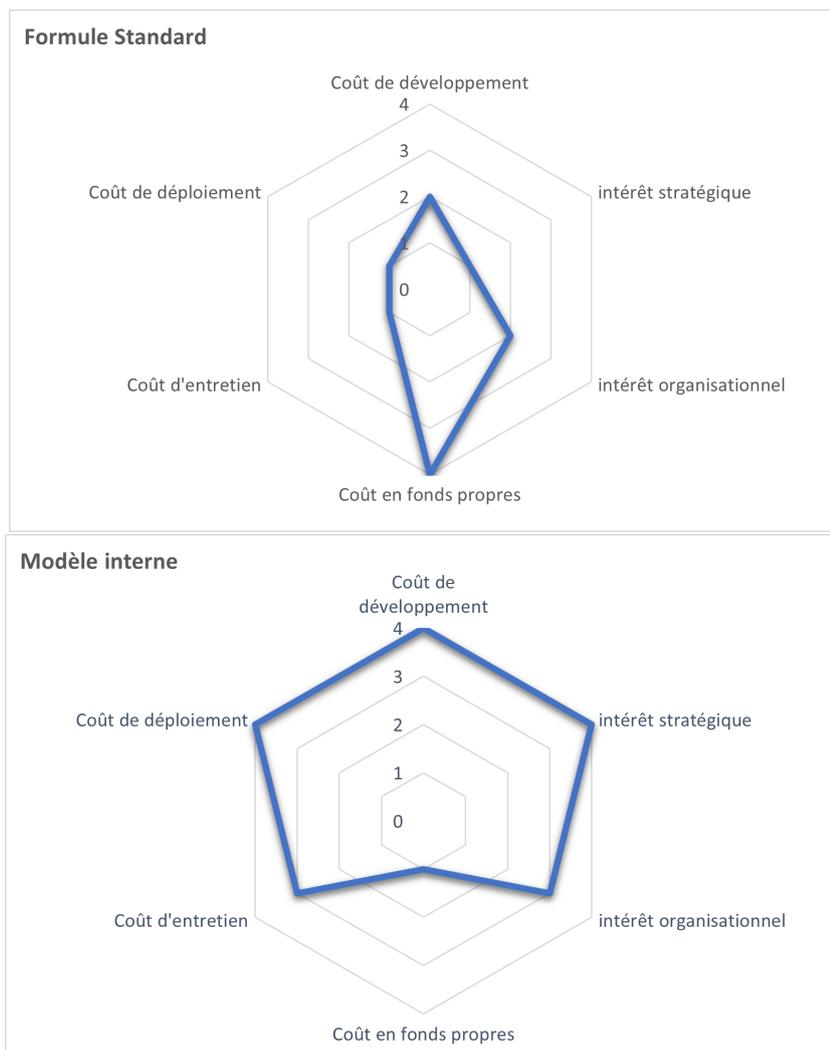


FIGURE 1.3: Avantages et inconvénients : Formule standard vs Modèle interne

Sia-Conseil, 2010

Portefeuille AXA France : les différents niveaux de risque

Dans le cadre du modèle interne AXA France, on désigne par STEC (Short Term Economic Capital) la métrique qui est équivalente au SCR dans le cas général. Conformément aux exigences quantitatives de la directive solvabilité 2, le STEC global est le résultat de plusieurs étapes de calcul. Il est calculé suivant une méthodologie qui prend en compte plusieurs niveaux de risques incluant des calculs probabilistes, des calibrations de modèle, des hypothèses qui font l'objet de test, des jugements d'expert, etc. Tous ces calculs se font dans un cadre comptable bien organisé avec un reporting régulier des données du portefeuille.

La figure 1.4 donne une vue d'ensemble de la procédure qui mène au STEC global. On

distingue différents blocs de risques que sont principalement le risque de marché, le risque de crédit, les risques vie et non vie et le risque opérationnel. Au sein de chaque bloc, on observe également des composantes qui à leur tour sont désagrégées jusqu'aux niveaux les plus fins du portefeuille. Cette structure révèle l'importance des méthodes d'agrégation qui ont un impact considérable dans la détermination du STEC global.

En pratique, une agrégation est faite à chaque niveau en tenant compte des dépendances entre les éléments qui constituent le bloc. Dans la suite, nous précisons le périmètre de l'étude et présenterons les techniques d'agrégation en analysant les effets de diversification ou impact sur le STEC.

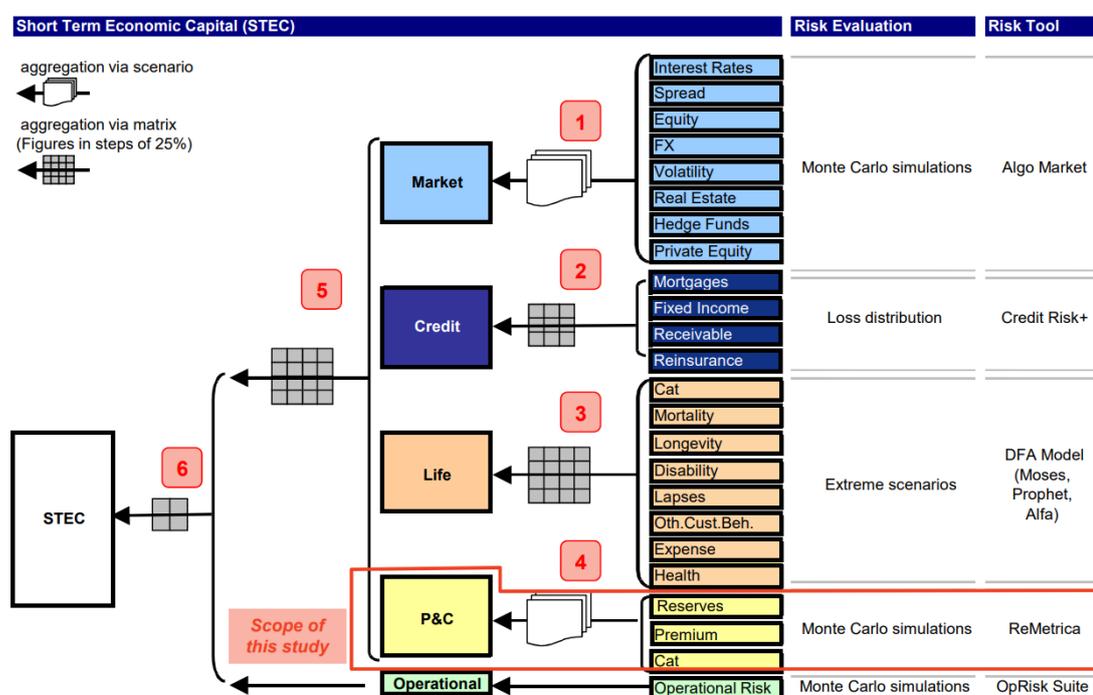


FIGURE 1.4: Vision globale du calcul du STEC

Appert-Raullin, Tann, and Richard, 2011

Dans la présente étude, nous nous intéressons aux branches du portefeuille P&C. L'accent sera mis sur le risque de réserve. L'objectif sera de montrer comment à partir des données sur les charges de sinistres (incluant les provisions) on peut arriver au calcul du STEC réserve. A travers ce processus, on identifie deux niveaux d'agrégation.

Dans un premier temps, une agrégation à l'intérieure de chaque Lob en tenant compte des corrélations entre années de survenance (ou années d'accident) permet de déterminer

le STEC réserve par Lob. Puis dans un second temps, le STEC reserve global est obtenu en estimant les dépendances entre Lobs.

Dans la suite, nous expliquerons en détails ces étapes dans le cadre du modèle interne. Il est important de rappeler qu'une méthodologie pour la première étape (corrélations entre années d'accident) a été déjà établie, validée et prise en compte dans le calcul du STEC en 2021. Il s'agit donc d'un changement de modèle interne qui se poursuit avec la seconde étape sur les corrélations entre Lobs. Les travaux du présent mémoire et les résultats qui en découlent seront utilisés à cette fin.

Chapitre 2

Modélisation de la dépendance entre risques

2.1 Concepts théoriques de dépendance statistique

Nous avons montré que l'effet de diversification que l'on cherche à quantifier à travers l'agrégation de risques est lié à la dépendance entre ces derniers. Dans la théorie, on trouve différentes méthodes permettant d'étudier ces dépendances en tenant compte d'un certain nombre de propriétés et d'hypothèses.

2.1.1 Caractérisation et propriétés

- L'indépendance entre deux variables X_1 et X_2 peut être caractérisée par la relation suivante :

$$F(x_1, x_2) = F_1(x_1) \cdot F_2(x_2)$$

F_1, F_2 et F étant respectivement les fonctions de répartition de X_1, X_2 et du couple (X_1, X_2)

D'une manière générale, on parle d'indépendance lorsque la probabilité de survenance de l'une des variables n'est pas affectée par la survenance de l'autre.

$$P(X_1 \leq x_1 | X_2 \in A) = P(X_1 \leq x_1) \quad \text{pour tout } A.$$

- Lorsqu'on considère un vecteur de dimension n , $X = (X_1, X_2, \dots, X_n)$, on parle d'indépendance multivariée ou par paire de variables.

$$F(x_1, x_2, \dots, x_n) = \prod_{i=1}^n F_i(x_i)$$

Dans le cas où il existe au moins une réalisation d'une des variables pour la quelle la précédente relation n'est pas vérifiée, on est en présence de dépendance. En général, d'un

point de vu statistique, c'est la significativité de la dépendance (relativement à un seuil) qui nous intéresse le plus. On peut également étudier le type, la structure et le niveau de dépendance.

Il existe des propriétés relatives à une mesure de dépendance δ :

- **Symétrie:** $\delta(X, Y) = \delta(Y, X)$
- $\delta(X, Y) = 1 \Leftrightarrow X$ et Y sont comonotones
tandis que $\delta(X, Y) = -1 \Leftrightarrow X$ et Y sont antimonotones
- Soit f une fonction croissante, et g une fonction décroissante,

$$\delta(f(X), Y) = \delta(X, f(Y)) = \delta(X, Y)$$

$$\delta(g(X), Y) = \delta(X, g(Y)) = -\delta(X, Y)$$

- δ normalisée telle que $|\delta(X, Y)| < 1$

On parlera de dépendance parfaite lorsque deux variables sont comonotones ou antimonotones. En cas d'indépendance, $\delta(X, Y) = 0$. Mais la réciproque n'est pas en général vérifiée, notamment quand on considère les corrélations comme mesure de dépendance. Ce point de vigilance sera vu en détail dans la suite.

2.1.2 La corrélation comme mesure de dépendance

A travers les définitions et propriétés précédentes, on peut dire que le concept de dépendance entre deux ou plusieurs variables n'est pas un concept mathématique simple. Par conséquent, il est assez difficile de communiquer des informations telles que le "degré", le "niveau" ou le "type" de dépendance. Le fait d'exprimer par un indicateur¹ les informations sur la structure de dépendance est une importante simplification. Cette approche se fait au prix d'une perte d'information et sous certaines hypothèses (*exemple* : dépendance linéaire), puisqu'elle correspond à un mapping de l'ensemble des copules vers les nombres réels (Mai and Scherer, 2014).

La corrélation est l'une des mesures de dépendance les plus utilisées en raison de sa simplicité de mise en oeuvre et d'interprétation. D'un point de vu statistique, la corrélation

¹un seul nombre qui mesure et quantifie le degré de dépendance

est représentée par un indice ou coefficient mesurant le degré de liaison ou d'association entre deux variables.

Corrélation linéaire de Pearson ρ

La corrélation de Pearson est l'une des mesures de dépendance statistique les plus courantes. Elle étudie la dépendance linéaire entre deux variables X et Y selon l'expression :

$$\rho^{X,Y} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

Cette quantité existe si les variables X et Y sont carré intégrables ($E[X^2] < \infty$).

Pour appliquer cette formule à des données réelles, il est d'usage de réaliser au préalable une représentation graphique (idéalement par paire et en dimension deux) pour mieux visualiser le lien entre les variables. Puis l'estimation empirique sur n observations peut être obtenue par l'expression :

$$\hat{\rho}_n = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

En rappelant que $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ et $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$

$-1 \leq \rho \leq 1$ et on dira que X et Y sont positivement corrélées (resp. négativement corrélées) si $\rho > 0$ (resp. $\rho < 0$) et non corrélées lorsque $\rho = 0$.

Il est important de préciser qu'en général, la non corrélation n'implique pas l'indépendance (l'inverse est tout de même vrai). Il faudrait être dans le cas de vecteurs gaussiens.

Pour illustrer ce fait, considérons une variable U qui suit une loi uniforme sur $[-1; 1]$ et $V = U^2$ qui prend ses valeurs dans l'intervalle $[0; 1]$. Il est clair qu'il existe une dépendance entre U et V qui sont liées selon une relation quadratique.

Cependant, $\rho^{U,V} = 0$, ce qui signifie que les variables U et V sont non corrélées. A travers cet exemple, on comprend mieux le fait que le coefficient de corrélation n'est pas suffisant pour étudier la structure de dépendance. Cet indicateur n'est adapté que dans le cas de distributions elliptiques.

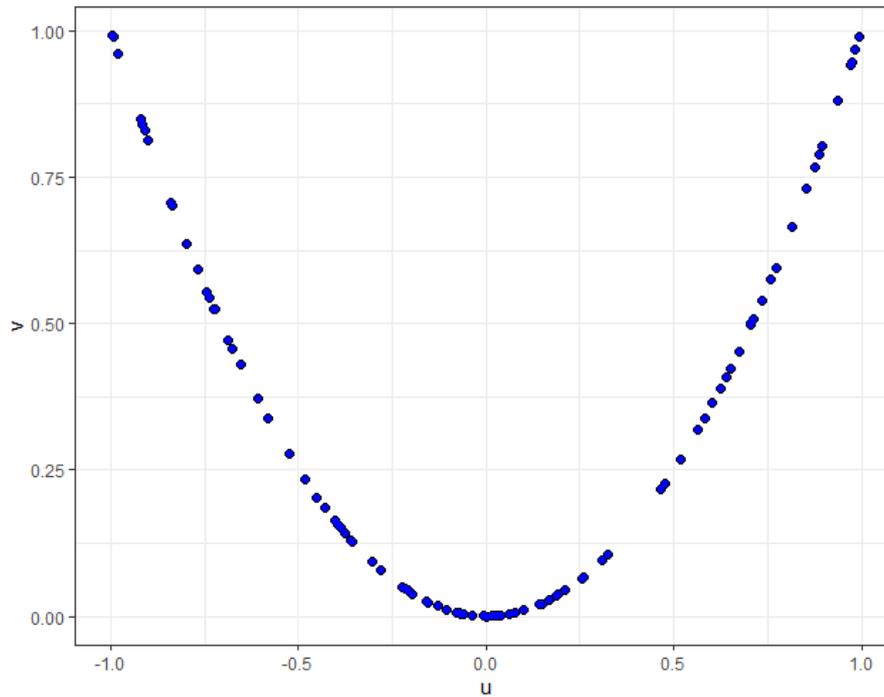


FIGURE 2.1: Plot entre U et V

Concernant les propriétés, on peut ajouter que la corrélation est invariante par transformations affines positives :

$$\rho(\alpha X + \beta, \gamma Y + \delta) = \text{signe}(\alpha\gamma)\rho(X, Y) \quad \alpha, \gamma \in \mathbb{R}^*, \text{ et } \beta, \delta \in \mathbb{R}$$

Par contre, l'invariance par transformation monotone n'est pas en général vérifiée.

En dimension $d > 2$, la corrélation de Pearson se présente sous forme de matrice dans laquelle chaque cellule représente le coefficient d'une paire de variable.

Corrélation de Rang

Avant de présenter les autres catégories de corrélation basées sur les rangs, on définit une notion appelée concordance (McNeil, Frey, and Embrechts, 2015). Considérons deux couples (X_1, Y_1) et (X_2, Y_2) .

- On dira que ces deux couples sont concordants si $[X_1 < X_2 \text{ et } Y_1 < Y_2]$ ou $[X_1 > X_2 \text{ et } Y_1 > Y_2]$
c'est à dire $\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] = 1$ (probabilité de concordance).

- Les couples sont discordants lorsque
 $[X_1 < X_2 \text{ et } Y_1 > Y_2]$ ou $[X_1 > X_2 \text{ et } Y_1 < Y_2]$
 c'est à dire $\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0] = 1$ (probabilité de disconcordance).

Les corrélations de rang se basent sur ce concept. Deux formules de corrélations sont les plus connues : tau de Kendall et rho de Spearman. Elles représentent des extensions de la corrélation de Pearson adaptées aux variables de rang. L'avantage d'utiliser ces types de corrélations est qu'elles sont moins sensibles aux valeurs aberrantes.

Le tau de Kendall τ

Le tau de Kendall peut être défini par la formule :

$$\tau^{X,Y} = P[(X_i - X_j)(Y_i - Y_j) > 0] - P[(X_i - X_j)(Y_i - Y_j) < 0]$$

Une estimation empirique de $\tau^{X,Y}$ est donnée par :

$$\frac{\sum_{i < j} \text{signe}((x_i - x_j)(y_i - y_j))}{\binom{n}{2}}$$

$\tau^{X,Y}$ est invariante par transformation strictement croissante (linéaire ou non linéaire). En effet, si f et g sont deux fonctions strictement croissantes, alors on a $\tau^{f(X),g(Y)} = \tau^{X,Y}$ (Nelsen, 2006)

Le rho de Spearman ρ_S

Considérons les couples de vecteurs $(X_1, Y_1), (X_2, Y_2)$ et (X_3, Y_3) , le rho des Spearman est donnée par :

$$\rho_S^{X,Y} = 3 [P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]]$$

L'estimation empirique peut être obtenue en se basant sur les séries de rang. Si on note par r_i^X (resp. r_i^Y) $i = 1, \dots, n$ la série des rangs de la variable X (resp. Y), on peut déterminer le $\rho^{X,Y}$ via la formule du coefficient de Pearson.

$$\hat{\rho}_{Sn} = \frac{\sum_{i=1}^n (r_i^X - \bar{r}^X)(r_i^Y - \bar{r}^Y)}{\sqrt{\sum_{i=1}^n (r_i^X - \bar{r}^X)^2} \sqrt{\sum_{i=1}^n (r_i^Y - \bar{r}^Y)^2}}$$

A la place des rangs, on peut aussi utiliser les fonctions de répartition empiriques. De même que le tau de Kendall, $\rho_S^{X,Y}$ est aussi invariant par transformation strictement croissante (linéaire ou non linéaire).

2.1.3 Dépendance de queue

Pour étudier les dépendances liées à la survenance d'événements extrêmes, la dépendance de queue (tail dependency) peut être une bonne mesure. Elle fournit une description au niveau des queues de distribution. La théorie distingue le coefficient de dépendance de queue supérieure² de celui de queue inférieure³

Les TDC (Tail dependency coefficient) peuvent être obtenus par les formules :

$$\lambda_{upper} = \lim_{t \rightarrow 1} P(Y > F_Y^{-1}(t) | X > F_X^{-1}(t))$$

$$\lambda_{lower} = \lim_{t \rightarrow 0} P(Y \leq F_Y^{-1}(t) | X \leq F_X^{-1}(t))$$

Avec $0 < \lambda_{upper} < 1$ et $0 < \lambda_{lower} < 1$

Le TDC s'interprète comme la probabilité conditionnelle qu'une marge excède un seuil (représenté par le quantile $F^{-1}(t)$) sachant que l'autre marge excède déjà ce seuil. On dira que les variables présentent une dépendance de queue supérieure (resp. inférieure) si $\lambda_{upper} > 0$ (resp. $\lambda_{lower} > 0$) au sens statistique.

2.1.4 Précision des corrélations empiriques et intervalle de confiance

Les corrélations empiriques sont des estimations de la dépendance entre variables. L'estimation se base sur un échantillon et l'objectif est de fournir une mesure qui s'approche le plus de la vraie dépendance. Pour cela, il faudrait une taille d'échantillon suffisamment élevée. Dans la présente étude, nous verrons que la quantité de données disponible est relativement faible. Dans ces conditions, les corrélations empiriques calculées sont susceptibles d'être biaisées.

²Upper tail dependance λ_{upper} : quadrant supérieur droit

³Lower tail dependance λ_{lower} : quadrant inférieur gauche

Il est donc nécessaire d'étudier la précision des estimations. En statistique, l'une des approches les plus connues est le test de significativité des coefficients. Pour cela, on peut réaliser un t -test (Ruscio, 2008):

$$t = \frac{\hat{\rho}}{\sqrt{(1 - \hat{\rho}^2)/(N - 2)}} \quad \text{avec } N - 2 \text{ degrés de liberté}$$

Toutefois, le test de significativité ne permet pas de vérifier l'écart entre l'estimateur empirique et la vraie mesure de dépendance.

Pour réaliser ce test, on peut construire des intervalles de confiance. A l'aide d'une transformation de fisher $z_{\hat{\rho}} = \tanh^{-1}(\hat{\rho})$ (voir Ruscio, 2008) une estimation usuelle de la variance de $Z_{\hat{\rho}}$ est $\frac{1}{N-3}$.

En faisant l'hypothèse que les erreurs sont normalement distribuées, l'intervalle de confiance peut être construit comme suit :

$$CI(\rho) = \tanh \left[Z_{\hat{\rho}} \pm q_{\frac{1+\alpha}{2}} \sqrt{\frac{1}{N-3}} \right]$$

$$\text{Rappelons que } \tanh^{-1}(x) = \frac{1}{2} \ln \frac{1+x}{1-x}$$

$q_{\frac{1+\alpha}{2}}$ est le quantile d'ordre $\frac{1+\alpha}{2}$ de la loi normale standard.

Par exemple, pour $\alpha = 0,95$; on a $q_{\frac{1+\alpha}{2}} = 1,96$.

Une autre technique permettant de construire des intervalles de confiance est d'utiliser la méthode bootstrap. Cette approche permet d'éviter les hypothèses sur la distribution de l'échantillon et l'estimation de la variance. Elle se base sur une estimation empirique par ré-échantillonnage. L'algorithme servant à la construction des intervalles de confiance des corrélations est présenté en annexe A.3.

2.2 La théorie des copules

2.2.1 Généralités, définitions et théorème de Sklar

Nous avons vu que les mesures de dépendance courantes n'étaient toujours pas adaptées pour décrire la dépendance entre les variables aléatoires. En effet, elles mesurent la dépendance par un scalaire et ne nous donnent aucune information sur celle-ci (dépendance dans les extrêmes par exemple ou des dépendances asymétriques). Il existe un outil pour palier à ces limites : les copules ! La notion de « copule » a été introduite par Sklar en 1959. A l'aide des copules, il est possible d'établir une représentation fonctionnelle de la dépendance et fournit une description complète.

On peut définir une copule de dimension n comme une fonction de répartition multivariée notée C définie sur $[0; 1]^n$ et dont les marginales sont des variables de loi uniformes sur $[0; 1]$.

La définition étant plus ou moins abstraite, nous présentons le théorème de Sklar et certaines propriétés sur les copules. Soit F une fonction de répartition de dimension n avec des marginaux F_1, F_2, \dots, F_n , alors il existe une copule C de dimension n telle que pour tout $x \in \mathbb{R}^n$, $F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$. (Cadoux and Loizeau, 2004)

Considérons n variables aléatoires continues avec comme fonctions de répartition marginales F_1, F_2, \dots, F_n . Si on note par F la fonction de répartition jointe, alors il existe une copule C qui vérifie que la relation précédente est unique. A travers cette définition et théorème, il est clair qu'avec une copule, on arrive à établir une relation entre la loi jointe et les lois marginales. En effet, on a en argument de la copule des réalisations de lois uniformes qui correspondent aux valeurs de la fonction de répartition de chaque marginal. En posant $u_1 = F_1(x_1)$, on peut écrire

$$C(u_1, \dots, u_n) = P(U_1 \leq u_1, \dots, U_n \leq u_n)$$

La densité d'une copule dans le cas où elle existe peut s'écrire :

$$c(u_1, \dots, u_n) = \frac{\partial^n C}{\partial u_1 \dots \partial u_n}(u_1, \dots, u_n)$$

Il est important de préciser que les copules généralisent les méthodes d'estimation de

dépendance vues dans la section précédente. Pour le coefficient linéaire de Pearson, il s'agit d'un coefficient basé sur la copule gaussienne que nous verrons par la suite. Les autres coefficients ou mesure de concordance sont basés sur les rangs et il est possible de récrire leurs expressions à l'aide des copules.

Soit un couple (X, Y) de variables aléatoires continues. Considérons une copule C de dimension 2 liée à ce couple. On notera par U_1 (resp. U_2), la loi uniforme associée à la fonction de répartition F_1 de la variable X (resp. F_2 pour Y)

- **Tau de Kendall**

$$\begin{aligned}\tau(X, Y) &= 4 \int \int_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2) - 1 \\ &= 1 - 4 \int \int_{[0,1]^2} \partial_1 C(u_1, u_2) \partial_2 C(u_1, u_2) du_1 du_2 \\ \tau(X, Y) &= 4\mathbb{E}[C(U_1, U_2)] - 1\end{aligned}$$

- **Rho de Spearman**

$$\rho_S(X, Y) = 12 \int \int_{[0,1]^2} u_1 u_2 dC(u_1, u_2) - 3 = 12 \int \int_{[0,1]^2} C(u_1, u_2) du_1 du_2 - 3$$

Pour la **dépendance de queue**, les coefficients TDC peuvent également être obtenus connaissant la fonction de répartition de la copule :

$$\lambda_{upper} = \lim_{t \rightarrow 1} \frac{1 - 2t + C(t, t)}{1 - t}$$

$$\lambda_{lower} = \lim_{t \rightarrow 0} \frac{C(t, t)}{t}$$

Nous verrons par la suite que les TDC peuvent être estimés dans le cas de familles spécifiques de copules (en général les copules archimédiennes).

2.2.2 Les familles de copule

Il existe plusieurs familles de copules, mais dans le cadre cette étude nous nous limitons à deux grandes familles : les copules elliptiques (Gaussienne et Student) et les copules archimédiennes (Clayton, Frank et Gumbel).

a. Copule Gaussienne

La copule gaussienne est une copule liée à la distribution gaussienne multivariée. C'est l'une des familles de copule la plus populaire utilisée en finance et en assurance pour traiter les problématique de dépendance entre les éléments d'un portefeuille. Notons par Φ la fdr de la loi normale standard et $\Phi_{\Sigma,n}$ celle de la loi normale multivariée de dimension n et de matrice de corrélation Σ . La copule gaussienne de dimension n est donnée par :

$$C_{\Sigma}(u_1, \dots, u_n) = \Phi_{\Sigma,n}(\Phi^{-1}(u_1, \dots, u_n))$$

Et la densité s'écrit :

$$c_{\Sigma}(u_1, \dots, u_n) = \frac{1}{\sqrt{\det \Sigma}} \exp \left(-\frac{1}{2} [\Phi^{-1}(u)]^t [\Sigma^{-1} - Id] [\Phi^{-1}(u)] \right)$$

En dimension 2, on peut écrire l'expression explicite de la copule bivariée:

$$C_{1,2}(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\Sigma^2}} \exp \left[-\frac{s^2 - 2st\Sigma + t^2}{1-\Sigma^2} \right]$$

Rappelons que si X est un vecteur gaussien, sa copule est gaussienne mais l'inverse n'est pas toujours vrai.

Dans le cadre d'une copule gaussienne, il n'existe pas de dépendance de queue, mais est appropriée pour les distributions symétriques (Embrechts, McNeil, and Straumann, 2001).

Les paramètres de la copule gaussienne sont les coefficients qui forment la matrice de corrélation qui la détermine entièrement. L'estimation d'une copule gaussienne passe donc par une estimation de ces paramètres. Donc plus la dimension est élevée, plus le nombre de paramètres à estimer devient important ($\frac{n(n-1)}{2}$ paramètres à estimer en dimension n)

b. Copule de Student

Faisant partie de la famille des copules elliptiques, la copule de student découle de la distribution de student multivariée. Elle est construite de manière similaire à la copule gaussienne. Cependant, à la différence de celle-ci, le coefficient de dépendance de queue n'est pas nul et on arrive à capter les dépendances extrêmes positives et négatives. Si on note par d la dimension, Σ la matrice de corrélation, ν le degré de liberté, la fonction de densité de la copule s'écrit

$$c_{\Sigma}(u_1, \dots, u_d) = \frac{f_{v,\Sigma}(t_v^{-1}(u_1), \dots, t_v^{-1}(u_d))}{\prod_{i=1}^d f_v(t_v^{-1}(u_i))}$$

Avec $f_{v,\Sigma}$ la densité de la loi de Student multivariée centrée réduite
Et f_v , la densité uni-variée de la loi de student centrée réduite.

En dimension 2, la densité est donnée par :

$$c(u_1, u_2) = \frac{\kappa}{2\sqrt{1-\rho^2}} \frac{\Gamma(\frac{\kappa}{2})^2}{\Gamma(\frac{\kappa+1}{2})^2} \frac{\left[1 + \frac{u_1^2 + u_2^2 - 2\rho u_1 u_2}{\kappa(1-\rho^2)}\right]^{-(\kappa+2)/2}}{\left[\left(1 + \frac{u_1^2}{\kappa}\right) \left(1 + \frac{u_2^2}{\kappa}\right)\right]^{-(\kappa+2)/2}}$$

Comme dans le cas gaussien, tous les paramètres de la copule de student doivent être estimés.

Les copules elliptiques s'utilisent en présence de distributions symétriques. En cas de données asymétriques, d'autres formes de copules sont mieux adaptées.

c. Les copules archimédiennes

Comme évoqué dans les sections précédentes, il existe différentes structures de dépendance que les copules elliptiques ne permettent pas de capter. C'est le cas de la dépendance asymétrique. Les copules archimédiennes permettent cependant de prendre en compte cette forme de dépendance. Dans cette famille on retrouve les copules de Clayton, Gumbel et Frank.

Considérons ϕ une fonction continue, convexe et décroissante définie de $[0; 1]$ dans $[0; +\infty]$ vérifiant $\phi(1) = 0$ et $\phi(0) = \infty$. On l'appellera fonction générateur. L'idée d'une copule archimédienne est d'appliquer la transformation $\omega(u) = \exp(-\phi(u))$ aux marginaux de sortes que "les composantes soient indépendantes" (**Planchet, 2021**):

$$\omega(C(u_1, \dots, u_n)) = \prod_{i=1}^n \omega(u_i)$$

La copule archimédienne de générateur ϕ est définie par :

$$C(u_1, \dots, u_n) = \phi^{-1}(\phi(u_1) + \dots + \phi(u_n))$$

dès lors que $\sum_{i=1}^n \phi(u_i) \leq \phi(0)$ et $C(u_1, \dots, u_n) = 0$ sinon.

Cette caractérisation permet de construire une grande variété de familles de copules, suivant l'expression de la fonction générateur. Les copules ainsi générées ont des formes analytiques qui dépendent d'un paramètre de dépendance θ .

Nom de la copule	$\phi(u)$	$C(u_1, u_2)$
Clayton	$u^{-\theta} - 1, \theta > 0$	$(u_1^{-\theta} + u_2^{-\theta})^{-\frac{1}{\theta}}$
Frank	$-\ln\left(\frac{e^{-\theta u} - 1}{e^{-\theta} - 1}\right), \theta \neq 0$	$-\frac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right)$
Gumbel	$(-\ln(u))^\theta, \theta \geq 1$	$\exp\left(-[(-\ln(u_1))^\theta + (-\ln(u_2))^\theta]^{\frac{1}{\theta}}\right)$

TABLE 2.1: Fonctions générateurs et expressions des copules archimédiennes dimension 2

- La copule de Clayton permet de capter les dépendances (positives) à faible intensité dans la queue inférieure. En dimension 2, il s'agit d'un nuage de points qui tend à croître lorsque l'on se déplace vers l'extrémité supérieur.
- La copule de Gumbel modélise une dépendance (positive) de l'extrémité supérieure de la distribution. Contrairement au cas des copules archimédiennes en général, la copule de Gumbel est symétrique.
- Avec la copule de Frank, il est possible de modéliser aussi bien les dépendances positives que négatives.

En dimension 2, on peut également obtenir l'expression de certains indicateurs tels que les coefficients de dépendance de queue (λ_L et λ_U) et le tau de Kendall (τ_θ).

Nom de la copule	λ_L	λ_U	τ_θ
Clayton	$2^{-\frac{1}{\theta}}$	0	$\frac{\theta}{\theta+2}$
Frank	0	0	$1 - \frac{4(1-D(\theta))}{\theta}$ où $D(\theta) = \frac{1}{\theta} \int_0^\theta \frac{t}{e^t - 1} dt$
Gumbel	0	$2 - 2^{-\frac{1}{\theta}}$	$1 - \frac{1}{\theta}$

TABLE 2.2: Dépendance de queue et tau de Kendall

Dans la théorie, on peut retrouver d'autres types de copules de la famille archimédienne telles que la copule de Joe et Ali-Mikhail-Haq. Il existe également des copules de valeurs

extrêmes et la copule de Marshall-Olkin, mais elles ne seront pas considérées dans le cadre de cette étude.

d. Copule empirique

En l'absence d'information sur la distribution des marginaux, la copule empirique peut être une alternative. Introduite par (Deheuvels, 1979), elle se base sur les fonctions de répartition empiriques estimées à travers les données de rangs.

Soit l'échantillon (x_1^k, \dots, x_n^k) avec $1 \leq k \leq K$

et (r_1^k, \dots, r_n^k) avec $1 \leq r_i^k \leq n$, la statistique de rang associée. La copule empirique pour les échantillons de taille K est donnée par :

$$\hat{C}_K \left(\frac{k_1}{K}, \dots, \frac{k_n}{K} \right) = \frac{1}{K} \sum_{k=1}^K \prod_{i=1}^n \mathbb{1}_{r_i^k \leq k_i}$$

\hat{C}_K converge asymptotiquement vers C pour K assez grand (**Faivre, 2002**).

2.2.3 Estimation paramétrique

Dans le cas où la loi jointe dépend d'un paramètre ou d'un vecteur de paramètres, la détermination de la copule passe par l'estimation de ceux-ci. Il s'agit de l'estimation paramétrique de la copule. Dans tout ce qui suit, on note par $\theta \in \Theta$ le paramètre de la copule (Θ représente l'espace des paramètres). Trois méthodes d'estimation seront présentées (**Cadoux and Loizeau, 2004**).

a. Méthode des moments

A partir d'une mesure de concordance telle que le tau de Kendall ou le rho de Spearman, on peut estimer les paramètres d'une copule par inversion. Prenons le cas du tau de Kendall (le plus souvent utilisé comme mesure de concordance en raison de sa simplicité de calcul). En dimension 2, on rappelle l'estimateur empirique et l'expression théorique du tau de Kendall :

$$\hat{\tau}(X, Y) = \frac{\sum_{i < j} \text{signe}((X_i - X_j)(Y_i - Y_j))}{\binom{n}{2}}$$

$$\tau_{X,Y}(\theta) = 4\mathbb{E}[C(F_X, F_Y, \theta)] - 1$$

C représente la copule bivariée.

L'objectif est de rechercher la valeur du paramètre θ qui égalise la valeur théorique à l'estimateur empirique (asymptotiquement sans biais):

$$\tau_{X,Y}(\hat{\theta}) = \hat{\tau}(X, Y)$$

$$\hat{\theta} = \tau_{X,Y}^{-1}(\hat{\tau}(X, Y))$$

b. Maximum de vraisemblance (ML)

Cette méthode se base sur la densité d'une copule (dans le cas où elle existe) pour exprimer la log-vraisemblance. En effet, considérons un vecteur (X_1, \dots, X_n) de densité jointe f et notons par f_i les densités marginales. On peut écrire :

$$f(x_1, \dots, x_n) = c(F_1(x_1), \dots, F_n(x_n); \theta) \prod_{i=1}^n f_i(x_i)$$

En rappelant que

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \text{ et } c(u_1, \dots, u_n) = \frac{\partial^n C}{\partial u_1 \dots \partial u_n}(u_1, \dots, u_n)$$

Selon le type de copule, le paramètre θ peut être un scalaire ou une matrice.

Soient $(x_1^k, \dots, x_n^k), 1 \leq k \leq K$ un échantillon d'observations de taille K , l'expression de la log-vraisemblance en fonction du paramètre peut être obtenue :

$$l(\theta) = \sum_{k=1}^K \ln \left[c(F_1(x_1^k), \dots, F_n(x_n^k)) \right] + \sum_{k=1}^K \sum_{i=1}^n \ln \left[f_i(x_i^k) \right]$$

L'estimateur se déduit en maximisant $l(\theta)$: $\hat{\theta} = \operatorname{argmax} l(\theta) ; \forall \theta \in \Theta$

Un tel calcul peut être complexe compte tenu de la dimension des données. Dans les logiciels une approche numérique est souvent utilisée pour l'estimation.

c. Inference Functions for Margins (IFM)

La méthode IFM ou en français fonctions d'inférence des marginales, se déroule en deux étapes. Elle consiste à estimer les paramètres des distributions marginales par maximisation de leurs vraisemblances. Puis, les paramètres estimés sont introduits dans la structure de dépendance représentée par la copule, à partir de laquelle on estime le paramètre final $\hat{\theta}$ par une seconde maximisation de la vraisemblance.

On notera par $\theta = (\theta_1, \dots, \theta_n, \rho)$, le vecteur des paramètres, ρ représente les paramètres de la copule et θ_i celui de chaque marginal i . La log-vraisemblance globale peut s'écrire :

$$l(\theta) = \sum_{k=1}^K \ln \left[c(F_1(x_1^k, \theta_1), \dots, F_n(x_n^k, \theta_n); \rho) \right] + \sum_{k=1}^K \sum_i^n \ln \left[f_i(x_i^k; \theta_i) \right]$$

Pour déterminer $\hat{\theta}_{IFM}$, on on estime dans un premier temps les paramètres

$$\hat{\theta}_i = \operatorname{argmax} \sum_{k=1}^K \ln \left[f_i(x_i^k; \theta_i) \right]$$

puis, à partir des ces estimateurs on détermine :

$$\hat{\rho} = \operatorname{argmax} \sum_{k=1}^K \ln \left[c(F_1(x_1^k, \hat{\theta}_1), \dots, F_n(x_n^k, \hat{\theta}_n); \rho) \right]$$

Comme dans le cas du ML, on peut être confronté à un problème de complexité des calculs. Cependant la méthode IFM, présente en général plus de souplesse pour la maximisation.

Dans le logiciel R, la librairie *copula* contient un ensemble de fonctions adaptées aux copules. Pour l'estimation, toutes ces méthodes sont également disponibles en option de la fonction *fitCopula*.

Rappelons qu'il est également possible d'estimer une copule suivant une approche non-paramétrique. Cette méthode se base sur les distributions.

2.2.4 Cas des copules hiérarchiques : RVine Tree Copula

Les familles de copules présentées permettent en général de modéliser la plus part des formes de dépendance qu'on pourrait rencontrer sur des données. Néanmoins elles présentent un certain nombre limites.

Dans la pratique, on est souvent confronté à la problématique d'agrégation de plusieurs vecteurs. Cette situation nécessite l'utilisation des copules multivariées. Il s'agit donc de trouver la copule (de dimension $n \geq 2$) la mieux adaptée à la structure de dépendance des données. En dimension 2, il existe des méthodes de sélection de copule (test du χ^2), mais dans le cas multivarié le choix devient de plus en plus compliqué.

Des travaux plus récents ont tenté de proposer une classe plus large et plus complexe de copules multivariées (voir Dißmann et al., 2012 pour plus de détails). Cette classe est construite en utilisant des copules bivariées pour l'agrégation de certaines paires de variables. Ces blocs sont appelés copules de paires (pair-copula) et sont utilisées pour construire des distributions multivariées. L'identification des paires de variables et les variables conditionnelles correspondantes est faite à l'aide d'une séquence d'arbres qu'on appelle *regular trees* ou arbre régulier (R-vine). La démarche est la suivante :

- Pour un vecteur de dimension n , l'application du R-vine consiste dans un premier temps à identifier $n - 1$ paires de variables et modéliser directement leur distribution
- La deuxième étape permet d'identifier $n - 2$ paires de variables dont la distribution est modélisée par une copule de paires conditionnellement à une variable
- A l'étape suivante, on identifie aussi des paires de variables et détermine leur distribution conditionnelle à l'aide d'une pair-copula. A ce niveau l'ensemble de conditionnement comprend deux variables.
- En procédant de manière similaire, le dernier arbre est déterminé par une paire de variables, dont la distribution conditionnellement aux variables restantes est déterminée par une pair-copula.

1. Formulation

Un R-vine est un ensemble imbriqué de $n - 1$ arbres tel que les arrêtes (edges, ou jonction entre deux noeuds) de l'arbre j deviennent les noeuds de l'arbre $j + 1$

Notation : $\mathcal{V} = (T_1, \dots, T_{n-1})$

La structure telle que définie dans la figure ci-après est appelée *Constraint set* avec comme notation : $\mathcal{CV} = (a, b|e_i)$

avec $e_i =$ l'ensemble des variables de conditionnement

Et le couple a, b représentant la paire de variable

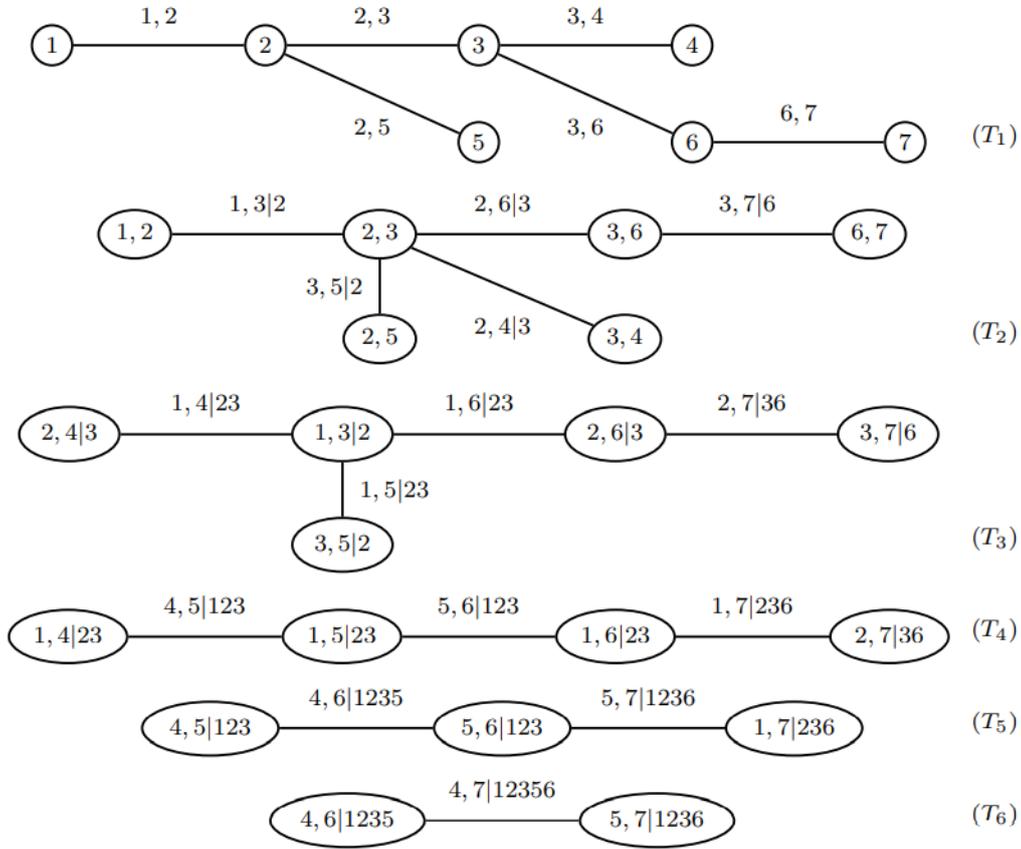


FIGURE 2.2: Exemple d'une structure R-vine de dimension $n=7$

2. Représentation matricielle

Pour faciliter l'implémentation d'algorithmes d'inférence statistique adapté à la structure R-vine, nous adopterons une représentation sous la forme d'une matrice triangulaire inférieure.

L'idée est de stocker le constraint set du R-vine dans les colonnes de cette matrice triangulaire.

Notations :

Soit la matrice triangulaire supérieure $M = (m_{i,j})_{i,j=1,\dots,n}$.

Le i^{eme} constraint set est donné par

$$C_M(i) = \{(m_{i,i}, m_{k,i}|D), \quad k = i + 1, \dots, n, \quad D = [m_{k+1,i}, \dots, m_{n,i}]\}, \quad i = 1, \dots, n - 1$$

Pour comprendre comment la matrice a été construite, on se place au niveau de la diagonale. Le premier élément de la diagonale noté $m_{i,i}$ est retenu pour former le couple de

variables dans le constraint set. Puis, on choisit un autre élément $m_{k,i}$, $k = i + 1, \dots, n$ dans la même colonne i . Le reste du constraint set est obtenu en considérant les éléments en dessous de $m_{k,i}$. Il s'agit de l'ensemble $D = [m_{k+1,i}, \dots, m_{n,i}]$. L'arbre contenant le constraint set correspond à celui obtenu à l'ordre $|D| + 1 = \text{Card}(D) + 1$

Considérons l'exemple avec la matrice ci-après qui reproduit la structure du R-Vine de la figure précédente.

$$\begin{pmatrix} 7 \\ 4 & 4 \\ 5 & 6 & 6 \\ 1 & 5 & 5 & 5 \\ 2 & 1 & 1 & 1 & 1 \\ 3 & 2 & 2 & 3 & 3 & 3 \\ 6 & 3 & 3 & 2 & 2 & 2 & 2 \end{pmatrix}$$

Au niveau de la deuxième colonne, on a $m_{2,2} = 4$ et à la troisième ligne de la même colonne, $m_{3,2} = 5$. En se basant sur la définition, on peut construire le constraint set $(\{4,5\}, \{1,2,3\})$ qu'on note aussi: 4,5|123.

Puisque $D = \{1,2,3\}$, $|D| + 1 = 4$, donc ce constraint set appartient à l'arbre T_4 .

3. Spécification de la densité jointe du R-Vine

La densité de la distribution R-Vine est donnée par le produit de densités des copules indexées par les constraint set définis précédemment. La structure complexe du R-Vine rend difficile un tel calcul. Par conséquent, un algorithme sera utilisé pour déterminer la densité. Pour illustrer, considérons le cas $n = 3$ variables. La densité jointe de la copule en se basant sur la structure R-Vine s'écrit :

$$c(u_1, u_2, u_3) = c_{12}(u_1, u_2) \cdot c_{13}(u_1, u_3) \cdot c_{23|1}(C_{2|1}(u_2, |u_1), C_{3|1}(u_3, |u_1))$$

L'algorithme prend en compte d'autres paramètres que sont les matrices carré $T = (t_{i,j})$ et $P = (p_{i,j})$ contenant respectivement les informations sur les types (Exemple : Normal, Student, Clayton, etc.) et les paramètres des copules bivariées. Étant donnée que certaines

copules peuvent nécessiter plusieurs paramètres, il est possible que P soit un tableau multi-dimensionnel au lieu d'une simple matrice.

La figure ci-après fait le récapitulatif des matrices utilisées comme inputs dans l'algorithme. On peut l'interpréter comme suit : la copule qui lie les variables indexées par $\{4, 5\}$ et avec comme variables de conditionnement $D = (1, 2, 3)$ est notée $c_{4,5|123}$, de type $t_{4,1}$ et ayant comme paramètre $p_{4,1}$. La copule $c_{7,6}$ est de type $t_{7,4}$ et de paramètre $p_{7,4}$.

$M^* =$	$T^* =$	$P^* =$
(4)	$t_{2,1}$	$p_{2,1}$
7 5	$t_{3,1} \quad t_{3,2}$	$p_{3,1} \quad p_{3,2}$
6 7 1	($t_{4,1}$) $t_{4,2} \quad t_{4,3}$	($p_{4,1}$) $p_{4,2} \quad p_{4,3}$
(5) 6 7 (7)	$t_{5,1} \quad t_{5,2} \quad t_{5,3} \quad t_{5,4}$	$p_{5,1} \quad p_{5,2} \quad p_{5,3} \quad p_{5,4}$
1 1 6 2 6	$t_{6,1} \quad t_{6,2} \quad t_{6,3} \quad t_{6,4} \quad t_{5,5}$	$p_{6,1} \quad p_{6,2} \quad p_{6,3} \quad p_{6,4} \quad p_{6,5}$
2 3 3 3 2 2	$t_{7,1} \quad t_{7,2} \quad t_{7,3} \quad (t_{7,4}) \quad t_{7,5} \quad t_{7,6}$	$p_{7,1} \quad p_{7,2} \quad p_{7,3} \quad (p_{7,4}) \quad p_{7,5} \quad p_{7,6}$
3 2 2 (6) 3 3 3		

FIGURE 2.3: Inputs de l'algorithme : matrice des constraint set, types et paramètres des copules

En annexe, on présente un algorithme récursif permettant de calculer la distribution conditionnelle (voir A.1).

4. Procédure de sélection de la structure R-Vine

Les différentes étapes permettant d'estimer une copule R-Vine sur des données sont

- Sélection de la structure de départ, c'est à dire les $n - 1$ paires de variables à considérer
- Le choix de la copule bivariée pour chaque paire considérée.
- Estimation des paramètres correspondants pour chaque copule

Les étapes cruciales pour la détermination du meilleur modèle parmi toutes les structures R-Vine possibles sont les étapes 2 et 3. En effet, le nombre de structures R-Vines possibles pour un vecteur de taille n est de l'ordre de $\frac{n!}{2} 2^{\binom{n-2}{2}}$ (Morales-Napoles, Cooke, and Kurowicka, 2010).

Etape 1 : La sélection du premier arbre (structure de départ) peut se baser sur des mesures de dépendance telles que le tau de Kendall. Le choix de ce critère s'explique par

le fait que la valeur tau de kendall entre deux variables ne dépend pas des hypothèses de distribution et s'adapte aux cas non linéaires et non gaussiens.

Dans cette étape, l'algorithme calcule dans un premier temps le tau de Kendall empirique $\hat{\tau}_{j,k}$ pour chaque paire de variables $\{j,k\}$ $j,k = 1, \dots, n$ et $j \neq k$. Il choisit l'arbre qui maximise la somme de la valeur absolue des tau de Kendall empirique.

Etape 2 : Après avoir sélectionné l'arbre de départ, on choisit la copule bivariable \hat{C}_{jk} correspondant à chaque paire. Ce travail peut être effectué à travers des méthodes graphiques (Chi-plots) ou autres critères statistiques.

Etape 3 : L'estimation des paramètres de la copule peut se faire aisément par la méthode du maximum de vraisemblance et méthodes des moments.

La même procédure sera répétée pour construire les autres arbres et définir les constraint set. En effet, on calcule les tau $\hat{\tau}_{j,k|D}$ pour toutes les paires de variables conditionnelles $\{j,k|D\}$. On utilise le même critère que précédemment pour sélectionner l'arbre, puis on détermine la copule conditionnelle $\hat{C}_{jk|D}$ et on estime les paramètres (voir A.2).

Cet algorithme est appelé maximum spanning tree (MST) et il est important de rappeler qu'il est possible d'utiliser d'autres indicateurs à la place du tau de kendall empirique.

Les copules hiérarchiques présentent aussi quelques limites, notamment le choix de la structure et du modèle. Lorsque la taille augmente, les algorithmes prennent du temps à fournir des résultats précis.

Dans ce chapitre, plusieurs mesures de dépendance entre risques ont été présentées. Qu'il s'agit de simple corrélations ou de dépendances via les copules, l'objectif est de trouver la structure de dépendance permettant d'agréger les risques marginaux et d'obtenir un risque global diversifié répondant aux exigences réglementaires.

Chapitre 3

Application au portefeuille AXA France IARD : méthodes initiales d'agrégation du STEC réserve

Dans ce chapitre, les méthodes d'agrégation présentes dans la documentation du modèle interne seront passées en revue. Dans un premier temps, un rappel sur la modélisation du risque de réserve sera présenté.

3.1 Rappels sur la modélisation du risque de réserve

Le STEC de réserve d'une Lob peut être défini comme le montant de capital nécessaire pour absorber le risque de réserve de cette Lob dans un horizon annuel avec une probabilité de 99,5%. Le calcul se base sur les données de sinistres relatives à tous les contrats souscrits pour cette Lob. On adoptera une structure par triangle de charges à partir duquel des méthodes de provisionnement à horizon un an permettront de modéliser le risque de réserve. Cependant, par rapport à ce risque, il existe une dépendance entre les années de survenance qui fait également l'objet d'une estimation.

Le calcul du STEC par Lob nécessite la détermination du volume des réserves qu'on notera EAXA (donnée déjà disponible) et la volatilité des réserves à horizon 1 an. Le focus sera sur mis la calibration de cette volatilité avec la méthodologie de Merz & Wuthrich utilisée dans le modèle interne et la calibration des corrélations entre années d'accident .

3.1.1 Le risque de réserve

Pour assurer la solvabilité de la compagnie, il est nécessaire de tenir compte d'un niveau de capital suffisant pour absorber les pertes significatives futures non attendues. La détermination de ces passifs se base sur une estimation des charges qui dépendent à leur tour des fréquences de survenance des sinistres et des coûts. Durant la période de développement des sinistres, l'assureur est souvent exposé aux risques liés aux incertitudes sur :

- les sinistres tardifs ou IBNR (Incurred but not yet reported)
- la nature aléatoire des montants de sinistre en cours
- la durée des règlements (reflétant le processus de traitement des sinistres et une éventuelle réouverture).
- les sinistres non encore manifestés dans le cas de la construction,
- etc.

La dynamique de vie des sinistres par année de survenance est représentée sous forme de triangle et cette représentation est fondamentale dans la modélisation du risque de réserve. Les dimensions sont : l'année de survenance en ligne, l'année de développement en colonne et l'année calendaire en diagonal. Il peut s'agir d'un triangle de paiements ou d'un triangle de charges. Tel est le point de départ de la calibration.

Nous adopterons les notations suivantes :

- les indices $i \in \{1, \dots, I\}$ et $j \in \{1, \dots, J\}$ en supposant que $I = J$ désignent respectivement les années de survenance et de développement;
- $i + j = n$ représente l'année calendaire. En effet, le triangle sera désigné par $N = I + J$ pour faire référence à l'année la plus récente pour laquelle les données sont disponibles. Le triangle $N - 1$ est donc celui obtenu en retirant la dernière diagonale, ainsi de suite. Ainsi, on note par D_n l'ensemble des données disponibles à la date n
- on notera par C_{ij} et P_{ij} respectivement les charges et paiements cumulées, X_{ij} les valeurs incrémentales et \hat{R}_{ij} les réserves estimées.

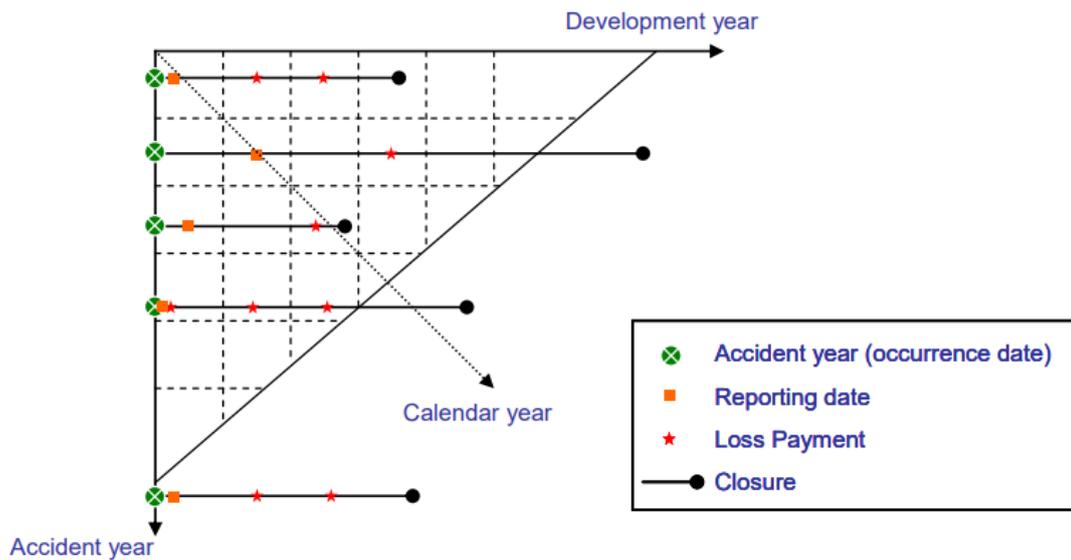


FIGURE 3.1: Dynamique de vie d'un sinistre

Source : Documentation modèle interne

3.1.2 Charges ultimes, CDR et volatilité des réserves

Les premières étapes du modèle consisteront à établir une méthodologie permettant d'estimer les charges ultimes, c'est à dire le montant total des charges à payer pour les sinistres survenus au cours d'une même année i . Autrement dit, il s'agira d'appliquer des méthodes de provisionnement permettant d'estimer les charges pour les années de développement futures. Ces méthodes peuvent être déterministes ou stochastiques et la plupart se basent sur l'hypothèse selon laquelle les cadences de paiements des charges du passé se reproduisent dans le futur.

Etant donné qu'on se place sur un horizon d'un an, le modèle de Merz et Wuthrich présente un cadre méthodologique qui s'adapte à cette problématique. Les hypothèses du modèle sont les suivantes :

- H1 : indépendance entre les paiements cumulés des différentes années d'accident
- H2 : Les processus $(C_{ij})_{j>1}$ sont markoviens, et il existe des réels $f_j > 0$ et $\sigma_j > 0$, $j = 1, \dots, J - 1$ tels que pour tout $i = 1, \dots, I$

$$\mathbb{E}[C_{i,j}|C_{i,j-1}] = f_{j-1}C_{i,j-1}$$

$$\text{Var}[C_{i,j}|C_{i,j-1}] = \sigma_{j-1}^2 C_{i,j-1}$$

A travers ces hypothèses et en s'inspirant de la méthode Chain-Ladder et Mack, il est démontré que les facteurs f_j sont sans biais et non corrélés. Ils peuvent être estimés par :

$$\hat{f}_j = \frac{\sum_{i=1}^{I-j-1} C_{i,j+1}}{\sum_{i=1}^{I-j-1} C_{i,j}}$$

Une estimation sans biais des charges ultimes peut donc être obtenue par la relation :

$$\hat{C}_{i,I} = C_{i,I-i-1} \prod_{j=I-i-1}^{I-1} \hat{f}_j$$

Et le montant des réserves correspondant à l'année de survenance i est estimé par :

$$\hat{R}_i = \hat{C}_{i,I} - C_{i,I-i-1}$$

En raison de l'incertitude liée à cette estimation, on cherchera à modéliser la volatilité de ces réserves. A ce niveau, il est important de rappeler que ces quantités sont estimées à l'ultime. Si on se place dans une vision à 1 an comme dans le cadre solvabilité 2, Merz et Muthrich ont introduit une notion appelée **Claim Development Result** ou CDR.

En considérant une vision en année calendaire et en année de survenance, une définition mathématique du $CDR_i(n)$ est la différence entre les estimations de deux années calendaires successives n et $n - 1$ du montant de charge à l'ultime et la formule est donnée par :

$$CDR_i(n) = \mathbb{E}[C_{i,n-1}|D_{n-1}] - \mathbb{E}[C_{i,n-1}|D_n]$$

Avec les données observables (estimées à partir des données du triangle) on a :

$$C\hat{D}R_i(n) = \hat{C}_{i,I}^{n-1} - \hat{C}_{i,I}^n$$

$\hat{C}_{i,I}^n$ (resp. $\hat{C}_{i,I}^{n-1}$) étant l'estimation des charges cumulées du triangle n (resp. $n - 1$).

Cette quantité peut être vue comme une estimation de la déviation de la charge ultime entre deux années calendaires. Idéalement, celle-ci devrait être nulle, c'est le cas où le développement futur des sinistres est parfaitement connu. Dès lors, plus le CDR en absolu est élevé, plus le risque de faire erreur sur les réserves est important.

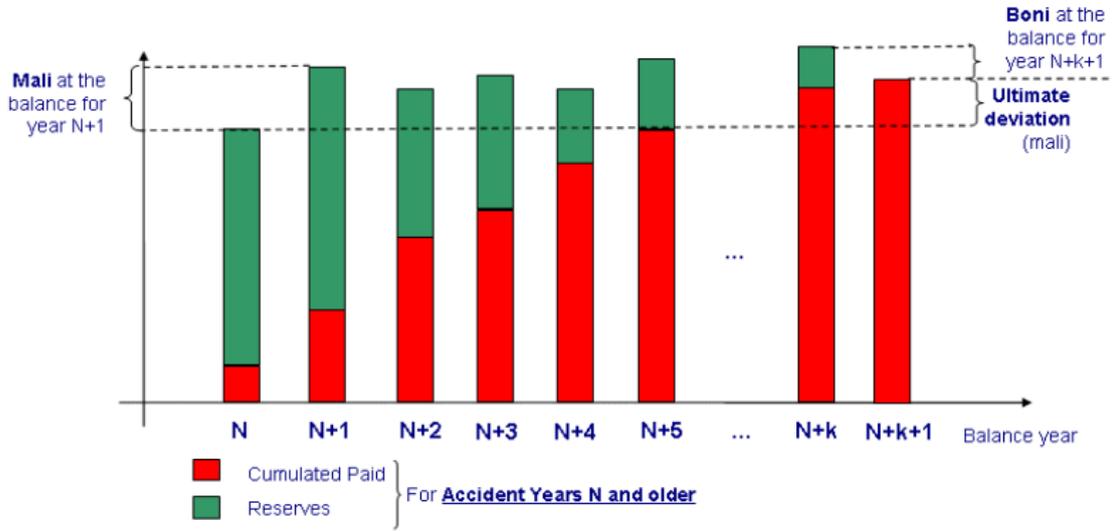


FIGURE 3.2: Illustration du CDR

Source : Documentation modèle interne

Quelques propriétés sur les CDR peuvent être déduites de l'hypothèse $H2$. En effet, il en découle que $\mathbb{E}[C_{i,n}|D_n]$ est martingale, ce qui entraîne que $\mathbb{E}[CDR_i(n)|D_{n-1}] = 0$.

Ceci s'interprète par le fait qu'en moyenne, la meilleure prédiction du CDR pour l'année en cours est 0. On s'intéressera alors aux déviations du CDR autour de la valeur 0. La mesure d'incertitude qui sera utilisée est le $MSEP$ (Mean Square Error Prediction). En d'autres termes, il s'agit de l'indicateur d'incertitude à horizon 1 an des ultimes.

$$MSEP_{\hat{C}\hat{D}R_i(n)|D_n(0)} = \mathbb{E}[(\hat{C}\hat{D}R_i(n) - 0)^2|D_n] = \mathbb{E}[(\hat{C}\hat{D}R_i(n))^2|D_n]$$

Avec cette méthodologie, les hypothèses faites et les données du triangle, on arrive à une formule fermée du $MSEP$. Dans la pratique, cette formule est implémentée dans le modèle interne pour la calibration du risque de réserve.

$$MSEP_{\hat{C}\hat{D}R_i(n)|D_{n-1}}(0) =$$

$$(\hat{C}_{i,I}^{n-1})^2 \times \left[\frac{\hat{\sigma}_{I-i+1}^2 / (\hat{f}_{I-i+1}^{n-1})^2}{C_{i,I-i+1}} + \frac{\hat{\sigma}_{I-i+1}^2 / (\hat{f}_{I-i+1}^{n-1})^2}{S_{I-i+1}^{n-1}} + \sum_{j=I-i+2}^{n-2} \left(\frac{C_{I-j+1,j}}{S_j^n} \right) \frac{\hat{\sigma}_j^2 / (\hat{f}_j^{n-1})^2}{S_j^{n-1}} \right]$$

$$\text{avec } S_j^n = \sum_{i=1}^{I-j} C_{i,j}$$

$$\text{Et } \hat{\sigma}_j^2 = \begin{cases} \frac{1}{I-j-1} \sum_{i=1}^{i-j} C_{i,j} \frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j^2 \\ \min \left[\frac{\hat{\sigma}_{I-2}^4}{\hat{\sigma}_{I-3}^2}, \min(\hat{\sigma}_{I-3}^2, \hat{\sigma}_{I-2}^2) \right] \end{cases}$$

Grâce cette formule, on estime la volatilité des réserves par le coefficient de variation du MSEP dont l'expression est donnée par :

$$cv_{ij} = \frac{SEP_{C\hat{D}R_i(i+j)|D_{n-1}}(0)}{R_{i,j+1}} = \frac{\sqrt{MSEP_{C\hat{D}R_i(i+j)|D_{n-1}}(0)}}{R_{i,j+1}}$$

3.1.3 Modélisation des Boni Mali

Dans ce qui précède, nous avons montré que le $CDR_i(n)$ est de moyenne nulle et sa variance est égale à $MSEP_{C\hat{D}R_i(n)|D_{n-1}}(0)$. L'objectif sera de construire une loi permettant d'aboutir à la distribution de pertes (loss distribution) nécessaire au calcul du STEC. Pour cela nous allons définir les taux de Boni Mali ou taux de CDR qui correspondent aux CDR normalisés (généralement par les montant de réserves). En effet, l'importance de travailler sur des taux (ou Boni Mali en pourcentage des réserves) au détriment des CDR absolus se justifie par le fait que les réserves peuvent être volatiles d'une année à l'autre. De plus, ceci permet de considérer une erreur de prédiction relative aux réserves de telles sortes qu'il soit possible de comparer les risques. Ils sont donnés par la relation :

$$BM_{i,j} = \frac{\hat{C}_{i,n}^{i+j+1} - \hat{C}_{i,n}^{i+j}}{\hat{R}_{i,j}} = \frac{\hat{R}_{i,j+1} + P_{i,j+1} - P_{i,j} - \hat{R}_{i,j}}{\hat{R}_{i,j}}$$

Dans le modèle interne AXA, on fait une hypothèse sur la distribution de ces taux de Boni Mali. Le modèle suppose que $1 + BM_{ij}$ est distribué suivant une loi Lognormale de moyenne 1 et de variance égale à cv_{ij}^2 (définie dans la section précédente).

$$1 + BM_{ij} \sim \mathcal{LN}(1, cv_{ij})$$

Désormais nous disposons de distributions marginales pour chaque année d'accident,

et donc on est en mesure de calculer le STEC en déterminant le quantile 99,5% de la distribution somme. Mais en raison de la dépendance, il est nécessaire de passer par une étape d'agrégation.

3.2 Dépendance inter Lob basée sur les paramètres de la FS et sur les Risk Driver

Dans ce qui précède, nous avons présenté la méthodologie utilisée dans le modèle interne pour la calibration du risque de réserve. Cette démarche permet de calculer le STEC réserve en agrégeant les années d'accident au sein de chaque Lob. Il s'agit là du premier niveau d'agrégation qui a fait l'objet d'un changement de modèle interne validé et a été appliquée dans les calculs du STEC en fin d'année 2021. L'impact STEC qui en découle a fourni un gain en diversification.

Comme nous l'avons évoqué dans la section 1.2.4, il existe un second niveau d'agrégation pour le STEC réserve. Il s'agit de l'agrégation inter Lob. Cette problématique est l'un des chantiers en P&C qui devra faire l'objet d'un changement de modèle. Les travaux sont en cours aussi bien pour le risque de réserve que le risque de prime. A l'instar des corrélations entre années d'accident, les méthodes initiales d'agrégation, bien qu'elles satisfont les exigences réglementaires, semblent être plus prudentes que nécessaire. Pour rappel, le but est de déterminer le STEC réserve qui est destiné à couvrir les risques liés aux erreurs d'estimation sur les charges de sinistres. Dans un tel contexte, l'idéal serait d'estimer des corrélations inter Lob qui reflètent au mieux ce risque. Cependant, les méthodes initiales ne permettent pas d'établir un lien entre les corrélations inter Lob et la sinistralité.

Ainsi, à travers cette nouvelle méthodologie, nous chercherons à explorer les techniques d'estimation des dépendances qui peuvent être appliquées dans le cas des Lobs du portefeuille non vie AXA France. Ces techniques devront satisfaire un certain nombre de critères :

- Applicables sur l'ensemble des Lobs du portefeuille
- En phase avec la réglementation solvabilité 2
- Un impact STEC qui fournit un gain de diversification

- La pertinence et l'interprétabilité des résultats
- Stabilité dans le temps et face aux variations de la sinsistralité dans les triangles
- Justification théorique et/ou pratique

Pour mieux aborder la problématique des dépendances inter Lob, nous allons dans un premier temps présenter comment est structuré le portefeuille P&C AXA France. Cette partie a pour but de montrer les spécificités qui existent dans le portefeuille et donner une première idée sur les caractéristiques des Lobs et les liens qui pourraient être observés à priori entre ces derniers.

3.2.1 Caractéristiques des Lobs du portefeuille d'AXA France IARD

Dans le marché de l'assurance, les compagnies exercent leur activité en proposant différents types de produits pour satisfaire la demande et générer un profil. Le choix du portefeuille et les risques à assurer est donc primordial chez les assurés. Chez AXA France IARD, la segmentation du portefeuille est plus fine qu'en formule standard. Actuellement, il compte 36 Lobs et plusieurs raisons expliquent cette importante taille du portefeuille. En effet, une segmentation est faite en fonction des types de produits, les garanties souscrites, le type d'assuré (selon qu'il s'agit des particuliers ou des entreprises) ou l'entité de AXA en charge du produit (AXA SA, Direct Assurance, CLP, etc.). Cette démarche devrait permettre une gestion plus adéquate des risques aux quels le portefeuille est exposé.

A travers cette segmentation, il est également établi une correspondance entre les 12 Lobs de la formule standard et celles du modèle interne (tableau 3.2). Cette approche est utile car elle permet des comparaisons et des contrôles sur les résultats du modèle interne. Par exemple, en responsabilité civile, bien que le modèle interne fait la distinction entre responsabilité civile médicale, entreprises, particuliers et construction, ces derniers peuvent être regroupés au sein de la même branche "General liability insurance" en formule standard. Ce mapping entre Lobs modèle interne et formule standard permet également de compléter certains processus de calculs. C'est le cas lorsque certaines Lobs ne présentent pas suffisamment de données ou si on souhaite utiliser des paramètres de la formule standard. Nous verrons dans la suite que ce mapping joue un rôle important dans la calibration des dépendances entre Lobs.

Il est important de préciser que le nombre de Lobs chez AXA France évolue dans le temps compte tenu des risques émergents et la nécessité d'ajuster le portefeuille aux réalités du marché. Ainsi, les Lobs n'ont pas les mêmes profondeurs de données et certaines sont plus importantes en volume que d'autres (en termes de primes acquises brutes de réassurance). Les données du tableau 3.1 révèlent que la branche auto dommage pour particuliers a la part la plus importante du portefeuille. Durant l'année 2021, l'entité a acquis plus d'un milliard de primes pour cette Lob, soit 14,4% du portefeuille. La branche MRH est également l'une des plus importantes en termes de volumes. Les primes acquises en 2021 pour cette Lob dépassent les 800 millions, ce qui représente 11,8%. Les autres Lobs ont toutes des parts inférieures à 8% et environ une quinzaine n'atteignent pas 100 millions de primes. En outre, on note des différences concernant leur exposition au risque, leur sensibilité à certains scénarios et chocs externes, ou selon que ce soit une branche longue ou courte. Ces disparités qui existent entre les Lobs peuvent plus ou moins influencer leurs corrélations et par ricochet l'impact sur le STEC.

Description	Lob AXA France	Volume primes 2021	Poids relatif
Assurance grêle en Agriculture	AGRI_GRELE	69 448 114	0,90%
Assistance	ASSISTANCE_DA	-	0,00%
Assistance	ASSISTANCE	10 289 789	0,10%
Auto entreprise dommage	AUTO_EN_DOM	440 899 762	5,90%
Auto entreprise Resp. civile	AUTO_EN_RC	305 305 582	4,10%
Auto particuliers dommage	AUTO_PP_DOM	1 070 503 439	14,40%
Auto particuliers Resp. civile	AUTO_PP_RC	461 140 130	6,20%
Autres entreprises	AUTRES_EN	38 384 982	0,50%
Autres particuliers	AUTRES_PP	4 549 084	0,10%
Catastrophe naturelle	CATNAT	223 185 897	3,00%
Caution	CAUTION	13 654 369	0,20%
Dommage aux biens en construction	DAB_CONST	34 953 934	0,50%
Divers dommages aux bien - entreprise	DIVERS_DAB_EN	125 640 584	1,70%
Divers dommages aux bien - particuliers	DIVERS_DAB_PP	87 949 813	1,20%
Dommage ouvrage	DO	104 894 623	1,40%
Dommage aux biens	HOD_DA	50 476 792	0,70%
Marine aviation et transport - entreprise	MAT_EN	47 831 955	0,60%
Marine aviation et transport - entreprise	MAT_PP	17 729 090	0,20%
auto	MBI_DA	263 580 347	3,50%
autres auto	MOD_DA	249 423 883	3,40%
Multirisques habitation	MRH	878 784 390	11,80%
Multirisques immeuble - entreprises	MRI_EN	322 959 890	4,30%
Multirisques immeuble - particuliers	MRI_PP	118 066 269	1,60%
Multirisques professionnel - entreprise	MRP_EN	35 242 037	0,50%
Multirisques professionnel - particuliers	MRP_PP	203 568 532	2,70%
Protection juridique	PJ	137 263 867	1,80%
Réassurance pour dommage aux biens	POOL_PROPERTY	158 388 199	2,10%
pertes diverses - entreprises et particuliers	PP_EN	564 057 569	7,60%
Quote part cédée	QPSAMUT	184 587 616	2,50%
Responsabilité civile en construction	RC_CONST	214 928 163	2,90%
Responsabilité civile - entrprises	RC_ENT	387 900 706	5,20%
Responsabilité civile médicale	RC_MED	-	0,00%
Responsabilité civile - particuliers	RC_PP	46 314 383	0,60%
Responsabilité civile décennale	RCD	193 050 774	2,60%
Risques industriels	RI	360 997 954	4,90%
santé	SANTE	190	0,00%

TABLE 3.1: Les Lobs du portefeuille P&C AXA France et leurs poids en volumes de primes

Nom Lob AXA	Nom Lob FS	Description Lob FS	Num
1. AUTO_PP_RC	Motor veh liab ins	Responsabilité civile auto	4
2. AUTO_EN_RC	Motor veh liab ins	Responsabilité civile auto	4
3. AUTO_PP_Dom	Oth motor ins	Autres automobile	5
4. MBI_DA	Motor veh liab ins	Responsabilité civile auto	4
5. MOD_DA	Oth motor ins	Autres automobile	5
6. MAT_PP	Mat ins	Maritime, aviation et transport	6
7. MAT_EN	Mat ins	Maritime, aviation et transport	6
8. AUTO_EN_Dom	Oth motor ins	Autres automobile	5
9. MRH	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
10. MRP_PP	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
11. MRI_PP	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
12. AGRI	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
13. DIVERS_DAB_PP	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
14. RI	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
15. MRP_EN	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
16. MRI_EN	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
17. DAB_const	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
18. DIVERS_DAB_EN	Misc fin loss ins	Pertes financières diverses	12
19. CATNAT	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
20. DO	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
21. HOD	Fire other dam to prop ins	Incendie et autres dommages aux biens	7
22. RC_PP	Gen liab ins	Responsabilité civile générale	8
23. RC_CONST	Gen liab ins	Responsabilité civile générale	8
24. RC_ENT	Gen liab ins	Responsabilité civile générale	8
25. RC_MED	Gen liab ins	Responsabilité civile générale	8
26. RCD	Gen liab ins	Responsabilité civile générale	8
27. PJ	Leg exp ins	Protection juridique	10
28. ASSISTANCE	Ass	Assistance	11
29. ASSISTANCE_DA	Ass	Assistance	11
30. AUTRES_EN	Non-prop casu reins	Réassurance dommage non proportionnelle	26
31. AUTRES_PP	Non-prop casu reins	Réassurance dommage non proportionnelle	26
32. CAUTION	Cred and suretyship ins	crédit et cautionnement	9
33. POOL_PROPERTY	Fire other dam to prop reins	Réassurance incendie, autres dommages	28
34. PP_EN	Misc fin loss ins	Pertes financières diverses	12
35. QPSAMUT	Non-prop prop reins	Réassurance non proportionnelle de biens	28
36. SANTE	Med exp ins	Frais médicaux	1

TABLE 3.2: Table de correspondance entre Lob AXA et Lob Formule Standard

3.2.2 Méthode d'agrégation initiale : matrice de corrélation avec les paramètres de la FS

Comme évoqué dans la section précédente, la correspondance établie entre Lob AXA France et Lob FS permet de déterminer de manière simple et prudente une matrice de corrélation stable permettant d'agréger le risque de réserve. Jusqu'à présent, dans le modèle interne, le STEC réserve est calculé en utilisant cette matrice. On s'intéressera dans cette partie à comment celle-ci a été construite à partir de la formule standard.

Dans le tableau 3.3, nous rappelons la matrice de corrélation obtenue via la méthodologie QIS5 de la formule standard. Les paramètres de corrélation varient entre 25 et 50%.

Corrélations/Lob	4	5	6	7	8	9	10	11	12	28	26	27
4. Motor veh liab ins	1											
5. Oth motor ins	0,5	1										
6. Mat ins	0,5	0,25	1									
7. Fire other dam	0,25	0,25	0,25	1								
8. Gen liab ins	0,5	0,25	0,25	0,25	1							
9. Cred and sur ins	0,25	0,25	0,25	0,25	0,5	1						
10. Leg exp ins	0,5	0,5	0,25	0,25	0,5	0,5	1					
11. Assistance	0,25	0,5	0,5	0,5	0,25	0,25	0,25	1				
12. Misc fin loss ins	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1			
28. Non-prop prop reins	0,25	0,25	0,25	0,5	0,25	0,25	0,25	0,5	0,25	1		
26. Non-prop casu reins	0,25	0,25	0,25	0,25	0,5	0,5	0,5	0,25	0,25	0,25	1	
27. reins. (MAT)	0,25	0,25	0,5	0,5	0,25	0,25	0,25	0,25	0,5	0,25	0,25	1

TABLE 3.3: Corrélations entre Lob Formule Standard

Au niveau du modèle interne, il s'agit d'une version adaptée de la matrice de corrélation de la méthodologie QIS5 en considérant une corrélation de 75% entre les Lobs AXA appartenant à la même Lob FS. L'hypothèse sous-jacente qui a conduit à ce choix méthodologique est de supposer que ces Lobs appartenant à la même Lob FS ne sont pas forcément affectées par les mêmes sources de risques. Cette approche est commune pour la grande majorité du marché, donc l'utilisation de paramètres communs tels que ceux de QIS5 peut être une bonne approximation. En guise d'exemple, la corrélation entre les Lobs RC entreprise et RC médicale est de 75%.

En ce qui concerne le cas des Lobs qui ne figurent pas dans la même Lob FS, on prendra comme coefficient de corrélation celui entre les Lobs FS concernées. Par exemple, toutes les Lobs de la branche fire and other damage to property insurance ont la même corrélation que toutes celles de General liability insurance qui est égale à 25%. Plus spécifiquement,

MRH (dans fire and other damage to proprety) et RC des particuliers (general liability) sont corrélées à 25%.

A travers cette démarche une matrice carrée de taille 36 a été construite, servant à l’agrégation du STEC réserve. Celle-ci est représentée partiellement dans le tableau 3.4. Cependant, pour des contraintes d’espace/affichage, on ne peut présenter la matrice qu’avec une vingtaine de Lobs.

Lob	AUTO_PP_RC	AUTO_EN_RC	AUTO_PP_Dom	MBI_DA	MOD_DA	MAT_PP	MAT_EN	AUTO_EN_Dom	MRH	MRP_PP	MRI_PP	AGRI_GRELE	DIVERS_DAB_PP	RI	MRP_EN	MRI_EN	DAB_CONST	DIVERS_DAB_EN	CATNAT	DO
AUTO_PP_RC	1	0,75	0,5	0,75	0,5	0,5	0,5	0,5	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
AUTO_EN_RC	0,75	1	0,5	0,75	0,5	0,5	0,5	0,5	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
AUTO_PP_Dom	0,5	0,5	1	0,5	0,75	0,25	0,25	0,75	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
MBI_DA	0,75	0,75	0,5	1	0,5	0,5	0,5	0,5	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
MOD_DA	0,5	0,5	0,75	0,5	1	0,25	0,25	0,75	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
MAT_PP	0,5	0,5	0,25	0,5	0,25	1	0,75	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
MAT_EN	0,5	0,5	0,25	0,5	0,25	0,75	1	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
AUTO_EN_Dom	0,5	0,5	0,75	0,5	0,75	0,25	0,25	1	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
MRH	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	1	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75
MRP_PP	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	1	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75
MRI_PP	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	0,75	1	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75
AGRI_GRELE	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	0,75	0,75	1	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75
DIVERS_DAB_PP	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	0,75	0,75	0,75	1	0,75	0,75	0,75	0,75	0,75	0,75	0,75
RI	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	0,75	0,75	0,75	0,75	1	0,75	0,75	0,75	0,75	0,75	0,75
MRP_EN	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	0,75	0,75	0,75	0,75	0,75	1	0,75	0,75	0,75	0,75	0,75
MRI_EN	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	0,75	0,75	0,75	0,75	0,75	0,75	1	0,75	0,75	0,75	0,75
DAB_CONST	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	1	0,75	0,75	0,75
DIVERS_DAB_EN	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	1	0,75	0,75
CATNAT	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	1	0,75
DO	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	1

TABLE 3.4: Corrélations Lobs modèle interne basées sur celles des Lobs FS

Le STEC réserve sera obtenu en appliquant la formule d’agrégation standard.

$$STEC_{res} = \sqrt{(STEC_i)_{i=1,\dots,36}^T \times \Sigma \times (STEC_i)_{i=1,\dots,36}} \tag{3.1}$$

i et j sont les indices des Lobs

$\Sigma = [corr(i, j)]_{i, j \in 1, \dots, 36}$ la matrice de corrélations construite

3.2.3 Méthode d'agrégation basée sur les scénarios Risk Driver

Les corrélations inter Lob peuvent être estimées en considérant les sources de risque. Il s'agit d'un ensemble de scénarios exogènes ou facteurs de risque qui peuvent avoir des effets directs ou indirects sur les Lobs. La méthode se base sur une cartographie des risques développée par le Risk Drivers Working Group (voir tableau 3.5). L'objectif est de challenger la matrice de corrélation QIS sur la base de laquelle est construite les corrélations inter Lob. Cette approche risk driver a l'avantage de s'adapter aux risques qui peuvent affecter le portefeuille et créer des déviations sur l'estimation des primes et charges de sinistres. A ce jour, 16 facteurs ont été identifiés, et ce nombre peut évoluer compte tenu des changements liés à l'environnement économique et financier.

Num	Facteurs de risque
1	Changements dans les structures sociales
2	Changements techniques
3	Action collective (indemnisation)
4	Augmentation de fréquence
5	Comportement défavorable des clients en matière d'assurance
6	Crise macro-économique ou récession économique grave
7	Surinflation (IPC)
8	Superinflation (surinflation sectorielle)
9	Risques émergents détectables
10	Facteur de risque inconnu
11	Problèmes politiques
12	Changements dans le traitement des sinistres
13	Événements naturels
14	Fait par l'homme (man made)
15	Accumulation
16	Scénarios défavorables pour les rentes

TABLE 3.5: Liste des 16 risk driver

Dans la pratique, cette approche suppose que le coefficient de corrélation entre deux lobs dépend du nombre de facteurs qui affectent simultanément ces dernières. Une matrice qualitative des impacts de chaque facteur sur les Lobs permet de réaliser le calcul des corrélations. Elle prend la valeur 1 si la Lob en question est impactée par un facteur et 0 sinon. Celle-ci a été établie par expert judgement en analysant cas par cas les facteurs de risques qui peuvent impacter ou non une Lob. Cette analyse se fait par expérience sur le portefeuille ou

suivant une vision économique. Les corrélations sont calculées par les formules suivantes :

$$\text{corr}(Lob_i, Lob_j) = \frac{\text{Card}(Lob_i \cap Lob_j)}{\text{Card}(Lob_i) + \text{Card}(Lob_j) - \text{Card}(Lob_i \cap Lob_j)}$$

On peut aussi considérer les corrélations empiriques linéaires :

$$\text{Corr}(Lob_i, Lob_j) = \frac{\text{Cov}(Lob_i, Lob_j)}{\sigma_{Lob_i} \sigma_{Lob_j}}$$

Numéro risk driver	Auto_PP_RC	Auto_EN_RC	Auto_PP_Dom	Auto_EN_DOM	MRH	MRP_PP	MRI_PP	Agri_Grele	Divers_Dab_PP	RI	MRP_EN	MRI_EN	DAB_Const	Divers_Dab_EN	RC_PP	RC_CONST	RC_ENT	RC_MED	MAT_PP	MAT_EN	CATNAT	DO
1	1	1	0	0	1	1	1	0	0	1	1	1	1	0	1	1	0	1	0	0	1	1
2	1	1	0	0	1	1	1	0	0	1	1	1	1	0	1	1	1	1	0	0	0	1
3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
4	0	0	1	1	1	1	1	0	1	0	1	0	1	1	0	0	1	1	1	1	0	1
5	0	1	0	1	0	0	1	0	1	1	0	1	0	1	0	0	1	1	1	1	0	0
6	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1
7	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0
8	1	1	1	1	1	0	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	1
9	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	1	0
14	0	0	1	1	0	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0	1
16	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0

TABLE 3.6: Matrice qualitative des impacts

En guise d'exemple, si on applique la formule pour les deux premières Lobs (Auto PP RC et Auto EN RC) :

$$\text{corr}(Lob_1, Lob_2) = \frac{\text{Card}(Lob_1 \cap Lob_2)}{\text{Card}(Lob_1) + \text{Card}(Lob_2) - \text{Card}(Lob_1 \cap Lob_2)} = \frac{5}{5 + 6 - 5} = 83\%$$

$$\text{Corr}(Lob_1, Lob_2) = \frac{\text{Cov}(Lob_1, Lob_2)}{\sigma_{Lob_1} \sigma_{Lob_2}} = 87\%$$

On remarque que suivant les deux formules, l'écart n'est pas important. Dans une logique de prudence et de stabilité des résultats, la méthode suggère d'arrondir au supérieur (en absolu) les paramètres de corrélations à 25% près. C'est à dire une corrélation de 87% sera arrondie à 100% et celle de 13% à 25%. Le choix d'un tel seuil se fait en tenant compte d'un certain nombre de tests statistiques et des analyses faites sur l'impact du STEC.

Ainsi, la matrice finale obtenue est présentée dans le tableau 3.7 (pour environ une vingtaine de Lobs). On pourra appliquer dessus la formule d'agrégation pour déterminer le STEC reserve diversifié.

Corrélations avec arrondi 25%	Auto_PP_RC	Auto_EN_RC	Auto_PP_Dom	Auto_EN_DOM	MRH	MRP_PP	MRI_PP	Agri_Grele	Divers_Dab_PP	RI	MRP_EN	MRI_EN	DAB_Const	Divers_Dab_EN	RC_PP	RC_CONST	RC_ENT	RC_MED	MAT_PP	MAT_EN	CATNAT	DO
Auto_PP_RC	1,00	1,00	0,25	0,25	0,50	0,25	0,50	0,50	-0,25	0,50	0,25	0,50	0,25	-0,25	0,75	0,50	0,50	0,50	-0,50	-0,50	0,50	0,50
Auto_EN_RC	1,00	1,00	0,25	0,25	0,50	0,25	0,50	0,50	-0,25	0,50	0,25	0,50	0,25	-0,25	0,75	0,50	0,50	0,50	-0,50	-0,50	0,50	0,50
Auto_PP_Dom	0,25	0,25	1,00	1,00	1,00	0,25	1,00	0,75	0,75	1,00	0,25	1,00	0,75	0,75	-0,25	0,50	0,25	0,25	0,50	0,50	-0,25	0,75
Auto_EN_DOM	0,25	0,25	1,00	1,00	1,00	0,25	1,00	0,75	0,75	1,00	0,25	1,00	0,75	0,75	-0,25	0,50	0,25	0,25	0,50	0,50	-0,25	0,75
MRH	0,50	0,50	1,00	1,00	1,00	0,50	1,00	0,75	0,75	1,00	0,50	1,00	1,00	0,75	0,25	0,75	-0,25	0,25	0,50	0,50	0,75	0,75
MRP_PP	0,25	0,25	0,25	0,25	0,50	1,00	0,50	-0,25	0,50	0,50	1,00	0,50	0,50	0,50	0,50	1,00	0,25	0,50	0,75	0,75	0,50	0,75
MRI_PP	0,50	0,50	1,00	1,00	1,00	0,50	1,00	0,75	0,75	1,00	0,50	1,00	1,00	0,75	0,25	0,75	-0,25	0,25	0,50	0,50	0,75	0,75
Agri_Grele	0,50	0,50	0,75	0,75	0,75	-0,25	0,75	1,00	-0,25	0,75	-0,25	0,75	0,50	-0,25	-0,25	-0,25	-0,25	-0,25	-0,25	-0,25	-0,25	0,50
Divers_Dab_PP	-0,25	-0,25	0,75	0,75	0,75	0,50	0,75	-0,25	1,00	0,75	0,50	0,75	0,50	1,00	-0,25	0,75	0,50	0,50	0,75	0,75	-0,25	0,50
RI	0,50	0,50	1,00	1,00	1,00	0,50	1,00	0,75	0,75	1,00	0,50	1,00	1,00	0,75	0,25	0,75	-0,25	0,25	0,50	0,50	0,75	0,75
MRP_EN	0,25	0,25	0,25	0,25	0,50	1,00	0,50	-0,25	0,50	0,50	1,00	0,50	0,50	0,50	1,00	0,25	0,50	0,75	0,75	0,50	0,75	0,75
MRI_EN	0,50	0,50	1,00	1,00	1,00	0,50	1,00	0,75	0,75	1,00	0,50	1,00	1,00	0,75	0,25	0,75	-0,25	0,25	0,50	0,50	0,75	0,75
DAB_Const	0,25	0,25	0,75	0,75	1,00	0,50	1,00	0,50	0,50	1,00	0,50	1,00	1,00	0,50	0,00	0,50	-0,25	0,25	0,25	0,25	0,50	1,00
Divers_Dab_EN	-0,25	-0,25	0,75	0,75	0,75	0,50	0,75	-0,25	1,00	0,75	0,50	0,75	0,50	1,00	-0,25	0,75	0,50	0,50	0,75	0,75	-0,25	0,50
RC_PP	0,75	0,75	-0,25	-0,25	0,25	0,50	0,25	-0,25	-0,25	0,25	0,50	0,25	0,00	-0,25	1,00	0,50	0,50	0,75	-0,25	-0,25	0,50	0,25
RC_CONST	0,50	0,50	0,50	0,50	0,75	1,00	0,75	-0,25	0,75	0,75	1,00	0,75	0,50	0,75	0,50	1,00	0,50	0,75	0,50	0,50	0,75	0,75
RC_ENT	0,50	0,50	0,25	0,25	-0,25	0,25	-0,25	-0,25	0,50	-0,25	-0,25	-0,25	-0,25	0,50	0,50	0,50	1,00	1,00	0,25	0,25	-0,25	0,25
RC_MED	0,50	0,50	0,25	0,25	0,25	0,50	0,25	-0,25	0,50	0,25	0,50	0,25	0,25	0,50	0,75	0,75	1,00	1,00	0,25	0,25	0,50	0,25
MAT_PP	-0,50	-0,50	0,50	0,50	0,50	0,75	0,50	-0,25	0,75	0,50	0,75	0,50	0,25	0,75	-0,25	0,50	0,25	0,25	1,00	1,00	-0,25	0,25
MAT_EN	-0,50	-0,50	0,50	0,50	0,50	0,75	0,50	-0,25	0,75	0,50	0,75	0,50	0,25	0,75	-0,25	0,50	0,25	0,25	1,00	1,00	-0,25	0,25
CATNAT	0,50	0,50	-0,25	-0,25	0,75	0,50	0,75	-0,25	-0,25	0,75	0,50	0,75	0,50	-0,25	0,50	0,75	-0,25	0,50	-0,25	-0,25	1,00	0,50
DO	0,50	0,50	0,75	0,75	0,75	0,75	0,75	0,50	0,50	0,75	0,75	0,75	1,00	0,50	0,25	0,75	0,25	0,25	0,25	0,25	0,50	1,00

TABLE 3.7: Matrice de corrélation suivant l'approche Risk Driver

Il est important de rappeler que cette méthode d'agrégation suppose que tous les risques ont le même impact financier. De plus, on admet que chaque facteur de risque impacte les queues de distribution des variables liées aux Lobs (distribution de pertes, provisions, erreur d'estimation des charges). Dans certaines entités AXA telles que AXA XL, des facteurs de pondération sont introduits pour prendre en compte l'ampleur des effets. Mais d'une manière générale, il s'agit d'une hypothèse forte qui peut créer un biais dans la l'agrégation du STEC.

Pour éviter ce biais, des réflexions ont été menées dans le but de proposer une autre méthode qui sera basée sur l'historique des données de sinistres et qui utilise éventuellement des techniques statistiques plus avancées. Cette nouvelle méthode devra aussi satisfaire les critères qui ont été définis précédemment et faire l'objet d'un changement de modèle.

Chapitre 4

Nouvelles méthodes de modélisation des dépendances inter Lob et d'agrégation du STEC réserve

4.1 Méthodes empiriques basées sur les taux de boni mali

4.1.1 Construction de la base de données des taux boni-mali : traitements et calculs

On a montré qu'il existe différentes approches permettant d'estimer les corrélations inter Lob. Malgré le fait que les justifications semblent être pertinentes, la question qu'on se pose c'est de savoir si ces méthodes sont en adéquation avec le risque de réserve. En analysant de près le concept de risque de réserve, la définition rappelée dans la section 1.2.3 fait référence aux engagements d'assurance et les possibles déviations de ceux-ci en raison d'un certain nombre de fluctuations. Le principal défi sera de trouver une méthode statistique qui permet d'estimer des corrélations pouvant s'adapter à un tel concept. Pour ce faire, il convient de s'inspirer de la modélisation du risque de réserve (déjà établie dans le modèle interne). On sera également amené à déterminer la série de données à partir de la quelle les dépendances seront estimées et éventuellement la structure de corrélation appropriée.

Triangles et traitements de données

Le risque de réserve comme présenté précédemment, est modélisé à partir des données sur les triangles de charges. Sur les 36 Lobs du portefeuille, on compte 27 pour lesquelles nous disposons de triangles. Cette situation rend impossible l'application "directe" des techniques statistiques basées sur des distributions et série de données. Pour contourner cette limite,

une approche mixte de choix méthodologique pourra être envisagée.

En considérant les 27 triangles, on dispose de deux variables :

- Les montants des règlements des sinistres survenus et déclarés (triangles paid).
- Les montants de charge de sinistres survenus comprenant les règlements et les provisions dossier/dossier (triangles incurred)

Nous allons particulièrement travailler sur les triangles incurred pour avoir une meilleure vision du risque de réserve. A l'instar, de la démarche adoptée dans l'agrégation entre années d'accident, les mêmes étapes seront suivies pour le développement des triangles de charges. Ce processus est implémenté dans un outil interne appelé Harp ou sur des maquettes de réplique Excel / VBA. Nous nous sommes servis de ces outils pour déterminer les charges ultimes et en déduire l'estimation des réserves et des CDR pour chacun des 27 triangles.

L'intérêt d'extraire les triangles des réserves et des CDR est de pouvoir déterminer une série des boni mali par Lob. Nous allons par la suite estimer les dépendances inter Lob sur la base de cette série de données. Il conviendra en terme d'historique de considérer une série de boni mali suivant les années calendaires. Dès lors, une transformation de la structure des triangles s'impose. On passera de la vision année d'accident (AY) - année de développement (DY) à la vision année d'accident - année calendaire (CY) en considérant les diagonales. Quelques exemples sont tirés de la Lob Auto PP dommage pour illustrer ce traitement. Seules les 5 dernières années ont été considérées pour cet exemple (par soucis d'affichage).

CDR (i,j)	DY					
AY	1	2	3	4	5	6
2015	-2 375,5	-1 084,3	-8 173,2	-499,8	-3 365,6	840,9
2016	-17 978,8	-14 914,3	-5 116,5	-1 870,0	1 537,2	
2017	-31 224,9	-11 523,6	-7 773,5	4 550,7		
2018	-36 577,7	-9 048,4	3 894,0			
2019	-31 218,0	10 626,2				
2020	-13 299,7					
2021						

↓

CDR (i,n)	CY					
AY	2015	2016	2017	2018	2019	2020
2015	-2 375,5	-1 084,3	-8 173,2	-499,8	-3 365,6	840,9
2016		-17 978,8	-14 914,3	-5 116,5	-1 870,0	1 537,2
2017			-31 224,9	-11 523,6	-7 773,5	4 550,7
2018				-36 577,7	-9 048,4	3 894,0
2019					-31 218,0	10 626,2
2020						-13 299,7
2021						

Réserves (i,j)	DY					
AY	1	2	3	4	5	6
2015	141 854,8	40 352,2	26 672,6	13 762,4	10 471,1	6 101,9
2016	168 846,7	51 416,9	25 423,6	14 973,7	10 209,5	
2017	177 663,0	46 181,9	26 857,1	14 580,0		
2018	168 693,1	37 516,2	19 940,7			
2019	157 960,8	26 962,5				
2020	110 800,3					
2021						

↓

Réserves (i,n)	CY					
AY	2015	2016	2017	2018	2019	2020
2015	141 854,8	40 352,2	26 672,6	13 762,4	10 471,1	6 101,9
2016		168 846,7	51 416,9	25 423,6	14 973,7	10 209,5
2017			177 663,0	46 181,9	26 857,1	14 580,0
2018				168 693,1	37 516,2	19 940,7
2019					157 960,8	26 962,5
2020						110 800,3
2021						

A partir de ces triangles, on effectue la somme des CDR pour chaque CY suivant les AY, idem pour les montants de réserves. On en déduit les taux de boni mali en pourcentage des réserves par CY.

$$BM_n = \frac{\sum_{i=1}^I CDR_{i,n}}{\sum_{i=1}^I R_{i,n}}, \quad n = 2021, 2020, \dots$$

Cette série de données a été calculée sur toutes les 27 Lobs. Pour la plupart, elle s'étend de 2020 à 2000 sauf pour les Lobs DO et RCD (2020-2001), la Lob PJ (2020-2002) et les Lobs MAT entreprise et particuliers qui ont une plus faible profondeur (2020-2011).

D'autres traitements sur les données et calculs ont été faits sur les Lob FS. Le principe reste le même que précédemment, la seule nouveauté, c'est de procéder à un certain nombre regroupements en se basant sur la table de correspondance (voir tableau 3.2 section 3.2.1). Sur les 27 Lobs modèle interne, on identifie 7 Lob FS selon cette table de correspondance.

On notera respectivement par CDR_n^l et R_n^l le CDR et le montant de réserves d'une Lob l pour l'année calendaire n . Le taux de boni mali en pourcentage des réserves d'une Lob

FS d'indice L contenant $\#L$ Lobs modèle interne correspondant à une année calendaire n est donc défini par :

$$BM_{L,n} = \frac{\sum_{l=1}^{\#L} CDR_n^l}{\sum_{l=1}^{\#L} R_n^l}$$

De même que pour les Lobs modèle interne, on obtient une série de taux de boni mali par année calendaire pour chaque Lob FS avec une historique allant de 2000 à 2020 pour la majorité. On note quelques exceptions avec les Lob FS legal expense insurance (2020-2001) et MAT insurance (2020-2011).

Le fait de déterminer une historique de données sur les taux de boni mali à maille Lob FS nous servira par la suite dans les calculs de corrélation et les choix méthodologiques à mettre en oeuvre pour prendre en compte les Lob modèle interne pour lesquelles on ne dispose pas d'historique de données.

4.1.2 Corrélations basées sur les séries de taux de boni mali

De manière similaire à la méthode appliquée pour la dépendance entre années d'accident, on se place dans un cadre gaussien pour estimer des corrélations empiriques linéaires. On néglige également les formes de dépendances non linéaires et la dépendance de queue. Dans un premier temps, les corrélations classiques de Pearson et de Spearman seront explorées.

Le calcul se fait directement par paire sur les séries des taux de boni mali des 27 Lobs ainsi que sur celles des 7 Lob FS correspondantes. Cependant, il existe un point d'attention en raison des écarts de profondeur. Dans le cas où les séries ont la même profondeur, le calcul est direct. Sinon, pour le coefficient de Pearson, on considère le nombre d'observations minimum entre les 2 séries. Par exemple, la corrélation entre Auto PP dommage et MAT entreprise se calculera sur 10 observations (2020-2011) au lieu de 21 (2020-2000) puisque la profondeur de la série de données de la Lob MAT entreprise est limitée à 2011.

En ce qui concerne le rho de Spearman, les corrélations se basent sur les rangs comme nous l'avons évoqué dans les sections précédentes. Cet écart de profondeur se traite en recalculant les rangs lorsqu'on réduit le nombre d'observations. Nous présentons dans le tableau 17 un exemple avec quelques Lobs n'ayant pas toutes les mêmes profondeurs pour illustrer

ces traitements. Le tableau du milieu contient les rangs des données initiales et sert à calculer les corrélations lorsqu'il n'y a pas d'écart de profondeur, sinon on se réfère au tableau de droite dans lequel les rangs ont été recalculés sur les 10 années.

Année	AUTO_PP_RC	AUTO_EN_RC	MAT_PP	MAT_EN		Année	AUTO_PP_RC	AUTO_EN_RC	MAT_PP	MAT_EN		Année	AUTO_PP_RC	AUTO_EN_RC	MAT_PP	MAT_EN
2000	-0,02	-0,03				2000	5	6				2000				
2001	-0,16	-0,14				2001	1	2				2001				
2002	-0,03	0,01				2002	4	12				2002				
2003	0,02	0,10				2003	15	19				2003				
2004	0,01	-0,05				2004	11	4				2004				
2005	-0,01	0,03				2005	8	14				2005				
2006	0,06	0,07				2006	19	18				2006				
2007	0,00	-0,01				2007	10	7				2007				
2008	0,17	0,06				2008	20	15				2008				
2009	-0,01	0,00				2009	6	9				2009				
2010	0,03	0,06				2010	16	16				2010				
2011	0,02	0,17	0,04	-0,24		2011	14	20	8	1		2011	7	9	8	1
2012	0,01	0,02	0,14	0,13		2012	12	13	9	8		2012	5	7	9	8
2013	0,06	0,07	-0,05	0,26		2013	18	17	5	10		2013	9	8	5	10
2014	-0,01	0,00	0,00	0,01		2014	7	8	7	3		2014	3	4	7	3
2015	0,05	0,01	-0,19	0,11		2015	17	11	2	6		2015	8	6	2	6
2016	-0,01	0,00	-0,04	-0,08		2016	9	10	6	2		2016	4	5	6	2
2017	0,01	-0,05	-0,18	0,13		2017	13	5	3	7		2017	6	3	3	7
2018	-0,07	-0,15	-0,21	0,15		2018	2	1	1	9		2018	1	1	1	9
2019	-0,05	-0,06	0,65	0,04		2019	3	3	10	4		2019	2	2	10	4
2020	0,18	0,27	-0,06	0,10		2020	21	21	4	5		2020	10	10	4	5

TABLE 4.1: Traitements de données sur les corrélations de rang

Les mêmes calculs et traitements de données ont été effectués pour les Lob FS afin de déterminer les coefficients de corrélation par paires. Ceci est possible étant donné que les séries de taux de boni mali ont été construites de manière similaire par rapport au niveau Lob AXA.

4.1.3 Construction de la matrice de corrélation inter Lob

L'une des méthodes les plus simples et facilement interprétables pour l'agrégation du STEC est celle qui se base sur une matrice de corrélation. Cette approche est beaucoup utilisée dans le modèle interne et à différents niveaux d'agrégation. Comme nous l'avons précisé dans les sections précédentes, cette technique suppose certaines hypothèses telles que des coefficients de corrélation fixes et une structure de dépendance linéaire basée sur la copule gaussienne. Idéalement, une démarche plus rigoureuse serait de déterminer la structure de corrélation la plus adéquate d'un point de vue statistique et d'estimer les dépendances basées sur cette structure. Cependant, dans le soucis de proposer une méthode applicable

sur tout le portefeuille, nous allons dans un premier temps nous placer dans le même cadre (copule gaussienne, corrélations fixes, etc.) et analyser l'impact sur le STEC.

De manière similaire à la section 3.2.2 dans laquelle nous avons présenté la méthode initiale, nous allons construire une nouvelle matrice selon un choix méthodologique qui fera intervenir aussi bien les corrélations inter Lob du modèle interne que celles entre Lob FS. La grande différence par rapport à la méthode initiale c'est que les coefficients sont issus d'un calcul basé sur une historique de données découlant de la modélisation du risque de réserve (taux de boni mali) pour les paires de Lobs dont on dispose de données. Ce choix méthodologique peut être décrit suivant trois cas de figure :

Cas 1 : Lobs correspondant à la même Lob FS et pour lesquelles une historique de données est disponible

Nous allons considérer tout simplement le coefficient de corrélation de la paire de Lobs concernée sur la base de leur série de boni mali. Dans le tableau 4.2, nous illustrons cette démarche à travers un exemple portant sur un certains nombre de Lobs. Pour chacune d'elles on précise également la Lob FS correspondante pour faciliter la compréhension du lecteur. Une catégorisation par couleur suivant les blocs va également dans ce sens. Rappelons que cette démarche ne s'applique que pour les Lobs à travers lesquelles on dispose d'une historique de taux de boni mali. Dans cet exemple, nous considérons les corrélations de rang de Spearman.

Corrélation. (spearman)	Lob QIS	Motor veh liab ins	Motor veh liab ins	Motor veh liab ins	Oth motor ins	Oth motor ins	Fire other dam to prop ins	Fire other dam to prop ins
Lob QIS	Lob AXA	AUTO_PP_RC	AUTO_EN_RC	MBI_DA	MOD_DA	AUTO_EN_D om	MRH	MRP_PP
Motor veh liab ins	AUTO_PP_RC	100%	77%	14%				
Motor veh liab ins	AUTO_EN_RC	77%	100%	19%				
Motor veh liab ins	MBI_DA	14%	19%	100%				
Oth motor ins	MOD_DA				100%	21%		
Oth motor ins	AUTO_EN_Dom				21%	100%		
Fire other dam to prop ins	MRH						100%	44%
Fire other dam to prop ins	MRP_PP						44%	100%

TABLE 4.2: Corrélations inter Lob au sein d'une même Lob FS avec historique de taux de boni mali

Cas 2 : Lobs correspondant à des Lobs FS différentes et pour lesquelles une historique de données est disponible

Dans une telle situation, les corrélations retenues sont celles obtenues entre Lob FS. Mais rappelons que sur les 12 Lob FS seules 7 disposent d'une historique de données (cellules bleues de la matrice 4.3). Nous allons donc compléter la matrice de corrélation en utilisant les paramètres de la formule standard. Il s'agit par là, l'une des principales motivations derrière la correspondance entre Lob modèle interne et Lob FS.

Grâce à cette matrice de corrélation on peut ainsi compléter celle des corrélations inter Lob en faisant l'hypothèse que la corrélation entre deux Lobs modèle interne qui ne correspondent pas à la même Lob FS est la même que celle entre les deux Lob FS concernées. Par exemple, les Lob Auto_PP_RC et MRH correspondent respectivement aux Lob FS Motor vehicle liability et Fire and other damage to property. Le coefficient de corrélation entre ces deux Lob FS est de -42% (4.3). On en déduit donc que

$$\text{corr}(\text{Auto_PP_RC} | \text{Motor vehicle}, \text{MRH} | \text{Fire other damage}) = -42\%$$

Ainsi, ces hypothèses et les corrélations inter Lob FS nous permettent de construire la matrice de corrélation inter Lob sur la base des historiques de taux de boni mali disponibles.

Correl Lob QJS	Fire other dam to prop ins	Gen liab ins	Mat ins	Misc fin loss ins	Motor veh liab ins	Oth motor ins	Leg exp ins	Ass	Non-prop casu reins	Cred and sur ins	Fire other dam to prop reins	Non-prop prop reins	Med exp ins
Fire other dam to prop ins	100%	7%	-37%	26%	-22%	-14%	-28%	50%	25%	25%	75%	50%	50%
Gen liab ins	7%	100%	-5%	27%	-31%	-11%	-13%	25%	50%	50%	25%	25%	50%
Mat ins	-37%	-5%	100%	-13%	21%	53%	68%	50%	25%	50%	25%	25%	50%
Misc fin loss ins	26%	27%	-13%	100%	10%	5%	-29%	50%	25%	25%	50%	25%	50%
Motor veh liab ins	-22%	-31%	21%	10%	100%	48%	-9%	25%	25%	25%	25%	25%	50%
Oth motor ins	-14%	-11%	53%	5%	48%	100%	36%	50%	25%	25%	25%	25%	50%
Leg exp ins	-28%	-13%	68%	-29%	-9%	36%	100%	25%	50%	50%	25%	25%	50%
Ass	50%	25%	50%	50%	25%	50%	25%	100%	25%	25%	50%	25%	50%
Non-prop casu reins	25%	50%	25%	25%	25%	25%	50%	25%	100%	50%	25%	25%	50%
Cred and sur ins	25%	50%	50%	25%	25%	25%	50%	25%	50%	100%	25%	25%	50%
Fire other dam to prop reins	75%	25%	25%	50%	25%	25%	25%	50%	25%	25%	100%	50%	50%
Non-prop prop reins	50%	25%	25%	25%	25%	25%	25%	50%	25%	25%	50%	100%	50%
Med exp ins	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	100%

TABLE 4.3: Corrélations entre Lob FS incluant les paramètres de la FS

Cas 3 : Lobs pour lesquelles on n'a pas d'historique de données

En absence de données, nous allons nous inspirer de la méthode initiale présentée dans la section 3.2.2. On supposera une corrélation de 75% entre Lob correspondant à la même Lob FS. Pour les autres Lob, ils seront retenues les paramètres de la formule standard.

i et j correspondent à la même Lob FS	Données	Corrélation
X	X	param FS
V	X	75%
V	V	$\text{corr}(\text{Lob}_i, \text{Lob}_j)$
X	V	$\text{corr}(\text{Lob FS } L, \text{Lob FS } K)$ avec $i \in L$ et $j \in K$

TABLE 4.4: Récapitulatif de la méthodologie de calcul des corrélations

En appliquant les méthodologies relatives aux 3 cas de figures, on arrive ainsi à construire une matrice de corrélation en nous basant sur un mixte de méthodes statistiques avec

les historiques de taux de boni mali et une approche plus ou moins qualitative en raison de l'absence de données. Dans la suite, il conviendra d'étudier l'adéquation de cette matrice à l'objectif d'agrégation du STEC réserve.

Par rapport à l'hypothèse de corrélations fixes, au critère de stabilité dans le temps et de cadre réglementaire et prudentiel, une étape importante est d'arrondir à un seuil supérieur les coefficients. On peut également fixer un seuil inférieur comme dans le cas de la dépendance entre années d'accidents. Ces paramètres de seuil d'arrondi et de limite inférieure seront déterminés en fonction de l'impact sur le STEC qui découle et la marge de prudence optimale.

La matrice prenant en compte toutes les étapes de calcul, les hypothèses et choix méthodologiques est représentée partiellement (10 Lobs) à travers la matrice 4.5.

Coeff de correl	Lob Level	Motor veh liab ins	Motor veh liab ins	Fire other dam to prop ins	Fire other dam to prop ins	Gen liab ins	Gen liab ins	Ass	Ass	Cred and sur ins	Non-prop prop reins
Lob Level	Lob internal model	AUTO_PP_RC	AUTO_EN_RC	MRH	MRP_PP	RC_MED	RCD	ASSISTANCE	ASSISTANCE_DA	CAUTION	QPSAMUT
Motor veh liab ins	AUTO_PP_RC	100%	78%	-34%	-34%	0%	0%	25%	25%	25%	25%
Motor veh liab ins	AUTO_EN_RC	78%	100%	-34%	-34%	0%	0%	25%	25%	25%	25%
Fire other dam to prop ins	MRH	-34%	-34%	100%	45%	-32%	-32%	50%	50%	25%	50%
Fire other dam to prop ins	MRP_PP	-34%	-34%	45%	100%	-32%	-32%	50%	50%	25%	50%
Gen liab ins	RC_MED	0%	0%	-32%	-32%	100%	26%	25%	25%	50%	25%
Gen liab ins	RCD	0%	0%	-32%	-32%	26%	100%	25%	25%	50%	25%
Ass	ASSISTANCE	25%	25%	50%	50%	25%	25%	100%	75%	25%	25%
Ass	ASSISTANCE_DA	25%	25%	50%	50%	25%	25%	75%	100%	25%	25%
Cred and sur ins	CAUTION	25%	25%	25%	25%	50%	50%	25%	25%	100%	25%
Non-prop prop reins	QPSAMUT	25%	25%	50%	50%	25%	25%	50%	50%	25%	100%

TABLE 4.5: Matrice de Corrélations inter Lob finale

A partir de cette matrice, nous pouvons calculer le STEC modèle interne et déterminer l'impact par rapport à la méthode initiale.

4.1.4 Agrégation et corrélation équivalente

Dans cette sous-section on définit le concept de niveau global de corrélation équivalente . Cet indicateur permet de fournir un résumé du niveau de corrélation qui peut être globalement représentative des niveaux de corrélation d'une matrice, d'un vecteur ou d'une copule.

Le calcul peut être fait par interpolation. L'idée est de déterminer le montant de STEC correspondant à chaque niveau de corrélation équivalente (scénarios de dépendance) et déterminer graphiquement la corrélation équivalente qui correspond au montant de STEC obtenu. L'objectif recherché à travers cette démarche est d'évaluer à quel niveau de corrélation équivalente on se situe et d'établir des scénarios de dépendance pour tester et challenger les résultats de l'agrégation.

Pour approfondir cette forme de relation, nous allons essayer de détailler la formule d'agrégation et d'en déduire une forme fonctionnelle.

On notera par $x = (x_1, x_2, \dots, x_n) = (STEC_i)_{i=1, \dots, n}$ le vecteur des montants de STEC réserve pour toutes les Lobs et ρ le paramètre de corrélation qui correspond à la corrélation équivalente.

La matrice de corrélation peut donc s'écrire :

$$\Sigma = \begin{pmatrix} 1 & & & \rho \\ & 1 & & \\ & & \cdot & \\ \rho & & & \cdot \\ & & & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \cdot & \\ 0 & & & \cdot \\ & & & & 1 \end{pmatrix} + \rho \begin{pmatrix} 0 & & & 1 \\ & 0 & & \\ & & \cdot & \\ 1 & & & \cdot \\ & & & & 0 \end{pmatrix}$$

$$= I_n + \rho \begin{pmatrix} 0 & & 1 & & \\ & 0 & & & \\ & & \cdot & & \\ 1 & & & \cdot & \\ & & & & 0 \end{pmatrix}$$

Rappelons que la formule d'agrégation est donnée par

$$STEC_{res} = \sqrt{x^T \times \Sigma \times x}$$

$$STEC_{res}^2 = \left[(x_1, x_2, \dots, x_n) + \rho (x_1, x_2, \dots, x_n) \begin{pmatrix} 0 & & 1 & & \\ & 0 & & & \\ & & \cdot & & \\ 1 & & & \cdot & \\ & & & & 0 \end{pmatrix} \right] \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$$

$$STEC_{res}^2 = \left[(x_1, x_2, \dots, x_n) + \rho \left(\sum_{i=1, i \neq 1}^n x_i, \dots, \sum_{i=1, i \neq n}^n x_i \right) \right] \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$$

$$STEC_{res}^2 = \sum_{i=1}^n x_i^2 + \rho \sum_{i=1}^n x_i S_{-i} \quad \text{avec } S_{-i} = \sum_{j=1, j \neq i}^n x_j$$

On en déduit qu'il existe une relation quadratique entre ρ et le $STEC_{res}$. Au niveau de la section résultats, cette relation permettra de déterminer plus facilement le niveau de corrélation équivalente et de réaliser des comparaisons entre les méthodes ou d'effectuer des tests suivants des scénarios.

4.2 Calibration des dépendances via différentes familles de copules

Dans la section précédente, la méthodologie appliquée suppose une structure de dépendance linéaire. Il s'agit d'une hypothèse qui néglige les autres formes de dépendance pouvant être observées. On pourrait étendre la méthode en utilisant les copules. Comme évoqué dans le chapitre 2, la dépendance ne se limite pas uniquement à une simple corrélation. Les copules donnent la possibilité de travailler dans un cadre plus général. L'agrégation du STEC réserve se fera en utilisant la formule de la VaR appliquée à la distribution somme des facteurs de risque de chaque Lob. Plus précisément, les montants de STEC réserve par Lob qu'on souhaite agréger sont des quantités déterministes, et puisque les copules portent sur des variables aléatoires, il convient d'utiliser un facteur de risque qui peut être modélisé par une loi statistique.

4.2.1 Données et paramètres des distributions marginales

Pour appliquer la théorie des copules, il est indispensable de connaître les distributions marginales pour chaque Lob. A défaut, il faudrait suffisamment d'observations pour travailler avec des distributions empiriques ou utiliser la copule empirique. Il est également nécessaire de disposer d'échantillons d'observations pour l'estimation des copules.

Comme dans le cas des corrélations empiriques, nous allons considérer les séries des taux de boni mali. Par rapport au portefeuille, on sait qu'il y'a 9 Lobs qui ne présentent pas de données. De plus, deux Lobs parmi les autres ont une faible profondeur de données. Il s'agit des Lobs MAT particuliers et entreprises. Compte tenu de ces limites, la calibration des copules se fera uniquement sur les 25 Lobs. On retiendra une profondeur de 2001 correspondant à 20 observations (série des taux de boni mali de 2001 à 2020). Ces données serviront essentiellement à l'estimation des paramètres des copules. De prime abord, notons que le nombre d'observation paraît insuffisant pour réaliser des estimations.

A travers la modélisation du risque de réserve, on a vu que $1 + BM_{ij}$ suit une loi lognormal de moyenne 1 et de variance cv_{ij}^2 qui représente la volatilité des réserves par année de survenance et année de développement. Cette quantité est définie par le coefficient de variation du MSEF (3.1.2). Dans le modèle interne, les travaux sur l'agrégation entre années de survenance fournissent également un MSEF agrégé permettant de déterminer un facteur de

volatilité au niveau Lob. Il en découle, l'hypothèse suivante : $1 + BM_n$ suit une lognormal de moyenne 1 et de variance

$$cv_n = \frac{\sqrt{MSEP_{C\hat{D}R_n|D_{n-1}}(0)}}{R_n}$$

Cette hypothèse est valable pour chaque Lob (en fonction des données disponibles).

$$1 + BM \sim \mathcal{LN}(\mu, \sigma^2)$$

$$\mu = -\frac{1}{2}\ln(1 + cv^2) \quad ; \quad \sigma = \sqrt{\ln(1 + cv^2)}$$

Là également, il s'agit d'une hypothèse du modèle interne, n'ayant pas suffisamment de données pour tester l'adéquation.

Lob	cv	mu	sigma	Volume_reserves
AUTO_PP_RC	0,042	-0,001	0,042	2112,439
AUTO_EN_RC	0,052	-0,001	0,052	867,749
AUTO_PP_Dom	0,091	-0,004	0,091	261,676
MBI_DA	0,026	0,000	0,026	783,623
MOD_DA	0,062	-0,002	0,061	37,450
AUTO_EN_Dom	0,050	-0,001	0,050	179,270
MRH	0,048	-0,001	0,048	504,386
MRP_PP	0,141	-0,010	0,140	186,151
MRI_PP	0,063	-0,002	0,063	71,814
AGRI	0,099	-0,005	0,099	37,825
DIVERS_DAB_PP	0,432	-0,086	0,414	1,575
RI	0,060	-0,002	0,060	240,252
MRP_EN	0,576	-0,143	0,536	7,940
MRI_EN	0,041	-0,001	0,040	264,835
DAB_const	0,069	-0,002	0,068	44,768
DIVERS_DAB_EN	0,100	-0,005	0,099	66,356
CATNAT	0,109	-0,006	0,109	299,699
DO	0,017	0,000	0,017	1,976
HOD	0,064	-0,002	0,064	36,475
RC_PP	0,181	-0,016	0,179	90,884
RC_CONST	0,063	-0,002	0,063	486,546
RC_ENT	0,053	-0,001	0,053	920,987
RC_MED	0,099	-0,005	0,099	619,525
RCD	0,046	-0,001	0,046	2,363
PJ	0,132	-0,009	0,132	4,548

TABLE 4.6: Paramètres des distributions marginales

4.2.2 Calibration des copules elliptiques et archimédiennes

La problématique est telle que les copules à estimer seront destinées à modéliser une dépendance multivariée. Au total, le modèle compte 25 risques marginaux, ce qui correspond à 300 paires.

Pour les copules elliptiques, les résultats porteront sur la copule de Student ou *tCopula*. En effet, les corrélations empiriques présentées dans la section 4.1 se basent sur le cas gaussien. Ainsi, les paramètres à estimer sont les coefficients de corrélation par paires et le degré de liberté.

L'estimation se fait par la méthode du maximum de vraisemblance (ou pseudo-vraisemblance). La méthode des moments avec inversion du tau de kendall a aussi été testée et les résultats sont quasi similaires. L'estimation fournit un degré de liberté de 4 et 300 coefficients de corrélations dont les valeurs sont présentées dans le tableau 4.7. La vraisemblance maximale atteinte est de -2212 obtenue après 1213 itérations.

Ces estimations représentent les dépendances par paires de Lob selon la copule de Student. On peut également en déduire la distribution jointe et déterminer celle de la somme (5.1). En dimension 2, nous avons récupéré les paramètres obtenus pour réaliser des représentation graphiques. 3 couples ou paires de Lobs ont été choisis arbitrairement pour illustrer cette démarche.

	Lob FS	Motor veh lia	Gen liab	Legal exp
Lob FS	Lob AXA	AUTO_PP_RC	RC_CONST	PJ
Fire and other dam	MRP_PP	0,49		
Oth motor ins	AUTO_EN_DOM			
Misc fin loss	DIVERS_DAB_EN			

Les graphiques sont présentés au niveau de la figure 4.1. Entre les Lobs RC Auto pour particuliers et Multirisques professionnelles le coefficient est de 49%. Le niveau de dépendance est presque égal avec celui obtenu entre le dommage auto pour entreprise et la RC en construction. Cependant la copule de Student révèle une dépendance négative $(-0,22)$ entre les Lobs Protection juridique et divers dommages aux biens pour entreprise (voir la matrice 4.7 pour plus de détails). Les coefficients de dépendance de queue ont également été estimés par paire et les résultats sont présentés en annexe (A.6).

		1	2	3	4	5	8	9	10	11	12	13	14
1	AUTO_PP_RC	1,00											
2	AUTO_EN_RC	0,71	1,00										
3	AUTO_PP_Dom	0,66	0,81	1,00									
4	MBL_DA	0,00	0,06	-0,07	1,00								
5	MOD_DA	-0,14	-0,07	-0,04	0,43	1,00							
8	AUTO_EN_Dom	0,11	0,43	0,54	0,01	0,16	1,00						
9	MRH	0,11	-0,02	-0,03	-0,08	-0,51	-0,28	1,00					
10	MRP_PP	0,50	0,03	0,11	-0,01	-0,33	-0,46	0,27	1,00				
11	MRI_PP	-0,01	-0,26	-0,30	0,16	-0,15	-0,51	0,77	0,35	1,00			
12	AGRI	0,03	-0,23	-0,27	0,22	-0,29	-0,48	0,77	0,49	0,84	1,00		
13	DIVERS_DAB_PP	-0,21	-0,16	-0,10	0,29	0,33	0,37	-0,25	-0,13	-0,10	-0,06	1,00	
14	RI	0,17	-0,11	-0,11	0,27	-0,09	-0,12	0,37	0,45	0,39	0,65	0,02	1,00
15	MRP_EN	0,40	0,14	0,20	0,30	-0,19	-0,25	0,04	0,49	0,11	0,01	-0,02	-0,10
16	MRI_EN	0,19	-0,02	-0,06	0,15	-0,29	-0,40	0,74	0,36	0,76	0,66	-0,08	0,34
17	DAB_const	0,02	-0,07	-0,08	0,12	0,31	-0,29	0,25	0,08	0,38	0,42	-0,37	0,60
18	DIVERS_DAB_EN	-0,17	-0,35	-0,36	-0,06	0,01	-0,16	0,28	0,04	0,26	0,10	-0,23	0,19
19	CATNAT	0,01	-0,12	0,16	-0,25	0,23	-0,21	-0,16	0,20	0,07	-0,19	0,03	-0,33
20	DO	-0,20	-0,33	-0,49	-0,05	0,20	-0,22	0,24	0,06	0,32	0,31	-0,09	0,32
21	HOD	0,00	-0,04	0,02	0,27	-0,35	-0,29	0,12	0,30	0,10	0,36	-0,29	0,26
22	RC_PP	0,39	0,39	0,50	-0,41	-0,18	0,12	-0,27	0,22	-0,34	-0,37	-0,15	-0,29
23	RC_CONST	-0,18	-0,01	0,03	-0,08	0,09	0,50	-0,05	-0,63	-0,13	-0,13	0,30	0,03
24	RC_ENT	-0,03	-0,22	-0,24	0,24	-0,11	-0,33	0,29	0,23	0,44	0,42	-0,25	0,63
25	RC_MED	-0,17	-0,16	-0,17	0,21	0,62	0,25	-0,29	-0,07	0,08	-0,01	0,41	0,07
26	RCD	0,00	-0,01	0,21	-0,32	-0,17	0,12	0,43	-0,06	0,26	0,21	-0,22	0,27
27	PJ	0,03	0,17	0,39	-0,27	0,15	0,64	-0,55	-0,19	-0,72	-0,68	0,37	-0,53

		15	16	17	18	19	20	21	22	23	24	25	26
1	AUTO_PP_RC												
2	AUTO_EN_RC												
3	AUTO_PP_Dom												
4	MBL_DA												
5	MOD_DA												
8	AUTO_EN_Dom												
9	MRH												
10	MRP_PP												
11	MRI_PP												
12	AGRI												
13	DIVERS_DAB_PP												
14	RI												
15	MRP_EN	1,00											
16	MRI_EN	0,25	1,00										
17	DAB_const	-0,34	0,15	1,00									
18	DIVERS_DAB_EN	0,10	0,33	0,14	1,00								
19	CATNAT	0,44	-0,02	-0,03	0,15	1,00							
20	DO	-0,48	0,00	0,38	0,34	-0,29	1,00						
21	HOD	0,35	0,20	0,12	0,00	0,00	-0,52	1,00					
22	RC_PP	0,09	-0,25	-0,35	-0,41	0,21	-0,12	-0,29	1,00				
23	RC_CONST	-0,39	-0,02	-0,02	0,04	-0,10	-0,29	0,02	-0,23	1,00			
24	RC_ENT	0,05	0,45	0,59	0,28	-0,18	0,01	0,40	-0,49	-0,03	1,00		
25	RC_MED	-0,28	-0,05	0,05	0,27	0,12	0,27	-0,22	-0,21	0,20	-0,08	1,00	
26	RCD	-0,44	0,36	0,35	0,01	-0,18	0,09	-0,12	0,04	0,20	0,43	-0,07	1,00
27	PJ	-0,03	-0,63	-0,58	-0,29	0,15	-0,27	-0,31	0,33	0,03	-0,60	0,17	-0,13

TABLE 4.7: Estimation des paramètres de la copule de Student

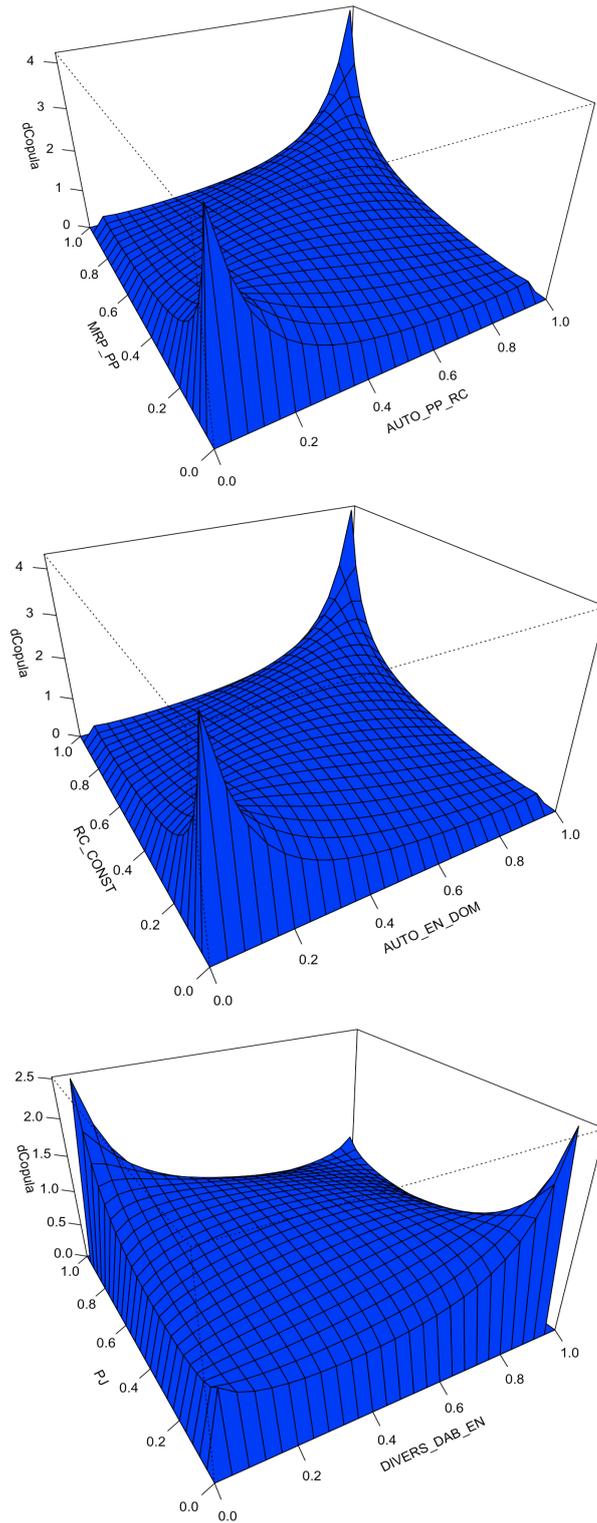


FIGURE 4.1: Densité de la copule de Student : exemple sur 3 paires de Lobs

L'estimation des copules archimédiennes requiert moins de paramètres puisque celle-ci se base sur les fonctions générateurs. Les résultats de l'estimation par la méthode du maximum de vraisemblance sont résumés dans le tableau 4.8. On en déduit également d'autres

paramètres tels que le tau de Kendall global et les coefficients de dépendance de queue. Les précisions des estimations sont également mises en évidence grâce aux indicateurs AIC et BIC ainsi que la vraisemblance maximale. Le tau de Kendall global est plus élevé pour la copule de Gumbel avec laquelle on observe une plus forte dépendance de queue supérieur.

Copule	theta	tau-K	LowerTailDep	UpperTailDep	BIC	AIC	LogLik
Clayton	0,228	0,102	0,048	0,000	-12,2	-13,2	7,6
Frank	0,222	0,025	0,000	0,000	1,7	0,7	0,6
Gumbel	1,131	0,116	0,000	0,154	-8,1	-9,1	5,5

TABLE 4.8: Estimation paramétrique des copules archimédiennes

4.2.3 Calibration de la copule hiérarchique RVine

La copule hiérarchique comme évoqué dans la sous-section 2.2.4 généralise la structure de dépendance. L'algorithme tel qu'il a été présenté estime les paramètres de manière séquentielle. Pour rappel, la structure RVine est représentée par des arbres contenant des estimations de copules par paires (ou paires conditionnelles). Sur les 25 Lobs, l'algorithme commence par déterminer la structure en calculant les tau de Kendall empiriques. Puis les 24 paires seront sélectionnées suivant ce critère pour former le premier arbre. De manière similaire, les autres arbres sont déterminés avec conditionnement sur les Lobs précédentes jusqu'au 24 ème arbre. La structure obtenue se présente sous forme de matrice (voir annexe A.7). Pour chaque paire (ou paire conditionnelle), une copule est sélectionnée et estimée. L'estimation a été réalisée avec le logiciel R , à l'aide de la librairie VineCopula contenant plusieurs fonctions prédéfinies.

Dans le tableau 4.9 un récapitulatif des résultats est présenté (partiellement). Il contient le numéro de l'arbre, les noeuds, les familles de copules et les paramètres.

Tree	edge	family	cop	par	par2	
1		18,19	3	C	1.10	0.00
1		13,17	3	C	0.65	0.00
1		8,13	4	G	1.62	0.00
1		8,21	2	t	-0.68	30.00
1		1,8	4	G	1.74	0.00
1		18,11	2	t	-0.50	30.00
1		2,18	2	t	-0.49	3.01
1		2,1	3	C	2.28	0.00
1		3,2	2	t	0.84	30.00
1		6,3	3	C	1.51	0.00
1		4,24	5	F	-3.73	0.00
1		25,6	4	G	1.77	0.00
1		14,25	2	t	-0.73	30.00
1		9,14	3	C	2.72	0.00
1		5,23	3	C	2.12	0.00
1		5,4	3	C	1.77	0.00
1		7,5	2	t	-0.56	11.05
1		10,7	3	C	4.41	0.00
1		10,9	1	N	0.87	0.00
1		12,1	3	C	1.59	0.00
1		12,15	1	N	0.69	0.00
1		20,16	2	t	-0.49	2.00
1		22,12	5	F	4.36	0.00
1		22,2	2	t	-0.57	2.47
.	
.	
.	
21	24,19 22,23,4,5,7,12,17,13,21,8,10,9,14,25,6,3,1,11,2,18	5	F	4.08	0.00	
21	15,17 24,22,23,4,5,7,12,10,9,11,14,25,18,6,3,2,21,1,8,13	4	G	1.26	0.00	
21	20,13 15,24,22,23,4,5,7,12,10,9,11,14,25,18,6,3,2,21,1,8	2	t	-0.21	30.00	
21	16,21 20,15,24,22,23,4,5,7,12,10,9,11,14,25,18,6,3,2,1,8	5	F	-0.62	0.00	
22	15,19 24,22,23,4,5,7,12,17,13,21,8,10,9,14,25,6,3,1,11,2,18	2	t	0.10	2.30	
22	20,17 15,24,22,23,4,5,7,12,10,9,11,14,25,18,6,3,2,21,1,8,13	4	G	1.16	0.00	
22	16,13 20,15,24,22,23,4,5,7,12,10,9,11,14,25,18,6,3,2,21,1,8	5	F	2.70	0.00	
23	20,19 15,24,22,23,4,5,7,12,17,13,21,8,10,9,14,25,6,3,1,11,2,18	3	C	0.03	0.00	
23	16,17 20,15,24,22,23,4,5,7,12,10,9,11,14,25,18,6,3,2,21,1,8,13	2	t	-0.23	30.00	
24	16,19 20,15,24,22,23,4,5,7,12,17,13,21,8,10,9,14,25,6,3,1,11,2,18	3	C	0.00	0.00	

TABLE 4.9: Estimation de la copule hiérarchique RVine

Les noeuds correspondent aux paires de Lobs avec les listes des Lobs dans le conditionnement. 5 familles de copules sont considérées. Elles sont représentées par un numéro et une lettre :

- 1: Gaussienne ou N
- 2: Student ou t

- 3: Clayton ou C
- 4: Gumbel ou G
- 5: Frank ou F

Le premier paramètre (par) représente la dépendance par paire de Lob pour les copules elliptiques et le theta d'une copule archimédienne. Le deuxième (par2) est spécifique à la copule de Student et représente le degré de liberté. Par exemple, au niveau de l'arbre 1, les Lobs 18 et 19 (divers dommages aux bien entreprises et CatNat) sont liées par une copule de Clayton de paramètre 1,1.

Chapitre 5

Résultats de l'agrégation du STEC réserve et analyse de la diversification

5.1 Résultats d'agrégation du STEC suivant les différentes méthodes

La diversification du portefeuille de AXA France P&C est un important levier pour améliorer le ratio de solvabilité. Dans la présente étude, les précédents chapitres ont permis de décrire les techniques d'agrégation. Il est ressorti que le STEC réserve du portefeuille peut être obtenu aussi bien à l'aide d'une matrice de corrélation empirique qu'en utilisant la théorie des copules. Dans ce chapitre, les résultats de chaque méthode seront présentés. Des comparaisons par rapport à la méthode initiale seront également effectuées pour évaluer le gain de diversification.

5.1.1 Agrégation avec une matrice de corrélation empirique

L'utilisation d'une matrice de corrélation empirique est l'une des techniques les plus simples permettant d'agréger le STEC. En solvabilité 2, cette approche est très répandue et s'utilise à différents niveaux. Elle met en évidence et évalue les dépendances par paire de Lobs. L'agrégation se fait avec la formule vue en 3.1 appliquée sur les données du tableau 5.1, en particulier le vecteur contenant les montants de STEC par Lob.

Lob AXA	Lob FS	STEC Res par Lob	Poids
AUTO_PP_RC	Motor veh liab ins	344109707	15,7%
AUTO_EN_RC	Motor veh liab ins	139832129	6,4%
AUTO_PP_Dom	Oth motor ins	105946723	4,8%
MBI_DA	Motor veh liab ins	48974428	2,2%
MOD_DA	Oth motor ins	6579177	0,3%
MAT_PP	Mat ins	4680042	0,2%
MAT_EN	Mat ins	9304428	0,4%
AUTO_EN_Dom	Oth motor ins	39938047	1,8%
MRH	Fire other dam to prop ins	71718596	3,3%
MRP_PP	Fire other dam to prop ins	15678354	0,7%
MRI_PP	Fire other dam to prop ins	12882750	0,6%
AGRI_GRELE	Fire other dam to prop ins	9669805	0,4%
DIVERS_DAB_PP	Fire other dam to prop ins	1023699	0,0%
RI	Fire other dam to prop ins	30190295	1,4%
MRP_EN	Fire other dam to prop ins	9327212	0,4%
MRI_EN	Fire other dam to prop ins	31145958	1,4%
DAB_CONST	Fire other dam to prop ins	7738291	0,4%
DIVERS_DAB_EN	Misc fin loss ins	23722325	1,1%
CATNAT	Fire other dam to prop ins	120391550	5,5%
DO	Fire other dam to prop ins	83980786	3,8%
HOD_DA	Fire other dam to prop ins	7098564	0,3%
RC_PP	Gen liab ins	35848787	1,6%
RC_CONST	Gen liab ins	86159809	3,9%
RC_ENT	Gen liab ins	129605093	5,9%
RC_MED	Gen liab ins	149136354	6,8%
RCD	Gen liab ins	281914738	12,9%
PJ	Leg exp ins	2490947	0,1%
ASSISTANCE	Ass	14217593	0,7%
ASSISTANCE_DA	Ass	2852025	0,1%
AUTRES_EN	Non-prop casu reins	85467481	3,9%
AUTRES_PP	Non-prop casu reins	634943	0,0%
CAUTION	Cred and sur ins	0	0,0%
POOL_PROPERTY	Fire other dam to prop reins	46844214	2,1%
PP_EN	Misc fin loss ins	128203294	5,9%
QPSAMUT	Non-prop prop reins	98770253	4,5%
SANTE	Med exp ins	26454	0,0%

TABLE 5.1: Montants du STEC réserve par Lob en 2021

La diversification est déterminée par la matrice de corrélation Σ (voir formule 3.1) qui dépend de l'approche utilisée. Le cas d'une parfaite dépendance correspond au STEC standalone et s'obtient par une somme directe des montants de STEC. Il est égal à 2,19 milliards et aucune diversification n'est observée. Les autres cas de figure correspondent aux matrices construites suivant les différentes méthodes présentées dans le chapitre 3. Le gain de

diversification est calculé en évaluant l'écart par rapport au STEC Standalone.

Le STEC initial représente le montant de STEC agrégé en utilisant la matrice de corrélation initiale qui s'approche des paramètres de la formule standard (3.2.2). Avec un montant de 1,53 milliards, on observe un gain de 29,9%. Il s'agit de la méthode officielle utilisée jusque là dans le modèle interne. Dans la suite, l'impact STEC ou gain en solvabilité sera calculé en évaluant l'écart par rapport à cette méthode initiale. En d'autres termes, il s'agit de la différence en pourcentage entre le montant de STEC initial et celui obtenu avec une autre méthode.

L'approche Risk driver fournit un gain de diversification de 27,7% avec un montant de STEC de 1,58 milliards. L'impact indique une perte de solvabilité par rapport à la méthode initiale. Autrement dit, le STEC obtenu est plus élevé de 3,1%.

Les résultats de la nouvelle méthode proposée dans le cadre de ce mémoire présente des montants de STEC plus faibles (environ 800 millions). Suivant le type de corrélation considéré on note une légère différence. Dans la suite, nous retiendrons la corrélation qui donne le STEC le plus élevé pour rester prudent. Selon les résultats du tableau, il s'agit de la corrélation de rang de spearman qui présente un gain de diversification de 62,3% et un impact STEC correspondant à un gain de solvabilité de 46,3%. En d'autres termes, la nouvelle méthode basée sur les boni mali permet de réduire le montant de STEC d'environ 760 millions par rapport à la méthode initiale.

Méthode	Standalone	STEC initial	Risk Driver	Correl pearson	Correl rang
STEC (milliards)	2,19	1,53	1,58	0,79	0,82
Impact	-	-	3,1%	-48,2%	-46,3%
Diversification	0,0%	29,9%	27,7%	63,7%	62,3%

TABLE 5.2: Résultats de l'agrégation sur le portefeuille global avec les méthodes de corrélation empiriques

Toutefois, il est important de rappeler que la détermination de la nouvelle matrice de corrélation empirique s'est faite avec des hypothèses fortes et simplificatrices. De plus, en raison de l'absence de données suffisantes, la précision des corrélations reste discutable. Compte tenu de cette situation, la matrice présente plusieurs corrélations faibles voir négatives ce qui réduit considérablement le montant de STEC agrégé. L'impact observé est ainsi très élevé et le niveau de capital qui en découle risque d'être insuffisant pour satisfaire les

exigences réglementaires.

Pour résoudre ce problème, il convient d'approfondir la technique d'agrégation et de réaliser des tests permettant de contrôler la cohérence et la robustesse des résultats.

5.1.2 Agrégation par les copules

La calibration des dépendances par l'intermédiaire des copules s'est faite sur une partie du portefeuille contenant 25 Lobs (voir 4.6), soit 82,1% du montant du STEC global (en standalone). Ce choix se justifie par les contraintes de données. Ainsi, les résultats qui seront présentés dans cette partie porteront sur ces 25 Lobs.

L'agrégation se fait en utilisant la formule du VaR 99,5%.

$$STEC_{res} = q_{0,995}[S] - \mathbb{E}[S]$$

Avec $S = \sum_{i=1}^n X_i$ obtenu par Monte Carlo.

$$X_i = \text{VolumeReserve}_i \times qlnorm(u[k, j], \text{meanlog} = \mu_i, \text{sdlog} = \sigma_i)$$

$$k = 1, \dots, N \text{ (nombre d'observations simulées)}$$

Il s'agit de la distribution de perte générée par la copule pour chaque Lob i . Cette quantité s'obtient en inversant les nombres $u[k, j]$ générés par la copule suivant la loi Lognormale (hypothèse faite en 4.2.1) et en multipliant par le volume des réserves.

Les résultats sont résumés dans le graphique 5.1. Tout d'abord, en considérant les 25 Lobs, le montant du STEC initial est de 1,26 milliards et toutes les autres méthodes présentent un niveau inférieur, donc potentiellement un gain de solvabilité. Néanmoins, les niveaux de STEC sont très faibles et peu prudents à l'exception de la copule de Gumbel. En effet, l'estimation de cette copule archimédienne ne prend pas en compte les dépendances négatives. De plus, les résultats de l'estimation de cette copule révèlent une dépendance de queue supérieure non négligeable.

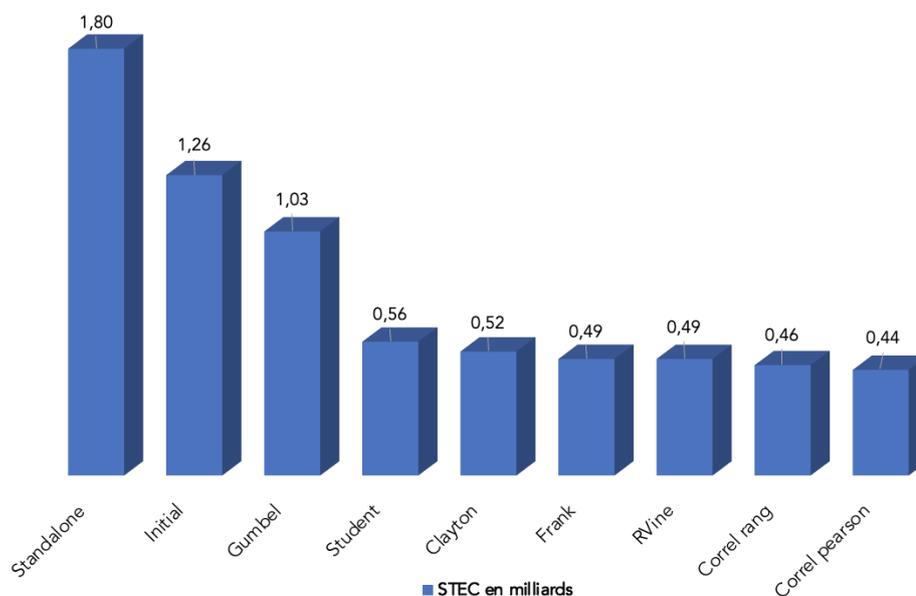


FIGURE 5.1: Montant du STEC réserve avec 25 Lobs suivant différentes méthodes

Methode	Impact en valeur abs	Diversification
STEC Standalone	-	-
STEC initial	-	29,6%
Gumbel	18,8%	42,8%
Student	55,5%	68,7%
Clayton	58,7%	71,0%
Frank	61,2%	72,7%
RVine	61,2%	72,7%
Correl rang	63,3%	74,2%
Correl pearson	64,8%	75,3%

TABLE 5.3: Impact STEC et Diversification sur les 25 Lobs

Il ressort de cette analyse que l'agrégation par chacune des copules génère un STEC plus élevé que les corrélations empiriques basées sur les boni mali. Ce résultat montre que cette méthode n'est pas assez prudente (montant de STEC pouvant être insuffisant).

En s'inspirant de la matrice de la formule standard, on peut arrondir les corrélations au supérieur et ajouter un seuil inférieur (corrélation minimale) pour limiter les anomalies sur les dépendances. Ces deux approches seront définies comme des paramètres de prudence et vont jouer un rôle important dans la suite.

5.2 Paramètres de prudence et corrélation équivalente

5.2.1 Optimisation des paramètres de prudence (arrondi et seuil inférieur)

Les résultats présentés jusqu'à maintenant montrent qu'on peut diversifier davantage et gagner en solvabilité par rapport à la méthode initiale. Cependant, les méthodes utilisées pour mesurer les dépendances souffrent d'une insuffisance de données et de précision. Pour palier à ces limites, on introduit les paramètres de prudence (arrondi au supérieur noté **Step_correl**⁴ et seuil minimal **limite_correl**⁵). Ces paramètres seront appliqués aux corrélations empiriques obtenues via les taux de boni mali.

Pour toute corrélation $Corr_{ij} \leq \text{limite_correl}$, on définit $Corr_{ij}^{Prud} = \text{limite_correl}$.

Et lorsque $Corr_{ij} \geq \text{limite_correl}$, on l'arrondit au multiple supérieur de Step_correl:

$$Corr_{ij}^{Prud} = \left(\left\lfloor \frac{Corr_{ij}}{\text{Step_correl}} \right\rfloor + 1 \right) \cdot \text{Step_correl}$$

$\lfloor \cdot \rfloor$ est le symbole partie entière.

En guise d'exemple, considérons les paramètres (Limite_correl , Step_correl) = (15,5).

$$Corr_{ij} = -10\% \longrightarrow Corr_{ij}^{Prud} = 15\%$$

$$Corr_{ij} = 33\% \longrightarrow Corr_{ij}^{Prud} = 35\%$$

On choisira les paramètres entre 0 et 30% par pas de 5%. A partir de ces valeurs, on peut construire une matrice indiquant pour chaque couple de paramètres le montant de STEC (sur les 25 Lobs) correspondant aux corrélations empiriques modifiées. Plus les paramètres sont élevés, plus le montant de STEC augmente. Par exemple, le couple (Limite_correl , Step_correl) = (15,5) génère un STEC de 955 millions. Si on considère le couple (30,30), on est à 1,156 milliards qui est le scénario le plus prudent mais génère moins de gain de solvabilité.

⁴un arrondi permet d'avoir des coefficients stables dans le temps

⁵en fixant un seuil, on fait l'hypothèse qu'il n'existe pas de dépendances négatives et d'indépendance parfaite

		Step_correl						
		0%	5%	10%	15%	20%	25%	30%
Limite_correl	0%	0,772	0,795	0,808	0,851	0,867	0,920	0,942
	5%	0,824	0,845	0,850	0,881	0,906	0,932	0,942
	10%	0,884	0,893	0,898	0,912	0,919	0,934	0,934
	15%	0,948	0,955	0,958	0,974	0,976	0,994	0,994
	20%	1,010	1,016	1,017	1,031	1,036	1,048	1,046
	25%	1,068	1,074	1,075	1,087	1,089	1,104	1,102
	30%	1,127	1,130	1,130	1,142	1,138	1,150	1,156

TABLE 5.4: Montant de STEC en milliards par les corrélations empiriques arrondies et avec seuil minimal

On peut également s'intéresser à la matrice des impacts dont les résultats vont dans le même sens (5.5).

		Step_Correl						
		0%	5%	10%	15%	20%	25%	30%
Limite_correl	0%	39%	37%	36%	33%	31%	27%	25%
	5%	35%	33%	33%	30%	28%	26%	25%
	10%	30%	29%	29%	28%	27%	26%	26%
	15%	25%	24%	24%	23%	23%	21%	21%
	20%	20%	20%	19%	18%	18%	17%	17%
	25%	15%	15%	15%	14%	14%	13%	13%
	30%	11%	11%	11%	10%	10%	9%	9%

TABLE 5.5: Impact solvabilité des corrélations empiriques arrondies et avec seuil minimal

L'objectif sera de trouver la combinaison de paramètres "optimale". Autrement dit, on cherche le couple $(\text{Limite_correl}^{opt}, \text{Step_correl}^{opt})$ permettant de maximiser le gain de solvabilité et qui garantit un niveau de fonds propres suffisant.

Dans le cadre de l'étude, on définit le concept de diversification prudentielle qui correspond à la mesure de dépendance qui génère un STEC agrégé supérieur à toutes les autres techniques d'agrégation et satisfaisant les tests. Pour cela, il faudrait comparer les différentes méthodes. Mais puisque celles-ci mesurent différemment les dépendances, il convient d'utiliser les corrélations équivalentes définies en 4.1.4.

5.2.2 Corrélations équivalentes

Pour rappel, la relation entre la corrélation équivalente et le montant de STEC obtenu par une quelconque méthode est donnée par :

$$STEC_{res}^2 = \sum_{i=1}^n x_i^2 + \rho \sum_{i=1}^n x_i S_{-i} \quad \text{avec } S_{-i} = \sum_{j=1, j \neq i}^n x_j \quad \text{et } x_i = STEC_i$$

$$\frac{\partial STEC_{res}}{\partial \rho} = \frac{1}{2} \left[\sum_{i=1}^n x_i S_{-i} \right] \left[\sum_{i=1}^n x_i^2 + \rho \sum_{i=1}^n x_i S_{-i} \right]^{-\frac{1}{2}} \geq 0$$

$$\frac{\partial^2 STEC_{res}}{\partial \rho^2} = -\frac{1}{4} \left[\sum_{i=1}^n x_i S_{-i} \right]^2 \left[\sum_{i=1}^n x_i^2 + \rho \sum_{i=1}^n x_i S_{-i} \right]^{-\frac{3}{2}} \leq 0$$

On montre que cette relation est croissante et concave. Cette trajectoire illustre la relation entre les scénarios de dépendance (représentés par la corrélation équivalente) et le montant du STEC agrégé en se basant sur les mêmes hypothèses que celles de la formule standard.

En l'appliquant sur les données des 25 Lobs on trouve:

$$\sum_{i=1}^n x_i^2 = 3,1 \times 10^{17}$$

$$\sum_{i=1}^n x_i S_{-i} = 29,1 \times 10^{17}$$

$$STEC_{res}^2 = (3,1 + 29,1\rho) \times 10^{17} \longrightarrow \rho = \frac{1}{29,1} \left(\frac{STEC_{res}^2}{10^{17}} - 3,1 \right)$$

Cette relation est définie si $\rho \in [-0,107; 1]$. Autrement dit, la matrice de corrélation correspondante est définie positive si ρ appartient à cet intervalle.

On dira par exemple qu'un scénario de dépendance équivalent à une corrélation de 50% entre Lobs fournit un montant de STEC de 1,33 milliards sur le portefeuille partiel. Le cas standalone ou absence de diversification correspond à un scénario de dépendance de 100%.

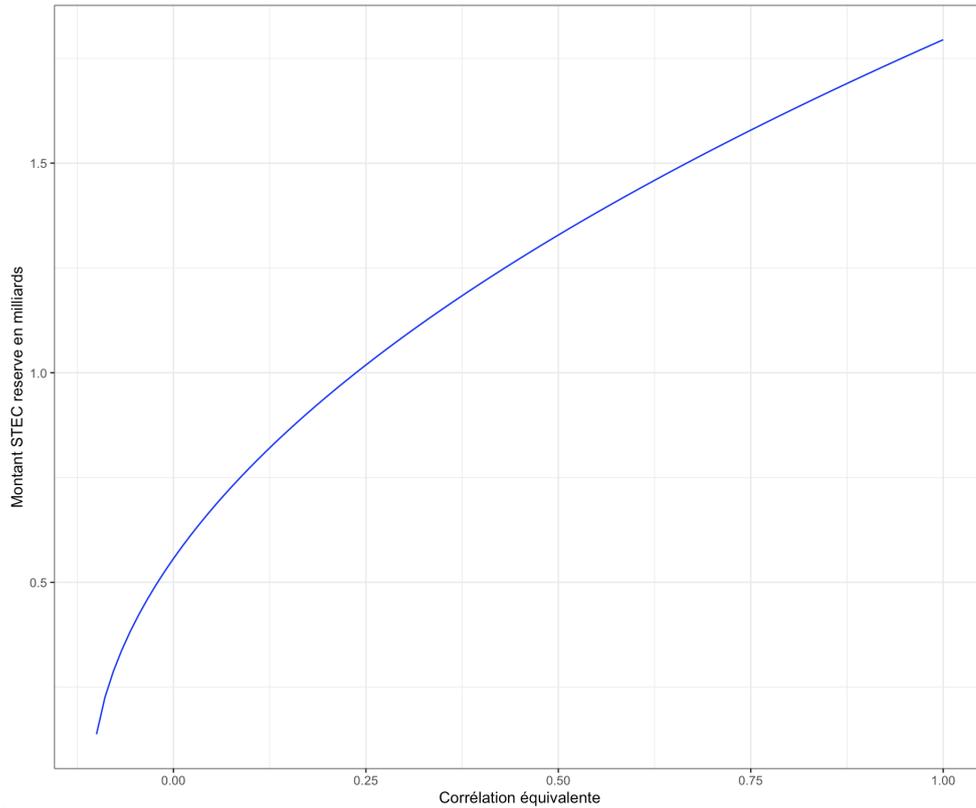


FIGURE 5.2: Montant de STEC en milliards en fonction du niveau de corrélation équivalente

La méthode initiale correspond à un scénario de dépendance de 44,2%. Par conséquent, toute méthode dont la corrélation équivalente est supérieure à ce niveau, ne génère pas de gain de solvabilité. Pour les corrélations empiriques basées sur les boni mali, on retiendra le max entre le coefficient de pearson et celui des rangs de spearman. Cette méthode correspond à un scénario de $-3,3\%$, alors que pour la copule de Gumbel, la corrélation équivalente est de 25,5%. En appliquant les paramètres d'arrondi et de seuil minimal, il est possible de passer de 9,8% (couple (0,0)) à 35% (couple (30,30)).

Methode	Correl Equiv
STEC Standalone	100,0%
STEC initial	44,2%
Max (Correl rang/pearson) param_prud (30,30)	35%
Gumbel	25,5%
Max (Correl rang/pearson) param_prud (0,0)	9,8%
Student	0,2%
Clayton	-1,3%
Frank	-2,4%
RVine	-2,4%
Max (Correl rang/pearson)	-3,3%

TABLE 5.6: Corrélations équivalentes en fonction des méthodes

Par ailleurs, une matrice des scénarios de dépendance en fonction des paramètres de prudence peut être construite. Celle-ci permettra de faciliter la recherche de la combinaison optimale.

		Step_Correl						
		0%	5%	10%	15%	20%	25%	30%
Limite_Correl	0%	10%	11%	12%	14%	15%	18%	20%
	5%	13%	14%	14%	16%	18%	19%	20%
	10%	16%	17%	17%	18%	18%	19%	19%
	15%	20%	21%	21%	22%	22%	23%	23%
	20%	24%	25%	25%	26%	26%	27%	27%
	25%	29%	29%	29%	30%	30%	31%	31%
	30%	33%	33%	33%	34%	34%	35%	35%

TABLE 5.7: Matrice des scénarios de dépendance de la méthode des corrélations empiriques arrondies et avec seuil minimale

5.3 Tests et Choix du niveau de dépendance

5.3.1 Backtests sur données historiques et données comptables

Données historiques

La nouvelle méthode de calibration des dépendances entre Lobs doit satisfaire le critère de stabilité dans le temps. Pour cela, il est indispensable de réaliser des backtests afin de prouver que la méthode est aussi conservatrice (voire plus conservatrice) que les années passées. Les tests seront réalisés sur les années 2019 et 2020 pour conserver la même taille du portefeuille (25 Lobs).

Tout d'abord, il faut recalculer les corrélations empiriques correspondant aux années 2020 et 2019 en reconstruisant les séries de boni mali. Pour cela, il suffit de partir des triangles de 2021 en retirant à chaque fois la dernière diagonale, puis calculer les charges ultimes et en déduire le triangles des CDR. A partir de ces triangles seront calculées les séries de taux de boni mali. On en déduit les corrélations empiriques linéaires de pearson et de rang de spearman.

En appliquant la formule d'agrégation sur les montants de STEC par Lobs pour les années 2019 et 2020, on trouve un STEC réserve agrégé pour ces années sur la base des corrélations empiriques estimées. Le test qui sera effectué devra permettre de vérifier si la matrice de corrélation empirique obtenue en 2021 pour la présente étude est-elle plus prudente que celle des années précédentes. Pour ce faire, il faudrait comparer les cellules correspondant à chaque paire de Lobs. Mais vu la taille élevée du portefeuille, une alternative serait d'utiliser les corrélations équivalentes.

Données comptables

Les données utilisées jusque là sont celles qui ont servi à la modélisation du risque de réserve en solvabilité 2. Au niveau de la comptabilité, on dispose également de données sur les charges ultimes correspondant aux Lobs du portefeuille en 2021. A partir de celles-ci on peut construire des séries de boni mali et estimer les corrélations empiriques. L'agrégation et les tests se feront de la même manière que précédemment.

5.3.2 Résultats des tests et choix optimal des paramètres

L'une des tâches les plus importantes de cette étude est de déterminer les paramètres d'arrondi et de seuil minimal permettant d'améliorer les corrélations empiriques. Ce choix devrait aussi bien prendre en compte les autres formes de dépendance (autres que linéaires) que satisfaire les tests.

Nous avons vu que la copule de Gumbel génère le montant de STEC le plus élevé parmi toutes les autres formes de dépendance. Par conséquent, le choix optimal devrait correspondre à un montant de STEC au moins égal à celui obtenu avec la copule de Gumbel. Ce niveau est équivalent à un scénario de dépendance de 25,5%.

$$\text{Step_correl}^{opt} \geq 15\% \quad \text{et} \quad \text{Limite_correl}^{opt} \geq 20\%$$

		Step Correl						
		0%	5%	10%	15%	20%	25%	30%
Limite Correl	0%							
	5%							
	10%							
	15%							
	20%				26%	26%	27%	27%
	25%	29%	29%	29%	30%	30%	31%	31%
	30%	33%	33%	33%	34%	34%	35%	35%

En considérant un tel critère, le choix optimal qui permettrait de générer un gain de solvabilité avec un scénario plus prudent que toutes les formes de dépendances qui ont été testées sur le portefeuille en fonction des données disponibles est la combinaison (20%,15%).

Pour vérifier ce choix, il convient d'analyser les résultats des tests sur les données historiques et comptables.

Range	Method	STEC en milliards	Impact	Equiv Correl	Résultat Test
2000-2021	Initial	1,263	0,0%	44,2%	
	⁵ Correl empirique param_prud (15,20)	1,031	-18,4%	25,8% ⁷	
	Correl empirique	0,464	-63,3%	-3,3%	
2000-2020	Initial	1,402	0,0%	42,8%	
	Correl empirique param_prud (15,20)	1,090	-22,2%	22,4%	PASS
	Correl empirique	0,541	-61,4%	-1,2%	PASS
2000-2019	Initial	1,060	0,0%	43,3%	
	Correl empirique param_prud (15,20)	0,811	-23,5%	20,7%	PASS
	Correl empirique	0,461	-56,5%	-0,7%	PASS

TABLE 5.8: Back test avec les corrélations estimées sur les données historiques

De prime abord, notons qu'en appliquant les paramètres de prudence (20%,15%) sur les corrélations empiriques, on trouve un STEC de 1,031 milliards correspondant à un scénario de dépendance de 25,8%. Les tests consisteront à comparer ce scénario à celui des autres cas de figure. Le tableau 5.8 révèle que tous les tests sont vérifiés. En effet, la corrélation équivalente est supérieure à celle des années passées. Ainsi, le test montre que les résultats sont plus prudents que les années passées malgré que l'impact en solvabilité par rapport à la méthode initiale soit moins important.

Le calcul du STEC 2021 agrégé par une matrice de corrélation empirique (max entre pearson et spearman) estimées par les données issues de la comptabilité donne un montant de 993 millions. Ce qui correspond à un scénario de dépendance de 23,2% moins prudent que le scénario de base, ce qui valide le test.

Method	STEC en milliards	Impact	Equiv Correl	Test
Initial	1,263	0,0%	44,2%	
Correl empirique	0,993	-21,4%	23,2%	PASS

TABLE 5.9: Test avec les corrélations estimées sur les données issues de la comptabilité

Un autre test consiste à vérifier si les corrélations obtenues avec les paramètres de prudence appartiennent aux intervalles de confiance de niveau 95% des corrélations empiriques.

⁵On considère ici le max entre les corrélations linéaires de pearson et de rang de spearman

⁷Scénario de base pour les tests

L'estimation par la méthode bootstrap a permis d'obtenir pour chaque paire de Lob, la borne inférieure et supérieure de la corrélation. Nous présentons un exemple avec 4 Lobs en nous basant sur les coefficients de rang de spearman (5.10).

Borne inférieure

Coef de corrélation	Lob QIS	Motor veh liab ins	Motor veh liab ins	Fire other dam to prop ins	Fire other dam to prop ins
Lob QIS	Lob AXA	AUTO_PP_RC	AUTO_EN_RC	MRH	MRP_PP
Motor veh liab ins	AUTO_PP_RC	100%	45%	-51%	-51%
Motor veh liab ins	AUTO_EN_RC	45%	100%	-51%	-51%
Fire other dam to prop ins	MRH	-51%	-51%	100%	-33%
Fire other dam to prop ins	MRP_PP	-51%	-51%	-33%	100%

Corrélation empirique avec paramètres de prudence

Coef de corrélation	Lob QIS	Motor veh liab ins	Motor veh liab ins	Fire other dam to prop ins	Fire other dam to prop ins
Lob QIS	Lob AXA	AUTO_PP_RC	AUTO_EN_RC	MRH	MRP_PP
Motor veh liab ins	AUTO_PP_RC	100%	90%	20%	20%
Motor veh liab ins	AUTO_EN_RC	90%	100%	20%	20%
Fire other dam to prop ins	MRH	20%	20%	100%	45%
Fire other dam to prop ins	MRP_PP	20%	20%	45%	100%

Borne supérieure

Coef de corrélation	Lob QIS	Motor veh liab ins	Motor veh liab ins	Fire other dam to prop ins	Fire other dam to prop ins
Lob QIS	Lob AXA	AUTO_PP_RC	AUTO_EN_RC	MRH	MRP_PP
Motor veh liab ins	AUTO_PP_RC	100%	90%	37%	37%
Motor veh liab ins	AUTO_EN_RC	90%	100%	37%	37%
Fire other dam to prop ins	MRH	37%	37%	100%	72%
Fire other dam to prop ins	MRP_PP	37%	37%	72%	100%

TABLE 5.10: Intervalles de confiance des corrélations : exemple avec 4 Lobs

En comparant cellule par cellule les 3 tableaux, il en découle que les corrélations obtenues avec les paramètres de prudence appartiennent aux intervalles de confiance. Néanmoins, en considérant tous les autres Lobs, on peut rencontrer des cas où la corrélation avec paramètres de prudence est supérieure à la borne supérieure (voir annexe A.5). Cette situation est en général due au seuil minimal de 20% de corrélation définie. Par conséquent, on s'intéressera aux corrélations équivalentes comme pour les autres tests.

Matrice	STEC en milliards	Impact	Correl Equiv
Borne Sup	1,185	-6,2%	37,6%
Borne Inf	#	#	#
Correl empirique param_prud (15,20)	1,031	-18,4%	25,8%

L'agrégation avec les bornes supérieures des corrélations fournit un STEC de 1,185 milliards correspondant à un scénario de dépendance de 37,6%. Pour les bornes inférieures, les corrélations sont presque toutes négatives et par conséquent, la matrice obtenue n'est pas définie positive. Le test consiste à comparer les corrélations équivalentes et il en découle que le scénario de base est bien dans l'intervalle de confiance.

On peut ainsi conclure que le choix du couple (20%,15%) est bien optimal au regard des critères définis.

5.4 Résultats de l'agrégation sur le portefeuille global

5.4.1 Présentation des résultats

Dans cette partie, nous allons reprendre les résultats de la section 5.1.1 en mettant à jour les corrélations empiriques sur les Lobs disposant de données. Cette mise à jour se fait avec les paramètres de prudence optimaux déterminés précédemment. Ainsi, une corrélation minimale de 20% sera retenue et toutes les autres corrélations seront arrondies à 15% supérieure.

A travers ce choix méthodologique, on obtient la matrice présentée ci-après. Les résultats concernent un certain nombre de Lobs (voir annexe A.4 pour plus de détails). Pour une interprétation de ces corrélations, on se basera sur les Lob FS correspondantes.

Num	Lob AXA	1	2	3	4	5	6	7	8	9	10	11	12
1	AUTO_PP_RC	100%											
2	AUTO_EN_RC	90%	100%										
3	AUTO_PP_Dom	60%	60%	100%									
4	MBL_DA	20%	20%	60%	100%								
5	MOD_DA	60%	60%	20%	60%	100%							
6	MAT_PP	30%	30%	60%	30%	60%	100%						
7	MAT_EN	30%	30%	60%	30%	60%	20%	100%					
8	AUTO_EN_Dom	60%	60%	75%	60%	30%	60%	60%	100%				
9	MRH	20%	20%	20%	20%	20%	20%	20%	20%	100%			
10	MRP_PP	20%	20%	20%	20%	20%	20%	20%	20%	45%	100%		
11	MRI_PP	20%	20%	20%	20%	20%	20%	20%	20%	90%	60%	100%	
12	AGRI	20%	20%	20%	20%	20%	20%	20%	20%	90%	60%	90%	100%
13	DIVERS_DAB_PP	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
14	RI	20%	20%	20%	20%	20%	20%	20%	20%	60%	60%	60%	75%
15	MRP_EN	20%	20%	20%	20%	20%	20%	20%	20%	30%	60%	45%	20%
16	MRI_EN	20%	20%	20%	20%	20%	20%	20%	20%	90%	60%	90%	75%
17	DAB_const	20%	20%	20%	20%	20%	20%	20%	20%	45%	30%	60%	60%
18	DIVERS_DAB_EN	20%	20%	20%	20%	20%	20%	20%	20%	30%	30%	30%	30%
19	CATNAT	20%	20%	20%	20%	20%	20%	20%	20%	20%	30%	30%	20%
20	DO	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
21	HOD	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	30%
22	RC_PP	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
23	RC_CONST	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
24	RC_ENT	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
25	RC_MED	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
26	RCD	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
27	PJ	20%	20%	45%	20%	45%	75%	75%	45%	20%	20%	20%	20%
28	ASSISTANCE	30%	30%	60%	30%	60%	60%	60%	60%	60%	60%	60%	60%
29	ASSISTANCE_DA	30%	30%	60%	30%	60%	60%	60%	60%	60%	60%	60%	60%
30	AUTRES_EN	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%
31	AUTRES_PP	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%
32	CAUTION	30%	30%	30%	30%	30%	60%	60%	30%	30%	30%	30%	30%
33	POOL_PROPERTY	30%	30%	30%	30%	30%	30%	30%	30%	75%	75%	75%	75%
34	PP_EN	20%	20%	20%	20%	20%	20%	20%	20%	30%	30%	30%	30%
35	QPSAMUT	30%	30%	30%	30%	30%	30%	30%	30%	60%	60%	60%	60%
36	SANTE	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%

TABLE 5.11: Matrice de corrélation empirique avec paramètres de prudence optimaux

L'agrégation finale génère un STEC de 1,28 milliards contre 1,53 milliards selon la méthode initiale. Ce montant correspond à un gain de solvabilité de 16,7% et il en résulte un gain de diversification de 41,6%. La méthode d'agrégation est équivalente à un scénario de dépendance de 28,9% entre les 36 Lobs du portefeuille.

STEC agrégé	Impact solva	Diversification	Correl Equiv
1,28	-16,7%	41,6%	28,9%

5.4.2 Bilan

Dans l'idéale, nous cherchons à estimer la vraie structure et le vrai niveau de dépendance entre Lobs sur la base des séries de boni mali. Néanmoins, dans la présente étude, les travaux

n'ont porté que sur les données existantes. L'approche adoptée se caractérise par un ensemble de choix méthodologiques et d'hypothèses faites en raison de la structure du portefeuille et le manque de données.

Premièrement, un calcul de corrélation sur les séries de boni mali et une hypothèse de corrélation sur les Lobs ne disposant pas de données a permis de construire des matrices de corrélations empiriques. L'hypothèse consiste à utiliser les paramètres de la FS en se basant sur la correspondance entre les Lobs AXA et Lobs FS. L'agrégation du STEC via cette méthode génère un faible montant qui s'avère insuffisant à première vue.

Ensuite, en explorant d'autres formes de dépendance via les copules, l'agrégation ne peut se faire que sur une partie du portefeuille contenant 25 Lobs.

Compte tenu de ces limites, l'approche méthodologique proposée consiste à modifier les corrélations empiriques. Il s'agit de fixer un seuil inférieur et d'arrondir au supérieur. Dès lors, le problème de l'estimation des dépendances se résume au choix des paramètres de seuil et d'arrondi. Le choix optimal de ces paramètres découle de certains critères et tests :

- Pour assurer un montant de capital suffisant, le nouveau STEC ⁸ doit être supérieur à celui obtenu avec les autres méthodes d'agrégation (autres formes de dépendances suivant les différentes copules calibrées sur les données)
- Pour générer un gain de solvabilité, le nouveau STEC devrait être inférieur à celui obtenu avec la méthode initiale.
- La nouvelle méthode doit être stable dans le temps pour être utilisée les années à venir. Des tests ont été effectués à cet effet, en comparant le niveau de dépendance avec celui des années passées.
- La précision est également un critère important. Celle-ci a été mesurée grâce aux intervalles de confiance.

A ces critères s'ajoutent l'interprétabilité des dépendances et le fait que la méthode soit applicable sur la globalité du portefeuille. En prenant en compte tous ces critères, les paramètres optimaux retenus sont un seuil de 20% et un arrondi à 15%. Cette nouvelle méthode permet de réduire le capital de solvabilité de 16,7%.

L'étude a montré que le montant du STEC est sensible aux dépendances entre Lobs. La diversification des risques est donc un important levier pour un assureur en solvabilité 2.

⁸obtenu via une agrégation utilisant la matrice des corrélations modifiées par les paramètres de seuil minimal et d'arrondi

Conclusion

En gestion des risques, la diversification est un moyen important pour les entreprises aussi bien du côté des actifs que des passifs. Chez les assureurs et réassureurs cette pratique est devenue indispensable pour une meilleure gestion du portefeuille. Le secteur de l'assurance fait face à un ensemble de risques, c'est pourquoi avec le cadre réglementaire solvabilité 2, des exigences quantitatives doivent être satisfaites pour justifier leur capacité à respecter tous les engagements d'assurance. Ceci explique l'intérêt de plus en plus croissant des assureurs à adopter des techniques de calculs pour bien estimer les métriques relatives à ces engagements dans le cadre de leur modèle interne.

C'est dans ce contexte qu'AXA France IARD à travers son modèle interne cherche à développer de nouvelles techniques permettant une meilleure estimation de ses passifs. Dans cette étude on s'est intéressé à l'agrégation du SCR ou STEC lié au risque de réserve pour l'année 2021 en mettant l'accent sur les dépendances inter Lobs.

Compte tenu des contraintes de données, l'objectif de l'étude a été de proposer une méthode alternative à la fois basée sur une approche statistique et qui répond aux exigences réglementaires. Différentes méthodes ont été testées pour prendre en compte différentes structures de dépendance. D'une part des corrélations empiriques sont calculées sur la base des séries de boni mali. D'autre part, des copules sont estimées sur une partie du portefeuille. Il s'agit de la copule de student, des copules archimédiennes (Gumbel, Clayton et Frank) et de la copule hiérarchique RVine. Cette dernière permet de contourner le problème de choix et d'estimation de copule en grande dimension. Puisque les méthodes d'agrégation via les copules ne peuvent pas s'appliquer sur tout le portefeuille, il convient d'utiliser les corrélations empiriques dans l'agrégation finale. Toutefois, ces paramètres devront être modifiés pour éviter une insuffisance des fonds propres. En utilisation le concept de corrélation équivalente, il est possible de comparer ces différentes méthodes. Ce concept permet de choisir le couple de paramètres de prudence (seuil minimal et arrondi au supérieur) à appliquer aux corrélations empiriques. De plus, des tests ont été réalisés sur les données historiques (2020

et 2019) et comptables pour déterminer le choix optimal.

La matrice de corrélation empirique qui en découle permet d'obtenir un montant de STEC de 1,28 milliards. Initialement ce montant est de 1,53 milliards, soit une baisse de 16,7%.

Les résultats de cette nouvelle méthode sont à l'heure actuelle satisfaisants au regard des hypothèses faites, du gain de diversification observé et du niveau de prudence choisi. L'objectif à venir sera de l'améliorer pour en faire un changement de modèle.

Cependant, la démarche est soumise à un certain nombre d'hypothèses et limites. Il s'agit principalement d'un manque de données. Rappelons que l'estimation des dépendances en grande dimension est une tâche difficile et il faut suffisamment de données pour des estimations précises.

Appendix A

Liste des annexes

A.1 Algorithme : Calcul de la densité de la copule RVine

Input : Les matrices M (structure de la copule RVine), T (contenant les types de copules) et P (les paramètres des copules correspondantes). Où $m_{k,k} = n - k + 1, \quad k = 1, \dots, n..$

Output : Densité de la distribution RVine au point (x_1, \dots, x_n) .

1. Soit $f = 1$.
2. On alloue $V^{direct} = (v_{i,k}^{direct} | i, k = 1, \dots, n)$.
3. On alloue $V^{indirect} = (v_{i,k}^{indirect} | i, k = 1, \dots, n)$.
4. $(v_{n,1}^{direct}, v_{n,2}^{direct}, \dots, v_{n,n}^{direct}) = (F_n(x_n), F_{n-1}(x_{n-1}), \dots, F_1(x_1))$
5. Soit $\mathbf{M} = (\mathbf{m}_{i,k} | i, k = 1, \dots, n)$ avec $\mathbf{m}_{i,k} = \max\{m_{i,k}, \dots, m_{n,k}\}$ pour tout $k = 1, \dots, n$ et $i = k, \dots, n$.
6. **for** $k = n - 1, \dots, 1$ **do** {itération sur les colonnes de M}
7. **for** $i = n, \dots, k + 1$ **do** {itération sur les lignes de M}
8. $z_{i,k}^{(1)} = v_{i,k}^{direct}$.
9. **if** $m_{i,k} = m_{i,k}$ **then**
10. $z_{i,k}^{(2)} = v_{i,(n-m_{i,k}+1)}^{direct}$.
11. **else**
12. $z_{i,k}^{(2)} = v_{i,(n-m_{i,k}+1)}^{indirect}$.
13. **end if**
14. $f = f \cdot c(z_{i,k}^{(1)}, z_{i,k}^{(2)} | t_{i,k}, p_{i,k})$.
15. $v_{i-1,k}^{direct} = h(z_{i,k}^{(1)}, z_{i,k}^{(2)} | t_{i,k}, p_{i,k})$ et $v_{i-1,k}^{indirect} = h(z_{i,k}^{(2)}, z_{i,k}^{(1)} | t_{i,k}, p_{i,k})$ où h est une fonction de densité conditionnelle.
16. **end for**
17. **end for**
18. **return** f

A.2 Algorithme : Méthode séquentielle de sélection du modèle RVine

Input : Données $(x_{l1}, \dots, x_{ln}), l = 1, \dots, N$ (réalisations d'une variable aléatoire i.i.d).

Output : Spécification de la copule RVine.

1. Calcul du tau de Kendall empirique $\hat{\tau}_{j,k}$ pour toute paire de variable possible $j, k, 1 \leq j < k \leq n$.
2. Selection du spanning tree qui maximise la somme des $\hat{\tau}_{j,k}$:

$$\max_{e=\{j,k\} \in \text{Spanning tree}} \sum |\hat{\tau}_{j,k}|$$

3. Pour tout noeud $\{j, k\}$ du spanning tree sélectionné, choisir la copule et estime les paramètres correspondants. Puis, transformer $\hat{F}_{j|k}(x_{lj}|x_{lk})$ et $\hat{F}_{k|j}(x_{lk}|x_{lj}), l = 1, \dots, N$ en utilisant la copule estimée \hat{C}_{jk}
4. **for** $i = 2, \dots, n - 1$ **do** { itération suivant les arbres }
5. Calcul du $\hat{\tau}_{j,k|D}$ pour toute paire de variable conditionnelle $\{j, k|D\}$ pouvant appartenir à l'arbre T_i ou suivant le critère de proximité.
6. Parmi tous les noeuds, sélectionner le spanning tree qui maximise la somme des $\hat{\tau}_{j,k|D}$:

$$\max_{e=\{j,k|D\} \in \text{Spanning tree}} \sum |\hat{\tau}_{j,k|D}|$$

7. Pour tout noeud $\{j, k|D\}$, dans le spanning tree sélectionné, choisir la copule conditionnelle et estimer les paramètres correspondants. Puis, transformer $\hat{F}_{j|k \cup D}(x_{lj}|x_{lk}, x_{lD})$ et $\hat{F}_{k|j \cup D}(x_{lk}|x_{lj}, x_{lD}), l = 1, \dots, N$ en utilisant la copule estimée $\hat{C}_{jk|D}$.
8. **end for**

A.3 Algorithme Bootstrap pour les intervalles de confiance avec R

```
## CI level by Bootstrap
```

```
B<-1000      # nb of simulations
```

```
boot.cor.all<-NULL # initialize correlation vector
```

```
n <- nrow(Data_Lob)
```

```
var1 = Data_Lob[,i]
var2 = Data_Lob[,j]

for (i in 1:B){
  index<-sample(1:n, replace=T)
  boot.var1<-var1[index,]
  boot.var2<-var2[index,]
  boot.cor<-cor(boot.var1, boot.var2,method = "spearman")
  boot.cor.all<-c(boot.cor.all, boot.cor)
}

## plot the bootstrap distribution
hist(boot.cor.all,prob=T,col = "darkblue")
lines(density(boot.cor.all))

## Bootstrap standard error
sd(boot.cor.all)

## percentile bootstrap CI 95% level
quantile(boot.cor.all, prob=c(0.025, 0.975))

## correlation CI 95% level
CorrCI = unname(quantile(boot.cor.all, prob=c(0.025, 0.975)))
```

A.4 Corrélations empiriques avec paramètres de prudence optimaux

Légende			
Lob QIS	Num Lob QIS	Lob AXA	Num Lob Axa
Motor veh liab ins	4	AUTO_PP_RC	1
Motor veh liab ins	4	AUTO_EN_RC	2
Oth motor ins	5	AUTO_PP_Dom	3
Motor veh liab ins	4	MBL_DA	4
Oth motor ins	5	MOD_DA	5
Mat ins	6	MAT_PP	6
Mat ins	6	MAT_EN	7
Oth motor ins	5	AUTO_EN_Dom	8
Fire other dam to prop ins	7	MRH	9
Fire other dam to prop ins	7	MRP_PP	10
Fire other dam to prop ins	7	MRI_PP	11
Fire other dam to prop ins	7	AGRI	12
Fire other dam to prop ins	7	DIVERS_DAB_PP	13
Fire other dam to prop ins	7	RI	14
Fire other dam to prop ins	7	MRP_EN	15
Fire other dam to prop ins	7	MRI_EN	16
Fire other dam to prop ins	7	DAB_const	17
Misc fin loss ins	12	DIVERS_DAB_EN	18
Fire other dam to prop ins	7	CATNAT	19
Fire other dam to prop ins	7	DO	20
Fire other dam to prop ins	7	HOD	21
Gen liab ins	8	RC_PP	22
Gen liab ins	8	RC_CONST	23
Gen liab ins	8	RC_ENT	24
Gen liab ins	8	RC_MED	25
Gen liab ins	8	RCD	26
Leg exp ins	10	PJ	27
Ass	11	ASSISTANCE	28
Ass	11	ASSISTANCE_DA	29
Non-prop casu reins	26	AUTRES_EN	30
Non-prop casu reins	26	AUTRES_PP	31
Cred and sur ins	9	CAUTION	32
Fire other dam to prop reins	28	POOL_PROPERTY	33
Misc fin loss ins	12	PP_EN	34
Non-prop prop reins	28	QPSAMUT	35
Med exp ins	1	SANTE	36

A.5 Intervalle de confiance des corrélations

Correl Rang	Lob QIS	4	4	5	4	5	6	6	5	7	7	7	7	7	7	7	7	12	7	7	7	8	8	8	8	8	10	
Lob QIS	Lob AXA	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
4	1	100%	90%	82%	47%	82%	78%	78%	82%	37%	37%	37%	37%	37%	37%	37%	37%	38%	37%	37%	37%	-14%	-14%	-14%	-14%	-14%	47%	
4	2	90%	100%	82%	54%	82%	78%	78%	82%	37%	37%	37%	37%	37%	37%	37%	37%	38%	37%	37%	37%	-14%	-14%	-14%	-14%	-14%	47%	
5	3	82%	82%	100%	82%	46%	96%	96%	84%	35%	35%	35%	35%	35%	35%	35%	35%	47%	35%	35%	35%	31%	31%	31%	31%	31%	75%	
4	4	47%	54%	82%	100%	82%	78%	78%	82%	37%	37%	37%	37%	37%	37%	37%	37%	38%	37%	37%	37%	-14%	-14%	-14%	-14%	-14%	47%	
5	5	82%	82%	46%	82%	100%	96%	96%	58%	35%	35%	35%	35%	35%	35%	35%	35%	47%	35%	35%	35%	31%	31%	31%	31%	31%	75%	
6	6	78%	78%	96%	78%	96%	100%	15%	96%	46%	46%	46%	46%	46%	46%	46%	46%	51%	46%	46%	46%	46%	70%	70%	70%	70%	100%	
6	7	78%	78%	96%	78%	96%	15%	100%	96%	46%	46%	46%	46%	46%	46%	46%	46%	51%	46%	46%	46%	70%	70%	70%	70%	70%	100%	
5	8	82%	82%	84%	82%	58%	96%	96%	100%	35%	35%	35%	35%	35%	35%	35%	35%	47%	35%	35%	35%	31%	31%	31%	31%	31%	75%	
7	9	37%	37%	35%	37%	35%	46%	46%	35%	100%	72%	91%	97%	24%	73%	49%	90%	67%	75%	37%	57%	50%	69%	69%	69%	69%	69%	4%
7	10	37%	37%	35%	37%	35%	46%	46%	35%	72%	100%	73%	81%	34%	80%	84%	74%	53%	75%	63%	43%	68%	69%	69%	69%	69%	69%	4%
7	11	37%	37%	35%	37%	35%	46%	46%	35%	91%	73%	100%	94%	38%	67%	61%	91%	66%	75%	58%	70%	55%	69%	69%	69%	69%	69%	4%
7	12	37%	37%	35%	37%	35%	46%	46%	35%	97%	81%	94%	100%	42%	84%	49%	89%	71%	75%	29%	66%	69%	69%	69%	69%	69%	69%	4%
7	13	37%	37%	35%	37%	35%	46%	46%	35%	24%	34%	38%	42%	100%	48%	53%	38%	22%	75%	41%	6%	22%	69%	69%	69%	69%	69%	4%
7	14	37%	37%	35%	37%	35%	46%	46%	35%	73%	80%	67%	84%	48%	100%	40%	72%	83%	75%	20%	46%	65%	69%	69%	69%	69%	69%	4%
7	15	37%	37%	35%	37%	35%	46%	46%	35%	49%	84%	61%	49%	53%	40%	100%	71%	15%	75%	75%	37%	75%	69%	69%	69%	69%	69%	4%
7	16	37%	37%	35%	37%	35%	46%	46%	35%	90%	74%	91%	89%	38%	72%	71%	100%	58%	75%	51%	69%	58%	69%	69%	69%	69%	69%	4%
7	17	37%	37%	35%	37%	35%	46%	46%	35%	67%	53%	66%	71%	22%	83%	15%	58%	100%	75%	51%	68%	48%	69%	69%	69%	69%	69%	4%
12	18	38%	38%	47%	38%	47%	51%	51%	47%	75%	75%	75%	75%	75%	75%	75%	75%	75%	100%	75%	75%	75%	76%	76%	76%	76%	76%	24%
7	19	37%	37%	35%	37%	35%	46%	46%	35%	37%	63%	58%	29%	41%	20%	75%	51%	51%	75%	100%	63%	48%	69%	69%	69%	69%	69%	4%
7	20	37%	37%	35%	37%	35%	46%	46%	35%	57%	43%	70%	66%	6%	46%	37%	69%	68%	75%	63%	100%	77%	69%	69%	69%	69%	69%	4%
7	21	37%	37%	35%	37%	35%	46%	46%	35%	50%	68%	55%	69%	22%	65%	75%	58%	48%	75%	48%	77%	100%	69%	69%	69%	69%	69%	4%
8	22	-14%	-14%	31%	-14%	31%	70%	70%	31%	69%	69%	69%	69%	69%	69%	69%	69%	69%	76%	69%	69%	69%	100%	35%	5%	31%	53%	41%
8	23	-14%	-14%	31%	-14%	31%	70%	70%	31%	69%	69%	69%	69%	69%	69%	69%	69%	69%	76%	69%	69%	69%	35%	100%	46%	63%	69%	41%
8	24	-14%	-14%	31%	-14%	31%	70%	70%	31%	69%	69%	69%	69%	69%	69%	69%	69%	69%	76%	69%	69%	69%	5%	46%	100%	42%	51%	41%
8	25	-14%	-14%	31%	-14%	31%	70%	70%	31%	69%	69%	69%	69%	69%	69%	69%	69%	69%	76%	69%	69%	69%	31%	63%	42%	100%	52%	41%
8	26	-14%	-14%	31%	-14%	31%	70%	70%	31%	69%	69%	69%	69%	69%	69%	69%	69%	69%	76%	69%	69%	69%	53%	69%	51%	52%	100%	41%
10	27	47%	47%	75%	47%	75%	100%	100%	75%	4%	4%	4%	4%	4%	4%	4%	4%	4%	24%	4%	4%	4%	41%	41%	41%	41%	100%	

Borne Sup

Correl Rang	Lob QIS	4	4	5	4	5	6	6	5	7	7	7	7	7	7	7	7	7	12	7	7	7	8	8	8	8	8	10
Lob QIS	Lob AXA	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
4	1	100%	45%	-6%	-52%	-6%	-71%	-71%	-6%	-51%	-51%	-51%	-51%	-51%	-51%	-51%	-51%	-51%	-50%	-51%	-51%	-51%	-74%	-74%	-74%	-74%	-74%	-54%
4	2	45%	100%	-6%	-41%	-6%	-71%	-71%	-6%	-51%	-51%	-51%	-51%	-51%	-51%	-51%	-51%	-51%	-50%	-51%	-51%	-51%	-74%	-74%	-74%	-74%	-74%	-54%
5	3	-6%	-6%	100%	-6%	-53%	-31%	-31%	3%	-55%	-55%	-55%	-55%	-55%	-55%	-55%	-55%	-55%	-45%	-55%	-55%	-55%	-50%	-50%	-50%	-50%	-50%	-14%
4	4	-52%	-41%	-6%	100%	-6%	-71%	-71%	-6%	-51%	-51%	-51%	-51%	-51%	-51%	-51%	-51%	-51%	-50%	-51%	-51%	-51%	-74%	-74%	-74%	-74%	-74%	-54%
5	5	-6%	-6%	-53%	-6%	100%	-31%	-31%	-33%	-55%	-55%	-55%	-55%	-55%	-55%	-55%	-55%	-55%	-45%	-55%	-55%	-55%	-50%	-50%	-50%	-50%	-50%	-14%
6	6	-71%	-71%	-31%	-71%	-31%	100%	-89%	-31%	-83%	-83%	-83%	-83%	-83%	-83%	-83%	-83%	-83%	-76%	-83%	-83%	-83%	-81%	-81%	-81%	-81%	-81%	14%
6	7	-71%	-71%	-31%	-71%	-31%	-89%	100%	-31%	-83%	-83%	-83%	-83%	-83%	-83%	-83%	-83%	-83%	-76%	-83%	-83%	-83%	-81%	-81%	-81%	-81%	-81%	14%
5	8	-6%	-6%	3%	-6%	-33%	-31%	-31%	100%	-55%	-55%	-55%	-55%	-55%	-55%	-55%	-55%	-55%	-45%	-55%	-55%	-55%	-50%	-50%	-50%	-50%	-50%	-14%
7	9	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	100%	-33%	42%	26%	-64%	-22%	-37%	41%	-28%	-8%	-61%	-32%	-46%	-26%	-26%	-26%	-26%	-26%	-68%
7	10	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	-33%	100%	-24%	-11%	-53%	-6%	-5%	-20%	-42%	-8%	-35%	-58%	-21%	-26%	-26%	-26%	-26%	-26%	-68%
7	11	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	42%	-24%	100%	64%	-54%	0%	-40%	30%	-8%	-8%	-44%	-34%	-40%	-26%	-26%	-26%	-26%	-26%	-68%
7	12	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	26%	-11%	64%	100%	-51%	29%	-50%	13%	3%	-8%	-66%	-30%	-9%	-26%	-26%	-26%	-26%	-26%	-68%
7	13	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	-64%	-53%	-54%	-51%	100%	-44%	-51%	-52%	-76%	-8%	-42%	-73%	-70%	-26%	-26%	-26%	-26%	-26%	-68%
7	14	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	-22%	-6%	0%	29%	-44%	100%	-60%	-23%	27%	-8%	-68%	-36%	-25%	-26%	-26%	-26%	-26%	-26%	-68%
7	15	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	-37%	-5%	-40%	-50%	-51%	-60%	100%	-34%	-68%	-8%	1%	-55%	-19%	-26%	-26%	-26%	-26%	-26%	-68%
7	16	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	41%	-20%	30%	13%	-52%	-23%	-34%	100%	-44%	-8%	-54%	-25%	-34%	-26%	-26%	-26%	-26%	-26%	-68%
7	17	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	-28%	-42%	-8%	3%	-76%	27%	-68%	-44%	100%	-8%	-56%	-24%	-33%	-26%	-26%	-26%	-26%	-26%	-68%
12	18	-50%	-50%	-45%	-50%	-45%	-76%	-76%	-45%	-8%	-8%	-8%	-8%	-8%	-8%	-8%	-8%	-8%	100%	-8%	-8%	-8%	-10%	-10%	-10%	-10%	-10%	-64%
7	19	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	-61%	-35%	-44%	-66%	-42%	-68%	1%	-54%	-56%	-8%	100%	-28%	-49%	-26%	-26%	-26%	-26%	-26%	-68%
7	20	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	-32%	-58%	-34%	-30%	-73%	-36%	-55%	-25%	-24%	-8%	-28%	100%	6%	-26%	-26%	-26%	-26%	-26%	-68%
7	21	-51%	-51%	-55%	-51%	-55%	-83%	-83%	-55%	-46%	-21%	-40%	-9%	-70%	-25%	-19%	-34%	-33%	-8%	-49%	6%	100%	-26%	-26%	-26%	-26%	-26%	-68%
8	22	-74%	-74%	-50%	-74%	-50%	-81%	-81%	-50%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-10%	-26%	-26%	-26%	100%	-68%	-84%	-64%	-50%	-44%
8	23	-74%	-74%	-50%	-74%	-50%	-81%	-81%	-50%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-10%	-26%	-26%	-26%	-68%	100%	-51%	-34%	-35%	-44%
8	24	-74%	-74%	-50%	-74%	-50%	-81%	-81%	-50%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-10%	-26%	-26%	-26%	-84%	-51%	100%	-55%	-38%	-44%
8	25	-74%	-74%	-50%	-74%	-50%	-81%	-81%	-50%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-10%	-26%	-26%	-26%	-64%	-34%	-55%	100%	-22%	-44%
8	26	-74%	-74%	-50%	-74%	-50%	-81%	-81%	-50%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-26%	-10%	-26%	-26%	-26%	-50%	-35%	-38%	-22%	100%	-44%
10	27	-54%	-54%	-14%	-54%	-14%	14%	14%	-14%	-68%	-68%	-68%	-68%	-68%	-68%	-68%	-68%	-68%	-68%	-64%	-68%	-68%	-68%	-44%	-44%	-44%	-44%	100%

Borne Inf

A.6 Coefficient de dépendance de queue à partir de la copule de student

Rappelons que la distribution de student est symétrique.

Par conséquent, $TDC_{upper} = TDC_{lower}$

Num	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
1	1																									
2	0,40	1																								
3	0,36	0,51	1																							
4	0,08	0,09	0,07	1																						
5	0,05	0,06	0,07	0,22	1																					
6	0,10	0,22	0,28	0,08	0,11	1																				
7	0,10	0,07	0,07	0,06	0,01	0,03	1																			
8	0,25	0,08	0,11	0,07	0,03	0,02	0,15	1																		
9	0,07	0,03	0,03	0,11	0,05	0,01	0,46	0,18	1																	
10	0,08	0,04	0,03	0,13	0,03	0,01	0,46	0,25	0,54	1																
11	0,04	0,05	0,06	0,16	0,18	0,19	0,04	0,05	0,06	0,06	1															
12	0,12	0,05	0,06	0,15	0,06	0,05	0,19	0,23	0,19	0,35	0,08	1														
13	0,20	0,11	0,13	0,16	0,04	0,03	0,09	0,25	0,10	0,08	0,07	0,06	1													
14	0,13	0,07	0,06	0,11	0,03	0,02	0,43	0,19	0,44	0,36	0,06	0,17	0,14	1												
15	0,08	0,06	0,06	0,10	0,16	0,03	0,15	0,10	0,19	0,21	0,02	0,31	0,02	0,11	1											
16	0,05	0,02	0,02	0,06	0,08	0,05	0,16	0,08	0,15	0,10	0,04	0,12	0,10	0,18	0,11	1										
17	0,08	0,05	0,12	0,03	0,14	0,04	0,05	0,13	0,09	0,04	0,08	0,03	0,21	0,07	0,07	0,11	1									
18	0,02	0,02	0,02	0,06	0,05	0,02	0,12	0,06	0,15	0,15	0,02	0,09	0,05	0,17	0,13	0,22	0,14	1								
19	0,08	0,07	0,08	0,15	0,02	0,03	0,10	0,16	0,10	0,19	0,03	0,15	0,17	0,13	0,10	0,08	0,08	0,24	1							
20	0,20	0,20	0,25	0,02	0,04	0,11	0,03	0,13	0,02	0,02	0,05	0,03	0,10	0,03	0,02	0,02	0,13	0,04	0,03	1						
21	0,04	0,07	0,08	0,06	0,10	0,26	0,06	0,01	0,05	0,05	0,16	0,08	0,02	0,07	0,07	0,08	0,06	0,14	0,08	0,04	1					
22	0,07	0,04	0,03	0,15	0,05	0,03	0,15	0,14	0,22	0,21	0,04	0,33	0,09	0,23	0,31	0,16	0,04	0,16	0,20	0,01	0,07	1				
23	0,05	0,05	0,04	0,13	0,33	0,14	0,03	0,06	0,10	0,07	0,21	0,09	0,03	0,07	0,09	0,15	0,11	0,09	0,04	0,04	0,13	0,06	1			
24	0,13	0,08	0,10	0,01	0,05	0,13	0,11	0,07	0,06	0,07	0,03	0,14	0,02	0,10	0,15	0,20	0,07	0,18	0,05	0,07	0,14	0,10	0,10	1		
25	0,10	0,13	0,17	0,06	0,23	0,26	0,00	0,04	0,01	0,01	0,10	0,02	0,06	0,00	0,06	0,04	0,15	0,02	0,05	0,09	0,10	0,02	0,15	0,08	1	

A.7 Matrice des arbres de la copule hiérarchique RVine

Légende (spécifique aux résultats de la copule RVine)

AUTO_PP_RC	AUTO_EN_RC	AUTO_PP_Dom	MBI_DA	MOD_DA	AUTO_EN_Dom	MRH	MRP_PP	
1	2	3	4	5	6	7	8	
MRI_PP	AGRI	DIVERS_DAB_PP	RI	MRP_EN	MRI_EN	DAB_const	DIVERS_DAB_EN	
9	10	11	12	13	14	15	16	
CATNAT	DO	HOD	RC_PP	RC_CONST	RC_ENT	RC_MED	RCD	PJ
17	18	19	20	21	22	23	24	25

19																							
16	17																						
20	16	13																					
15	20	16	21																				
24	15	20	16	8																			
22	24	15	20	16	11																		
23	22	24	15	20	16	18																	
4	23	22	24	15	20	16	1																
5	4	23	22	24	15	20	16	2															
7	5	4	23	22	24	15	20	16	3														
12	7	5	4	23	22	24	15	20	16	24													
17	12	7	5	4	23	22	24	15	20	16	6												
13	10	12	7	5	4	23	22	24	15	20	16	25											
21	9	10	12	7	5	4	23	22	24	15	20	16	14										
8	11	9	10	12	7	5	4	23	22	6	15	20	16	23									
10	14	11	9	10	12	7	5	4	23	22	22	15	20	16	4								
9	25	14	11	9	10	12	7	5	4	25	23	22	15	20	16	5							
14	18	25	14	11	9	10	12	7	5	14	4	23	22	15	20	16	7						
25	6	18	25	14	14	9	10	12	7	9	5	4	23	22	15	20	16	9					
6	3	6	18	25	25	14	9	10	12	12	7	5	4	9	22	15	20	16	10				
3	2	3	6	18	6	25	14	9	10	10	12	7	5	12	9	22	15	20	16	15			
1	21	2	3	6	3	6	25	14	9	23	10	12	7	10	12	9	22	15	20	16	16		
11	1	21	2	3	1	3	6	25	14	7	9	10	12	7	10	12	9	22	15	20	12	12	
2	8	1	1	2	2	1	3	6	25	5	14	9	10	4	7	10	12	22	22	22	20	20	
18	13	8	8	1	18	2	2	3	6	4	25	14	9	5	5	7	10	12	22	22	20	22	22

Matrice constraint set avec numéro de Lob correspondants à la légende

A.8 Matrice des types de copules et paramètres de la copule hiérarchique RVine

Légende types de copule

Numéro	1	2	3	4	5
Nom copule	Gaussienne	Student	Clayton	Gumbel	Frank

Bibliographie

- Appert-Raullin, Y., P. Tann, and L. Devineau (2014). "Corrélations et agrégation des risques non-vie : Aspects théoriques et réalité opérationnelle". In: *Congrès des actuaires*.
- Appert-Raullin, Y., P. Tann, and J.F. Richard (2011). "Solvency II Technical documentation : PS11, PS12, PS13 - Aggregation". In.
- Artzner, P. et al. (1999). "Coherent measures of risk". In: *Mathematical Finance*, 203–228.
- Baradel, Nicolas (2022). "Langage R en actuariat". In: *ENSAE*. URL: https://www.nicolasbaradel.fr/enseignement/ressources/cours_langage_r_en_actuariat.pdf.
- Basel-Committee (Oct. 2010). "Developments in Modelling Risk Aggregation". In: *Bank for international settlements*. URL: <https://www.bis.org/publ/joint25.htm>.
- Bernard, C. and S. Vanduffel (Aug. 2016). "Risk Aggregation and Diversification". In: *Canadian Institute of Actuaries (CIA)* 36. URL: <http://www.cia-ica.ca/docs/default-source/2016/216037e.pdf?sfvrsn=0>.
- Cadoux, David and Jean-Marc Loizeau (2004). "Copules et Dépendances : Application pratique à la détermination du besoin en fonds propres d'un assureur non vie". In: *BULLETIN FRANÇAIS D'ACTUARIAT* 6.12.
- "Calculation of the Solvency Capital Requirement" (2011). In: URL: https://www.eiopa.europa.eu/rulebook/solvency-ii/article-2324_en.
- Deheuvels, Paul (1979). "La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance". In: *Bulletin de la Classe des Sciences*.
- Dißmann, J. et al. (2012). "Selecting and estimating regular vine copulae and application to financial returns". In.
- Embrechts, P., A.J. McNeil, and D. Straumann (2001). "Correlation and Dependency in Risk Management : Properties and Pitfalls". In: *Dempster, M., ed., Value at Risk and Beyond (Cambridge University Press)*, pp. 173–208.
- Faivre, F. (2002). "Copules, besoin en fonds propres et allocation de capital pour une compagnie d'assurance non vie multibranches". In: *Ressources actuarielles*.
- Mai, Jan-Frederik and Matthias Scherer (2014). "How to Measure Dependence?" In: *Financial Engineering with Copulas Explained*. London: Palgrave Macmillan UK, pp. 35–48. ISBN:

- 978-1-137-34631-5. DOI: 10.1057/9781137346315_3. URL: https://doi.org/10.1057/9781137346315_3.
- McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts (2015). *Quantitative risk management: Concepts, techniques and tools: Revised edition*. English. Princeton University Press. ISBN: 9780691166278.
- Morales-Napoles, O., R. Cooke, and D. Kurowicka (2010). "About the number of vines and regular vines on n nodes". In.
- Nelsen, Roger (Jan. 2006). *An Introduction to Copulas (Springer Series in Statistics)*. ISBN: 0387286594.
- Partrat, Christian (2003). "« Corrélation » entre risques d'assurance non vie : une introduction". In: *BULLETIN FRANÇAIS D'ACTUARIAT* 6.10, pp. 79–91.
- Pestano, Joshua S. (Aug. 2017). "Understanding Diversification of Risk and Three Ways Insurance Companies Achieve It". In: *Insurance Industry Knowledge Resources*. URL: <https://www.riskreinsurance.com/understanding-diversification-of-risk-and-three-ways-insurance-companies-achieve-it/>.
- Planchet, Frédéric (2021). "Introduction à la théorie des copules". In: *Ressources actuarielles*.
- "Risk diversification benefits in insurance" (n.d.). In: *The Actuarial Profession* (). URL: <https://www.actuaries.org.uk/system/files/documents/pdf/riskdiversificationbenefitsininsurances.pdf>.
- Ruscio, John (Nov. 2008). "Constructing Confidence Intervals for Spearman's Rank Correlation with Ordinal Data: A Simulation Study Comparing Analytic and Bootstrap Methods". In: *Journal of Modern Applied Statistical Methods* 7, pp. 416–434. DOI: 10.22237/jmasm/1225512360.
- Sia-Conseil (2010). "Le Point sur Solvabilité II, Mise en place opérationnelle du projet : Le Pilier 1". In: p. 5.