

**Mémoire présenté devant le**  
**Conservatoire National des Arts et Métiers**  
**pour l'obtention du diplôme du Master Actuariat**  
**et l'admission à l'Institut des Actuares**

**le 4 juillet 2018**

Par : Théo SERMET

Titre: Allocation du chargement de réassurance et du coût du capital à la tarification  
des catastrophes naturelles

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

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

*Membre présent du jury de l'Institut  
des Actuares :*

Florence PICARD

Pierre PETAUTON

Vincent RUOL

*Membres présents du jury du Master  
Actuariat du CNAM :*

François WEISS

David FAURE

Nathanaël ABECERA

*Signature :      Entreprise :*

*Nom : AXA Global P&C (AXA Global Re)*

*Signature :*

*Directeur de mémoire en entreprise :*

*Nom : Simon BLAQUIERE*

*Signature :*

*Président du jury : Alexis COLLOMB*

*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)***

*Signature du responsable entreprise :*

*Secrétariat :*

*Bibliothèque :*

*Signature du candidat :*



## Résumé

**Mots-clés :** catastrophe naturelle, tarification, réassurance, AEP, OEP, ouragan, modèle stochastique, coût du capital, segmentation.

Avec le développement de modèles stochastiques de catastrophes naturelles, les assureurs ont précisé leur besoin en réassurance afin de se protéger contre ce type de risque.

Aujourd'hui, la précision et la qualité des modèles devient suffisante pour permettre de tarifier spécifiquement les garanties de catastrophes naturelles site par site. Cette tarification Cat permet d'intégrer les événements extrêmes à une tarification auparavant limitée aux pertes attritionnelles.

Néanmoins, cette tarification ne prend pas en compte le chargement supplémentaire engendré par les accumulations de risques à un même endroit. Ces accumulations conduisent l'assureur à se protéger à l'aide de la réassurance et à immobiliser du capital règlementaire pour pouvoir continuer son activité. C'est notamment le cas des grands risques (entreprises) qui génèrent davantage de volatilité que les risques de plus faible somme assurée.

L'objectif de ce mémoire est de proposer une méthode d'allocation du chargement de réassurance et du coût du capital dans la tarification Cat effectué site par site.

*Pour des raisons de confidentialité, les montants de résultats de tarification ont été modifiés et il est conseillé au lecteur de se focaliser sur la méthodologie plus que sur les tarifs chiffrés.*

## Abstract

**Key words:** natural catastrophes, pricing, reinsurance, AEP, OEP, hurricane, stochastic model, cost of capital, segmentation.

Thanks to a strong development of natural catastrophe stochastic models, insurers have been able to precise their reinsurance need in order to get a better protection against this type of risk. The current precision and quality of these models now allows pricing specifically the natural catastrophes risk for each and every site location. This Cat pricing makes it possible to take into account extreme events into insurance pricing that used to be limited to attritional losses.

However, this Cat pricing does not integrate additional loadings caused by risk geographical accumulations. These accumulations lead the insurer to buy reinsurance protection and allocate solvency capital to be able to grow. This is notably the case of large risks (commercial) as they generate more portfolio volatility than risks with a smaller sum insured.

The aim of this paper is to present a methodology of allocation of reinsurance loading and cost of capital into Cat pricing per location.

*For confidentiality reasons, pricing results have been modified and the reader is advised to focus on the methodology rather than the tariff figures.*

## Note de synthèse

### 1. Contexte et problématique

L'entrée en vigueur de la directive européenne Solvabilité 2 (2009/138/EC) au 1<sup>er</sup> janvier 2016 contraint les entreprises à respecter, entre autres, une exigence quantitative de fonds propres détenus pour se protéger contre la perte la plus importante susceptible de se produire avec une probabilité d'une chance sur 200 dans l'année. Quelle que soit la branche étudiée, cette exigence de capital règlementaire intègre les catastrophes naturelles, pour lesquelles des modèles stochastiques ont été développés dans les années 1990.

Ces modèles stochastiques produisent une distribution des pertes et permettent non seulement d'obtenir la valeur du quantile de pertes à 200 ans pour le portefeuille de l'assureur, mais aussi l'ensemble de la distribution statistique de ces pertes. Dès lors, il est intéressant pour l'assureur d'utiliser ces résultats sur l'ensemble de la chaîne de valeur, de la souscription au coût du capital.

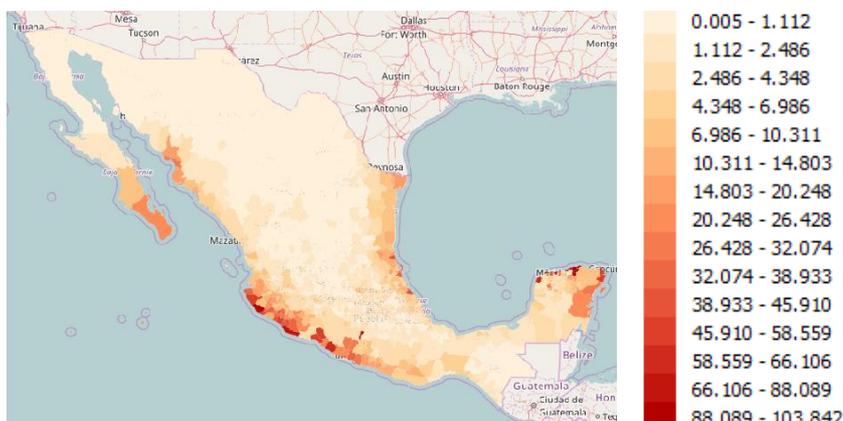
En utilisant les modèles stochastiques site par site, on peut dériver une tarification des catastrophes naturelles et l'intégrer dans la tarification de l'assureur (tarification Cat). Afin de proposer une tarification complète du site qui intègre non seulement son risque individuel mais aussi l'interaction avec le reste du portefeuille (accumulations, besoin en capital), on cherche à proposer une méthode d'allocation du chargement de réassurance et du coût du capital dans la tarification Cat du site.

### 2. Approche proposée

Le périmètre de l'étude est défini par un portefeuille d'assureur européen présent au Mexique, pour le risque ouragan uniquement. Le modèle utilisé est issu du modèle interne d'AXA pour les ouragans.

Afin de pouvoir proposer une méthode d'allocation, il a été nécessaire d'analyser les résultats du modèle Cat et leur utilisation pour une tarification. Différentes méthodes d'allocation ont été ensuite comparées et testées avant d'effectuer une proposition de régression sur les résultats modélisés afin de pas avoir besoin de relancer la totalité de la chaîne de modélisation pour avoir les résultats pour un nouveau site souscrit.

#### a. Tarification Cat



**Figure 1 :** carte de la tarification Cat obtenue avec le modèle. La légende est en échelle relative par rapport à une valeur maximale. Les résultats ont été moyennés par municipalité.

La tarification Cat est issue de la moyenne modélisée des pertes liées aux ouragans pour chaque site. Les résultats de la tarification montrent les différentes relativités utilisées dans le modèle, en fonction de variables qui ont un impact sur la vulnérabilité du site aux ouragans comme par exemple la localisation géographique, l'occupation, la structure du bâtiment ou son nombre d'étages.

Cette méthode permet aussi de montrer la segmentation géographique de zones de risques (cf. Figure 1).

#### b. Comparaison de méthodes d'allocation

Différentes méthodes d'allocation sont étudiées, avec un cadre théorique lié à la théorie des jeux coopératifs. On compare notamment les méthodes proportionnelles, la méthode marginale, la méthode des bénéfiques résiduels, la méthode de Shapley et du Nucléole.

Si les méthodes de Shapley et du Nucléole sont les plus satisfaisantes d'un point de vue théorique, la difficulté technique de mise en œuvre conduit plutôt à sélectionner entre deux méthodes : la méthode des bénéfiques résiduels et la méthode marginale.

La première méthode répond à une logique d'allocation sur un stock de sites préexistants dans un portefeuille donné, alors que la seconde méthode répond plutôt à l'introduction d'un nouveau site en supposant le portefeuille déjà constitué.

En plus de la sélection des méthodes d'allocation à étudier en profondeur, il convient aussi de choisir la métrique qui sera utilisée dans ces méthodes. Par exemple, la méthode la plus simple d'allocation du chargement de réassurance serait d'utiliser la perte moyenne modélisée pour l'ouragan de chaque site. Cette méthode ne prend pas en compte le fait que certains sites sont plus susceptibles de faire jouer la réassurance que d'autres, indépendamment de leur perte moyenne.

Les deux métriques retenues sont donc l'impact sur les récupérations de réassurance en vision modèle pour l'allocation du chargement de réassurance et la TVaR au-delà du quantile de perte à 200 ans pour le coût du capital.

#### c. Proposition de méthodes de régression

Contrairement à la perte moyenne Cat (tarification Cat) modélisée par site, l'allocation du chargement de réassurance et du coût du capital nécessitent de modéliser la totalité du portefeuille à chaque mise à jour du calcul. En raison de la complexité de la chaîne de modélisation pour un portefeuille total, il est proposé une méthode de prédiction des résultats de prime pure de réassurance afin d'éviter d'effectuer le calcul à chaque nouvelle souscription.

Après avoir testé les modèles linéaires généralisés, ainsi qu'un genre de régression géographique très simple, il est proposé une méthode de forêt aléatoire pour prédire le chargement de réassurance associé à chaque site.

Cette méthode n'est cependant pas proposée pour l'allocation du coût du capital en raison de sa forte volatilité et de sa forte segmentation. Au contraire, il est proposé une méthode d'allocation du coût du capital à un niveau géographique plus large que site par site, à savoir la municipalité (équivalent de la commune en France).

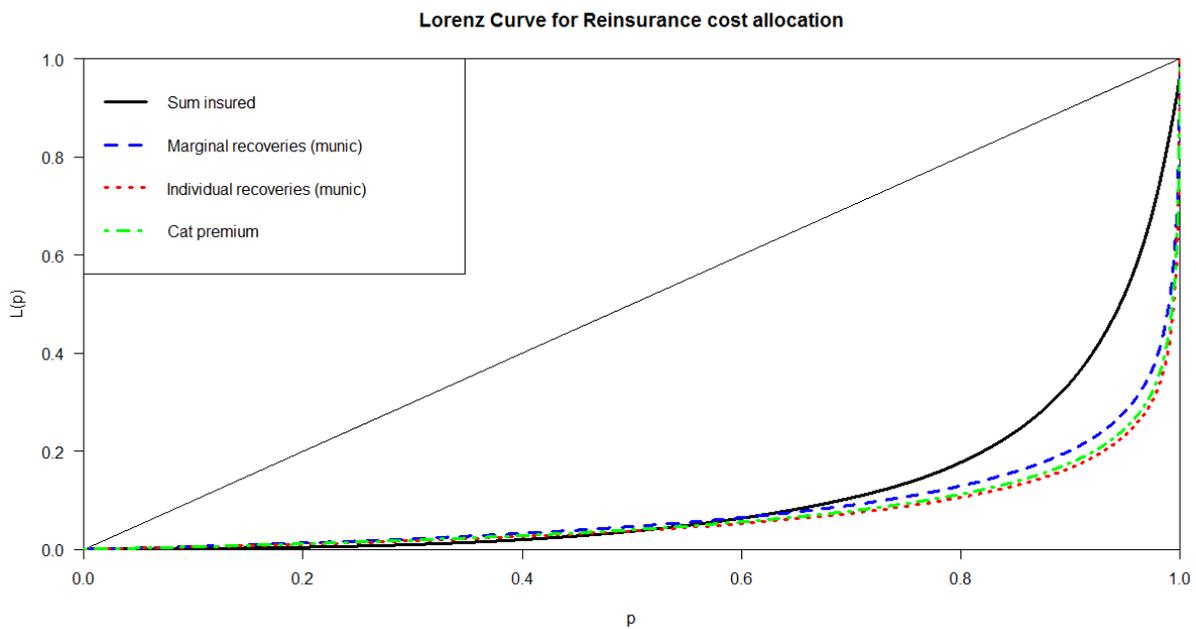
### 3. Résultats obtenus

De façon générale, on constate que les méthodes étudiées présentent davantage d'intérêt pour les grands risques (risques entreprises) que pour les petits, étant donné leur impact plus élevé sur la volatilité du portefeuille.

#### a. Segmentation du chargement de réassurance

Les résultats obtenus montrent un degré élevé de segmentation du chargement de réassurance sur les sites. L'intérêt d'une forte segmentation réside dans une meilleure sélection entre les « bons » et les « mauvais » risques du point de vue du péril considéré. Néanmoins, il n'est pas certain que la segmentation soit toujours la même selon les assureurs et leur portefeuille. Ce critère est donc à prendre selon l'objectif de l'assureur de mutualiser plus ou moins fortement le chargement de réassurance sur son portefeuille.

La méthode des bénéfices résiduels (récupérations individuelles) présente la plus forte segmentation selon l'indice de Gini. Cette méthode est basée sur l'allocation des récupérations de réassurance par site pour chaque événement, à l'aide de la contribution de chaque site à la perte totale de l'événement. On calcule ensuite la moyenne par année des récupérations par site.

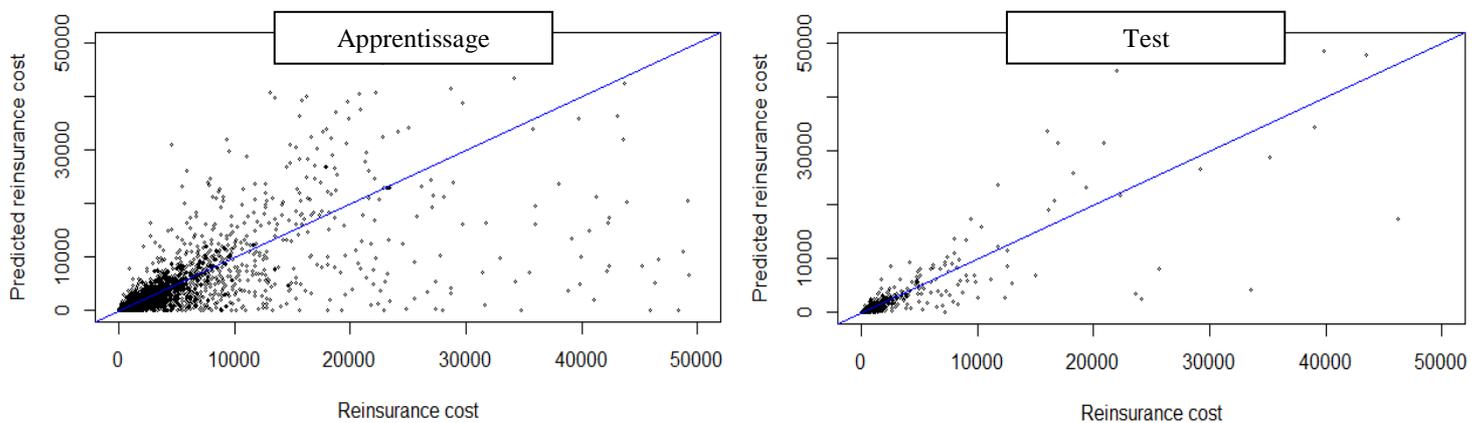


Allocation individuelle	1% des sites	10% des sites	20% des sites	50% des sites	Coefficient de Gini
Somme assurée	30,6%	69,1%	82,1%	94,6%	77,4%
Récupérations marginales	16,5%	73,1%	89,9%	99,0%	83,6%
Récupérations individuelles	19,8%	78,6%	93,0%	99,4%	86,5%
Prime Cat	18,4%	77,0%	92,2%	99,3%	85,6%

#### b. Qualité de la régression du chargement de réassurance

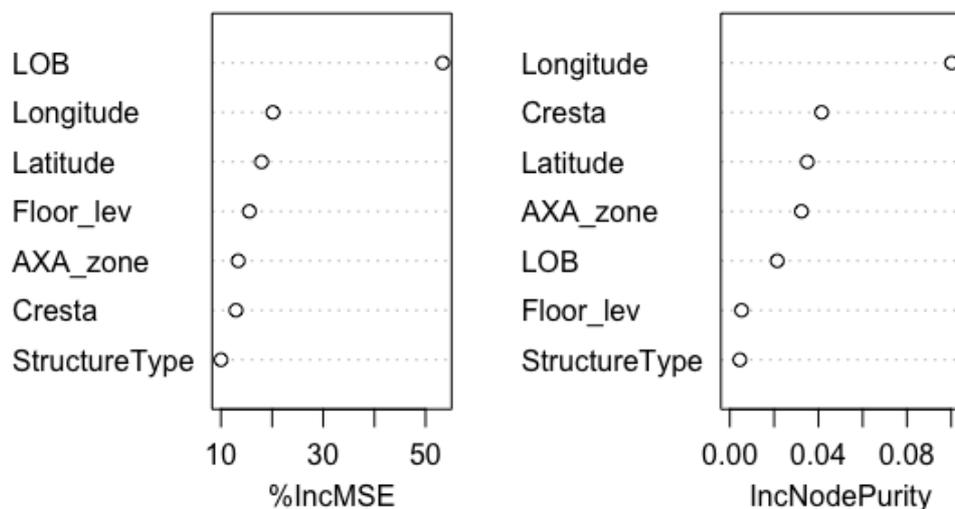
Pour la méthode des récupérations individuelles (bénéfices résiduels), on propose ensuite une méthode de prédiction qui permette d'approcher de façon satisfaisante le chargement de réassurance par site. La méthode la plus satisfaisante est la forêt aléatoire. Grossièrement, il s'agit d'une moyenne de nombreux arbres de régression.

Cette méthode permet d'expliquer 64% de la variance totale, ce qui est une approximation correcte de la valeur modélisée. Un contrôle visuel permet en outre de confirmer cette méthode.



**Figure 2 :** Graphes des coûts de réassurance prédits en fonction des coûts à prédire. A gauche, l'échantillon représenté est celui d'apprentissage. A droite, il s'agit de l'échantillon de test.

Les variables les plus importantes sont les indicateurs géographiques (latitude, longitude, AXA\_zone, Cresta) et la ligne de métier (LOB). La structure et le nombre d'étages viennent ensuite.



**Figure 3 :**

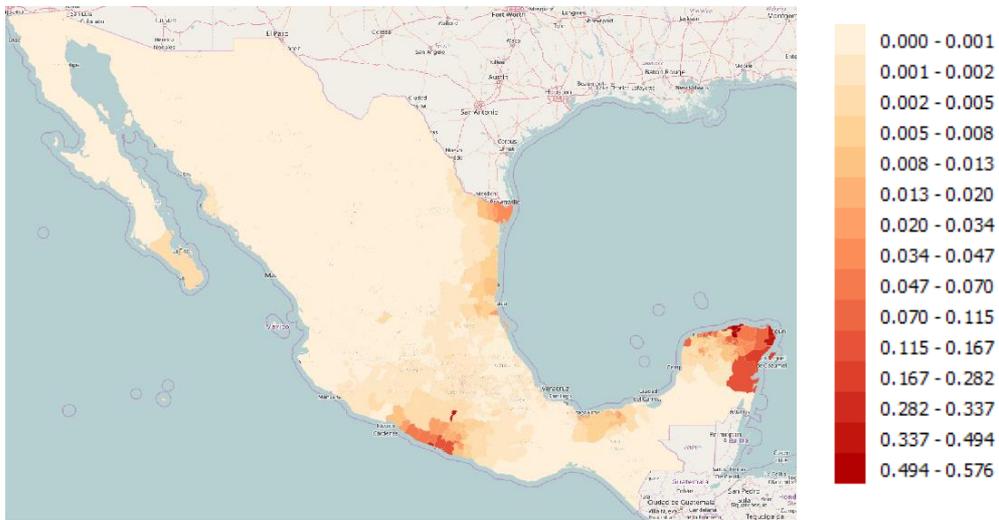
Classement des variables explicatives par ordre d'importance dans les arbres de régression. Le graphe de gauche présente l'augmentation d'erreur quadratique moyenne associée à la permutation aléatoire des modalités de chaque variable. Le graphe de droite présente le gain de pureté, mesuré en somme du carré des résidus, associé à chaque variable prédictive.

### c. Coût du capital

La méthode d'allocation du coût du capital repose sur l'allocation des pertes au-delà du montant correspondant à l'événement deux-centenaire sur le portefeuille.

Bien que la logique règlementaire veuille que l'on se base uniquement sur le quantile à 200ans, cette méthode permet de réduire un peu la volatilité liée à la définition d'un quantile. De plus, en raison de la segmentation très élevée d'une allocation site par site (indice de Gini à 99,3%),

il est proposé de répartir ce coût sur la somme assurée des sites assurés par municipalité afin d'éviter des effets « gros site » trop importants.



**Figure 4 :** Carte du coût du capital proposé pour la tarification, en taux de somme assurée. Les deux zones les plus sensibles correspondent aux zones de plus forte sévérité de l'aléa des ouragans. Légende en % somme assurée.

## Executive summary

### 1. Context and objectives

The Solvency 2 directive (2009/138/EC) entry into force on the 1st of January 2016 puts pressure on insurers to respect several quantitative measures. Among them, a solvency capital must be held to cover the company against the highest possible loss that occurs once every 200 years. Whatever insurance branch is considered, this solvency capital requirement takes into account natural catastrophes. Therefore, these have been increasingly studied and stochastic models have been developed by specialized companies in the 1990s.

These stochastic models produce a loss distribution. Not only do they allow extracting the 200 years quantile of insurer's portfolio loss, but it also gives the full loss statistical distribution. Therefore, insurer is able to use these results on the full underwriting value chain, from local underwriting to reinsurance and cost of capital.

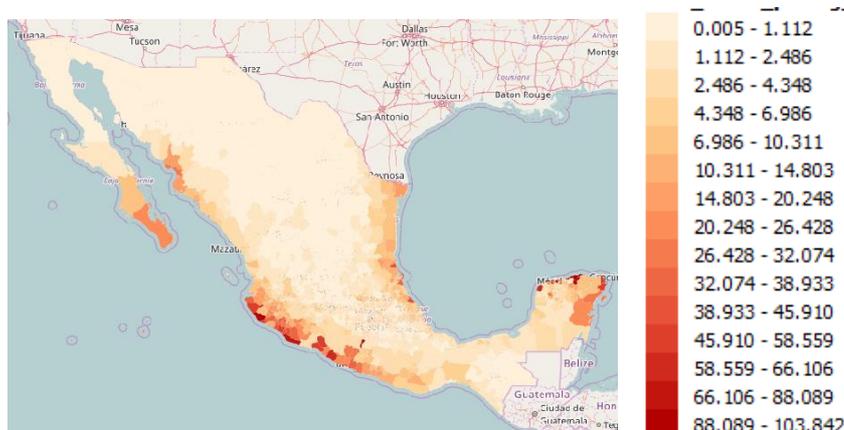
Using stochastic models on a site per site basis, one can derive a pricing (Cat pricing) for natural catastrophes and integrate it into its classical pricing. In order to propose a full pricing that would not only integrate individual risk but also interaction of this site with the rest of the portfolio (accumulation, capital requirement), this thesis aims at proposing a methodology to allocate both costs of reinsurance and capital within site per site Cat pricing.

### 2. Solving approach

The study is focusing on the Mexican insurance portfolio for a European insurer, for hurricane risk only. The hurricane model that has been used comes from AXA internal model.

In order to propose an allocation methodology, Cat model results have been analysed and used to establish the Cat pricing. Different methodologies of cost allocation have been compared and tested before looking for a prediction algorithm on modelled results. This prediction aims at avoiding to re-run the whole modelling chain for each new underwritten site.

#### a. Cat pricing



**Figure 5:** Map of the Cat pricing produced with the stochastic model result.

Figures are based on relative scale compared to a maximal value at 100. They are averaged per municipality.

Cat pricing is computed based on hurricane modelled losses per site and is averaged over modelled years. Pricing results underline the various relativities used in the model depending on variables that have an impact on vulnerability of each site to hurricane. For example, these variables can be geolocation, occupancy type, structure type of a building or its number of stories.

This methodology also shows the geographical segmentation of different risk level zones (cf. Figure 5)

#### b. Allocation methodologies comparison

Different allocation methods have been studied based on a theoretical framework related to the cooperative game theory. More precisely, the proportional method, the marginal method, the residual profits, the method of Shapley and of the nucleolus has been compared.

If the methods of Shapley and the nucleolus are more satisfactory from a theoretical point of view, the technical difficulty of their implementation leads to select between two methods: the residual income method and the marginal method.

The first method has an allocation logic which is more adapted to a stock of pre-existing sites in a given portfolio, while the second method rather corresponds to the introduction of a new site assuming the existing portfolio.

In addition to the selection of allocation methods that will be more deeply studied, it is also important to choose the metrics that will be used in these methods. For example, the simplest method of allocating the cost of reinsurance would be to use the modeled average loss due to hurricanes at each site. This method does not take into account the fact that some sites are more likely to play reinsurance than others, regardless of their average loss.

The two metrics that were selected are:

- the impact on vision model in reinsurance recoveries for the allocation of the cost of reinsurance
- the TVaR of losses beyond the 200 years loss quantile for the cost of capital

#### c. Regression methodologies

Unlike the average Cat loss (Cat pricing) modelled site per site, reinsurance and capital cost allocation need to model the entire portfolio all over again at each portfolio update. Due to the complexity of modelling chain for a total portfolio, this paper suggests a methodology to predict reinsurance costs results based on stochastic model in order to avoid the full calculation for each new underwritten site.

After testing the generalized linear models, and a kind of simple geographical regression, a random forest method is presented to predict the reinsurance cost associated to each site.

However, this type of methodology is not proposed to allocate the cost of capital due to its higher volatility and its strong segmentation.

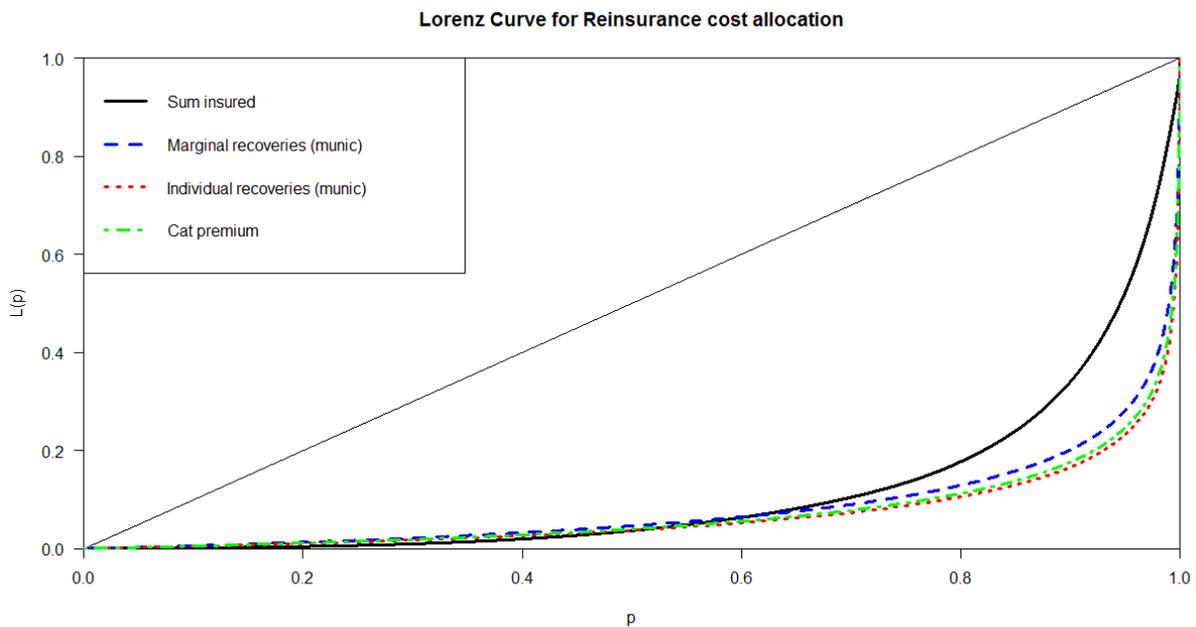
On the contrary, this paper provides a method of cost of capital allocation at a geographical level that is wider than site by site, that is to say the municipality (comparable to postcode).

### 3. Results

Generally speaking, the results presented in this thesis are more interesting for large risks (commercial) than for small ones as their impact on the portfolio volatility is stronger.

#### a. Reinsurance cost segmentation

The results show a high degree of segmentation in the cost of reinsurance site per site. The point of a strong segmentation relies in a better selection of “good” risks against “bad” risks for a given peril. However, that segmentation may not always be the same for every insurer and every portfolio as it depends on the portfolio structure. That criterion is therefore to be used depending on whether the insurer wants to strongly mutualize reinsurance cost on the portfolio or not. The residual profits method (individual recoveries) has the best segmentation according to the Gini index. This method is based on the allocation of reinsurance recoveries per site for each event, with the contribution of each site to the total loss of the event. Then the average annual recoveries per site is calculated and used as the allocation key.

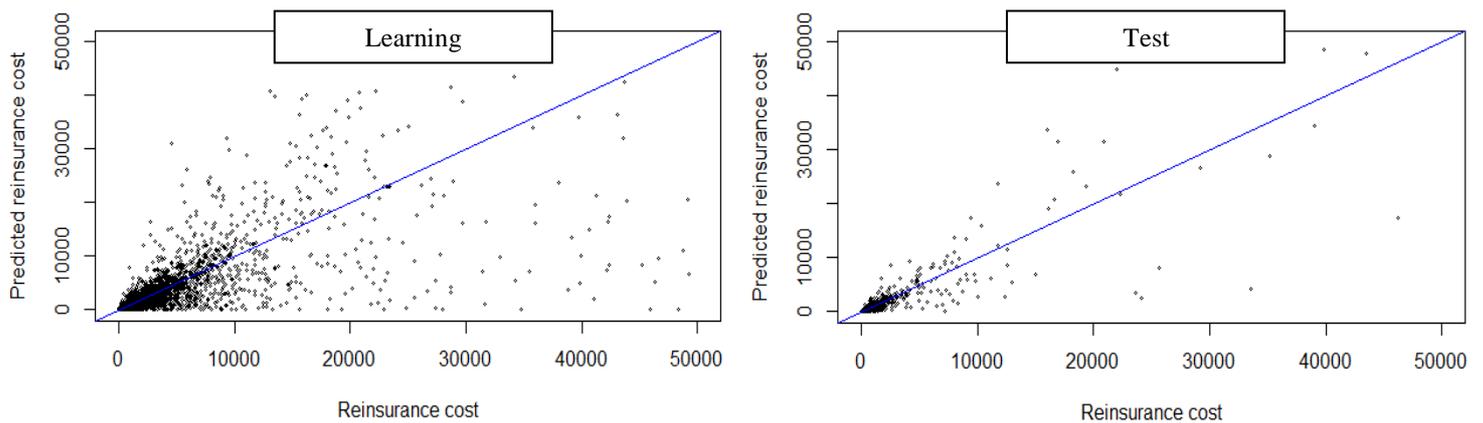


Individual allocation	1% of sites	10% of sites	20% of sites	50% of sites	Gini coefficient
Sum insured	30,6%	69,1%	82,1%	94,6%	77,4%
Marginal recoveries	16,5%	73,1%	89,9%	99,0%	83,6%
Individual recoveries	19,8%	78,6%	93,0%	99,4%	86,5%
Cat premium	18,4%	77,0%	92,2%	99,3%	85,6%

#### b. Reinsurance cost prediction quality

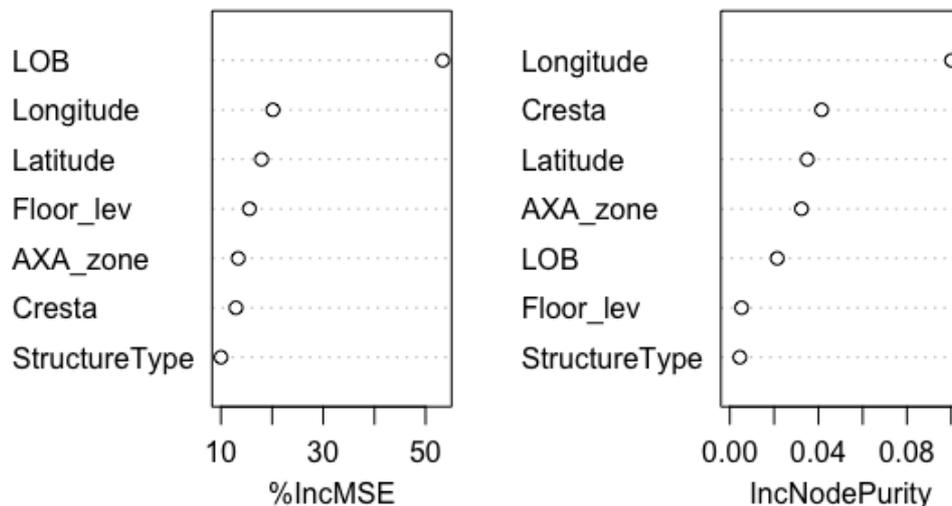
For the individual recoveries method (residual profits), different prediction methods have been tested to propose a good approach site by reinsurance cost. The most satisfactory method is the random forest. Roughly explained, it is an average of many regression trees.

This method can explain 64% of the total variance, which is a good approximation of the modeled results. Furthermore, a visual inspection seems to confirm this method.



**Figure 6 :** graphs of predicted reinsurance costs over costs to predict. On the left, the data used is the learning sample. On the right, it is the test sample.

The most important variables are geographical parameters (latitude, longitude, AXA\_zone, Cresta) and the line of business. The structure and the number of stages come next.



**Figure 7 :**

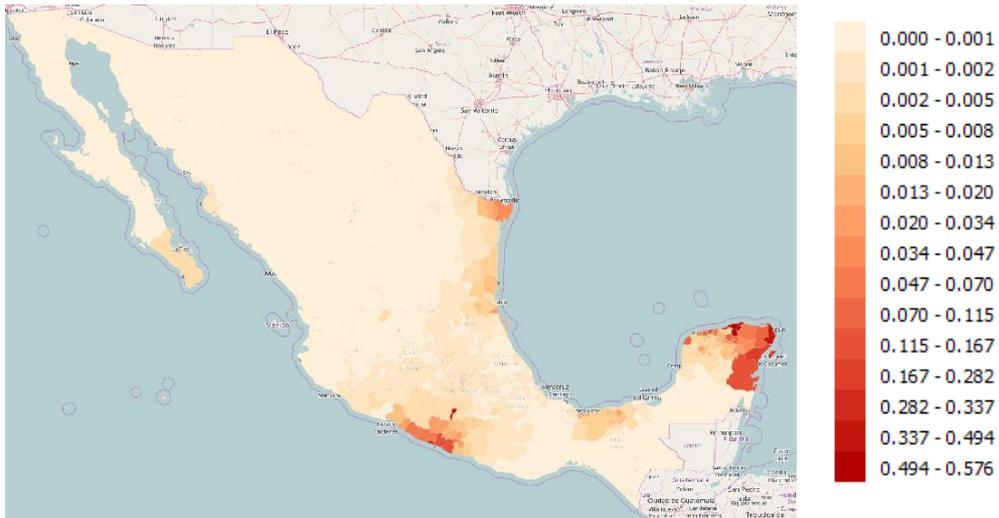
Explanatory variables ranking in order of importance for the regression trees. The left graph shows the MSE increase associated with the random permutation of the terms of each variable. The right graph shows the gain of purity, measured in square sum of residuals associated with each predictor variable.

c. Cost of capital

The method of cost of capital allocation is based on the allocation of losses in excess of the loss amount corresponding to the event occurring on average once every 200 years on the portfolio.

Although the regulatory logic would expect the methodology to be based on the quantile at 200 years, this TVaR method reduces the volatility related to the definition of a quantile.

In addition, due to the high segmentation of the allocation based on a site per site approach (Gini index 99.3%), it is proposed to allocate the cost of the insured amount of the sites provided by municipality to avoid effects of "big site" (big part of the cost allocated to one or two sites only).



**Figure 8 :** Map of the proposed cost of capital pricing, in sum insured rate. The two most sensitive areas correspond to zones of higher severity of the hazard of hurricanes. The figures are shown in % of sum insured.

## **Remerciements**

Je remercie le directeur du département Actuariat, IT et Reporting, Amaury Dufétel, pour son soutien dans mon projet de poursuivre la formation du CNAM.

Je remercie aussi mon manager Simon Blaquière et mon collègue Robin Chiche pour m'avoir guidé et donné du temps sur ce mémoire.

Je remercie enfin l'ensemble de l'équipe actuariat réassurance d'AXA Global P&C pour les échanges, conseils techniques et relectures, ainsi que pour l'excellente atmosphère qui y règne.

<b>INTRODUCTION .....</b>	<b>18</b>
<b>I- PRESENTATION D'AXA GLOBAL P&amp;C .....</b>	<b>20</b>
<b>a. Le groupe AXA .....</b>	<b>20</b>
<b>b. AXA Global P&amp;C .....</b>	<b>20</b>
<b>II- MODELES CAT, TARIFICATION ET CADRE DE L'ETUDE .....</b>	<b>21</b>
<b>a. Modèles Cat .....</b>	<b>21</b>
i. Structure d'un modèle Cat.....	21
ii. Résultat, intérêts et limites des modèles Cat .....	28
<b>b. Tarification Cat .....</b>	<b>32</b>
i. Rappels de tarification en assurance non-vie .....	32
ii. Tarification Cat.....	32
<b>c. Définitions et cadre de l'étude.....</b>	<b>36</b>
i. Notions de réassurance .....	36
ii. Une approche spécifique aux traités XL .....	37
iii. Introduction du chargement de réassurance dans la tarification Cat.....	40
iv. Principes théoriques d'allocation de coût ( <i>cost allocation</i> ) .....	41
v. Méthodes d'allocation de coût considérées.....	46
<b>III- ALLOCATION DU CHARGEMENT DE REASSURANCE .....</b>	<b>56</b>
<b>a. Méthode proportionnelle .....</b>	<b>56</b>
i. Présentation de la méthode.....	56
ii. Analyse des résultats et conclusions .....	56
<b>b. Méthode des bénéficiaires résiduels.....</b>	<b>59</b>
i. Présentation et analyse des résultats d'allocation.....	59
ii. Construction de modèles prédictifs simples du chargement de réassurance.....	61
iii. Construction d'un modèle de prédiction du chargement de réassurance de type arbre de régression.....	64
<b>c. Méthode du coût marginal .....</b>	<b>69</b>
i. Présentation .....	69
ii. Analyse des résultats .....	69
<b>d. Comparaison de la segmentation des méthodes principales .....</b>	<b>72</b>
i. Définitions .....	72
ii. Comparaisons individuelles .....	73
iii. Comparaisons à l'échelle des municipalités.....	74
iv. Comparaisons à l'échelle des zones CRESTA.....	74
v. Méthodes à retenir .....	75
<b>e. Valeur de Shapley et méthode du nucléole : première approche.....</b>	<b>77</b>
i. Principe.....	77
ii. Résultats .....	77
<b>f. Prise en compte de l'extension potentielle de la portée du traité à la souscription d'un risque .....</b>	<b>79</b>
i. Principe et définitions.....	79
ii. Effet marginal par site et par municipalité .....	79

iii. Capacité restante par municipalité .....	80
<b>IV- EXTENSION AU COUT DU CAPITAL .....</b>	<b>82</b>
<b>a. Notion de coût du capital .....</b>	<b>82</b>
i. Le capital règlementaire dans l'assurance.....	82
ii. Le coût du capital .....	83
<b>b. Application de l'allocation du CoC dans la tarification Cat .....</b>	<b>85</b>
i. Choix d'un indicateur .....	85
ii. Allocation du CoC.....	85
<b>V- LIMITES ET PISTES D'AMELIORATION DE L'ETUDE .....</b>	<b>88</b>
<b>a. Tarification Cat .....</b>	<b>88</b>
i. Les limites de la segmentation .....	88
ii. La dépendance à une unique vision du risque .....	89
iii. Limitation du périmètre de péril.....	89
<b>b. Allocation de la réassurance et coût du capital .....</b>	<b>91</b>
i. Traités spécifiques et réassurance non traditionnelle .....	91
ii. Calculs de chargement de réassurance global et extension de capacité .....	91
iii. Un approfondissement des méthodes d'allocation et de prédiction .....	92
iv. Vers un véritable outil de machine learning.....	92
<b>VI- CONCLUSION .....</b>	<b>94</b>
<b>VII- BIBLIOGRAPHIE .....</b>	<b>95</b>
<b>VIII- LISTE DES FIGURES .....</b>	<b>96</b>
<b>IX- ANNEXES .....</b>	<b>98</b>
<b>a. Rappels sur la théorie des valeurs extrêmes .....</b>	<b>98</b>
<b>b. Démonstration de la propriété sur la méthode des bénéfices résiduels.....</b>	<b>99</b>
<b>c. Démonstration du nombre de calculs pour la méthode de Shapley .....</b>	<b>100</b>
<b>d. Résultats de la construction des modèles de prédiction (GLM).....</b>	<b>101</b>
<b>e. Rappels sur les arbres de régression et de décision.....</b>	<b>102</b>

## INTRODUCTION

Les catastrophes naturelles représentent un risque d'extrême sévérité et rareté. Bien qu'imprévisible, de nombreux acteurs du marché de l'assurance, de la réassurance et plus globalement de la gestion du risque ont investi sur une meilleure compréhension, prévention et compensation de ces risques.

D'un point de vue réglementaire, la directive Solvabilité 2, entrée en vigueur au 1<sup>er</sup> janvier 2016, rend nécessaire la prise en compte des catastrophes de manière spécifique par les assureurs et réassureurs. Elle oblige notamment ces acteurs à disposer d'un montant de capital suffisant pour se couvrir contre un quantile extrême de perte (99,5%).

La mesure de ce risque s'est développée dans les années 1990 par la production de modèles stochastiques créés par des entreprises spécialisées (RMS, AIR, EQE, etc.). Ces modèles physiques ont été affinés au cours des 25 dernières années en se basant notamment sur les capacités de calcul et les données plus nombreuses de pertes assurancielles disponibles. En effet, plusieurs catastrophes importantes telles que les tempêtes Lothar et Martin (Europe, 1999), la grêle ELA (Europe, 2014), les ouragans Wilma et Katrina (Mexique et USA, 2005), Sandy (USA, 2012), Odile (Mexique, 2014), les inondations européennes de 2002 et 2013, les tremblements de terre d'Izmit (Turquie, 1999), l'Aquila (Italie, 2009), etc., ont permis d'obtenir des données conséquentes afin d'améliorer le comportement des modèles.

Plusieurs acteurs de l'assurance et de la réassurance ont d'ailleurs investi dans le développement de modèles propriétaires, éventuellement intégrés à leur modèle interne.

Aujourd'hui, la maîtrise du risque Cat est devenue un enjeu de compétitivité pour les acteurs du marché de l'assurance, non seulement d'un point de vue de leur besoin en capital et de couverture de réassurance, mais aussi d'un point de vue de la prévention et de la tarification des assurés.

En effet, ces pertes ont augmenté de façon importante sur les 30 dernières années sous les effets combinés de l'accroissement démographique, la hausse du taux de pénétration de l'assurance (notamment dans les marchés émergents), et potentiellement du changement climatique (Munich Re, 2015). Afin de mieux segmenter, d'offrir des services toujours plus importants aux assurés et de progresser sur des sujets de prévention et d'atténuation en plus des sujets d'estimation et de protection, les assureurs cherchent donc à intégrer ce risque dans leur chaîne de valeur.

Ce type d'étude est particulièrement nécessaire aux entreprises d'assurance pour lesquelles les catastrophes naturelles représentent une part non négligeable de leur sinistralité. Dans le cadre de ce mémoire, on s'intéresse à un assureur mexicain dont la charge liée aux catastrophes naturelles représente près de 60% de la prime dommage aux biens (source : AXCO, 2012) et une part encore plus élevée de la prime de réassurance dommage qu'il paye à ses réassureurs.

Pour un assureur de ce type, la première question qui se pose est alors de savoir comment ajuster la prime de l'assuré pour refléter le risque de catastrophe auquel il est exposé. Un risque situé sur les côtes sera par exemple plus vulnérable aux tempêtes qu'un risque localisé à l'intérieur des terres. La prime pure mérite donc d'être ajustée.

Une seconde question qui se pose est de s'assurer que la souscription d'un nouveau risque dans une zone dans laquelle s'accumulent déjà de nombreux assurés et qui est sujette aux catastrophes ne risque pas de générer un surcoût de réassurance ou de capital supérieur à la prime d'assurance perçue à la souscription du contrat. Le sujet se présente plus particulièrement dans le cas de grands risques (par exemple des hôtels de luxe situés sur la côte). La prime technique, qui contient souvent (au moins implicitement) un chargement pour les coûts de réassurance et du capital, mérite donc d'être ajustée pour prendre en compte ces chargements supplémentaires.

Cette seconde question fait néanmoins émerger un sujet de choix de mutualisation de la part de l'assureur. En effet, alors que la question de l'ajustement de la prime pure de l'assuré est intrinsèque à celui-ci, celle de l'intégration des chargements de réassurance et du capital dépend de paramètres extérieurs à l'assuré, comme le reste du portefeuille, le choix de réassurance acheté par l'assureur, etc. Afin de rester compétitif, l'assureur peut ne pas avoir intérêt à segmenter au maximum ses risques vis-à-vis des coûts de réassurance et du capital et à garder un certain niveau de mutualisation, selon les autres acteurs du marché.

La réponse à cette problématique consiste donc à intégrer dans la tarification classique une partie liée au risque Cat. Cela permettrait ainsi aux assureurs de mieux segmenter, en intégrant dans la prime de chaque assuré une partie supplémentaire qui serait, dans l'idéal et comme proposé par ce mémoire, composée de :

- Une prime Cat : cette prime serait calculée de façon individuelle en tenant compte des caractéristiques de l'assuré.
- Une prime de réassurance : cette prime nécessiterait de prendre en compte l'impact de l'assuré sur le chargement de réassurance global de l'assureur.
- Une prime de capital : cette prime viserait à refléter le coût du capital représenté par l'assuré pour l'assureur.

De cette manière, on pourrait imaginer que la souscription de l'assureur soit liée aux contraintes de réassurance et de capital, en plus du risque de catastrophe naturelle.

## **I- PRESENTATION D'AXA GLOBAL P&C**

### **a. Le groupe AXA**

Le groupe AXA est la conséquence d'une longue série de regroupements par fusion acquisition dont la première entreprise était l'ancienne mutuelle d'assurance de Rouen créée en 1817. A la suite d'une croissance interne au XIX<sup>e</sup> et sur la première moitié du XX<sup>e</sup> siècle, cette entreprise poursuit son développement par une croissance externe. Cette croissance, d'abord nationale puis internationale, a été accentuée sous la présidence de Claude Bébéar.

Le groupe prend le nom AXA en 1985 et exerce désormais sur trois grands secteurs d'activités : l'assurance vie, l'assurance non-vie et la gestion d'actifs.

Avec un chiffre d'affaire de 110 Mds\$ en 2015, AXA se situe au deuxième rang mondial des assureurs. Son résultat net en 2014 était de 5,3Mds€.

### **b. AXA Global P&C**

Au sein d'Axa, Axa Global Property and casualty (AGPC, désormais AXA Global Re) est responsable de la réassurance et du support technique des entités. AGPC est la plateforme de placement des traités de réassurance pour l'ensemble du Groupe AXA.

Afin de décider au mieux de la réassurance du groupe, AGPC a développé une expertise importante sur la modélisation des sinistres liés à la réassurance, notamment en matière de catastrophes. L'approfondissement des compétences en modélisation catastrophe est donc essentiel pour définir une vision du risque propre au groupe.

Par ailleurs, AGPC a aussi un rôle de définition de la stratégie dommages du groupe, à savoir la définition, l'élaboration et la mise en œuvre des éléments de stratégies ainsi que l'harmonisation des pratiques du groupe.

Enfin, AGPC est une filiale dynamique et rentable, comme le montre le résultat net de 2012 à 30,5m€, en croissance par rapport à l'année précédente.

## II- MODELES CAT, TARIFICATION ET CADRE DE L'ETUDE

Un modèle de catastrophe (Cat) vise à construire une vision du risque en construisant une distribution de pertes liées à un péril modélisé en fonction du portefeuille d'assurance donné en entrée. Il permet notamment d'élaborer :

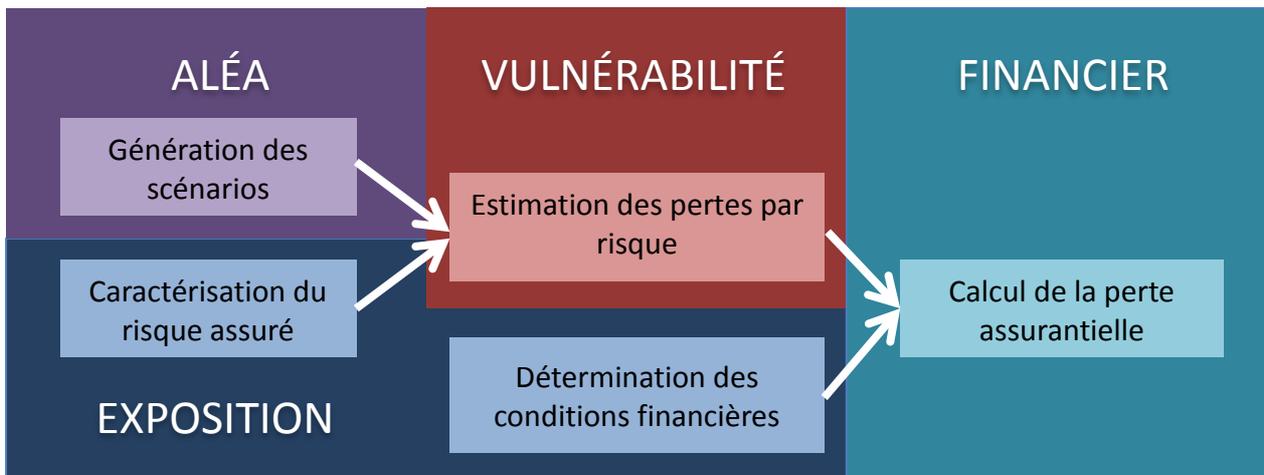
- Un tarif d'assurance (prime Cat) : en utilisant la distribution de pertes site par site, l'assureur est capable de tarifier la garantie catastrophe naturelle du péril associé - ici, les ouragans - pour l'ensemble des polices de son portefeuille (partie II.b).
- Un tarif de réassurance : en utilisant la distribution de pertes agrégées au niveau du portefeuille, le réassureur est capable de tarifier le traité de réassurance. C'est ce tarif, ou plus précisément la part qui n'est pas déjà incluse dans la tarification d'assurance (chargement), que nous cherchons à allouer dans le III via les méthodes de la partie II.c.

### a. Modèles Cat

Avant d'évoquer la tarification d'assurance et de réassurance, il convient de présenter d'abord la construction d'un modèle de catastrophe naturelle, ou modèle Cat.

#### i. Structure d'un modèle Cat

Un modèle Cat se décompose généralement en quatre modules : le module exposition, le module aléa, le module vulnérabilité et le module financier.



#### 1. *Module exposition*

Le module exposition est constitué de ce qui définit les risques présents dans le portefeuille. Ces risques sont caractérisés par différents éléments comme la ligne de métier, la localisation, l'altitude, l'âge, l'occupation du risque, le type de structure, le type de toit (si existant), la valeur assurée totale du risque, le numéro d'étage, etc.

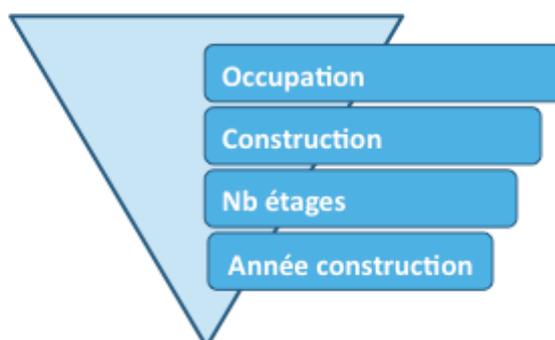
Ces paramètres sont de criticité variable pour le modèle selon la précision et le péril de celui-ci. Ainsi, pour un modèle d'inondation, il est critique d'avoir une exposition précisément localisée à la coordonnée. En effet, une précision de l'ordre de quelques mètres est nécessaire pour faire la distinction entre les zones inondées ou non. Dans le cas d'un modèle séisme, les zones de dommage sont bien moins précises que pour l'inondation et une précision de l'ordre du kilomètre peut suffire.

Par ailleurs, certaines variables ont plus d'importance sur la vulnérabilité du risque que d'autres. Par exemple, le type de toit a plus d'importance pour les tempêtes que pour les inondations. On imagine bien qu'un toit en paille sera plus vulnérable aux vents violents qu'un toit en tuiles. A l'inverse, pour une inondation, il est plus important de savoir si un risque est situé en étage ou en rez-de-chaussée que pour une tempête.

Dans le cas d'un modèle d'ouragan, on cherchera en priorité à avoir une information d'un niveau géographique au niveau code postal ou plus précis. En effet, les ouragans présentent des zones d'aléa très sensibles à la localisation géographique (comme on le verra par la suite).

La hiérarchie des autres variables par ordre de criticité peut être vue ci-dessous :

*Par ordre décroissant d'importance :*



(Source : Aon Benfield)

Dans le cadre de l'étude, on s'intéresse à un portefeuille mexicain d'un assureur pour la garantie dommage aux biens, comprenant notamment une couverture contre les catastrophes naturelles.

Le portefeuille contient 108829 sites qui seront considérés comme des localisations uniques, sur quatre lignes de métier différents (résidentiel, industriel, commercial, municipal) pour un total de près d'une centaine de milliards de dollars d'actifs assurés. Ces actifs représentent à 55% la somme assurée des bâtiments, à 30% du contenu et à 15% de potentielles pertes d'exploitation (lignes de métier commerciale et industrielle principalement). La répartition par ligne de métier est la suivante :

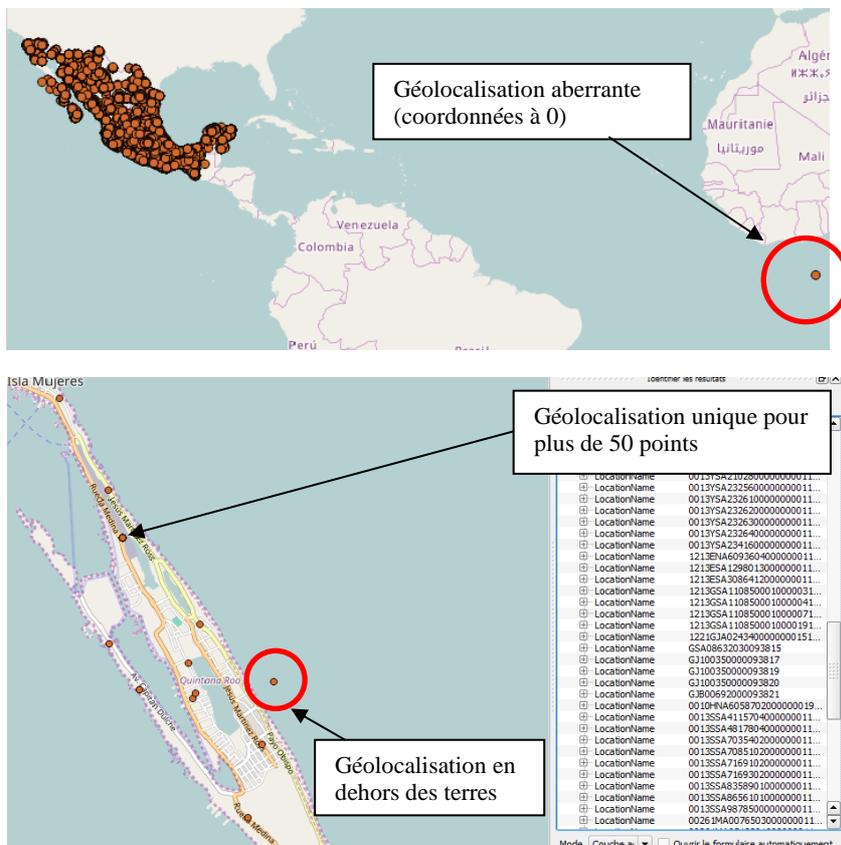
Ligne de métier	Somme assurée (MdsUSD)	% Total	Nombre de sites	% Total
<b>COMMERCIAL</b>	48,1	56%	38 623	35%
<b>INDUSTRIEL</b>	10,5	12%	6 241	6%
<b>MUNICIPAL</b>	4,9	6%	5 940	5%
<b>RESIDENTIEL</b>	22,0	26%	58 025	53%
<b>Total général</b>	85,5	100%	108 829	100%

Le portefeuille est géocodé à 100%. Cependant, la présence d'une trentaine de sites clairement aberrants, de sites géocodés à l'extérieur des terres ou encore des agrégations de sites géocodés au même endroit incite à la prudence quant à la qualité de la géolocalisation (cf Figure 9).

En ce qui concerne les variables les plus critiques pour la modélisation des ouragans, le portefeuille contient partiellement leur information :

- Occupation : 34% de la somme assurée totale.
- Construction : 100% de la somme assurée totale.
- Nombre d'étages : 100% de la somme assurée totale.
- Année de construction : 0% de la somme assurée totale.

En comparant la répartition du type de construction avec des données marché, on constate que le portefeuille semble représentatif de la répartition marché.

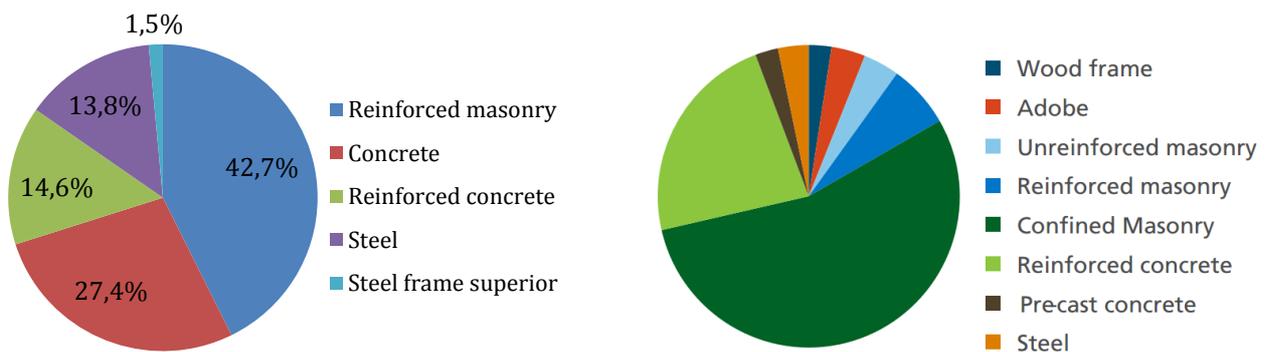


**Figure 9 :** carte des sites géolocalisés à l'aide de leurs coordonnées.

Certains points présentent des anomalies :

- coordonnées à 0 (haut)
- localisation dans la mer (bas)
- localisation de nombreux sites au même endroit (bas)

### Part de somme assurée totale



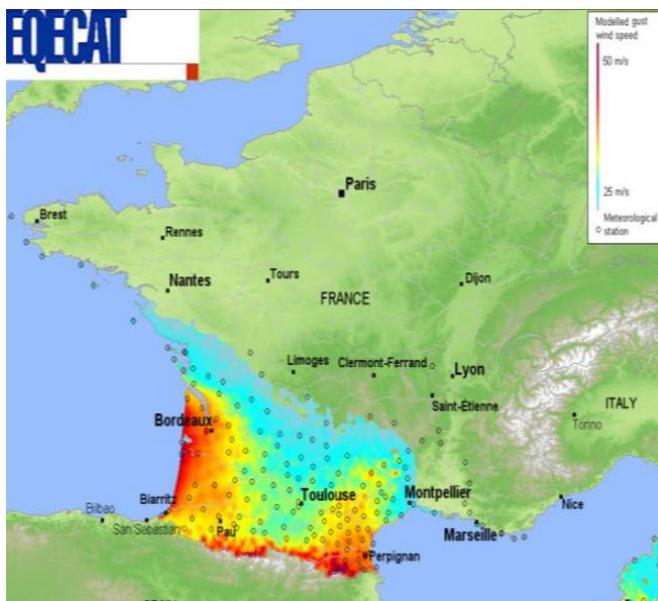
**Figure 10 :** comparaison de la répartition du type de structure entre le portefeuille utilisé pour la modélisation (gauche) et un portefeuille marché (droite, source : AIR, 2016). On constate que les bâtiments en maçonnerie (*masonry*) sont les plus présents dans les deux cas, avec une surreprésentation des bâtiments métalliques (*steel* et *steel frame superior*) dans le portefeuille utilisé par rapport au portefeuille marché. Ceci pourrait s'expliquer par une part plus importante de la ligne de métier industrielle dans le portefeuille étudié par rapport aux autres portefeuilles du marché. A l'inverse, les bâtiments en béton (*concrete*) ou en adobe sont sous-représentés.

Au vu de la qualité de l'information du portefeuille et en considérant la hiérarchie de criticité vérifiée, on peut raisonnablement considérer les résultats du modèle comme acceptables.

## 2. Module aléa

Le module aléa sert à construire la représentation physique du péril modélisé. Il produit des scénarii d'aléa à chaque simulation, une simulation représentant en général une année d'événements Cat. Chaque scénario d'aléa prend le plus souvent la forme d'une empreinte physique dont la variable est une grandeur physique à l'origine des pertes économiques (et assurantielles). On obtient ainsi un catalogue d'événements possibles, calibrés sur des données historiques, qui cherchent à extrapoler ce que pourraient être des événements extrêmes. Le modèle d'ouragan utilisé pour l'étude génère par exemple 10 000 années pendant lesquelles surviennent plus de 40 000 événements pouvant causer des pertes.

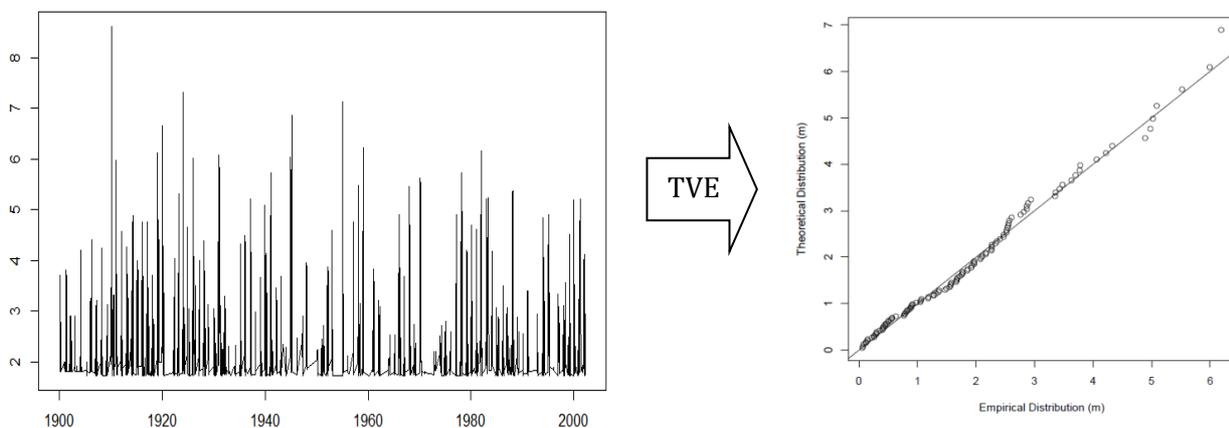
Pour l'inondation, un scénario d'aléa prend la forme d'une empreinte d'inondation dont chaque point est affecté d'une hauteur d'eau différente. Pour une tempête, un scénario est représenté par une empreinte (ou trajectoire) dont chaque point est affecté d'une vitesse de vent. La figure ci-dessous montre une vision du risque tempête pour un événement historique (Klaus) survenu en 2009.



**Figure 11** : exemple de carte du risque tempête d'un point de vue de l'aléa selon un modèle marché (EQE), pour l'événement Klaus survenu le 24 janvier 2009. (source : EQECAT, 2013)

La génération d'un scénario d'aléa peut se décomposer en deux grandes étapes :

- **Tirage de valeurs extrêmes** : il s'agit d'un tirage aléatoire de lois de distribution de grandeurs physiques en des points spécifiques dont on connaît l'historique et sur lesquels on a calibré au préalable des distributions de valeurs extrêmes.



**Figure 12** : Fit d'une distribution de valeurs extrêmes sur les maxima d'une série temporelle de hauteur d'eau pour une station de mesure hydrologique. A gauche, on observe la série temporelle de hauteur d'eau en m entre 1900 et 2005. A droite, on observe le diagramme quantile-quantile (QQ plot) de la distribution de valeurs extrême théorique calibrée à partir des données historiques. (source : étude interne AGPC)

La calibration se base sur des séries temporelles d'une grandeur physique et sur la théorie des valeurs extrêmes (TVE) qui propose un cadre statistique pour la construction de distribution de valeurs extrapolées (cf. Annexe a.).

Pour un ouragan, on génère des vitesses de vent à partir de stations de mesure. Pour une inondation, on génère des hauteurs d'eau en divers points de fleuves (cf. exemple ci-dessus).

- **Propagation de l'événement** : les valeurs tirées sont utilisées comme input d'un modèle de propagation, généralement basé sur la résolution d'équations aux dérivées partielles de modèle physique.

Dans l'idéal, la calibration de ces modèles de propagation est basée sur des événements historiques dont on connaît les valeurs physiques d'input et l'empreinte finale de l'événement. Par exemple, les trajectoires historiques des tempêtes et l'évolution de la vitesse de vent le long de ces trajectoires sont autant d'informations utiles à la calibration des scénarii « possibles » de propagation modélisée.

Pour une inondation, l'eau qui déborderait du lit principal d'un cours d'eau est propagée à l'aide des équations de saint venant. Pour un séisme, on peut utiliser les équations de Navier-Stokes pour représenter la propagation de l'onde sismique à partir de l'épicentre.

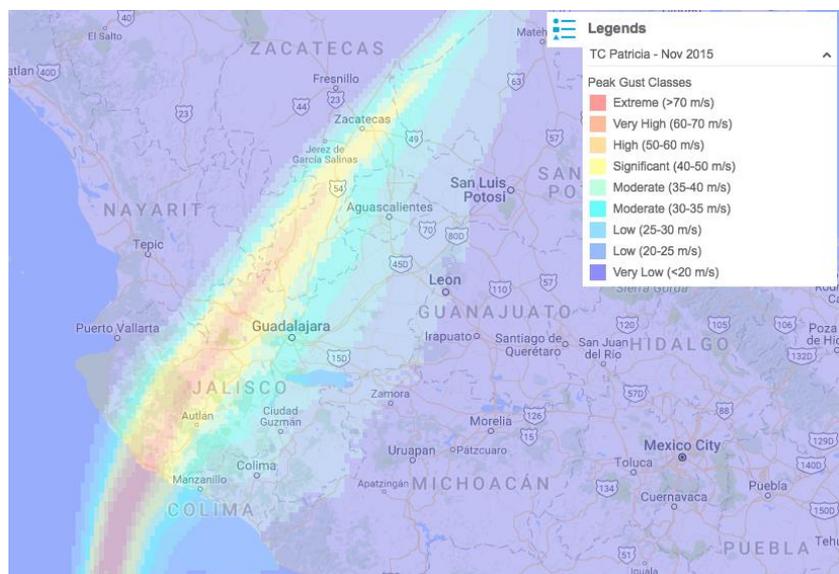
A l'issue de la propagation, on a une valeur physique d'aléa en tout point de la zone de propagation et donc en tout point du portefeuille de l'entité d'assurance modélisée. Chaque site est ainsi associé à une grandeur physique, ou intensité d'aléa, par événement.

**Figure 13 :**

Propagation de l'ouragan Patricia sur la période du 20 au 24 Octobre 2015.

Les niveaux de couleur indiquent les vitesses de vent liées à l'événement.

(Source : SwissRe CATNET)



Eventuellement, certains modèles effectuent aussi le calcul d'aléas secondaires associés à l'aléa principal. Par exemple, il peut s'agir d'incendies qui se déclencheraient à la suite d'un tremblement de terre (*fire following*), ou encore d'inondation liées au déclenchement non souhaité des système d'extinction des incendies (*sprinklers leakage*).

Dans le modèle qui est utilisé pour l'étude, les inondations liées aux pluies torrentielles causées par les ouragans sont modélisées au même titre que les vitesses de vent.

### 3. Module vulnérabilité

Le module vulnérabilité a pour objectif de calculer des pertes brutes de toute condition financière de police d'assurance locale. Il permet de passer des valeurs physiques à chaque site à une perte financière brute.

Une des méthodes les plus pratiquées consiste à construire des courbes de vulnérabilité qui donnent un taux de destruction du site en fonction de l'intensité de l'aléa (cf. Figure 14).

Ces courbes de vulnérabilité sont construites à l'aide de bases de données de sinistres pour lesquels on connaît l'intensité d'aléa qui a généré chaque sinistre.

On imagine bien qu'avec 20 cm d'eau pour une inondation, les dégâts ne seront pas les mêmes qu'avec 4 m d'eau. Le contenu affecté au rez-de-chaussée sera totalement sous l'eau avec 4 m, tandis que le premier étage restera au sec avec 20 cm.

Dans l'idéal, les courbes de vulnérabilité sont construites avec une segmentation assez fine, basée sur les paramètres d'aléa mais aussi d'exposition pour mieux capter leur influence sur les pertes.

Par exemple, un toit à quatre pans résistera mieux à des vents extrêmes qu'un toit à deux pans (New Jersey Institute of Technology, 2007). D'autres exemples de vulnérabilité relative sont disponibles ci-dessous pour un modèle d'inondation disponible sur le marché.

Figure 14 :

Exemple de courbe de vulnérabilité.  
Le taux de destruction est calculé en fonction de l'intensité de la mesure utilisée pour quantifier l'aléa.

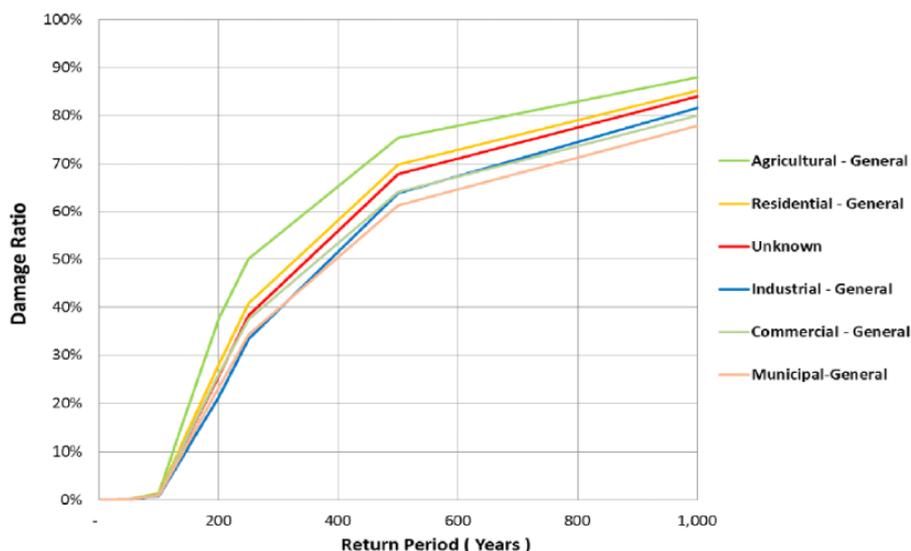
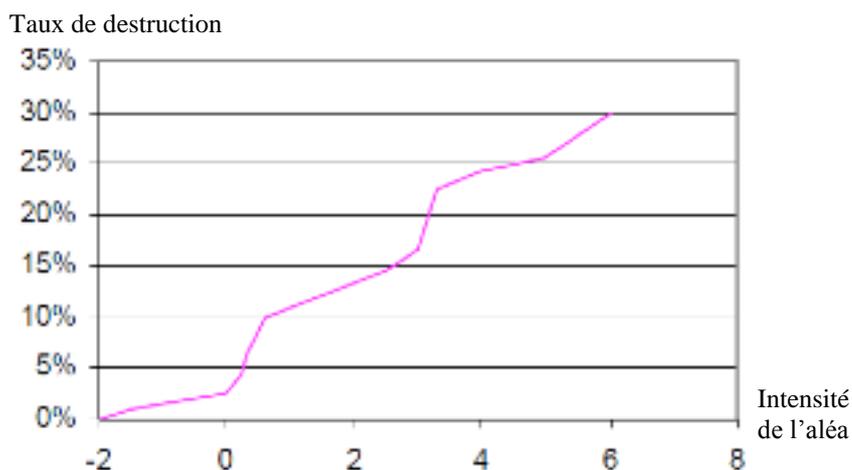


Figure 15 :

Exemple de vulnérabilité relative en fonction de la ligne de métier. (source : Willis Re)

On constate que les sites industriels résistent mieux à l'inondation que les sites résidentiels d'après le modèle.

Le modèle de l'étude utilise des courbes de vulnérabilité pour plusieurs types de constructions et d'occupations différentes. Le nombre d'étages et l'âge du bâtiment sont aussi pris en compte, ce qui est une pratique courante sur ce type de modèle.

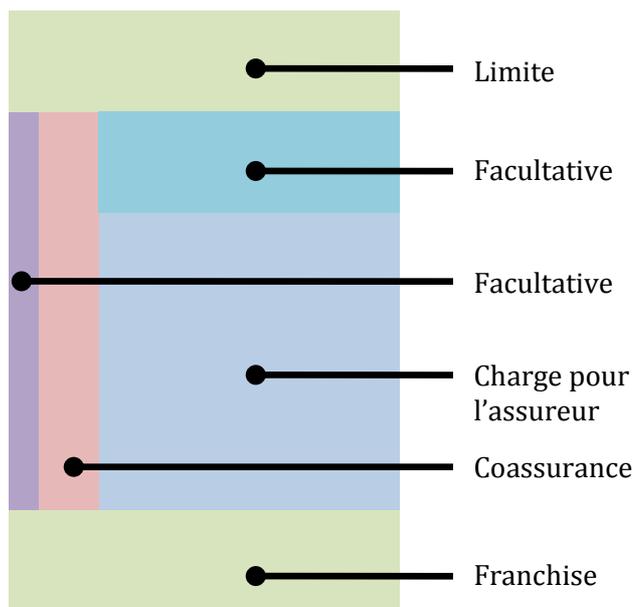
#### 4. Module financier

Comme son nom l'indique, le module financier effectue l'application des conditions financières des polices d'assurance à chaque perte brute. Ces conditions financières s'appliquent en général au niveau de chaque site mais certains cas peuvent complexifier le calcul.

En effet, dans certaines situations, le site peut être couvert avec d'autres au sein d'une même police (multi site). La police peut aussi être partagée entre plusieurs assureurs (coassurance) et le site peut aussi faire l'objet d'une réassurance facultative, proportionnelle ou non.

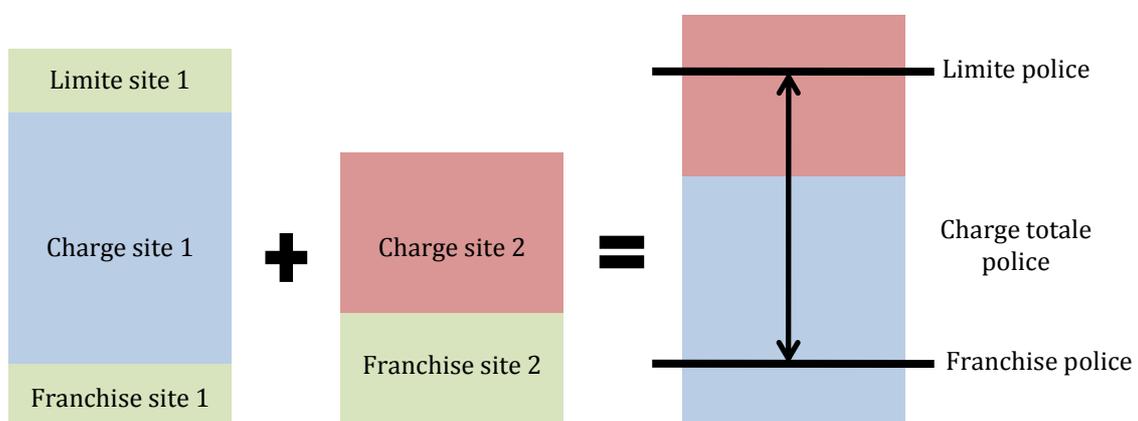
Dans le cas simple, on applique uniquement la part de la police assurée par l'entité d'assurance modélisée (retrait de la coassurance) et on applique franchise et limite au site ainsi que la réassurance facultative si elle existe (cf. Figure 16).

Dans le cas d'une police multi site, on calcule les pertes site à site que l'on agrège ensuite à la police pour appliquer les conditions au niveau police.



**Figure 16 :**

Application des conditions financières à une perte brute d'assurance.

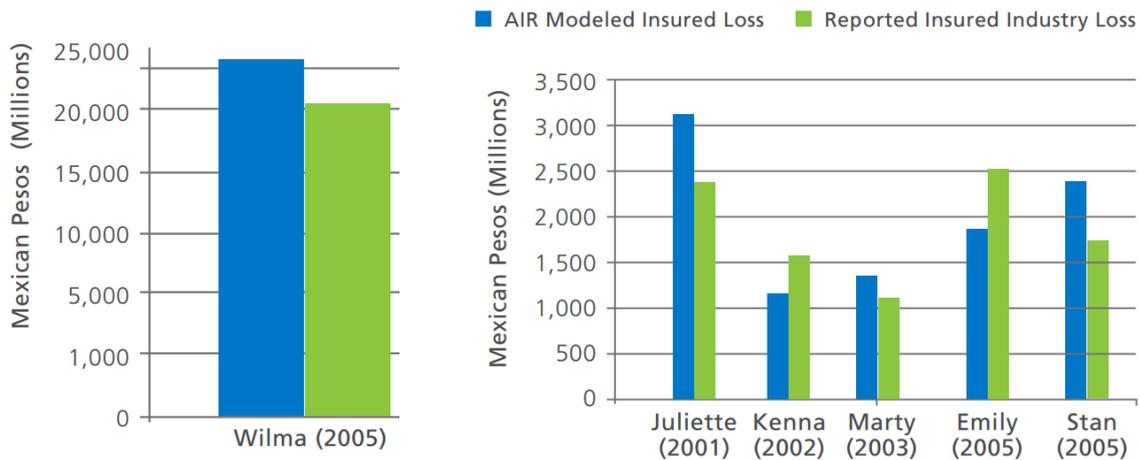


**Figure 17 :** application des conditions financières à une perte brute d'assurance sur une police multi-site, pour deux sites affectés.

Dans l'exemple ci-dessus, on considère des pertes brutes sur deux sites appartenant à une même police, avec des sous-limites et franchises par site.

Dans le portefeuille utilisé, moins de 5% des polices présentent des caractéristiques spécifiques.

Dans le modèle ouragan étudié, les résultats du modèle ont été calibrés et vérifiés avec les montants de pertes des événements survenus dans le passé. Cela permet de valider la qualité du modèle pour reproduire les événements passés et donc de supposer que les événements extrapolés prédisent fidèlement les pertes potentielles.



**Figure 18** : Comparaison des pertes modélisées (en bleu) pour un portefeuille marché par rapport aux pertes historiques (en vert) lors d'événements survenus dans le passé. Source : AIR, 2016

## ii. Résultat, intérêts et limites des modèles Cat

### 1. *Résultat d'un modèle Cat*

Le résultat d'un modèle Cat consiste en une distribution de pertes, généralement présentée sous la forme d'une courbe OEP (Occurrence exceeding probability). L'OEP représente la probabilité annuelle de dépasser un certain montant de perte pour un seul événement. L'AEP représente la probabilité annuelle de dépasser un certain montant de pertes. En notant  $F_C$  la fonction de répartition des pertes cumulées et  $F_M$  la fonction de répartition de la perte maximale, et T la période de retour, on a :

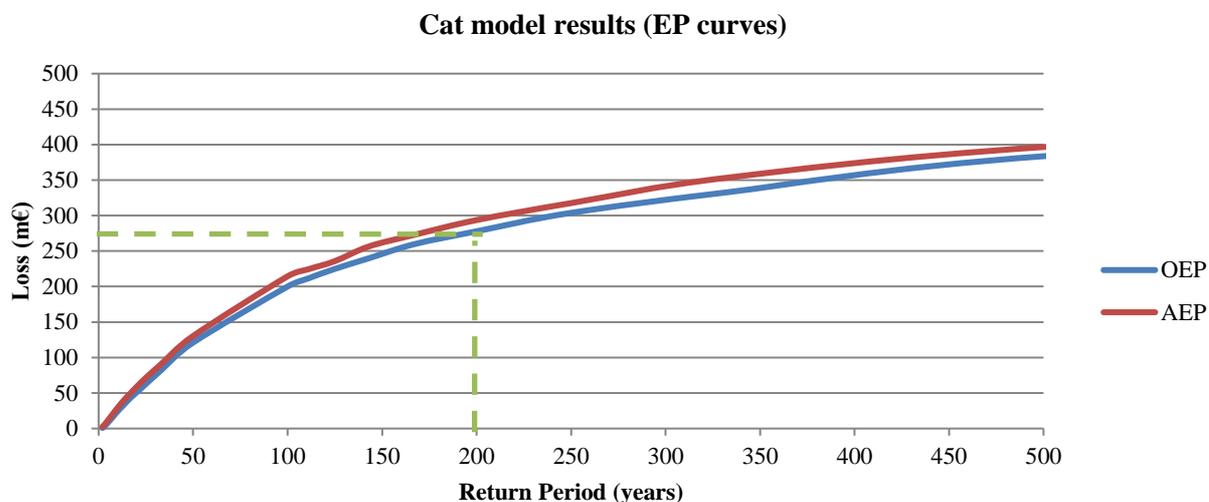
$$AEP(T) = F_C^{-1} \left( 1 - \frac{1}{T} \right)$$

$$OEP(T) = F_M^{-1} \left( 1 - \frac{1}{T} \right)$$

La période de retour est définie par l'inverse de la probabilité de survenance de la mesure considérée (perte maximale, perte cumulée, etc.).

Pour un modèle ouragan Mexique, la distribution se présente comme sur le graphe ci-après (Figure 19).

Dans cet exemple, un événement générant une perte de 275m€ sur l'ensemble du portefeuille de polices d'assurance se produit en moyenne une fois tous les 200 ans (quantile à 99,5%). Autrement dit, le portefeuille modélisé a, chaque année, une chance sur 200 de subir une perte liée à un unique événement de 275m€.



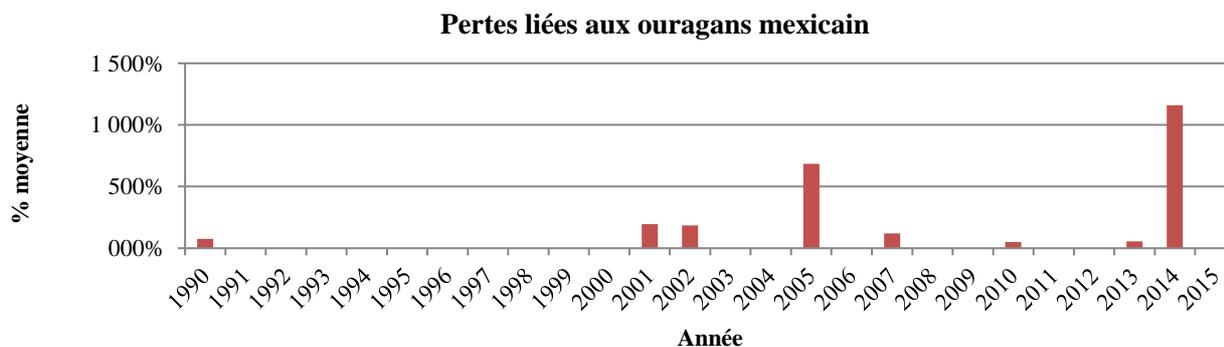
**Figure 19 :** Courbes d'OEP (en bleu) et d'AEP (en rouge) pour un modèle d'ouragan mexicain sur le portefeuille utilisé pour l'étude.

Remarque : On constate que l'AEP n'est que légèrement supérieure à l'OEP. Cela est lié au fait que le péril ouragan n'est que très rarement constitué de plus d'un très gros événement.

## 2. Intérêts des modèles Cat

Le premier avantage des modèles Cat par rapport à d'autres méthodes de définition d'une vision du risque est de permettre une approche d'analyse du risque **par exposition**. En effet, un modèle Cat est indépendant du portefeuille d'origine et des sinistres sur lesquels il a été calibré. Quel que soit le portefeuille, le modèle est en mesure de fournir une évaluation de la vision du risque. Avec un modèle fréquence coût classique, les évolutions de portefeuilles posent le problème de la réévaluation des sinistres historiques à la date de modélisation.

Les modèles Cat proposent une **approche physique** des événements, ce qui les rend a priori beaucoup plus aptes à capter les spécificités des périls modélisés qu'une modélisation statistique. En effet, l'extrapolation statistique de valeurs extrêmes s'effectue au niveau de la génération des variables d'aléa (hauteurs d'eau des fleuves, intensité de séismes, vitesses de vent, etc.) à partir de données historiques. Ceci permet d'être au plus près du comportement du péril au lieu de chercher à capter cela dans des séries historiques de sinistres.



Remarque : « % moyenne » correspond au pourcentage de la moyenne annuelle représenté par la perte. Par exemple, pour l'année 2005, le montant de l'événement était de 6,8 fois la moyenne annuelle.

Ainsi, le portefeuille étudié a fortement évolué au cours du temps et seules 9 tempêtes ont généré des pertes élevées sur les 26 dernières années. Générer une distribution sur des probabilités très faibles (périodes de retour élevées) à partir de 9 montants de sinistres sur 26 ans paraît donc limité par rapport à un modèle Cat.

La vision du risque définie par un modèle Cat sur un portefeuille d'assurance se traduit par une distribution de pertes agrégées à l'échelle du portefeuille d'assurés. Cette distribution permet de construire un tarif pour les traités de réassurance associés (cf. partie II.c.).

Enfin, un modèle Cat peut être utilisé pour la **souscription et la tarification** d'assurance en plus de la vision du risque, de sa gestion et de sa tarification de réassurance. En effet, les caractéristiques de l'aléa associé à un péril peuvent être utilisées pour savoir dans quelles zones les événements sont les plus susceptibles de se produire. Ces caractéristiques peuvent être par exemple utilisées pour construire des cartes de risque ainsi que mettre en place une tarification Cat (cf. II.b.).

### 3. *Limites des modèles Cat*

Plusieurs limites peuvent néanmoins être soulevées quant aux modèles Cat.

D'abord, la **quantité de données** nécessaires pour la construction est conséquente et la qualité du modèle final dépend bien sûr de celle des données utilisées à la construction.

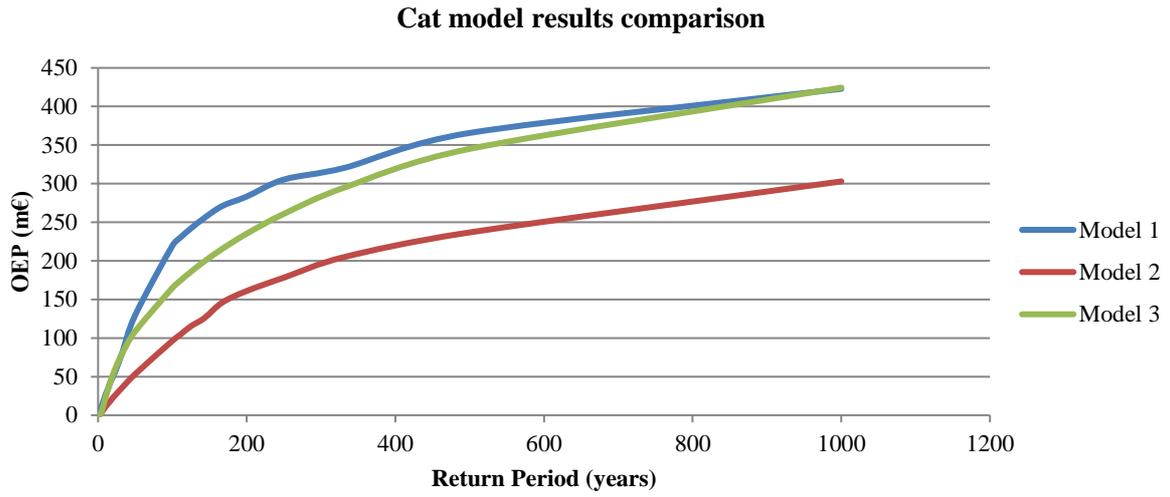
Par exemple, la résolution de la grille géographique sur laquelle les calculs de localisation et propagation des événements seront effectués est crucial pour la précision du modèle. Utiliser une grille (ou raster) de 150m de résolution n'est pas suffisant pour un modèle d'inondation, mais raffiner à une précision de 5m implique des coûts de récupération de données, de retraitement, et de calculs beaucoup plus importants.

Autre exemple, la longueur des séries temporelles de variables physiques utilisées pour la génération de distribution de valeurs extrêmes doit être suffisante et exister en plusieurs endroits sur le territoire à modéliser. Rares sont les pays qui ont mis en place des stations de mesure depuis suffisamment longtemps pour offrir une telle précision de modélisation.

Ensuite, la **calibration** nécessite quasi systématiquement d'avoir des bases de données de sinistres conséquentes pour permettre la calibration des courbes de vulnérabilité. Dans de nombreux pays modélisés par les logiciels disponibles sur le marché, des approximations sont effectuées en raison du manque de sinistres. La conséquence est que les courbes de vulnérabilité varient d'un logiciel à l'autre, ce qui indique une confortable marge de progression.

Ainsi, la quantité de degrés de liberté présents lors de la conception d'un modèle Cat implique nécessairement une volatilité non négligeable sur les résultats. Cette volatilité est observable en comparant les résultats fournis par différents modèles disponibles sur le marché (cf. Figure 20).

Par exemple, sur le modèle ouragan mexicain, on peut observer ci-dessous que l'écart relatif entre différents modèles marchés peut s'élever à plus de 75% sur l'événement de période de retour 200 ans (une chance sur 200 de survenir dans l'année).



**Figure 20 :** Comparaison des OEP de trois modèles vendeurs pour un même portefeuille d'assurés, pour le péril tempête mexicaine.

## **b. Tarification Cat**

Comme indiqué plus haut, les résultats d'un modèle Cat peuvent être utilisés pour la tarification et la souscription des polices d'assurance. Ceci est un pré requis nécessaire à toute analyse plus approfondie de la tarification et d'allocation du chargement de réassurance.

### **i. Rappels de tarification en assurance non-vie**

De façon générale, la prime tarifée par l'assureur peut se décomposer par :

- La prime pure : il s'agit du coût moyen estimé généré par les pertes de la police d'assurance. Il est en général calibré à partir d'un modèle de tarification.
- Des frais : frais de gestion des sinistres, frais d'administration des contrats, frais d'acquisition. Ils sont en général estimés globalement et appliqués sur la tarification comme un montant proportionnel à la prime pure.
- Une marge : il peut s'agir d'une marge pour risque qui prend en compte le fait que la prime chargée des frais est volatile, d'une marge intégrant le coût du capital et le chargement de réassurance engendré par la souscription du risque ou simplement un gain espéré sur le contrat.

Le tarif est établi de façon à être suffisant pour payer les sinistres et assurer la pérennité de l'assureur, compétitif par rapport à un marché concurrentiel et adapté au type de risque. Il est calculé selon une segmentation de « risques homogènes » qui dépend des assureurs à partir des sinistres historiques.

En France, le tarif en assurance non-vie est globalement libre malgré des dispositions spécifiques en assurance automobile (bonus/malus, surprime jeune conducteur) ou l'interdiction de discrimination sur le genre (art. L11-7 Code des Assurances). Cet encadrement n'existe pas non plus au Mexique.

L'introduction d'une charge liée aux catastrophes naturelles n'est donc pas limitée et semble même essentielle pour calculer proprement la prime pure de contrats situés dans des zones régulièrement et fortement sinistrées.

En France, les catastrophes naturelles sont en grande partie couvertes par le biais du régime Cat Nat de la CCR, alimenté par un taux de prime additionnel d'assurance sur l'ensemble des contrats :

- Dommage hors véhicule à moteur : 12% de la prime
- Vol et incendie pour les véhicules à moteur : 6% de la prime

Ce taux unique est justifié par un principe de solidarité entre les assurés. L'objet de ce mémoire, sans nécessairement remettre en cause ce principe de solidarité, vise à aller vers davantage de segmentation du tarif Cat et d'intégrer en outre le chargement de réassurance au niveau site.

### **ii. Tarification Cat**

La méthode de tarification d'assurance présentée ici est basée sur le même modèle qui servira ensuite à effectuer les calculs de chargement de réassurance. C'est un modèle stochastique qui a été lancé à la granularité la plus fine possible de façon à avoir des résultats site par site.

## 1. Présentation des résultats

Les résultats de la modélisation se présentent sous la forme d'un tableau comme indiqué ci-dessous, où :

- YearID (nombre entier) : année simulée
- EventID (nombre entier) : identifiant de l'événement survenu lors de l'année concernée. Noté k par la suite.
- LocationID (nombre entier) : identifiant du site touché par l'événement. Noté i par la suite.
- Loss : perte associée au site touché par l'événement (dans la monnaie de simulation, ici en dollar).

YearID	EventID	LocationID	Loss
1	130	45696	23
1	130	34665	765
1	346	9778	24457
2	543	2345	31
2	543	232211	287
...	...	...	...
10 000	34	56867	465

Pour chaque scénario d'une année, on a donc tiré aléatoirement des événements dont les empreintes ont été croisées avec les sites du portefeuille et les pertes calculées site par site. Ce type de tableau est nommé *Year Loss Table* (YLT) ou *event set*. Dans cet exemple de modèle, 10 000 années sont simulées pour obtenir une distribution de perte (méthode de Monte-Carlo).

## 2. Analyse des résultats

La tarification Cat des polices d'assurance nécessite d'avoir des résultats à un niveau de granularité site par site. Les résultats présentés dans la partie précédente ont donc été agrégés par site en calculant la perte moyenne sur le nombre total d'événements simulés par le modèle. Une fois les pertes agrégées en moyennes par site, on calcule un taux de destruction moyen site par site :

$$TD(i) = \frac{1}{Nsim} \times \frac{\sum_{k=1}^{Nsim} P(i, k)}{VA(i)}$$

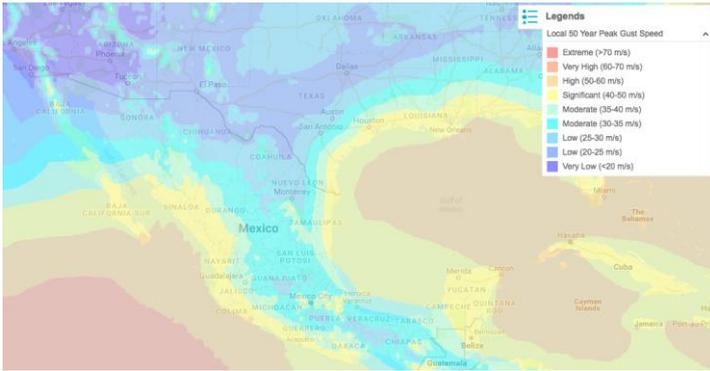
Où :

- $P(i, k)$  : perte de l'événement k pour le site i
- $VA(i)$  : valeur assurée du site i
- $Nsim$  : nombre total de simulations du modèle
- $TD(i)$  : taux de destruction moyen pour le site i sur la totalité des simulations. Parfois noté DR (*destruction rate*)

De façon assez logique, on retrouve une corrélation spatiale entre les taux de destruction les plus élevés et les zones de vitesses de vent les plus importantes (au niveau des côtes du pays, cf. Figure 21).

Afin d'effectuer des comparaisons sur un segment de portefeuille à peu près homogène, à la fois en termes d'aléa que de caractéristiques de portefeuille. On a ainsi sélectionné un segment de portefeuille correspondant aux deux critères suivants :

- Localisation dans la zone administrative de Mexico city.
- Somme assurée comprise entre 100 000 € et 1 000 000 €.



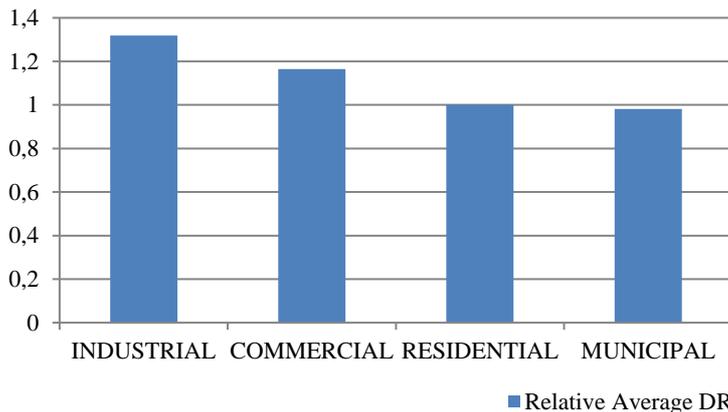
**Figure 21 :** Comparaison des sites les plus détruits en moyenne par le modèle (premier décile des taux de destruction, à droite) avec la carte des vitesses de vent moyennes (à gauche, les zones jaune-orange sont les zones aux vitesses de vent les plus élevées).

A partir des résultats sur ce segment de portefeuille, il a été possible de retrouver certains éléments de vulnérabilité du modèle en comparant les taux de destruction (DR) :

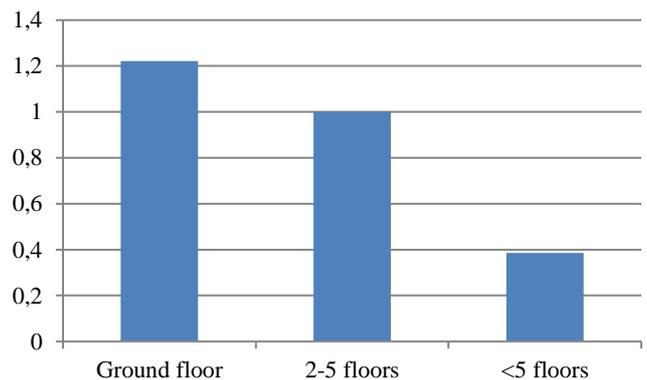
- Au niveau des lignes de métier, on constate ainsi que le résidentiel est la ligne de métier la plus vulnérable aux ouragans.
- Sur le plan des structures de bâtiment, on trouve assez logiquement que le béton armé (*reinforce concrete*) est plus résistant que le béton (*concrete*), lui-même plus résistant que les ouvrages en maçonnerie (*masonry*). Les ouvrages de structure métallique (*steel*) sont les moins résistants (bâtiments provisoires, plus légers, etc.)
- Enfin, plus un risque est situé en hauteur (<5 floors par exemple), moins il sera vulnérable à une tempête, ce qui n'apparaît pas nécessairement intuitif.
- L'absence du type de toit dans le portefeuille utilisé ne permet pas d'étudier ce paramètre.

Ces résultats seront vérifiés statistiquement dans la suite de l'étude.

**Comparison per line of business**



**Comparison per floor level**



**Comparison per structure type**



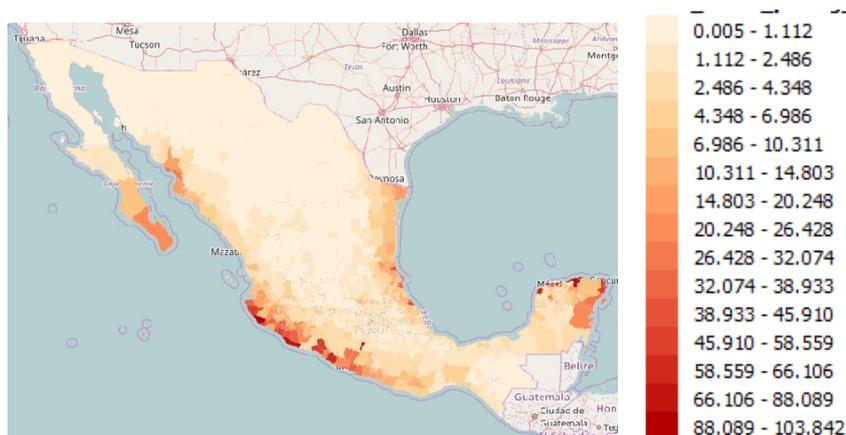
### 3. Tarification sur les résultats du modèle

Avant de pouvoir réfléchir à l'allocation du chargement de réassurance Cat, il convient de connaître au préalable le coût des Cat site par site, autrement de connaître la part de la prime qui correspond aux catastrophes concernées (ici ouragan). Cette part est appelée « prime Cat » et peut se définir plus simplement comme le tarif d'assurance pour la garantie catastrophe naturelle ouragan.

Cette prime Cat est établie à partir des résultats modélisés en calculant la moyenne des pertes modélisées sur le site considéré. Ceci fournit une prime pure Cat qui correspond exactement au montant moyen que le site subit en raison des catastrophes naturelles, comme indique dans l'équation ci-dessous.

$$Prime\ Cat(i) = \frac{1}{Nsim} \times \sum_{k=1}^{Nsim} P(i, k)$$

En traçant les primes Cat en pourcentage de somme assurée par municipalités mexicaine, on observe assez logiquement que les taux de primes les plus élevés se situent au niveau des côtes du Mexique, ce qui est cohérent avec les caractéristiques physiques des ouragans.



**Figure 22 :** carte de la tarification d'assurance Cat en pourcentage de somme assurée, obtenue avec le modèle. La légende est en échelle relative par rapport à une valeur maximale. Les résultats ont été moyennés par municipalité.

Une fois cette tarification d'assurance effectuée, on peut ajouter à la partie du risque qui est purement individuelle (prime Cat) une partie liée à l'accumulation de risques dans une même zone. Cette partie correspond au chargement de réassurance et au coût du capital car c'est notamment l'accumulation de risques qui entraîne l'achat de réassurance et le besoin de capital requis dans le cadre de la norme Solvabilité 2.

## c. Définitions et cadre de l'étude

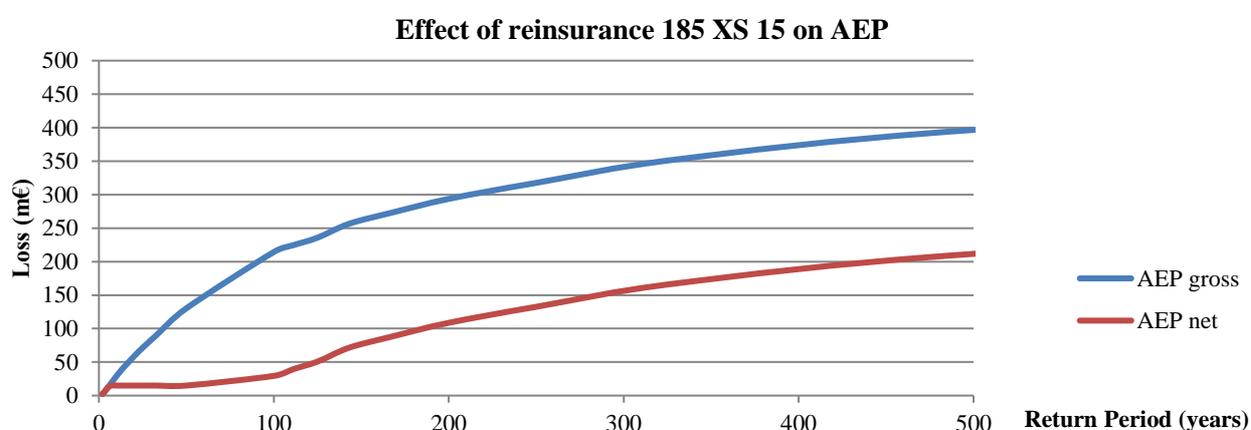
Après avoir construit une tarification d'assurance pour la garantie catastrophe naturelle ouragan, on peut s'intéresser à la partie qui concerne la réassurance, sa tarification, et l'allocation de son chargement.

### i. Notions de réassurance

#### 1. *Définitions*

La réassurance permet à l'assureur (cédante) de céder une partie des risques qu'il a souscrits à un réassureur (cessionnaire) pour plusieurs raisons :

- Se couvrir contre des **pics de sinistralité** exceptionnels : par exemple en cas de collision entre deux porte-conteneurs en assurance marine.
- Réduire la **charge en capital** d'un portefeuille : la réassurance permet de réduire le besoin règlementaire en capital (*Solvency Capital Requirement*, pilier 1 de la directive Solvabilité 2) et donc d'offrir plus de liberté au niveau de la souscription.  
Dans l'exemple présenté dans le graphe ci-dessous, l'AEP – charge cumulée annuelle - 200ans passe d'environ 295m€ à 110m€ avec un traité 185m€ XS 15m€.
- **Lisser le résultat** : permet de réduire la volatilité du résultat et donc d'avoir un message plus stable vis-à-vis des parties prenantes (régulateurs, marché, actionnaires, etc.).



#### 2. *Modes et formes de réassurance*

La réassurance se classifie selon trois modes principaux :

- La **réassurance facultative** : la cédante souhaite céder un risque spécifique et propose l'affaire à un réassureur. L'affaire n'est conclue qu'avec l'accord des deux parties.
- La **réassurance obligatoire** (traité) : le traité de réassurance définit une typologie de risque (une ligne de business, un portefeuille complet) et tous les risques correspondants sont obligatoirement couverts par le traité.
- La **réassurance facultative-obligatoire** (fac-ob) : assez rare, se comporte comme un traité dans lequel la cédante à le choix de céder ou non les risques souscrits et le réassureur s'engage à l'avance à tous les accepter.

Pour chaque mode, deux formes de réassurance existent :

- La réassurance **proportionnelle** : une proportion donnée de la prime et des pertes est cédée au réassureur pour chaque site. Il y a un partage du sort clair entre assureur et réassureur.
- La réassurance **non proportionnelle** : la perte est cédée au réassureur au-delà d'un seuil appelée priorité et dans la limite d'un montant total appelée portée, en échange d'une prime fixée par le traité.

Dans la réassurance non proportionnelle, on trouve entre autres les traités Cat. Ils couvrent les événements catastrophiques qui ont pour conséquence un grand nombre de sinistres. Ces sinistres sont donc agrégés par événement et la perte totale est présentée au traité.

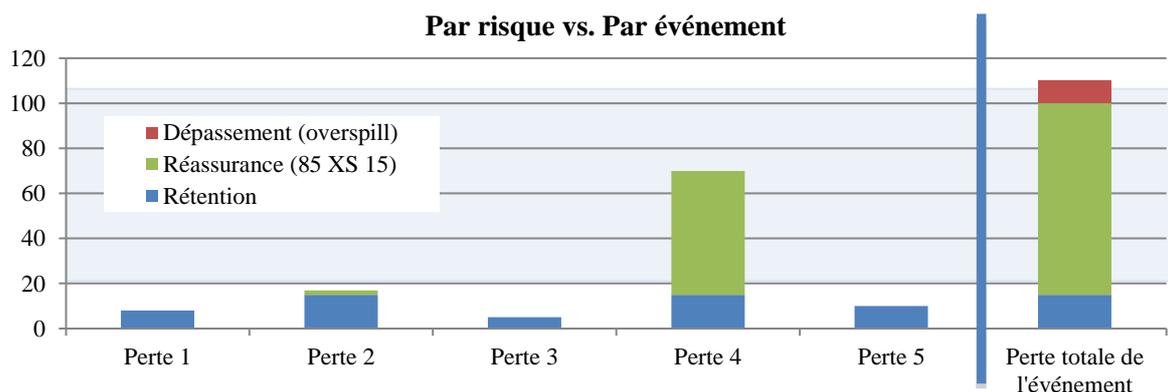
## ii. Une approche spécifique aux traités XL

### 1. *Un intérêt limité pour les autres formes de réassurance*

L'étude présentée se concentre spécifiquement sur les traités par événement non proportionnels (excédent de sinistre, ou XL, par événement).

En effet, les traités (ou facultatives) proportionnels n'offrent qu'un intérêt limité à l'étude étant donné que les pertes réassurées sont directement proportionnelles aux pertes site par site. Chercher à allouer le prix d'un traité proportionnel revient donc à extraire de la prime pure Cat site par site (tarif d'assurance) un montant de réassurance qui lui est proportionnel et qui sera cédée par l'assureur au réassureur.

En ce qui concerne les traités non proportionnels par risque (ou facultatives), une approche similaire à celle sur les traités par événement pourrait être employée. L'intérêt est néanmoins assez limité dans la mesure où les traités par risque ne jouent que rarement lors de catastrophes naturelles. Celles-ci sont en effet plutôt caractérisées par un nombre important de pertes de petits montants.



**Figure 23** : Comparaison du fonctionnement d'un traité par risque (à gauche) par rapport à un traité par événement (à droite). Pour le traité par risque, chaque perte est considérée individuellement. Pour le traité par événement, la somme de toutes les pertes liées à l'événement est considérée pour les récupérations.

Pour les traités non proportionnels, la présence d'une rétention et d'une limite introduit des effets de seuil au niveau des récupérations de réassurance qui rendent intéressante l'étude d'allocation du chargement d'un tel traité.

## 2. *Prise en compte des spécificités des traités XL*

Plusieurs spécificités peuvent apparaître au niveau des traités XL. On distingue en général :

- **Les reconstitutions** : il s'agit du paiement d'une fraction de la prime d'origine en fonction de la proportion de portée consommée. L'idée est de payer pour la « reconstitution » de la portée qui a été consommée par un événement de façon à permettre de couvrir plusieurs événements qui surviendraient sur le programme de réassurance XL considéré. Cette prime de reconstitution peut s'effectuer à un taux différent de la prime d'origine et le nombre de reconstitutions possibles est généralement limité.

Par exemple, soit un événement de 100 m€ sur un traité 100 XS 50 m€ de prime 10 m€ et de prime de reconstitution 1@80% (1 reconstitution complète possible à un taux de 80%). L'événement entraîne 50 m€ de récupération et consomme 50% de la portée. Dans ce cas, 50% de la prime sera payée en tant que prime de reconstitution, à un taux de 80%, soit  $50\% \times 80\% \times 10 = 4\text{m€}$ .

Les primes de reconstitutions introduisent un effet non linéaire dans le chargement de réassurance car l'assureur ne paye pas de prime de reconstitution sur la dernière portée (puisque non reconstituée). Dans l'exemple précédent, un second événement de 50m€ causerait la consommation totale de la première portée et la fin du paiement de prime de reconstitution. La portée serait alors reconstituée à 100m€, mais tout événement supplémentaire consommerait cette dernière portée sans paiement de prime de reconstitution.

Dans le cadre de l'étude, il a été vérifié que les primes de reconstitution ont un impact négligeable sur les résultats pour le traité considéré. Ceci s'explique par deux éléments.

Le premier est lié aux caractéristiques physiques des ouragans. En effet, les ouragans, de même que les séismes et contrairement à d'autres périls comme la grêle, sont des événements à très forte sévérité mais à très faible fréquence. Il est donc très peu probable d'avoir plusieurs événements dépassant la rétention du traité par événement la même année (cf. Tableau 1 en page suivante).

Le second est lié au fait que la prime de reconstitution est en général petite par rapport à la portée du traité. Ainsi, le paiement de la prime de reconstitution représente 8% des récupérations dans l'exemple cité plus haut (événement de 100m€).

Dans le cas de tranches de réassurance dites « travaillantes » (touchée régulièrement en raison d'une rétention basse) le prix est un taux important de la portée. Néanmoins, bien que les primes de reconstitutions puissent être élevées au regard des récupérations de réassurance, leur nombre est généralement plus élevé que sur des tranches de réassurance de rétention plus haute pour éviter l'épuisement total de la capacité du traité. Il est donc peu probable d'arriver sur la dernière portée du traité.

Pour ces raisons, il est donc possible de négliger l'effet des primes de reconstitution sur les résultats de l'allocation du chargement de réassurance de la suite de l'étude.

Période de retour	Premier événement	Deuxième événement	Troisième événement	Quatrième événement
2	1,51	0,64	0,40	0,27
3	2,87	0,90	0,55	0,36
5	4,98	1,33	0,71	0,49
10	9,83	2,20	1,00	0,63
11	10,73	2,41	1,04	0,66
14	13,41	2,87	1,19	0,71
17	15,97	3,27	1,29	0,75
20	18,75	3,56	1,38	0,79
25	23,43	4,01	1,53	0,85
50	47,29	5,58	2,12	1,09
100	77,37	8,33	3,13	1,36
125	85,37	8,86	3,44	1,46
167	96,66	10,11	3,69	1,66
200	104,58	10,69	3,98	1,77
250	113,74	12,07	4,27	1,89
500	157,37	17,02	5,41	2,70
1000	206,46	23,14	6,81	3,53

**Tableau 1 :** Montant des événements de rang 1 à 4 dans une année en m€, par période de retour. Chaque montant d'événement de rang donné suppose que les pertes de rang inférieur étaient supérieures ou égales à ce montant. On constate que la période de retour de deux événements supérieurs à un certain montant est bien supérieure à celle d'un seul événement de ce même montant.

Par exemple, un événement de 23m€ survient en moyenne une fois tous les 25 ans (colonne « Premier événement »), mais la survenance de deux événements supérieurs ou égaux à 23m€ est de l'ordre de 1000 ans (colonne « Deuxième événement »). Autre exemple, alors qu'un événement de 3m€ a une chance sur 3 de se produire par an (colonne « Premier événement »), avoir 4 événements de ce montant ou plus grand n'a qu'une chance sur 700 environ de se produire sur une année (colonne « Quatrième événement »).

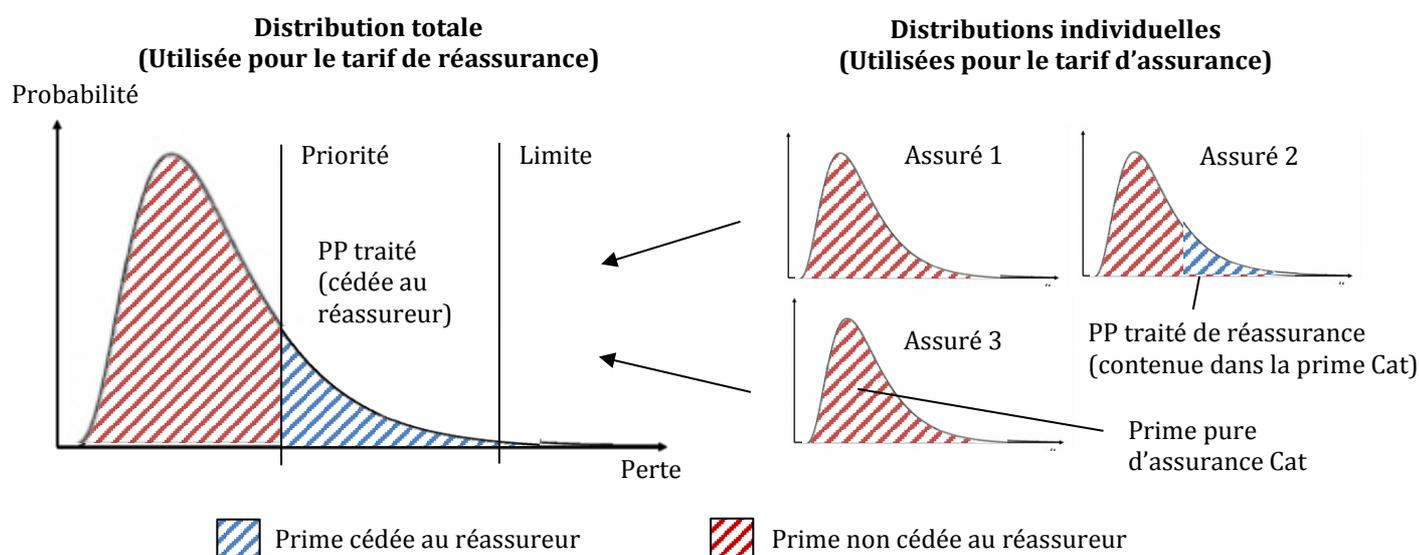
- **Les *annual aggregate deductible (AAD) ou limit (AAL)* :** un AAD est une rétention qui s'applique sur l'ensemble des pertes présentées au traité à la fin de l'année. L'AAL est l'équivalent pour la limite. Pour le calcul de ces éléments, on calcule les récupérations sur le traité de façon normale, puis on les somme à la fin de l'année et on applique l'AAD et l'AAL sur ces récupérations agrégées.

Dans l'exemple précédent, si l'on suppose que deux événements respectifs de 100m€ et 80m€ surviennent sur le traité 150 XS 50 et qu'un AAD de 20 m€ est présent, alors les événements entraîneront respectivement 50 et 30 m€ de récupérations. L'application de l'AAD donne un montant final de récupérations de 60m€.

Événement	Perte brute	Récupérations – 150 XS 50
Événement 1	100m€	50m€
Événement 2	80m€	30m€
TOTAL hors AAD	180m€	80m€
TOTAL avec AAD (20m€)	180m€	60m€

Dans le cadre de ce mémoire, le traité étudié ne contient pas d'AAD ou d'AAL. Néanmoins, il est tout à fait possible d'intégrer ces conditions financières dans le calcul du chargement de réassurance.

### iii. Introduction du chargement de réassurance dans la tarification Cat



Comme évoqué plus haut, les résultats modélisés contiennent des pertes site par site, par année simulée. Année par année, on peut donc calculer des récupérations liées à un traité de réassurance et donc en déduire un « chargement de réassurance ». Néanmoins, dans un cadre de tarification Cat, il convient de bien définir ce qui est tarifé pour éviter un double comptage de la prime de réassurance cédée (en bleu ci-dessus) dans les primes d'assurance tarifées (à droite, distributions complètes).

En effet, lorsque l'on calcule une prime d'assurance Cat, on fait payer à l'assuré sa prime pure Cat qui contient l'ensemble de la distribution des pertes possible de la police. Par exemple, dans le schéma ci-dessus, la distribution de l'Assuré 1 correspond à la prime d'assurance pour le risque de catastrophe.

Une partie de cette distribution est associée à des événements extrêmes qui entraînent des récupérations de réassurance. C'est la partie bleue de la prime Cat de l'assuré 2 dans l'exemple ci-dessus. Dans cet exemple, l'Assuré 2 paye l'ensemble de la distribution (bleue et rouge) dans sa prime d'assurance mais l'assureur en cède une partie (en bleu) et conserve le reste (en rouge). Ainsi, la prime pure de réassurance (notée PP dans la suite) est déjà contenue dans la tarification d'assurance Cat. Il n'est pas nécessaire de l'allouer.

La partie qui n'est pas contenue dans cette tarification et que l'on cherche à allouer dans ce mémoire est la marge de réassurance. Plus précisément, il s'agit de la différence entre la prime pure et la prime commerciale. Cette marge contient plusieurs choses que l'on sépare en deux parties :

- La prime de risque : il s'agit du chargement introduit par le réassureur pour rémunérer la variabilité du risque accepté. On considèrera cette prime de risque comme un pourcentage de la volatilité des récupérations de réassurance  $\sigma$  avec un facteur  $\alpha$ .
- Les frais de gestion : il s'agit d'un chargement rémunérant la gestion du traité, des sinistres associés, etc. On considèrera ces frais comme proportionnels à la prime commerciale avec un facteur  $\beta$ .

Une étude interne sur les traités Cat montre ces deux facteurs égaux à 15%.

Ainsi, on écrit alors la prime tarifée par le réassureur (prime commerciale  $P_c$ ) comme :

$$P_c = \frac{PP + \alpha \cdot \sigma}{1 - \beta}$$

La partie de cette prime qui n'est pas contenue dans la prime Cat de l'assuré est donc :

$$CoR = P_c - PP = \frac{\beta \cdot PP + \alpha \cdot \sigma}{1 - \beta}$$

C'est ce montant que l'on définit comme chargement de réassurance (CoR) et que l'on cherche à allouer entre les différents assurés. On constate qu'il est constitué d'une partie liée à la prime pure et une partie liée à la volatilité des récupérations du traité.

D'un point de vue de la prime d'assurance, ce chargement de réassurance est habituellement compté dans la partie technique de la prime d'assurance, de même que les frais de gestion et d'acquisition de l'assureur. Ce chargement se traduit généralement par un pourcentage global appliqué sur l'ensemble des polices d'assurance. L'objectif de la suite de ce mémoire est d'allouer ce chargement de réassurance en fonction de la contribution des polices individuelles dans celui-ci.

#### iv. Principes théoriques d'allocation de coût (*cost allocation*)

La partie qui suit a pour objectif de synthétiser le cadre conceptuel en matière d'allocation de coût existant dans la littérature (Monin, 2008 et Denault, 2011) et de l'appliquer à l'allocation du chargement de réassurance. Une analogie avec les méthodes de la théorie des jeux coopératifs est présentée.

La principale différence entre les jeux coopératifs et le chargement de réassurance tient à ce que les jeux coopératifs visent à allouer un gain collectif, tandis que le chargement de réassurance vise au contraire à allouer un coût collectif.

En effet, un exemple simple de jeu collectif est le covoiturage, dans lequel les participants génèrent un gain avec le covoiturage par rapport à la situation dans laquelle chacun parcourt le trajet de façon solitaire.

Dans le cas de la réassurance (ainsi que du coût du capital), l'accumulation de risques entraîne a priori une volatilité accrue contre laquelle une cédante souhaite se couvrir à l'aide de réassurance. C'est le coup lié à ce transfert de volatilité qui vient s'ajouter à la prime pure Cat.

##### 1. *Notations et définitions*

En théorie des jeux, on définit un jeu comme une situation dans laquelle le coût pour un joueur dépend aussi des autres joueurs.

Dans le cas du chargement de réassurance, ce dernier dépend en effet des autres sites assurés. Un site assuré peut être comparé à un joueur au sens de la théorie des jeux.

On note :

- **N** l'ensemble des joueurs. Avec  $n$  sites, on a :  $N = \{1, 2, \dots, n\}$ .
- **S** une coalition de joueurs. C'est un sous ensemble de  $N$  qui comprend  $s$  joueurs.
- **c** la fonction de coût (coût de catastrophe et de réassurance) qui peut être associée à un joueur ou à une coalition. On notera  $c(i)$  le coût individuel du joueur  $i$ .
- **CM(i)** le coût marginal du joueur  $i$  :  $CM(i) = c(N) - c(N \setminus i)$ . Il s'agit du coût supplémentaire lié à l'introduction de ce joueur dans l'ensemble  $N$  privé de  $i$ .

- $\mathbf{X} = (x_1, x_2, \dots, x_n)$  le vecteur d'allocation des coûts (prime Cat incluant le chargement de réassurance).
- $\mathbf{G} = (g_1, g_2, \dots, g_n)$  le vecteur de répartition du gain total associé au jeu de l'ensemble N.
- $\mathbf{R} = (r_1, r_2, \dots, r_n)$  le vecteur de répartition du chargement de réassurance total (analogie avec le vecteur G), avec :

$$\sum_i r_i = \sum_k c(k) - c(N) = CoR$$

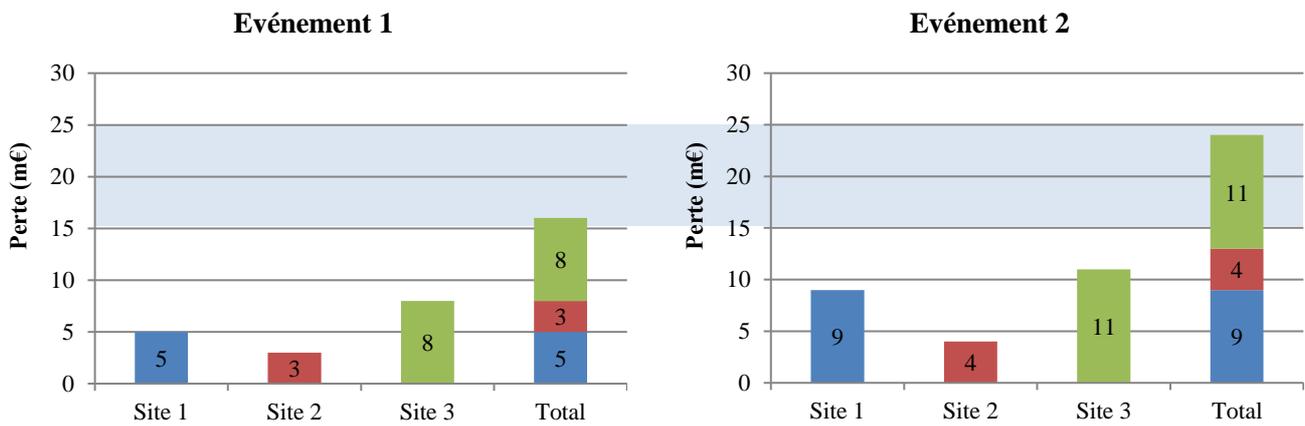
- $\mathbf{R}_s$  le chargement de réassurance total de la coalition S et qui est la différence entre le coût total et la somme des coûts individuels des joueurs de la coalition :  $\sum_{k \in S} c(k) - c(S) = CoR_S$

Une méthode d'allocation est une application qui associe à une fonction de coût c un vecteur d'allocation X. Elle est dite stable lorsqu'elle est dans le noyau du jeu (cf. paragraphe suivant).

## 2. Exemple

On considère un exemple simplifié de modèle Cat dans lequel deux scénarios d'année sont possibles avec la même probabilité, avec pour chacune d'elle un seul événement. Trois sites sont présents en portefeuille et le traité de réassurance est un traité en excédent de perte par événement 10 XS 15, de prime 7.5, sans reconstitution ni *annual aggregate*.

Chaque événement est associé à une perte totale du portefeuille : 16 m€ pour l'événement 1, et 24 m€ pour l'événement 2. Les récupérations associées sont donc 1 m€ pour l'événement 1 et 9 pour l'événement 2, ce qui donne une prime pure de réassurance de 5m€. Le chargement de réassurance est fixé dans l'exemple à 50% de la prime pure, soit 2.5m€. C'est ce montant que l'on cherche à allouer sur les trois sites.



Dans ce cas, on a donc, pour  $N = \{1, 2, 3\}$

La fonction de coût est ici la somme des primes Cat individuelles (moyenne des pertes sur l'ensemble des événements) et du chargement de réassurance. Dans l'exemple ci-dessus, le coût est donc de 22.5 m€ (20m€ de perte moyenne par événement et 2.5 m€ de chargement de réassurance).

$$c(i) = \frac{1}{Nsim} \times \sum_{k=1}^{Nsim} P(i, k) + r_i$$

Où :

- $P(i, k)$  : perte de l'événement  $k$  pour le site  $i$ , par exemple égal à 5 m€ pour l'événement 1 et le site 1 de l'exemple simplifié.
- $r_i$  : chargement de réassurance du site  $i$  recherché
- $N_{sim}$  : nombre total de simulations du modèle, égal à 2 dans l'exemple simplifié.

A partir de cette fonction de coût, on peut calculer les coûts individuels et marginaux. Le coût individuel de chaque site correspond au résultat de la fonction de coût du site s'il était seul en portefeuille. Dans notre exemple, il est égal à la prime pure d'assurance pour la garantie catastrophe car il ne génère pas de récupérations de réassurance. Le coût marginal correspond au coût supplémentaire (prime Cat et chargement de réassurance supplémentaire) généré par l'introduction du site dans l'ensemble privé de celui-ci. Dans notre exemple, il s'agit donc de la différence entre le coût total (22.5 m€) et le coût sans le site,  $C(N \setminus \text{site})$ .

Cet exemple sera utilisé pour exposer les différentes méthodes d'allocation du chargement de réassurance dans les paragraphes suivants.

	Site 1	Site 2	Site 3
<b>Coûts individuels</b>	<b>7</b>	<b>3,5</b>	<b>9,5</b>
Récupérations de $N \setminus \text{site}$	0	2,5	0
Chargement de réassurance de $N \setminus \text{site}$	0	1,25	0
$C(N \setminus \text{site})$	13	17,75	10,5
<b>Coûts marginaux</b>	<b>9,50</b>	<b>4,75</b>	<b>12,00</b>

Remarque : les coûts individuels sont directement les primes Cat car les sites ne génèrent jamais de récupérations de réassurance.

### 3. Propriétés et noyau du jeu

Les méthodes d'allocation peuvent présenter différentes propriétés présentées ci-dessous. Ce n'est cependant pas toujours le cas.

#### Propriété 1 : rationalité individuelle

Si un joueur fait partie d'un jeu collectif, alors il ne peut payer plus que son coût individuel.

$$x_i \leq c(i)$$

En effet, si un joueur paye plus que son coût individuel, il n'a plus intérêt à faire partie du jeu.

Traduite en termes d'allocation de réassurance, on inverse la proposition : si un site fait partie d'un portefeuille réassuré, alors il ne peut payer moins que son coût individuel.

$$x_i \geq c(i)$$

Remarque : on considère dans ce mémoire que le marché de la réassurance n'est pas en situation de vente à perte dans laquelle les primes commerciales (payés par les cédantes) des traités de réassurance sont plus faibles que les primes pures modélisées associées.

#### Propriété 2 : rationalité collective

Si un joueur fait partie d'un jeu collectif, alors il ne peut payer moins que son coût marginal.

$$x_i \geq CM(i)$$

En effet, si un joueur paye moins que son coût marginal, cela signifie que les autres joueurs sont pénalisés par l'entrée de ce nouveau joueur. Pour les autres joueurs, il n'y a plus d'intérêt à former une coalition incluant le nouvel entrant.

En termes d'allocation de réassurance, on inverse ici aussi la proposition : si un site fait partie d'un portefeuille réassuré, alors il ne peut payer plus que son coût marginal.

$$x_i \leq CM(i)$$

En effet, cela signifierait que ce nouveau site paierait plus que sa contribution au chargement de réassurance de la coalition. Sans asymétrie d'information, cette propriété paraît théoriquement justifié. Cependant, ce principe pourrait ne pas être respecté en pratique étant donné que l'assuré ne connaît ni sa prime pure ni son chargement éventuel.

### Noyau du jeu

On définit le noyau (ou cœur) d'un jeu l'ensemble des allocations possibles vérifiant à la fois la rationalité individuelle et la rationalité collective. Il n'existe pas toujours.

L'intérêt de cette notion est qu'elle définit un espace en dehors duquel les joueurs, s'ils sont rationnels, refusent d'aller en tant qu'individu ou en tant que collectivité.

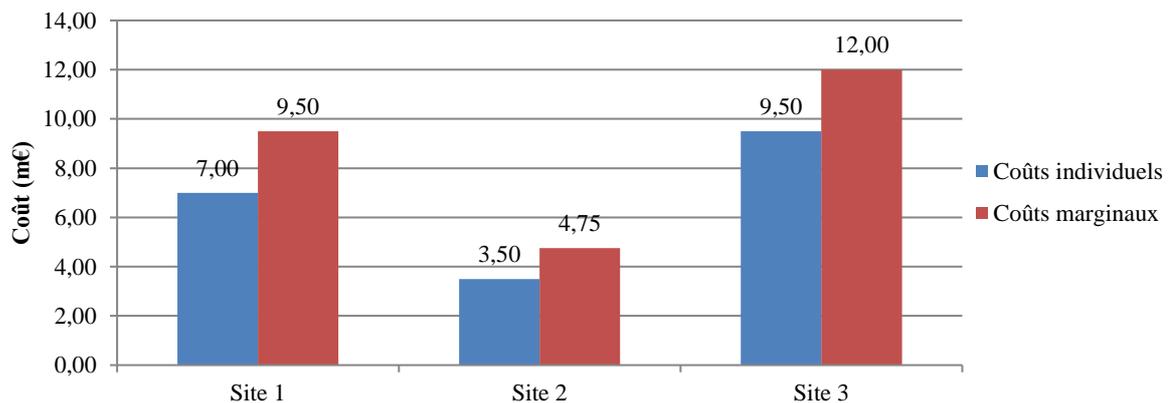
Le noyau du jeu est donc l'ensemble des allocations situées entre le coût individuel et le coût marginal.

Dans notre exemple, le coût individuel est la prime d'assurance Cat et est déjà fixée. Il convient donc d'allouer le chargement de réassurance (2.5 m€) entre les trois sites sans que le coût total du site (somme de la prime Cat et du chargement de réassurance) dépasse son coût marginal.

Par exemple, allouer la totalité du chargement de réassurance au Site 1 respecterait le noyau du jeu (mais serait injuste).

Néanmoins, l'allouer au Site 2 ne respecterait plus le noyau du jeu. Son coût total serait alors de 6 m€, ce qui est plus grand que son coût marginal de 4.75 m€.

Noyau du jeu

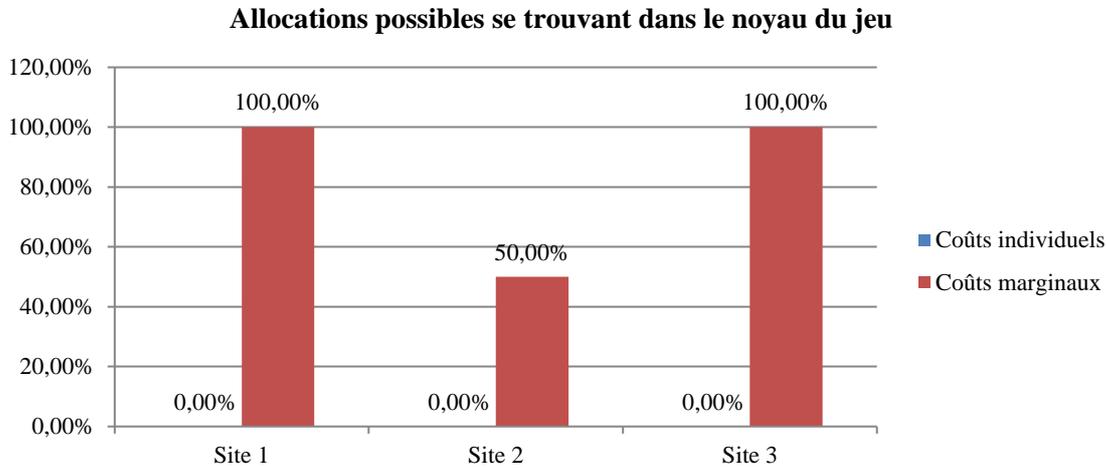


Si l'on exprime le noyau du jeu uniquement en termes d'allocation du chargement de réassurance, le pourcentage que chaque assuré doit payer est compris entre :

- 0% : l'assuré paye son coût individuel
- Un pourcentage correspondant au pourcentage de chargement de réassurance (2.5m€) payé par l'assuré lorsque celui-ci est à son coût marginal.

L'intervalle de pourcentage d'allocation du chargement de réassurance par site respectant le noyau du jeu qui correspond à l'exemple simplifié est montré ci-après.

Pour chaque site, le pourcentage du chargement de réassurance alloué doit être compris entre le coût individuel (0%) et le coût marginal. La somme de ces pourcentages doit évidemment s'élever à 100%.



Les propriétés suivantes sont conservées pour le chargement de réassurance.

### Propriété 3 : symétrie

Si, en rejoignant une coalition  $M \setminus \{i, j\}$ , chaque joueur  $i$  et  $j$  contribuent au gain total de la coalition de la même manière, alors leurs facteurs de répartition sont égaux.

$$\forall i, j, c(M \setminus i) = c(M \setminus j) \Rightarrow x_i = x_j$$

Cela revient à dire que deux assurés qui auraient exactement le même impact sur le chargement de réassurance auront la même allocation de chargement de réassurance.

### Propriété 4 : joueurs inutiles

Si un joueur n'apporte pas de gain à la coalition, alors il est dit inutile et son coût marginal est égal au coût individuel.

$$\forall i, CM(i) = c(i) \Rightarrow x_i = c(i)$$

Dans le cadre de l'étude, cela revient à dire qu'un site assuré qui ne serait jamais touché par un événement Cat, ou qui ne serait touché que par des événements de montant de perte inférieur à la rétention du traité de réassurance, est inutile pour l'allocation et recevra donc un chargement de réassurance nul.

### Propriété 5 : additivité

Soient deux jeux sans interaction caractérisés par le même ensemble  $N$  et deux fonctions de coût  $c$  et  $c'$ . Alors l'allocation associée au coût des deux jeux est égale à la somme des allocations de chaque jeu, pour tout joueur  $i$ .

$$x_i(c + c') = x_i(c) + x_i(c')$$

Cette propriété permet par extension de sommer par exemple une allocation de prime de réassurance pour les ouragans et une autres pour les tremblements de terre, qui constituent deux jeux a priori sans interaction.

### Propriété 6 : monotonie des coûts

Soient deux fonctions de coût  $c$  et  $c'$  associées à un même ensemble  $N$ . Si  $c'(N) > c(N)$  et, pour tout  $S$ ,  $c'(S) > c(S)$ , alors, pour tout  $i$ ,  $x'_i > x_i$ .

## v. Méthodes d'allocation de coût considérées

### 1. Méthode proportionnelle : égale répartition du coût total

La première méthode, qui est aussi la plus simple, consiste à répartir le chargement total de réassurance de façon égale entre chaque participant. Tout se passe comme si chaque assuré payait un montant constant de chargement de réassurance en plus de sa prime d'assurance Cat.

$$x_i = c(i) - \frac{1}{n} \left[ \sum_{k=1}^n c(k) - c(N) \right]$$

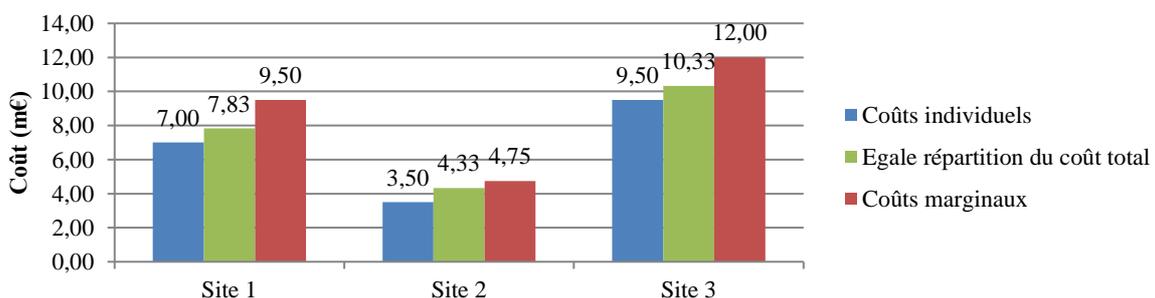
Le vecteur  $R$  est alors toujours égal à  $(1/n, \dots, 1/n)$ . Avec l'exemple précédent, chaque site paye un tiers du chargement de réassurance.

#### Allocation avec méthode proportionnelle : égale répartition



Avec l'exemple précédent, chaque assuré paye un tiers de 2.5m€ (0.83m€) en plus de son coût individuel et on obtient les résultats suivants :

#### Méthode proportionnelle : égale répartition



Cette méthode, outre sa simplicité, vérifie les propriétés de symétrie et de monotonie. En revanche, elle n'est pas toujours dans le noyau du jeu, contrairement à l'exemple. Il suffit en effet d'un joueur inutile, c'est-à-dire ici un joueur qui ne génère pas de chargement de réassurance (le coût marginal est alors égal au coût individuel), pour que ce joueur ne soit pas dans le noyau du jeu.

Cette méthode fait l'hypothèse que le chargement de réassurance est généré de façon homogène par l'ensemble des sites, ce qui est une vision approximative qui ne prend pas en compte les accumulations potentielles de risques dans les zones de fort aléa.

Par ailleurs, elle distribue le coût de façon complètement indépendante de la prime Cat, ce qui désavantage les petits sites et avantage les plus gros. Dans l'exemple, le site 2 paye plus de 96% de son coût marginal alors que le site 3 n'en paye que 86%.

Cette méthode n'est pas retenue car elle ne permet pas de créer un avantage compétitif en ciblant spécifiquement les sites qui génèrent le plus de volatilité dans le portefeuille et pour la réassurance.

## 2. Méthode proportionnelle : répartition proportionnelle du coût total

Dans cette méthode, on répartit le chargement de réassurance proportionnellement au coût individuel de chaque site. Cette méthode répond de façon simple au problème d'égalité répartition. L'allocation se définit comme suit :

$$x_i = c(i) - \frac{c(i)}{\sum_{j=1}^n c(j)} \left[ \sum_{k=1}^n c(k) - c(N) \right] = \frac{c(i)}{\sum_{j=1}^n c(j)} c(N)$$

De même que la méthode précédente, la répartition proportionnelle vérifie les propriétés de symétrie et de monotonie tout en restant simple à calculer et à expliquer. Elle n'est pas forcément dans le noyau du jeu.

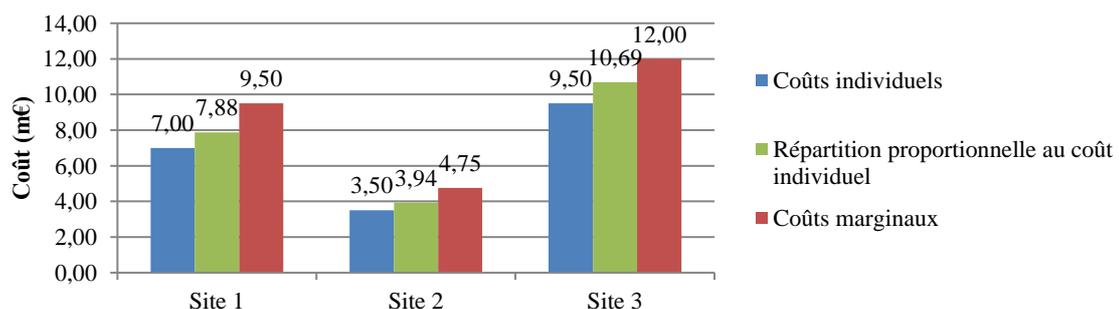
Avec l'exemple simple, l'allocation du chargement de réassurance est visible ci-après.

### Allocation avec méthode proportionnelle : répartition proportionnelle



Avec cette allocation, les résultats en termes de coût total par site sont présentés ci-dessous.

### Méthode proportionnelle : répartition proportionnelle



Par rapport aux méthodes suivantes, elle a plutôt tendance à impacter plus fortement les gros sites par rapport aux petits. Le site 3 de l'exemple concentre ainsi 47.5% du chargement de réassurance avec cette méthode là où les méthodes suivantes se situent entre 31 et 40%.

Dans le problème qui nous intéresse ici, cette méthode ne fait pas de distinction entre les sites qui font plus jouer la réassurance que les autres, en raison de leur localisation dans des zones plus susceptibles de causer de gros événements qui toucheront la priorité du traité. Par exemple, si des sites sont touchés souvent mais par de petits événements, ils peuvent se voir allouer un chargement équivalent à des sites qui sont touchés par de gros événements rares.

Autrement dit, l'hypothèse sous-jacente est que le chargement de réassurance est complètement proportionnel à la moyenne des pertes Cat de chaque site. Ceci n'est pas vérifié a priori puisque la réassurance introduit des non-linéarités en raison de la portée et la priorité du traité.

### *3. Méthodes proportionnelles : égale répartition et répartition proportionnelle des coûts non-marginaux*

De la même manière que les méthodes précédentes ajoutaient au coût individuel un chargement de réassurance, ces deux méthodes utilisent le coût marginal en lieu et place du coût individuel dans les méthodes précédentes.

Ces méthodes vérifient aussi les propriétés de symétrie, monotonie. Elles ne sont toujours pas nécessairement dans le noyau mais sont simples à calculer.

Ainsi, on peut retirer de façon égale la différence entre la somme des coûts marginaux et le coût total aux coûts marginaux respectifs de chaque site. C'est la méthode d'égale répartition des coûts non-marginaux.

$$x_i = CM(i) + \frac{1}{n} \left[ c(N) - \sum_{k=1}^n CM(k) \right]$$

Dans le cadre de l'étude de réassurance, le coût marginal est constitué du coût individuel auquel s'ajoute le surplus de réassurance généré par l'ajout du site en portefeuille. Cette méthode revient donc à effectuer une égale répartition du chargement de réassurance ajustée de la différence entre le chargement de réassurance supplémentaire généré par l'ajout du site en portefeuille et la moyenne des chargements de réassurance supplémentaires de l'ensemble des sites.

Ainsi, un site inutile (qui ne génère pas de chargement de réassurance supplémentaire) paiera moins que dans l'exemple 1 ci-dessus. Il paiera même un coût inférieur à son coût individuel et ne sera pas dans le noyau du jeu.

A l'inverse, un gros site paiera davantage que dans le cas de l'égale répartition du coût total, ce qui paraît plus juste étant donné qu'il est à l'origine d'un chargement de réassurance moyen plus élevé.

Au global, cette méthode avantage plutôt les petits sites, qui peuvent sortir du noyau du jeu si leur coût marginal est suffisamment plus faible que celui de la moyenne des autres sites.

On peut aussi allouer le coût total de l'ensemble proportionnellement aux coûts marginaux : c'est la méthode de répartition proportionnelle au coût marginal.

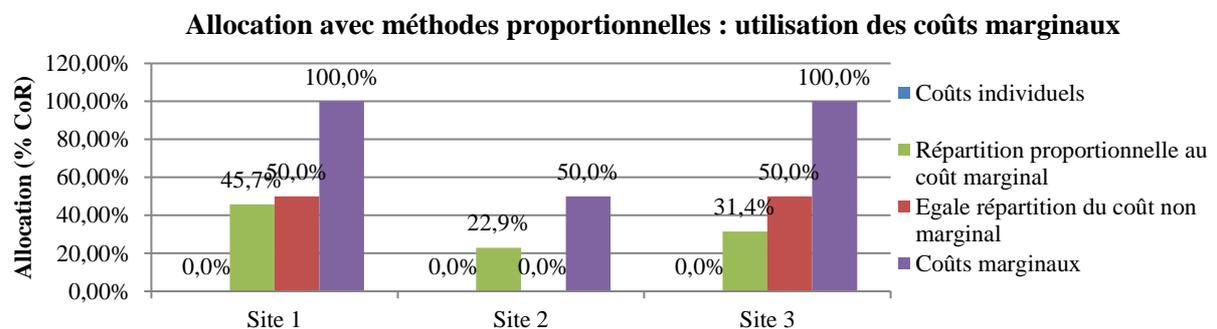
Dans le cadre assurantiel étudié, cette méthode revient à se baser sur le coût (assurance et réassurance) supplémentaire généré par l'ajout d'un site en portefeuille.

$$x_i = CM(i) + \frac{CM(i)}{\sum_{j=1}^n CM(j)} \left[ c(N) - \sum_{k=1}^n CM(k) \right] = \frac{CM(i)}{\sum_{j=1}^n CM(j)} c(N)$$

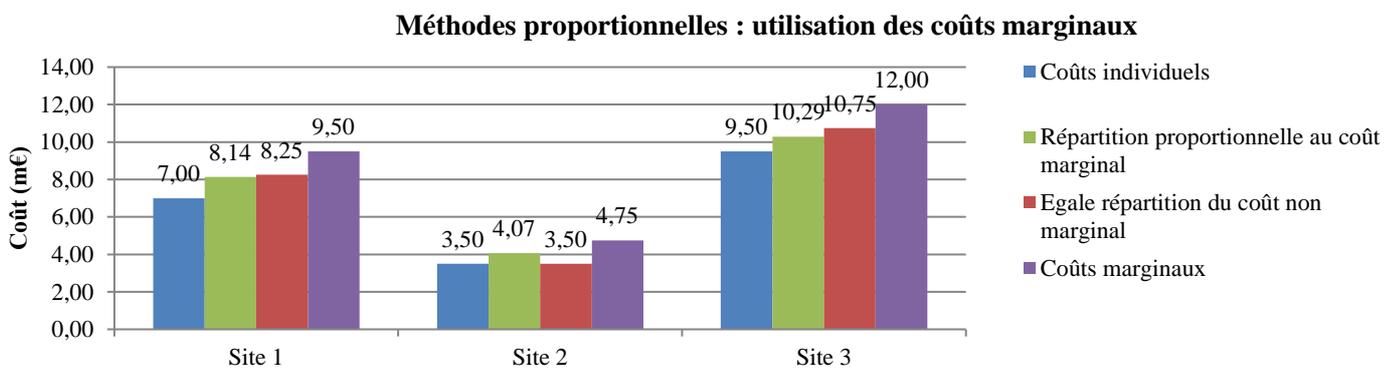
En utilisant le coût marginal plutôt que le coût individuel, on suppose que le site vient s'ajouter à un portefeuille déjà constitué. Ceci se rapproche alors plutôt d'une approche « flux » par rapport à une approche « stock ».

Avec l'exemple cité en début de partie, on obtient une allocation avec chacune des deux méthodes présentées dans le graphe ci-dessous.

Dans cet exemple les allocations sont dans le noyau du jeu (entre le coût individuel et le coût marginal), mais on constate par exemple qu'avec la méthode d'égale répartition du coût non marginal, le site 2 aurait été en dehors du jeu s'il avait généré des récupérations de réassurance plus faibles.



L'allocation du chargement de réassurance donne les résultats suivants en coût total (prime individuelle Cat et chargement de réassurance).



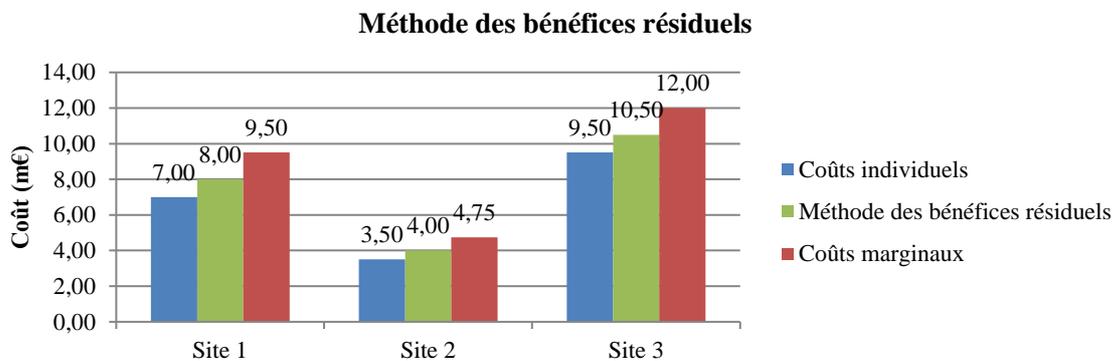
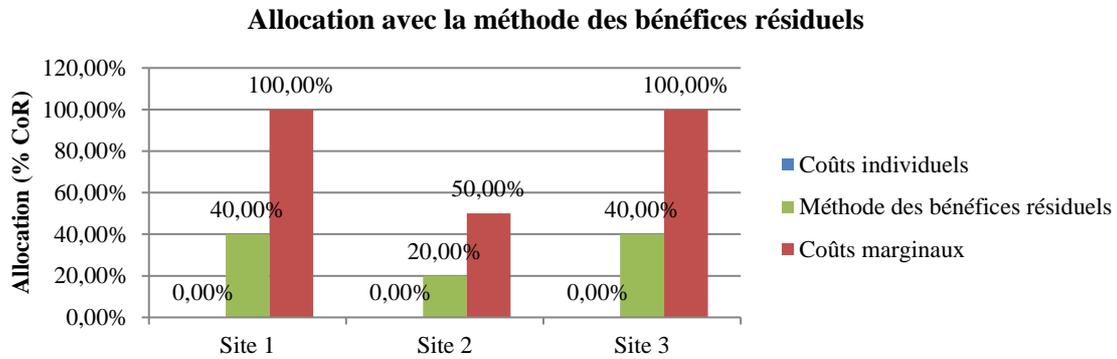
#### 4. Méthode des bénéfices résiduels

La méthode des bénéfices résiduels est une méthode hybride qui utilise l'écart entre le coût individuel et le coût marginal comme pondération pour l'allocation du chargement de réassurance. Il s'agit donc de pondérer par le coût supplémentaire de chaque joueur à la collectivité, ce qui correspond dans notre cas au chargement de réassurance supplémentaire généré par l'ajout de chaque site à la coalition des deux restants.

$$x_i = c(i) - \frac{c(i) - CM(i)}{\sum_{j=1}^n (c(j) - CM(j))} \left[ \sum_{k=1}^n c(k) - c(N) \right]$$

A l'instar des méthodes précédentes, celle-ci vérifie symétrie et monotonicité sans être systématiquement dans le noyau du jeu. Elle reste néanmoins simple à calculer.

Avec l'exemple proposé plus haut, on obtient les résultats ci-dessous en termes d'allocation et de répartition du chargement.



Cette méthode permet de prendre en compte uniquement le chargement de réassurance supplémentaire généré par le site. Cela permet de régler le problème lié par exemple à un site très gros générant très peu de chargement de réassurance.

En outre, cette méthode a la particularité de vérifier la rationalité collective si la fonction de coût vérifie les deux conditions ci-après

$$(1) \forall i, c(i) > CM(i)$$

$$(2) \sum_{k=1}^n CM(k) \leq c(N)$$

Une démonstration de cette propriété se trouve en annexe b.

### 5. Valeur de Shapley

La valeur de Shapley permet de prendre en compte une limitation des méthodes précédentes apportée par le coût marginal. Le coût marginal suppose que le joueur arrive en dernier dans une coalition déjà formée de tous les autres joueurs. Il est donc lié à l'ordre de formation de la coalition. Cependant, rien n'indique que le coût de l'introduction de ce joueur dans la coalition serait la même si le joueur entraînait dans une coalition plus petite, avant d'autres joueurs.

La valeur de Shapley considère ainsi l'ensemble des ordres de formation de coalition possibles, en calculant à chaque fois le chargement de réassurance supplémentaire (ou gain en théorie des jeux) apportée par le joueur à cet ordre de formation. Elle accorde la même importance à chaque ordre de formation en attribuant  $\frac{1}{n!}$  à cet ordre, comme indiqué dans la formule suivante :

$$x_i = c(i) - \frac{1}{n!} \sum_s (s-1)! (n-s)! [c(S\{i\}) + c(i) - c(S)]$$

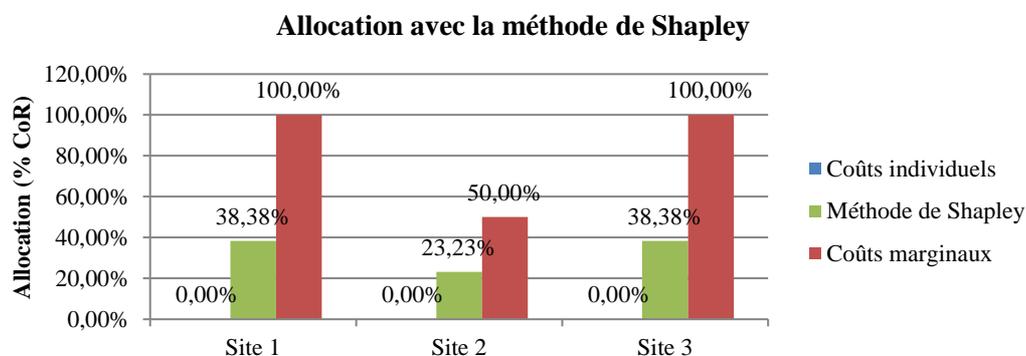
Cette méthode apparaît donc comme la plus juste car elle prend en compte toutes les possibilités d'ordonnement de l'arrivée des joueurs et ce qu'ils apportent à la coalition déjà formée. De plus, elle est dans le noyau du jeu dès qu'il est non vide, vérifie l'additivité, la symétrie et se calcule par une formule fermée.

1er arrivé	2e arrivé	3e arrivé	Site 1	Site 2	Site 3
1	2	3	0,00	0,00	2,50
1	3	2	0,00	1,25	1,25
2	1	3	0,00	0,00	2,50
2	3	1	2,50	0,00	0,00
3	1	2	1,25	1,25	0,00
3	2	1	2,50	0,00	0,00
		c(i)	7,00	3,50	9,50
		<b>xi</b>	<b>8,04</b>	<b>3,92</b>	<b>10,54</b>

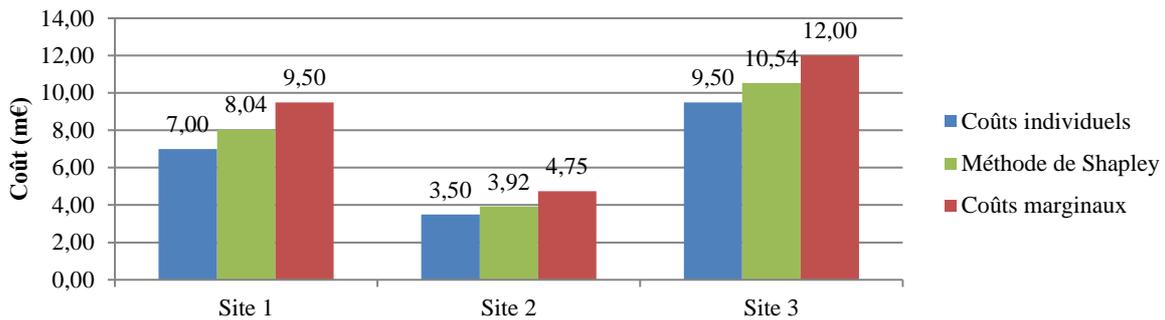
**Figure 24 :** Calcul de la valeur de Shapley pour les trois sites de l'exemple. A chaque ligne correspond la constitution du groupe de trois éléments dans une configuration possible. Pour chaque ligne, on calcule par site le coût de réassurance engendré par l'arrivée du site dans la configuration existante. Par exemple, sur la deuxième ligne, la valeur du Site 3 correspond au coût de réassurance engendré par l'arrivée du site 3 dans la coalition composée du seul Site 1.

Dans le cadre de l'exemple utilisé, on a calculé les résultats de la valeur de Shapley par site en distinguant la totalité des configurations possibles par site. Ils sont regroupés dans le tableau ci-dessus et les valeurs des premières lignes du tableau (Figure 24) correspondent au chargement de réassurance supplémentaire apporté par chaque site à la coalition formée avant lui.

Au final, on obtient les allocations et valeurs présentées dans les graphes suivants.



### Méthode de Shapley



Cette méthode semble être la meilleure d'un point de vue théorique mais pose un sérieux sujet de quantité de calcul. En effet, le nombre d'opérations nécessaires au calcul de valeurs de Shapley pour un ensemble de  $n$  sites évolue en  $2^n$  (cf. démo en annexe c).

Dans le cas qui nous intéresse, le portefeuille étant constitué d'environ  $10^5$  sites, cela correspond à un nombre de  $2^{100000}$ , soit environ  $10^{30103}$ , ce qui est impossible à calculer avec une machine normale (la fonction puissance elle-même ne permet pas de calculer ce nombre avec un logiciel classique type excel ou R).

### 6. Méthode du nucléole

La méthode du nucléole est une méthode itérative qui cherche à calculer un maximum du « degré de satisfaction » pour l'ensemble des coalitions possibles. Plus précisément, on définit l'excédent de la coalition  $S$  par :

$$e(X, S) = c(S) - \sum_{i \in S} x_i$$

Ce nombre est la mesure du « degré de satisfaction » de la coalition  $S$ . La coalition  $S$  peut être n'importe quelle combinaison de joueurs. L'ensemble de ces mesures est noté  $e(X)$ , de la plus petite à la plus grande. Le nucléole est alors la répartition  $X^*$  qui maximise lexicographiquement la valeur  $e(X)$ .

Dans le cadre de l'étude, le degré de satisfaction est négatif car le chargement de réassurance augmente avec la taille de la coalition et l'allocation proposée (dans les  $x_i$ ) est celle du chargement de la coalition composée de tous les joueurs.

La comparaison lexicographique est définie comme une comparaison, entre deux vecteurs, des premiers éléments. S'ils sont égaux, on compare les suivants (c'est aussi le principe d'ordonnement du dictionnaire).

Par exemple, une comparaison lexicographique des allocations données par les méthodes de Shapley et la méthode proportionnelle donnerait les étapes suivantes :

- Les  $x_i$  sont simplement les allocations données par les deux méthodes, coût d'assurance et chargement de réassurance inclus.

$x_i$	Proportionnel	Shapley
<b>Site 1</b>	8,14	7,96
<b>Site 2</b>	4,07	4,08
<b>Site 3</b>	10,29	10,46

- On construit les différentes coalitions possibles, c'est-à-dire les combinaisons du jeu.
- Pour chaque combinaison, on calcule le coût associé  $c(S)$  qui contient coût d'assurance et chargement de réassurance.

S (n° de site)	c(S)
<b>1</b>	7,00
<b>2</b>	3,50
<b>3</b>	9,50
<b>1,2</b>	10,50
<b>1,3</b>	17,25
<b>2,3</b>	13,00
<b>1,2,3</b>	22,50

- On soustrait à ces coûts la somme des valeurs allouées par les méthodes proportionnelle et de Shapley.

S (n° de site)	c(S)	$e(X,S)_{\text{Shapley}}$	$e(X,S)_{\text{Proportionnelle}}$
<b>1</b>	7,00	-0,96	-1,14
<b>2</b>	3,50	-0,58	-0,57
<b>3</b>	9,50	-0,96	-0,79
<b>1,2</b>	10,50	-1,54	-1,71
<b>1,3</b>	17,25	-1,17	-1,18
<b>2,3</b>	13,00	-1,54	-1,36
<b>1,2,3</b>	22,50	0,00	0,00

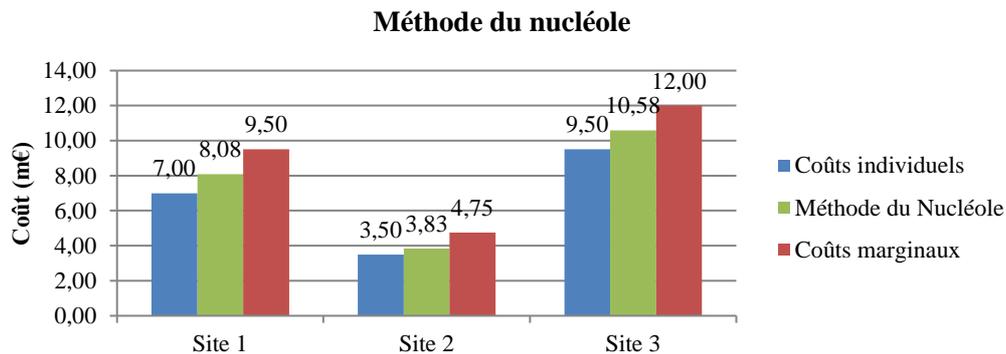
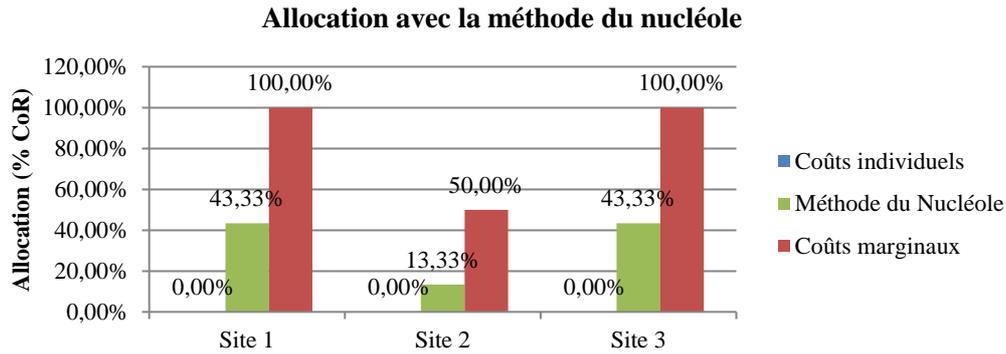
- Finalement, on ordonne les  $e(X,S)$  de chacune des méthodes et on les compare dans l'ordre lexicographique.

	$e(X,S)_{\text{Shapley}}$	$e(X,S)_{\text{Proportionnelle}}$
<b>rang 1</b>	0,00	0,00
<b>rang 2</b>	<b>-0,58</b>	<b>-0,57</b>
<b>rang 3</b>	-0,96	-0,79
<b>rang 4</b>	-0,96	-1,14
<b>rang 5</b>	-1,17	-1,18
<b>rang 6</b>	-1,54	-1,36
<b>rang 7</b>	-1,54	-1,71

- Dans notre exemple, le premier terme est égal pour chacune des deux méthodes (0). La comparaison lexicographique se fait donc au deuxième rang (deuxième plus grande valeur) et la méthode proportionnelle (-0,57) est donc considérée comme meilleure au sens lexicographique que la méthode de Shapley.

Un des inconvénients de la méthode du nucléole est qu'il nécessite aussi le calcul de toutes les configurations possibles puis un processus d'optimisation, ce qui est encore plus lourd que la méthode de Shapley. Il nécessite une résolution numérique, contrairement à la méthode de Shapley qui se calcule à l'aide d'une formule fermée.

Dans le graphe ci-dessous, on a calculé les résultats en utilisant la méthode du nucléole pour l'exemple de ce chapitre.



On peut vérifier que cette allocation est meilleure que celle donnée par Shapley au sens lexicographique. Sur la base de l'exemple précédent, on construit le tableau des  $e(X,S)$  ordonné pour la comparaison lexicographique. On constate que la comparaison se fait aussi à la deuxième valeur, et la méthode du Nucléole est meilleure que les deux autres avec ce critère (-0,33 plus grand que les deux autres valeurs).

	$e(X,S)_{\text{Shapley}}$	$e(X,S)_{\text{Proportionnelle}}$	$e(X,S)_{\text{Nucléole}}$
<b>rang 1</b>	0,00	0,00	0,00
<b>rang 2</b>	-0,58	-0,57	-0,33
<b>rang 3</b>	-0,96	-0,79	-1,08
<b>rang 4</b>	-0,96	-1,14	-1,08
<b>rang 5</b>	-1,17	-1,18	-1,42
<b>rang 6</b>	-1,54	-1,36	-1,42
<b>rang 7</b>	-1,54	-1,71	-1,42

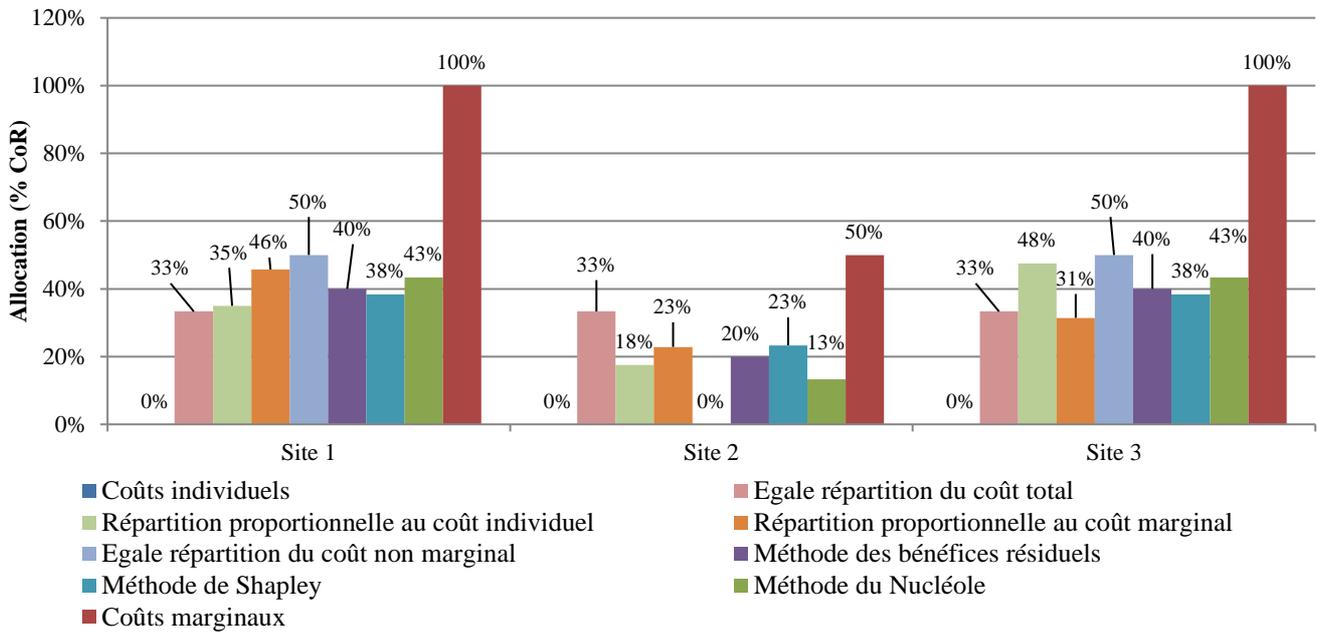
Un inconvénient de la méthode est néanmoins que la notion de justesse de l'allocation finale ne paraît pas intuitive car elle se base uniquement sur des comparaisons lexicographiques.

## 7. Synthèse des méthodes

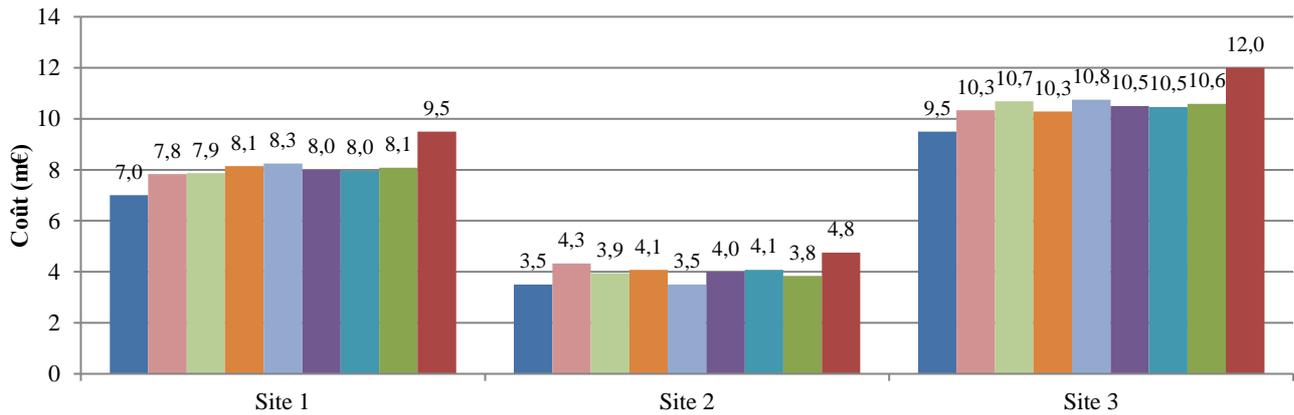
Pour conclure cette partie plus théorique, on constate que les méthodes qui paraissent les plus justes sont les méthodes de Shapley et la méthode du nucléole. Néanmoins, la complexité du calcul ainsi que des explications nécessaires à la mise en place de telles méthodes plaident plutôt pour l'utilisation des méthodes plus simples que sont les méthodes précédentes (bénéfices résiduels, méthode des coûts marginaux).

Dans les graphes ci-dessous, on a synthétisé l'ensemble des résultats des différentes méthodes.

### Allocation avec les différentes méthodes présentées



### Coût avec les différentes méthodes présentées



Dans cet exemple simplifié, on constate que la méthode de répartition proportionnelle au coût individuel et la méthode d'égale répartition du coût non marginal avantagent le petit site (site 2), alors que la méthode de répartition proportionnelle au coût marginal et l'égale répartition du coût total avantagent le gros site (site 3).

La méthode des bénéfices résiduels semble relativement proche de la valeur de Shapley, ce qui tend à plaider pour cette méthode d'un point de vue théorique, plus simple que la méthode de Shapley mais plus proche que les autres méthodes d'une approche équitable pour les assurés.

### III- ALLOCATION DU CHARGEMENT DE REASSURANCE

L'objectif de cette partie est de proposer une méthode de répartition du chargement de la réassurance au niveau de la tarification Cat, site par site.

#### a. Méthode proportionnelle

##### i. Présentation de la méthode

La méthode proportionnelle consiste à considérer uniquement les coûts individuels des sites pour allouer le chargement de réassurance. Les coûts individuels ayant été calculés par la modélisation, cela revient donc à allouer le chargement de réassurance selon la prime Cat du site par rapport au reste du portefeuille.

Plus précisément :

- On calcule la perte moyenne du site, site par site  $i$ , sur l'ensemble des années  $j$  et des événements  $k$ .

$$Perte(i) = \frac{1}{10\,000} \sum_{j=1}^{10\,000} \sum_{k=1}^{Nb\_event(j)} Perte_k(i)$$

- On répartit le chargement de réassurance selon les pertes individuelles dans le reste du portefeuille.

$$CoR(i) = CoR * \frac{Perte(i)}{\sum_{k=1}^{Nb\_sites} Perte(k)}$$

Cette méthode ne prend pas en compte les récupérations potentielles de réassurance des sites car leurs coûts individuels sont toujours inférieurs à la rétention du traité considéré.

##### ii. Analyse des résultats et conclusions



**Figure 25 :** Cartes de la répartition des pertes individuelles par municipalité mexicaine (% du total). Les résultats site par site ont été agrégés par municipalité pour faciliter la visualisation.

En observant la carte de répartition des pertes individuelles par municipalité, on constate qu'un petit nombre de municipalités concentrent une part importante des pertes totales. Ceci

s'explique par plusieurs facteurs : la somme assurée qui est très importante dans ces zones, mais aussi des zones exposées à un fort aléa.

En effet, comme l'indique le tableau ci-dessous, 15 municipalités rassemblent 76% des pertes modélisées ainsi que près d'un quart de la somme assurée sur un total de 1229 municipalités.

Municipalité	CoR (% total)	Somme assurée (% total)
Benito Juarez	20,74%	1,890%
Cozumel	13,17%	2,207%
La Paz	8,55%	2,593%
Acapulco de Juarez	7,48%	1,581%
Isla Mujeres	5,22%	0,503%
Guadalajara	3,55%	2,509%
Felipe Carrillo Puerto	3,47%	0,817%
Zapopan	3,01%	2,738%
Iguala de la Independencia	2,85%	0,193%
Puerto Vallarta	1,70%	1,388%
Ahome	1,50%	0,444%
Compostela	1,26%	1,102%
Altamira	1,17%	0,876%
Azcapotzalco	1,16%	4,925%
Colima	1,15%	0,376%
<b>TOTAL</b>	<b>76,0%</b>	<b>24,1%</b>

**Tableau 2** : classement des 15 principales municipalités expliquant plus des trois quarts des pertes moyennes (et donc du chargement de réassurance pour cette méthode) et près d'un quart de la somme assurée.

En normalisant les pertes moyennes par les sommes assurées, on s'affranchit de la somme assurée comme facteur explicatif pour ne garder que l'aléa modélisé du péril ouragan. Ainsi, on observe que les taux de perte les plus élevés sont localisés sur les zones côtières du Mexique (cf. Figure 26).



**Figure 26** : Cartes des taux moyen de pertes individuelles par municipalité mexicaine en % de somme assurée). Les résultats site par site ont été moyennés par municipalité pour faciliter la visualisation.

Cette méthode permet de répartir simplement le chargement de réassurance en ajoutant un petit montant proportionnel à la perte moyenne (prime Cat). Elle est donc simple à appliquer. Néanmoins, plusieurs critiques peuvent être effectuées :

- La méthode fait l'hypothèse que chaque site possède un impact proportionnel à sa perte moyenne sur le chargement de réassurance. Ceci est a priori faux car l'impact dépend au moins de la forme de la distribution de la perte de ce site. En effet, un site dont la queue de distribution des pertes est plus épaisse, autrement dit qui est impacté plus rarement mais plus sévèrement, aura plus d'impact qu'un site présentant de fréquentes mais petites pertes. A moyennes égales, les effets des deux sites peuvent donc différer.
- La méthode ne tient pas compte des phénomènes d'accumulation. En effet, l'accumulation d'un grand nombre de sites (ou de sites de forte somme assurée) dans une zone donnée devrait avoir plus d'impact que si les sites sont répartis sur plusieurs zones du territoire. Cela se traduit indirectement par un chargement de réassurance plus important (car le réassureur prendra en compte cette information) mais pas directement sur la répartition en elle-même.
- Au global donc, la segmentation proposée de l'exercice semble limitée en ne ciblant pas vraiment les sites les plus responsables du chargement de réassurance.

## **b. Méthode des bénéfices résiduels**

### i. Présentation et analyse des résultats d'allocation

#### *1. Présentation*

La première méthode d'allocation du chargement de réassurance cherche à attribuer à chaque site son chargement de réassurance à partir d'une répartition des récupérations de réassurance proportionnelle à la perte de chaque site par événement. Le chargement de réassurance étant proportionnel aux récupérations de réassurance (prime pure de réassurance), leur utilisation est justifiée. Plus précisément, la démarche utilise l'*event set* de résultats de façon suivante :

- Agrégation des résultats par événement sur 10 000 années simulées : cela permet de calculer la perte totale de l'événement  $k$  de l'année  $j$  à partir des pertes individuelles de chaque site  $i$ .

$$Loss = (Loss_1^1, Loss_j^k, \dots, Loss_{10\,000}^1), \quad \text{avec } Loss_j^k = \sum_{i \in k} Loss_j^k(i)$$

- Calcul des récupérations de réassurance associées à chaque événement  $k$  :

$$Récupérations_k = \text{Min}(\text{Max}(Loss_j^k - \text{Rétention}; 0); \text{Limite})$$

- Répartition de ces récupérations de réassurance par site  $i$ , proportionnellement à la part de perte du site dans la perte totale des sites touchés :

$$Récupérations_k(i) = Récupérations_k * \frac{Loss_j^k(i)}{Loss_j^k}$$

- Calcul de la moyenne des récupérations site par site  $i$ , pour chaque année  $j$  ( $Nb\_event(j)$  correspond au nombre d'événements de l'année  $j$ ) :

$$Récupérations(i) = \frac{1}{10\,000} \sum_{j=1}^{10\,000} \sum_{k=1}^{Nb\_event(j)} Récupérations_k(i)$$

- Calcul de l'écart type des récupérations  $SDRécup$  site par site  $i$  :

$$SDRécup(i) = \sqrt{\frac{1}{10\,000} \sum_{j=1}^{10\,000} \left( \sum_{k=1}^{Nb\_event(j)} Récupérations_k(i) - Récupérations(i) \right)^2}$$

Ces récupérations et écart type seront ensuite utilisés selon différentes méthodes.

La première est la plus simple et consiste à répartir le chargement calculé au III.A.iii à l'aide des récupérations de réassurance site par site :

$$CoR(i) = CoR * \frac{Récupérations(i)}{\sum_{k=1}^{Nb\_sites} Récupérations(k)}$$

Bien que légèrement différente de la méthode des bénéfices résiduels du a, on n'utilise pas ici la prime Cat pour allouer le CoR et on se focalise sur les récupérations de réassurance.

## 2. Analyse des résultats

De même que pour les pertes Cat lors de la tarification Cat, on observe assez logiquement une répartition des récupérations de réassurance site par site localisée au niveau des côtes mexicaines, au niveau des zones où frappent les ouragans (cf. Figure 29).



**Figure 27 :** Cartes de la répartition des récupérations de réassurance par municipalité mexicaine (% du total). Les résultats site par site ont été agrégés par municipalité pour faciliter la visualisation.

De plus, on constate aussi que les zones dans lesquelles les récupérations sont les plus élevées en proportion de la somme assurée se trouvent aussi dans les zones côtières, ce qui montre qu'il s'agit bien à la fois d'une concentration de risque et d'aléa dans ces zones (cf. Figure 30).



**Figure 28 :** Cartes des taux moyen de récupérations de réassurance par municipalité mexicaine en % de somme assurée). Les résultats site par site ont été moyennés par municipalité pour faciliter la visualisation.

Municipalité	CoR méthode bénéf. résiduels (% total)	CoR méthode propor. (% total)
Benito Juarez	22,34%	20,74%
Cozumel	13,92%	13,17%
Acapulco de Juarez	8,22%	7,48%
La Paz	8,16%	8,55%
Isla Mujeres	5,55%	5,22%
Felipe Carrillo Puerto	3,67%	3,47%
Iguala de la Independencia	3,17%	2,85%
Guadalajara	2,71%	3,55%
Zapopan	2,36%	3,01%
Ahome	1,49%	1,50%
Puerto Vallarta	1,42%	1,70%
Azcapotzalco	1,14%	1,16%
Altamira	1,13%	1,17%
Compostela	1,12%	1,26%
José Azueta	1,03%	1,04%
<b>TOTAL</b>	<b>77,4%</b>	<b>75,9%</b>
<b>TOTAL top 4</b>	<b>52,6%</b>	<b>49,9%</b>

**Tableau 3 :** comparaison de l'allocation du chargement de réassurance en pourcentage du total entre la méthode proportionnelle et la méthode des bénéfiques résiduels sur les 15 municipalités qui rassemblent les plus gros pourcentages.

Comme présenté dans le tableau ci-dessus, on constate que la méthode des bénéfiques résiduels cible plus particulièrement les municipalités qui se distinguaient déjà comme cumulant les plus gros pourcentages du total des pertes moyennes modélisées.

En effet, alors qu'avec la méthode proportionnelle ces 15 municipalités regroupaient légèrement moins de 76%, ce pourcentage s'élève à 77,4% avec la méthode des bénéfiques résiduels. L'écart est plus élevé pour les quatre plus grosses municipalités, avec 53% pour la méthode des bénéfiques résiduels au lieu de 50% avec la méthode proportionnelle.

La méthode des bénéfiques résiduels permet ainsi de corriger les défauts de la méthode proportionnelle, et cela s'observe sur une répartition du chargement de réassurance plus importante sur les municipalités les plus risquées en termes d'aléa et d'accumulations (8% du portefeuille total pour les quatre premières municipalités).

ii. Construction de modèles simples de prédiction du chargement de réassurance

Les résultats sur l'ensemble du portefeuille présentent une première méthode pour allouer le chargement de la réassurance sur le portefeuille existant. Néanmoins, cela ne permet pas de calculer le chargement qui serait associé à la souscription d'un nouveau risque dans le portefeuille existant.

Dans l'idéal, il conviendrait de refaire tourner la chaîne de modélisation sur le nouveau portefeuille contenant le site supplémentaire. Ceci introduit une série de contraintes matérielles, notamment en matière de temps et de complexité de calcul. Différentes méthodes de prédiction sont donc testées ici afin d'obtenir rapidement un résultat approximatif quant au chargement de réassurance d'une nouvelle souscription.

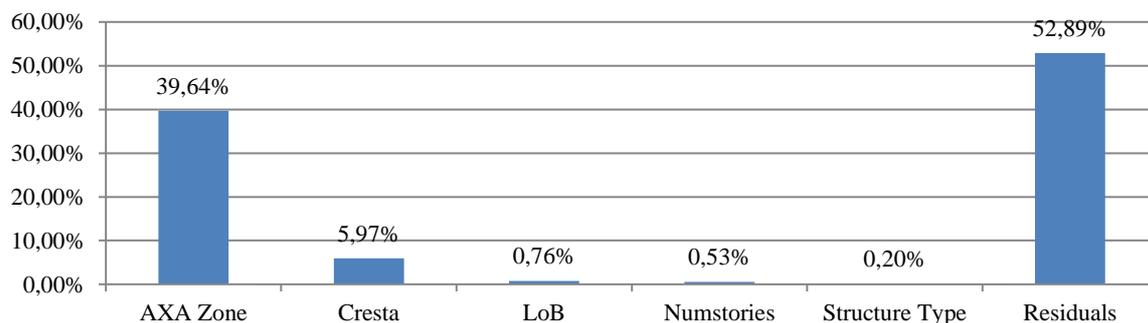
## 1. Régression simple

On effectue une régression linéaire simple sur les résultats avec les variables présentes dans le portefeuille. On cherche à prédire la variable `reins_rate`, qui est le taux de somme assurée que représente le chargement de réassurance moyen.

Parmi ces variables, cinq se démarquent comme étant explicatives (cf. Annexe d). Ce sont :

- AXA Zone (variable catégorielle) : l'indicateur de zone de risque ouragan dans laquelle se trouve le site. Cet indicateur se présente sous la forme de catégories basées sur des cartes de vitesses de vent moyennes mesurées dans ces zones.
- Cresta (variable catégorielle) : développé par Swiss Re, les Cresta (*Catastrophe Risk Evaluation and Standardizing Target Accumulations*) sont un découpage géographique des territoires utilisé par le secteur assurantiel. Au Mexique, ce découpage correspond aux zones administratives des municipalités (~2400 municipalités).
- Ligne de métier (LoB) : fait la distinction entre les lignes commerciale, résidentielle, industrielle et municipale (collectivités).
- Nombre d'étage (Numstories) : varie de 1 à 57.
- Type de structure (Structure type) : type de structure du site affecté, notamment ouvrages en maçonnerie, en béton, béton armé ou encore structure en acier.

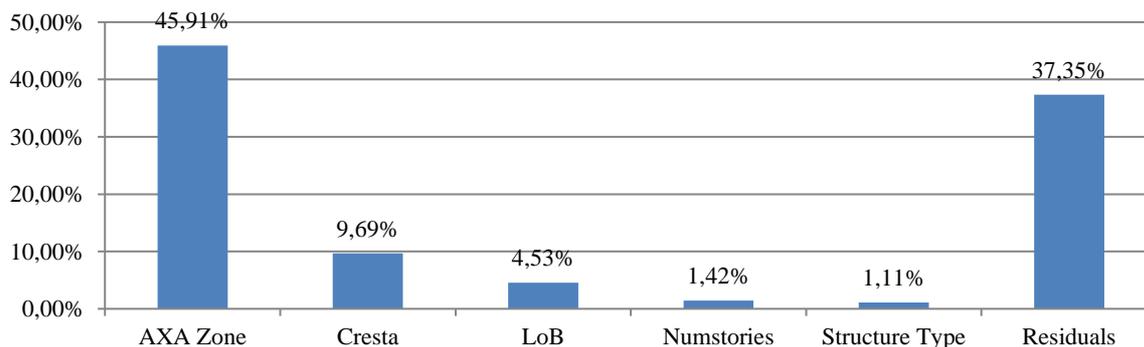
**Variance expliquée par le modèle linéaire**



Il semble cependant possible d'améliorer les résultats étant donné la forte part de variance résiduelle (53%). Une première piste est de supprimer les sites ne générant que de faibles pertes. En effet, les taux de réassurance pour les sites de petite somme assurée introduisent un bruit plus élevé en étant facilement sensibles à une petite variation du chargement de réassurance alors qu'ils ne représentent qu'une petite partie du chargement total à allouer.

Ainsi, en supprimant les sites qui présentent des pertes moyennes inférieures à 100€, on conserve 97% du chargement de réassurance (et 96% de la perte moyenne).

**Variance expliquée par le modèle linéaire**



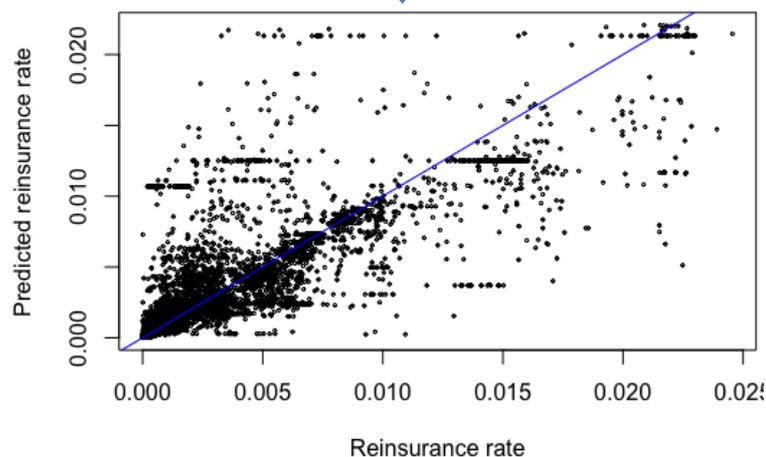
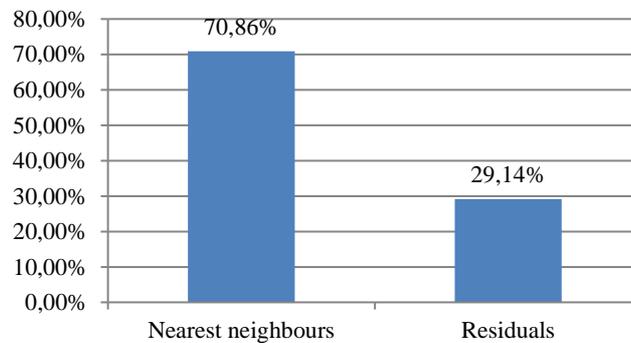
La variance non expliquée descend à 37%, ce qui est un gain de plus de 15 points par rapport à la méthode précédente. Néanmoins, il convient alors d'être capable de prédire les faibles pertes. Ce modèle ne convient donc que pour les sites de forte valeur assurée dont on est sûr à l'avance qu'ils généreront des pertes moyennes supérieures à 100 €.

## 2. Moyenne géographique par plus proches voisins

Au vu de l'importance des zones de vitesse de vent dans les résultats (AXA zone), on a cherché une méthode simple pour prendre en compte plus précisément l'information géographique sur le portefeuille filtré sur les pertes supérieures à 100€.

Pour chaque site, on sélectionne les 5 sites les plus proches et on calcule la moyenne des taux de réassurance de ces 5 sites. L'objectif ici est alors de tarifier un nouveau site à partir de la moyenne de ses plus proches voisins.

**Variance expliquée par plus proches voisins**



En attribuant ainsi à chaque site un taux de réassurance moyen basé sur les plus proches voisins, on améliore encore la variance expliquée en montant celle-ci à **71%**. Ceci représente un gain de 8 points par rapport à la régression linéaire classique.

Ceci indique que prendre en compte l'information géographique de façon plus fine que les cartes de vitesse de vent moyenne permet d'améliorer encore la qualité de la prédiction. Par ailleurs les résultats de prédiction des taux de réassurance site par site apparaissent visuellement plus satisfaisants.

Remarque : une optimisation sur le nombre de plus proches voisins à sélectionner a été effectuée mais le gain de variance lié à cette optimisation par rapport à un choix arbitraire de 5 est limité.

### iii. Construction d'un modèle de prédiction du chargement de réassurance de type arbre de régression

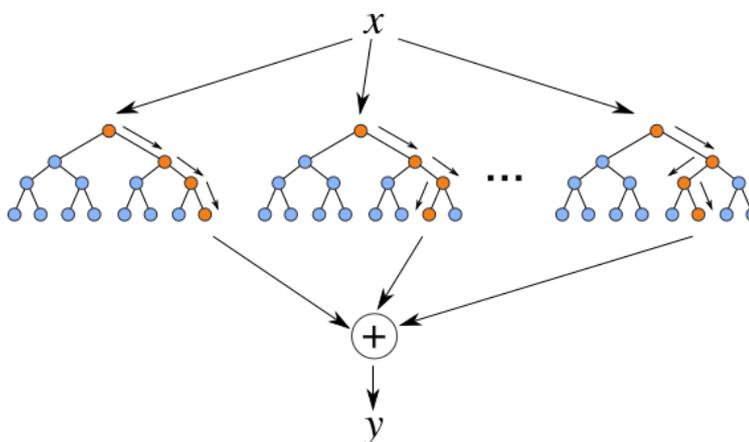
Afin de prendre en compte l'information géographique directement à l'aide des coordonnées et de garder les autres variables, on utilise une méthode de *random Forest*.

#### 1. *Arbre de régression : forêts aléatoires*

La notion de forêt aléatoire a été développée au début des années 2000 comme un approfondissement des méthodes d'arbre de régression pour les variables continues ou de décision pour des variables discrètes (cf. annexe IX-e). Le principe de la méthode consiste à construire un grand nombre d'arbres basés sur des sous-ensembles des données de départ, de la même manière que le bootstrap, de façon à construire un résultat « moyen » de nombreux arbres.

**Figure 29 :**

Illustration du fonctionnement d'une forêt aléatoire. Plusieurs arbres de décision sont construits, et chacun d'entre eux est utilisé pour la prédiction, en moyennant les résultats obtenus à partir de  $x$  pour obtenir la valeur prédite  $y$ .



Cela permet de s'affranchir de l'instabilité liée à la nature des arbres de régression/décision. En effet, de légères variations des données peuvent avoir des conséquences importantes sur les nœuds de l'arbre. La multiplication du nombre d'arbres permet d'avoir un ensemble de modalités possibles, augmentant ainsi la stabilité.

De plus, cette méthode permet aussi d'améliorer la performance par rapport à un arbre simple sur l'échantillon de test en réduisant le sur-apprentissage potentiel lié à un unique arbre.

Par ailleurs, cette méthode permet de conserver certaines qualités des arbres de régression :

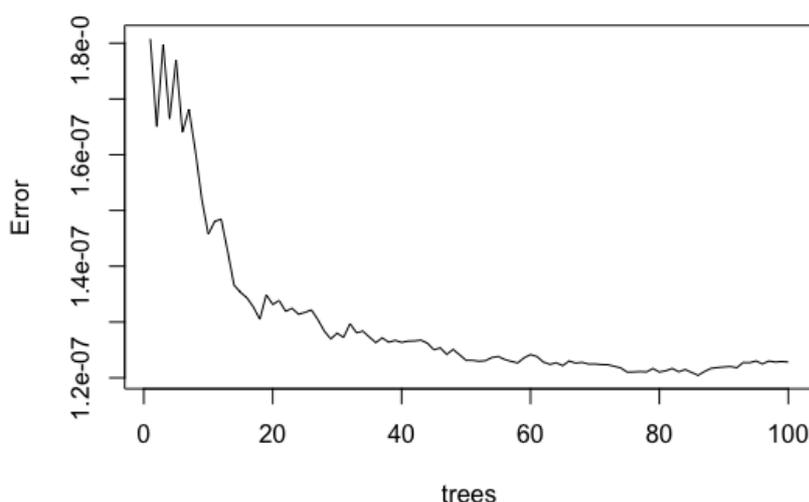
- La nature des prédicteurs : il est possible d'utiliser des prédicteurs qualitatifs ou quantitatifs.
- Distribution : aucune hypothèse de distribution n'est effectuée sur les prédicteurs.

- Hiérarchisation et sélection : la méthode permet de hiérarchiser et de sélectionner les prédicteurs les plus pertinents pour l'étude.

Cette méthode présente néanmoins l'inconvénient majeur de perdre la qualité visuelle d'un arbre simple, et donc complexifie l'interprétation des résultats.

Dans notre cas, cela permet aussi de prendre en compte la segmentation géographique en générant une série de seuils sur les coordonnées latitude et longitude. Effectuer un grand nombre d'arbres permet de construire ainsi des zones géographiques, ce qui est impossible avec les méthodes de régression classiques.

En ce qui concerne le nombre d'arbres, on a sélectionné un nombre suffisant (100) pour faire converger l'erreur de prédiction. On constate sur la Figure 30 qu'à partir de 80 arbres environ, on n'améliore plus la qualité de la prédiction.



**Figure 30 :**

Evolution de l'erreur de prédiction au sens de la somme du carré des résidus en fonction du nombre d'arbres de régression choisis.

## 2. Construction des données

Pour ce type d'algorithme de prédiction, le risque de sur-apprentissage est systématique et il est donc nécessaire d'effectuer une séparation des données en deux parties :

- Un échantillon d'apprentissage (« *train* ») : cet échantillon est utilisé pour calibrer l'algorithme de classification. Il doit regrouper une fraction importante du jeu de données de départ. Dans notre cas, on effectue la sélection d'un échantillon représentant 90% du jeu de données d'origine.
- Un échantillon de test (« *test* ») : cet échantillon permet de tester l'algorithme de prédiction calibré sur l'échantillon *train*. En général (et logiquement), les résultats obtenus sur l'échantillon *test* sont un peu moins bons que ceux obtenus sur l'échantillon *train*. Dans le cas qui nous concerne, on a sélectionné 10% des données du jeu d'origine.

Pour des raisons de capacités de calcul, les arbres de régression de type forêts aléatoires prennent en compte les variables catégorielles dans un nombre limité de modalités. En effet, si une variable possède  $n$  valeurs possibles, alors il existe  $2^n - 2$  combinaisons possibles pour l'algorithme. Ainsi, il peut être nécessaire d'effectuer des regroupements de ces modalités si leur nombre est trop important.

Dans notre cas, le nombre d'étages des bâtiments peut s'élever à plus de 50, ce qui pose un double problème de nombre modalités ainsi que de représentativité. Les risques de plus de 10 étages sont en effet présents dans le portefeuille en nombre très limité et la séparation en deux échantillons représentatifs du portefeuille d'origine est alors délicate.

Pour pallier ces deux limitations, on a regroupé les risques de plus de 10 étages dans une même catégorie, de façon à ce qu'elle contienne plus de 300 sites. Enfin, on s'est intéressé aux cinq variables du modèle linéaire ainsi qu'aux coordonnées (latitude, longitude) des sites afin de mieux capter l'information géographique.

### 3. Résultats du modèle sur les taux de réassurance

Les résultats de cette méthode sont beaucoup plus satisfaisants que les modèles linéaires. En effet, sur l'échantillon d'apprentissage, la variance expliquée ( $R^2$ ) des taux de réassurance s'élève à 95% et est de même ordre de grandeur sur l'échantillon test. Par ailleurs, la séparation supervisée entre grandes et petites pertes (<100€) n'est plus nécessaire.

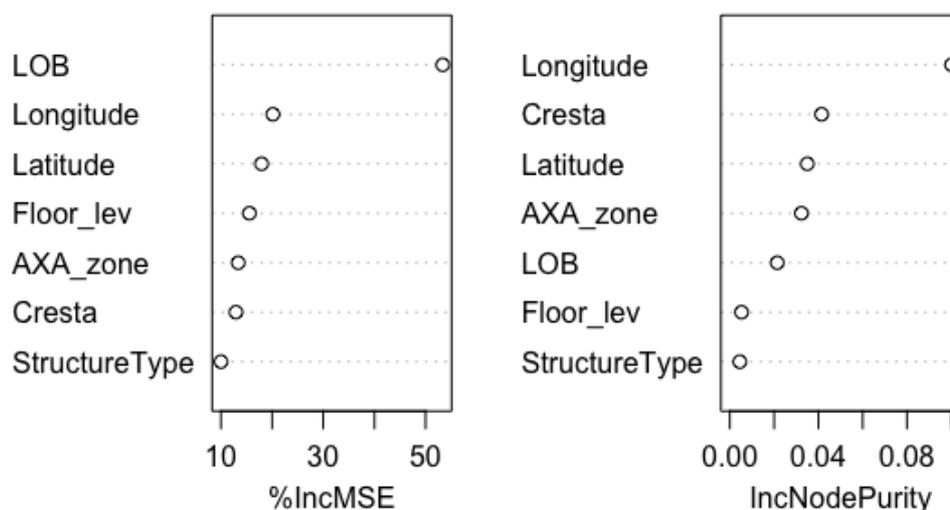


Figure 31 :

Classement des variables explicatives par ordre d'importance dans les arbres de régression. Le graphe de gauche présente l'augmentation d'erreur quadratique moyenne associée à la permutation aléatoire des modalités de chaque variable. Le graphe de droite présente le gain de pureté, mesuré en somme du carré des résidus, associé à chaque variable prédictive.

Les variables les plus explicatives changent d'ordre et deviennent la ligne de métier (LoB) et la position géographique du site, ce qui était attendu au vu des régressions linéaires (cf. Figure 31).

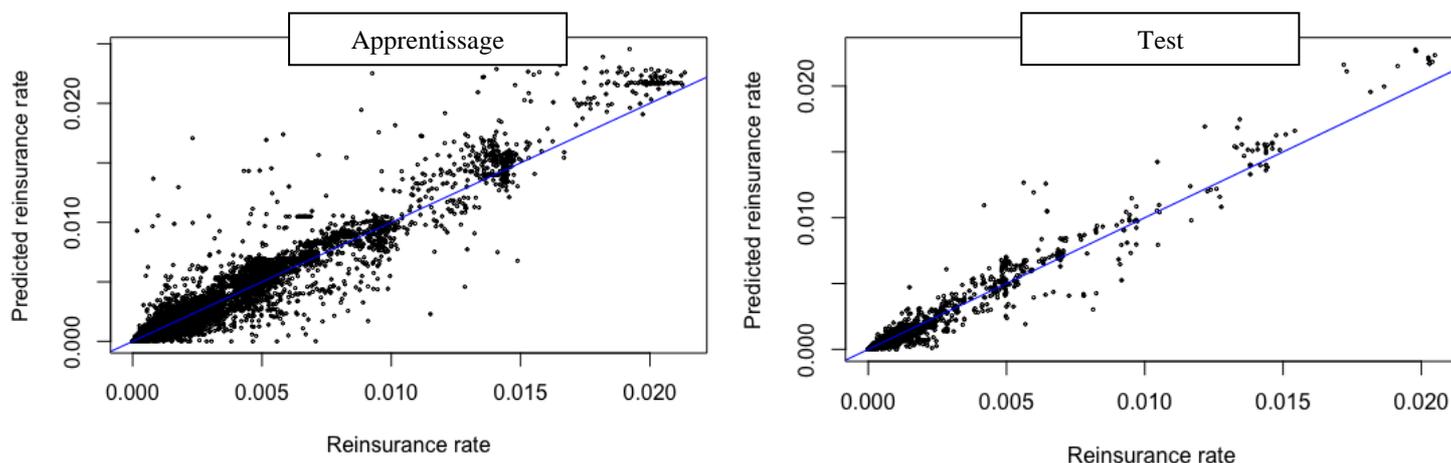
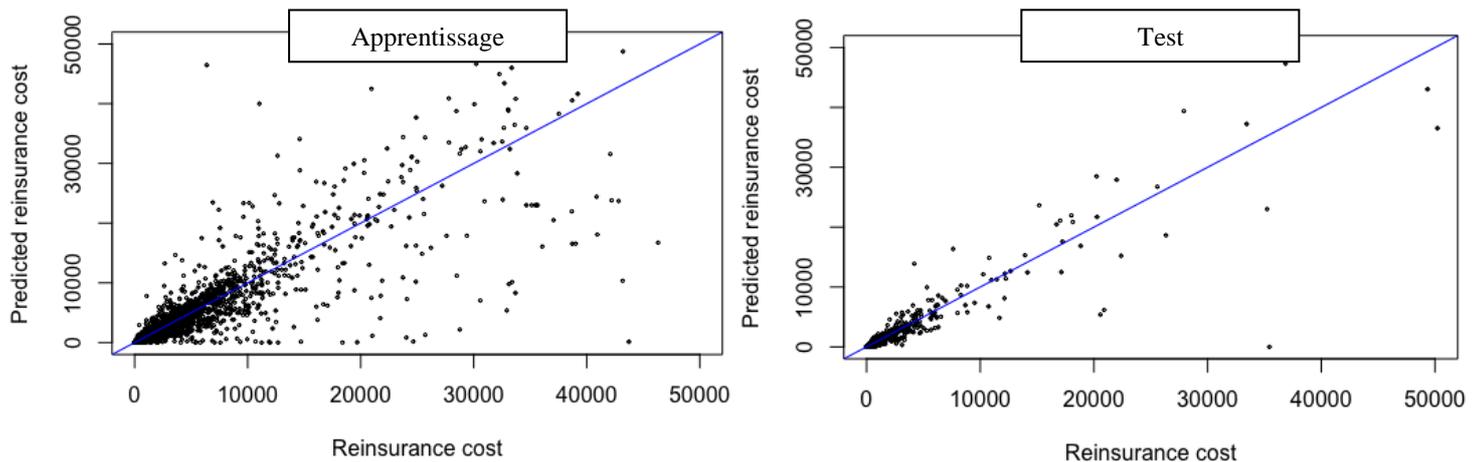


Figure 32 : Graphes des taux de réassurance prédits en fonction des taux à prédire. A gauche, l'échantillon représenté est celui d'apprentissage. A droite, il s'agit de l'échantillon de test.

Ainsi, on constate que les graphes des valeurs prédites en fonction des valeurs à prédire sont beaucoup plus intéressants que ceux des régressions simples. Ce modèle de régression est donc un bon proxy au modèle Cat stochastique.

Au niveau des pertes prédites, on observe une variance expliquée plus faible, de l'ordre de 61%, mais qui reste raisonnable pour un modèle proxy de prédiction du chargement de réassurance aussi bien sur l'échantillon d'apprentissage que de test.

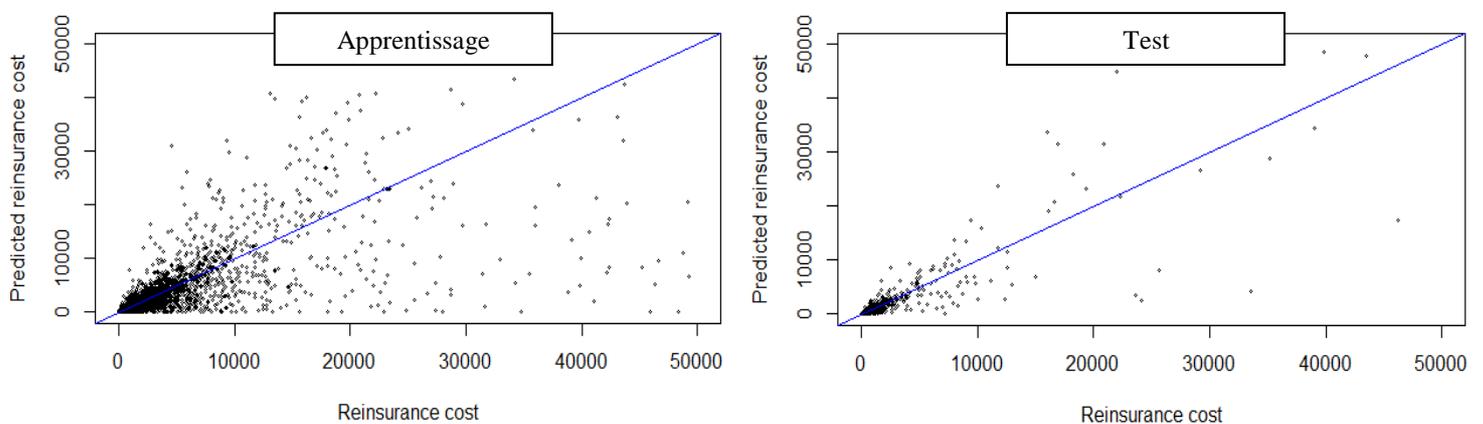


**Figure 33 :** Graphes des coûts de réassurance prédits en fonction des coûts à prédire. A gauche, l'échantillon représenté est celui d'apprentissage. A droite, il s'agit de l'échantillon de test.

#### 4. Résultats du modèle sur les récupérations de réassurance

Plutôt que de passer par les taux de réassurance, on a cherché à prédire directement les récupérations. En introduisant la somme assurée et la franchise comme variables explicatives, on obtient des résultats légèrement plus intéressants qu'avec les taux de réassurance, contrairement à ce qui était observé sur les modèles linéaires.

En effet, la variance expliquée augmente à 64% sur le jeu de test, ce qui est meilleur que la variance expliquée à l'aide de la prédiction sur les taux de réassurance.



**Figure 34 :** Graphes des coûts de réassurance prédits en fonction des coûts à prédire. A gauche, l'échantillon représenté est celui d'apprentissage. A droite, il s'agit de l'échantillon de test.

Avec ce dernier modèle de prédiction, on prend en compte les spécificités de chaque site, la localisation géographique, et on calcule directement un chargement de réassurance pour un traité donné site par site.

## c. Méthode du coût marginal

### i. Présentation

Pour cette méthode, on a cherché à connaître l'effet sur le chargement de réassurance de l'ajout de chaque site dans le portefeuille. Le déroulement de la méthode a été le suivant :

- Suppression du site considéré dans l'*event set* : l'hypothèse faite ici est que le site n'influe pas sur les résultats des autres sites du portefeuille.
- Calcul des récupérations moyennes globales :

$$\text{Récupérations} = \frac{1}{10\,000} \sum_{j=1}^{10\,000} \sum_{k=1}^{Nb\_event(j)} \text{Récupérations}_k, \text{ où}$$

$$\text{Récupérations}_k = \text{Min}(\text{Max}(\text{Loss}_k - \text{Rétention}; 0); \text{Limite}), \text{ avec}$$

$$\text{Loss}_k = \sum_{j \in k} \text{Loss}_k(j)$$

- Calcul des pertes en agréant par événement, privé de  $i$  :

$$\text{Loss}(\setminus i) = (\text{Loss}_1(\setminus i), \dots, \text{Loss}_{10\,000}(\setminus i)), \quad \text{avec } \text{Loss}_k(\setminus i) = \sum_{j \in k \setminus i} \text{Loss}_k(j)$$

- Calcul des récupérations de réassurance privée de  $i$  :

$$\text{Récupérations}(\setminus i) = \frac{1}{10\,000} \sum_{j=1}^{10\,000} \sum_{k=1}^{Nb\_event(j)} \text{Récupérations}_k(\setminus i), \text{ où}$$

$$\text{Récupérations}_k(\setminus i) = \text{Min}(\text{Max}(\text{Loss}_k(\setminus i) - \text{Rétention}; 0); \text{Limite})$$

- Calcul de l'impact sur les récupérations de réassurance du retrait de chaque site :

$$\text{Impact}(i) = \text{Récupérations} - \text{Récupérations}(\setminus i)$$

- Allocation du chargement de réassurance proportionnellement au coût marginal :

$$\text{CoR}(i) = \text{CoR} * \frac{\text{Impact}(i)}{\sum_{k=1}^{Nb\_sites} \text{Impact}(k)}$$

### ii. Analyse des résultats

En agréant les résultats à un niveau CRESTA (équivalent au département), on compare les valeurs obtenues en valeur absolue et en taux de réassurance.

Sur les valeurs, on constate que la méthode proportionnelle aux récupérations de réassurance du b. concentre la majorité du chargement de réassurance dans l'état de Quintanaroo, à la pointe Est du pays. Cette zone concentre une grande partie des trajectoires de tempête ainsi qu'une forte accumulation d'hôtels de luxe, ce qui explique bien l'effet sur la réassurance.



Etat mexicain	Chargement de réassurance (récup. Indiv., %)	Chargement de réassurance (récup. margin., %)	% cumulé (récup. Indiv.)	% cumulé (récup. margin.)
Quintana Roo	46,9%	34,2%	46,9%	34,2%
Guerrero	14,4%	11,8%	61,3%	46,0%
Baja California Sur	8,6%	11,6%	69,9%	57,5%
Jalisco	7,5%	12,4%	77,4%	69,9%
Tamaulipas	3,3%	4,5%	80,8%	74,4%
Mexico	3,0%	3,7%	83,8%	78,1%
Distrito Federal	3,0%	3,7%	86,8%	81,8%
Sinaloa	2,8%	3,8%	89,5%	85,6%
Yucatan	2,2%	2,5%	91,8%	88,1%
Colima	1,6%	2,8%	93,3%	90,9%
Veracruz	1,3%	1,8%	94,6%	92,7%
Nayarit	1,2%	1,8%	95,9%	94,5%
Sonora	0,9%	1,2%	96,7%	95,6%
Puebla	0,5%	0,6%	97,2%	96,2%
Michoacan	0,4%	0,5%	97,6%	96,8%
Aguascalientes	0,3%	0,4%	97,9%	97,2%
Hidalgo	0,3%	0,4%	98,2%	97,6%
Queretaro	0,3%	0,4%	98,5%	98,0%
Guanajuato	0,3%	0,3%	98,8%	98,3%
Nuevo Leon	0,3%	0,4%	99,1%	98,7%

**Tableau 4 :** Tableau de l'allocation du coût de réassurance par état, suivant les deux méthodes présentées plus tôt, pour les 20 premiers états qui concentrent 99% du coût de réassurance total.

## d. Comparaison de la segmentation des méthodes principales

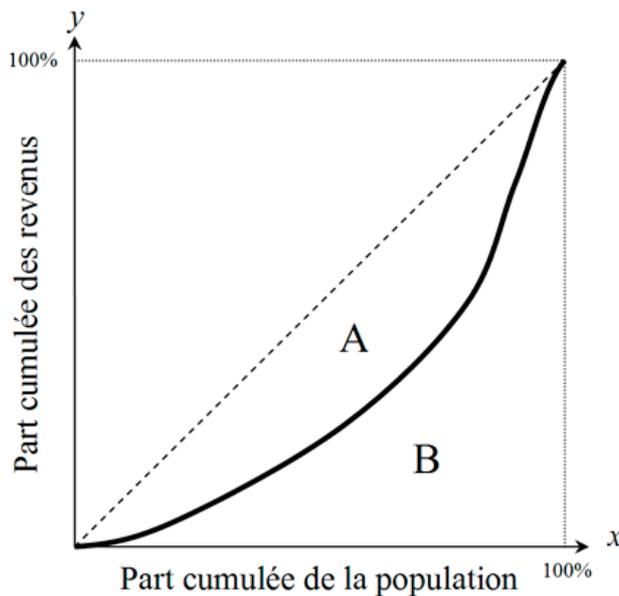
Afin de rendre plus objectif le début de comparaison observé dans la partie précédente, on a cherché à s'appuyer sur la théorie des inégalités, souvent utilisée pour comparer la répartition des richesses au sein de populations.

### i. Définitions

#### 1. *Courbe de Lorenz*

La courbe de Lorenz est la représentation graphique de la fonction qui associe à chaque part d'une population la part de l'indicateur (par exemple la richesse) associé à cette population.

Autrement dit, elle permet d'observer comment un indicateur est réparti au sein de cette population. Dans notre cas, elle permet de savoir si le chargement de réassurance se répartit de façon homogène entre les sites assurés ou si un petit nombre de sites explique la majeure partie du chargement de réassurance.



**Figure 37 :** Illustration de la courbe de Lorenz. En abscisse se trouve la part de la population auquel on associe un indicateur (ici les revenus), rangée par ordre croissant de cet indicateur. En ordonnée se trouve la part cumulée de l'indicateur considéré.  
(source : wikipedia)

Plus la courbe se rapproche de la première bissectrice (aire A petite), plus l'indicateur est réparti de façon homogène entre les individus. A l'inverse, plus la courbe est convexe (aire B petite), plus l'indicateur se concentre sur une petite partie des individus.

En matière d'assurance, si l'on observe qu'une petite partie de la population d'assuré supporte une grande partie des primes ou d'un chargement (de réassurance par exemple), cela indique que la segmentation entre assurés est forte et que la mutualisation se fait à moindre échelle.

#### 2. *Coefficient de Gini*

Le coefficient de Gini est une mesure d'inégalité qui donne une valeur entre 0 et 1, 0 correspondant à une parfaite égalité entre les individus et 1 une parfaite inégalité.

Formellement, si on note  $L_c$  la fonction de Lorenz qui sert à décrire la courbe précédente, l'indice de Gini donne une indication sur la globalité de la courbe, et l'on a :

$$Gini = 1 - 2 \int_{x_{min}=0}^{x_{max}=1} L_c(x) dx$$

Pratiquement, si l'on considère n sites assurés qui coûtent  $CoR_i$  pour le chargement de réassurance, i allant 1 à n et les  $CoR_i$  étant rangés par ordre d'indice croissant, le calcul du coefficient devient (Weiner, 2000):

$$Gini = \frac{2 \sum_{i=1}^n i \cdot CoR_i}{n \sum_{i=1}^n CoR_i} - \frac{n+1}{n}$$

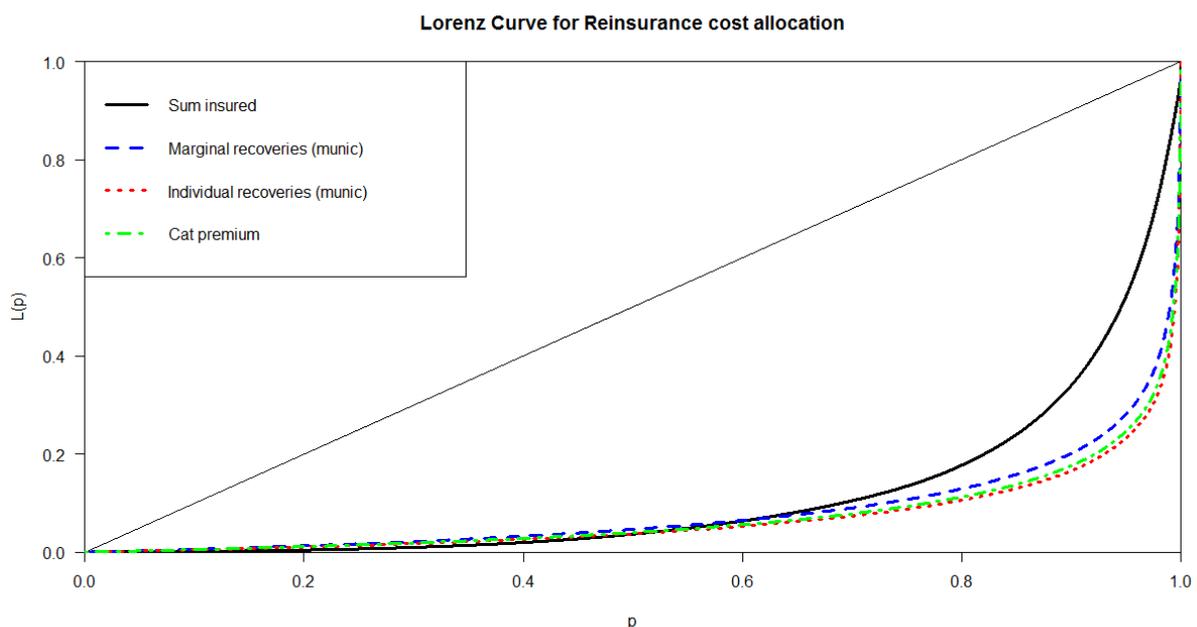
Graphiquement, le coefficient de Gini est la part de l'aire du triangle inférieur (A+B) représenté par A. L'aire du triangle inférieur étant égale à  $\frac{1}{2}$ , on a :

$$Gini = \frac{A}{A+B} = 2A = 1 - 2B$$

Et on retrouve l'intégrale formelle vue plus haut.

## ii. Comparaisons individuelles

En traçant les courbes de Lorenz des trois méthodes principales, à savoir la méthode d'allocation basée sur les récupérations de réassurance individuelles, celle basée sur l'impact marginal par site sur les récupérations de réassurance et celle basée sur la moyenne des pertes modélisée (prime pure Cat), on observe que l'allocation est très inégalitaire.



Allocation individuelle	1% des sites	10% des sites	20% des sites	50% des sites	Coefficient de Gini
Somme assurée	30,6%	69,1%	82,1%	94,6%	77,4%
Récupérations marginales	16,5%	73,1%	89,9%	99,0%	83,6%
Récupérations individuelles	19,8%	78,6%	93,0%	99,4%	86,5%
Prime Cat	18,4%	77,0%	92,2%	99,3%	85,6%

En effet, 10% des sites concentrent 73%, 79% et 77% respectivement pour les méthodes des récupérations marginales, individuelles et la méthode de la prime Cat. Cette relation se vérifie par ailleurs sur les courbes de Lorenz et on observe que la méthode des récupérations marginales segmente légèrement moins que les deux autres méthodes qui sont assez proches.

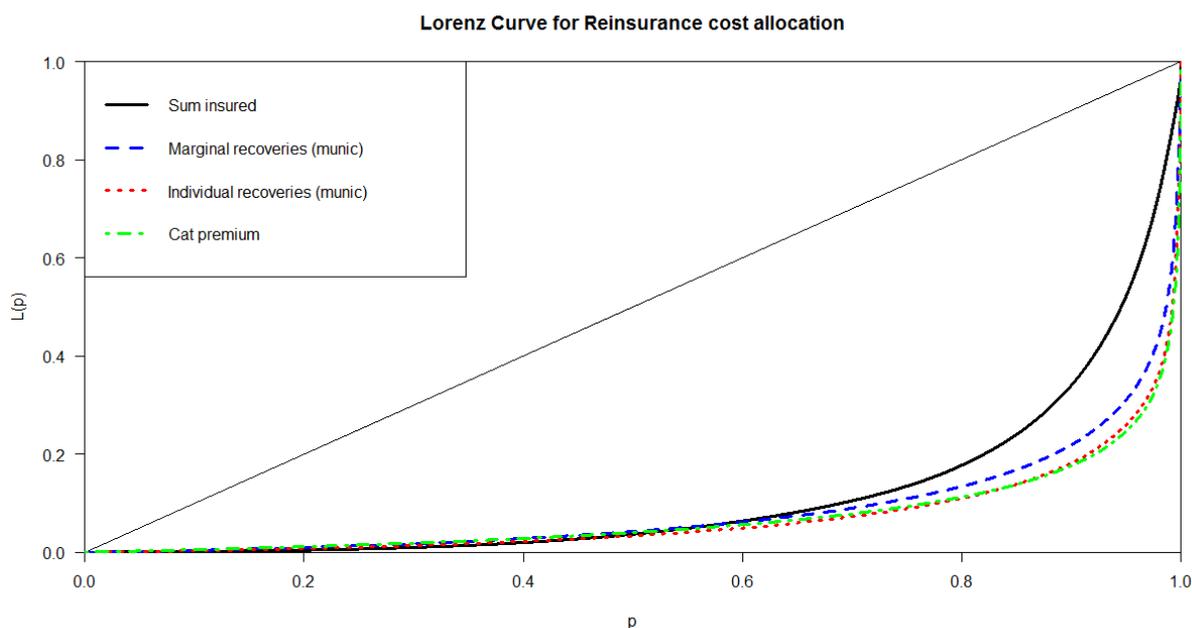
Sur les coefficients de Gini, on observe des coefficients supérieurs à 80 pour les trois méthodes et on retrouve la relation d'ordre. Ce que l'on observait a priori sur l'analyse géographique se retrouve statistiquement. La méthode des récupérations individuelles est celle qui segmente le plus le portefeuille d'un point de vue du chargement de réassurance, et celle des récupérations marginales celles qui segmente le moins.

### iii. Comparaisons à l'échelle des municipalités

Afin de tenter de réduire la segmentation du chargement de réassurance au niveau des municipalités mexicaine, on a cherché à répartir le chargement de chaque municipalité entre les différents sites en fonction de leur somme assurée, tel qu'indiqué ci-dessous.

$$CoR(i) = \frac{\text{Somme Assurée}(i)}{\sum_{k \in \text{municipalité}} \text{Somme Assurée}(k)} * \sum_{k \in \text{municipalité}} CoR(k)$$

On constate que l'allocation du chargement de réassurance est « pondérée » par la répartition des sommes assurées sur les municipalités. Cela permet de réduire légèrement le coefficient de Gini mais augmente la segmentation sur le premier percentile des sites assurés. L'effet reste néanmoins limité.



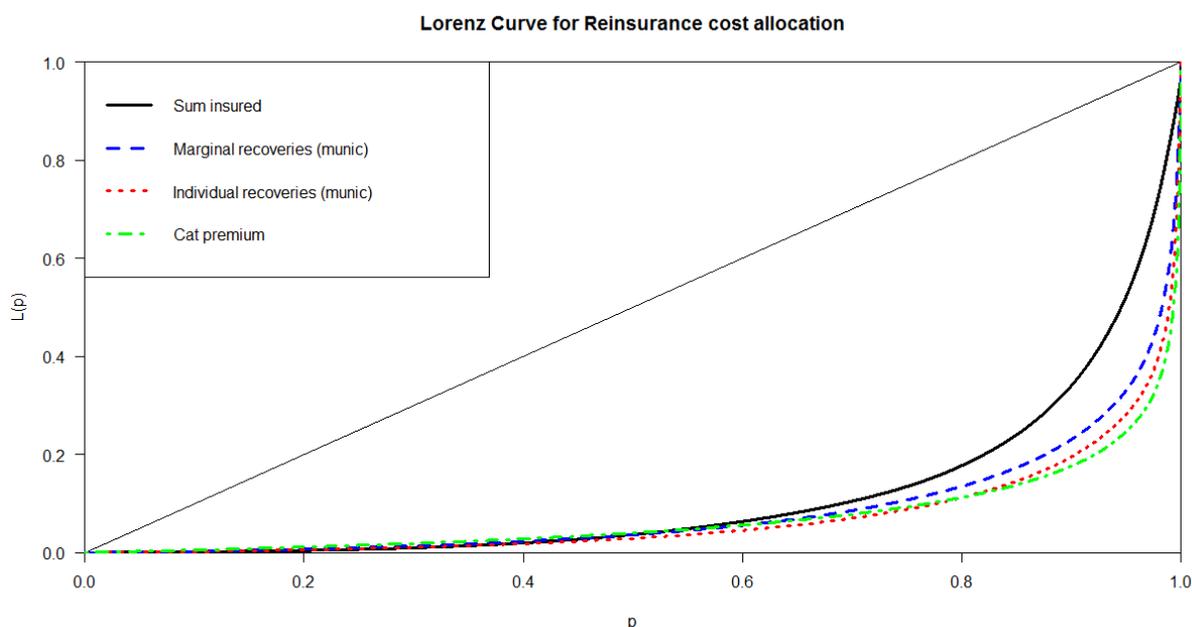
Allocation par municipalité	1% des sites	10% des sites	20% des sites	50% des sites	Coefficient de Gini
Somme assurée	30,6%	69,1%	82,1%	94,6%	77,4%
Récupérations marginales (munic)	20,9%	72,8%	88,4%	98,6%	83,2%
Récupérations individuelles (munic)	24,8%	77,9%	91,5%	99,2%	86,1%
Prime Cat	18,4%	77,0%	92,2%	99,3%	85,6%

### iv. Comparaisons à l'échelle des zones CRESTA

Le même exercice est ensuite effectué au niveau CRESTA (zones géographiques équivalentes aux départements français), plus large que les municipalités. L'effet de la segmentation est un peu plus atténué sauf sur le premier percentile. La formulation de cette méthode donne :

$$CoR(i) = \frac{Somme\ Assurée(i)}{\sum_{k \in CRESTA} Somme\ Assurée(k)} * \sum_{k \in CRESTA} CoR(k)$$

L'intérêt du passage à un niveau géographique plus large semble néanmoins limité du fait de la faible modification de la répartition du chargement de réassurance.



Allocation par CRESTA	1% des sites	10% des sites	20% des sites	50% des sites	Coefficient de Gini
Somme assurée	30, 6%	69, 1%	82, 1%	94, 6%	77, 4%
Récupérations marginales (CRESTA)	25, 3%	73, 5%	87, 6%	98, 3%	83, 1%
Récupérations individuelles (CRESTA)	29, 0%	77, 6%	90, 4%	98, 9%	85, 8%
Prime Cat	18, 4%	77, 0%	92, 2%	99, 3%	85, 6%

#### v. Méthodes à retenir

En conclusion de ces méthodes, on peut retenir que :

- Au final, le chargement de réassurance est réparti de façon plus inégale avec la méthode des récupérations individuelles qu'avec la méthode de la prime Cat, mais de façon moins inégale avec la méthode des récupérations marginales.

Ceci paraît contre-intuitif car on aurait plutôt pensé que répartir un chargement basé sur la partie des distributions de perte de chaque site correspondant aux événements extrêmes aurait été toujours plus inégal qu'une vision moyenne des pertes Cat.

- Cette assertion est toutefois nuancée par le fait que le premier percentile des pertes avec l'une des deux méthodes d'allocation des récupérations contient toujours une part significativement supérieure du chargement de réassurance par rapport à une allocation purement basée sur la moyenne des pertes (prime Cat).

D'un point de vue métier, il paraît plus logique d'appliquer la méthode des récupérations marginales si l'on souhaite évaluer le chargement de réassurance d'un nouveau site à souscrire. La méthode des récupérations individuelles paraît plutôt adaptée si l'on souhaite répartir le chargement de réassurance d'un stock de sites, par exemple un portefeuille. Cela est donc plus utile d'un point de vue « gestion des risques » que d'un point de vue « souscription ».

D'un point de vue modèle, un assureur pourra préférer une méthode qui mutualise plus s'il n'a pas une confiance aveugle dans un modèle de catastrophes naturelles. En effet, il serait dommageable pour un assureur de se priver d'une partie du marché potentiel de clients en raison d'un modèle d'allocation de coût dont les hypothèses peuvent être fortes. Dans ce cas-là, son intérêt se portera par exemple sur la méthode des récupérations marginales.

D'un point de vue commercial, un assureur pourra enfin choisir d'ajuster le résultat d'un modèle purement statistique par une réflexion sur la qualité subjective que la souscription d'un nouveau risque apporte au portefeuille. Dans ce cas, il peut décider de ne pas allouer à cette souscription son chargement de réassurance et de le répartir sur d'autres affaires. Cela peut être le cas d'une affaire qui offre des perspectives de *cross-selling* (une assurance chantier qui débouche sur l'assurance du bâtiment final par exemple) ou de réputation (l'assureur d'une grande entreprise).

Finalement, contrairement à la tarification classique où il est crucial de construire une segmentation la plus forte possible, l'indicateur de Gini n'est pas le seul paramètre « objectif » à étudier pour l'allocation du chargement de réassurance et du coût du capital. En effet, utiliser le modèle qui segmente le plus revient à supposer que cette segmentation sera la même quel que soit l'assureur, et que le même petit nombre de sites sera toujours à l'origine d'une grande partie du chargement de réassurance. En ayant montré que cela dépend aussi de l'accumulation des sites déjà présents en portefeuille, la méthode d'allocation n'est pas forcément dictée par la segmentation.

Ainsi, le choix de l'assureur se portera plus sur le degré de mutualisation qu'il souhaite appliquer à son portefeuille. Dans le cas de sites de faible valeur assurée qui n'ont que peu de chance d'impacter fortement le chargement de réassurance, l'assureur peut avoir intérêt à garder une mutualisation importante (à un niveau CRESTA par exemple). En revanche, dans le cas de sites à forte valeur assurée (segment de portefeuille Commercial), l'assureur a intérêt à intégrer une segmentation plus fine du chargement de réassurance et de capital dans son tarif.

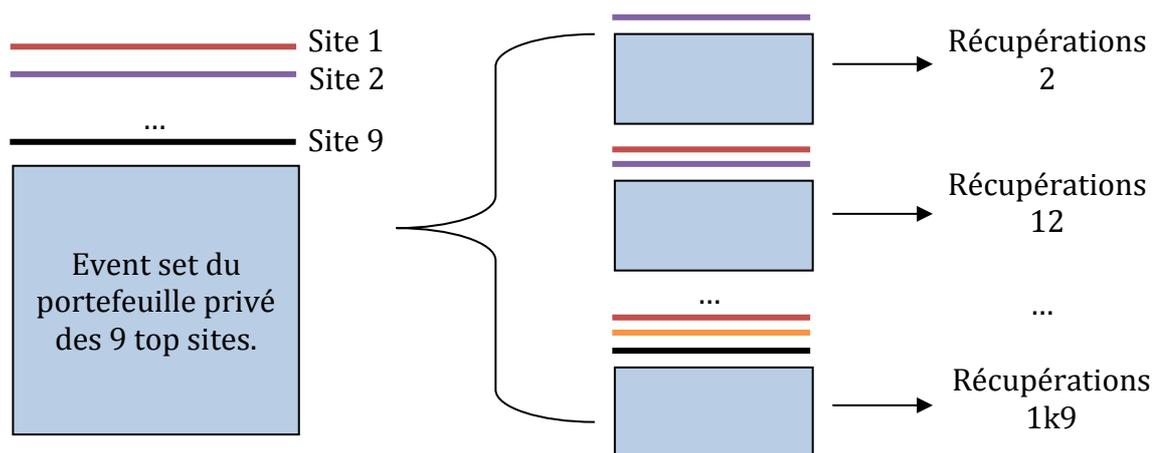
## e. Valeur de Shapley et méthode du nucléole : première approche

### i. Principe

Les méthodes qui semblent les plus satisfaisantes sur le plan théorique ont été partiellement testées mais leur utilisation est impossible en l'état actuel de la puissance de calcul offerte par les ordinateurs. En effet, au-delà de quelques dizaines de sites, le nombre de combinaisons possibles entre les sites devient trop grand pour permettre un calcul de récupérations de traité.

On peut néanmoins effectuer le calcul en considérant les plus gros sites et en supposant que le reste du portefeuille a déjà été constitué, comme indiqué sur la Figure 38.

On considère l'ensemble de l'*event set* privé des neuf top sites comme indissociable. On ajoute ensuite à cet *event set* toutes les permutations possibles présentées dans la méthode de Shapley d'événements associés à chaque site. Pour chaque permutation, on calcule les récupérations qui serviront de base à l'allocation de l'écart de récupérations de réassurance avec et sans les 9 sites.



**Figure 38** : Schéma du procédé de calcul de la valeur de Shapley et de la méthode du nucléole. Chaque permutation possible de l'event set fait l'objet d'un calcul des récupérations.

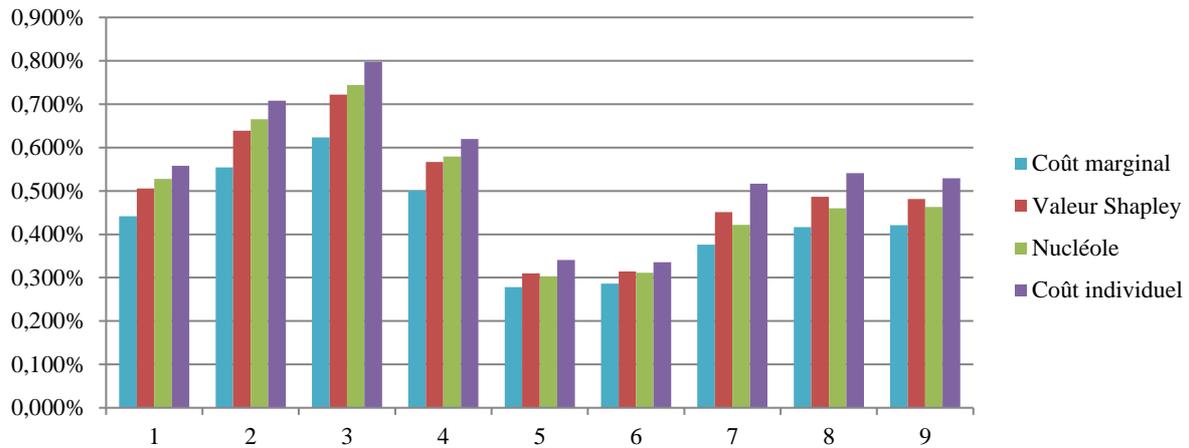
Les coûts individuels sont alors l'impact sur les récupérations de l'ajout du site au portefeuille privé des 9 sites. Les coûts marginaux sont les mêmes que ceux calculés dans la précédente.

### ii. Résultats

On obtient ainsi, pour les 9 premiers sites en termes de primes Cat, l'allocation des récupérations avec la méthode de Shapley et la méthode du Nucléole. Cette allocation est comparée avec les coûts marginaux et les récupérations individuelles (Figure 39).

Remarque : On constate que les coûts marginaux sont ici plus faibles que les coûts individuels, ce qui pourrait s'expliquer par une mutualisation entre les 9 sites dans l'événement deux-centenaire qui existe moins au niveau individuel.

### Comparaison des résultats de l'allocation des récupérations de réassurance par méthode



**Figure 39 :** Comparaison des résultats d'allocation de récupération par méthode en pourcentage de somme assurée. On constate que les méthodes de Shapley et du nucléole produisent des résultats situés entre le coût marginal et le coût individuel (plus précisément les récupérations de réassurance).

Comme évoqué précédemment, les deux méthodes étant très gourmandes en temps de calcul, elles ont été écartées au profit de méthodes plus simples présentées dans les parties précédentes, notamment la méthode des bénéfices résiduels.

## **f. Prise en compte de l'extension potentielle de la portée du traité à la souscription d'un risque**

### **i. Principe et définitions**

Jusqu'à présent, il a été évoqué l'allocation de la réassurance existante, en supposant que le traité ne varie pas en fonction du portefeuille étudié et de la souscription du site que l'on souhaite tarifier en termes de prime Cat et de prime de réassurance. L'exercice peut être poussé un peu plus loin en introduisant l'extension éventuelle de la portée du traité lors de la souscription d'un nouveau site.

On considère que la limite du traité (portée + priorité) est fixée au montant de la perte maximale survenant avec une probabilité de 0,5% (événement deux-centenaire). Cette perte est notée  $OEP_{200}$

Dans ce cas-là, la méthode la plus simple consiste à supposer le prix de réassurance de la dernière tranche du traité (celle qui arrive à la limite) constant en pourcentage de portée (aussi appelé *rate on line*, ou *RoL*). Le chargement lié à l'extension de portée du traité est donc une fonction CoR de cette extension, notée  $\Delta P$  :

$$\text{CoR}(\Delta P) = \text{RoL} * \Delta P - PP(\Delta P)$$

Où  $PP$  est la fonction associant à l'extension de portée sa prime pure.

Pour simplifier le calcul, le chargement est approximé comme proportionnel au chargement de réassurance de la tranche du haut (notée Top Portée) :

$$\text{CoR}(\Delta P) = \text{CoR}(\text{Top Portée}) * \Delta P$$

La méthode la plus propre consiste à ajouter le site souscrit au portefeuille existant et à relancer la modélisation pour en déduire la nouvelle distribution des pertes de ce nouveau portefeuille. Néanmoins, comme évoqué plus haut, effectuer cette modélisation prend du temps et deux moyens ont été testés pour approcher au mieux ce calcul.

### **ii. Effet marginal par site et par municipalité**

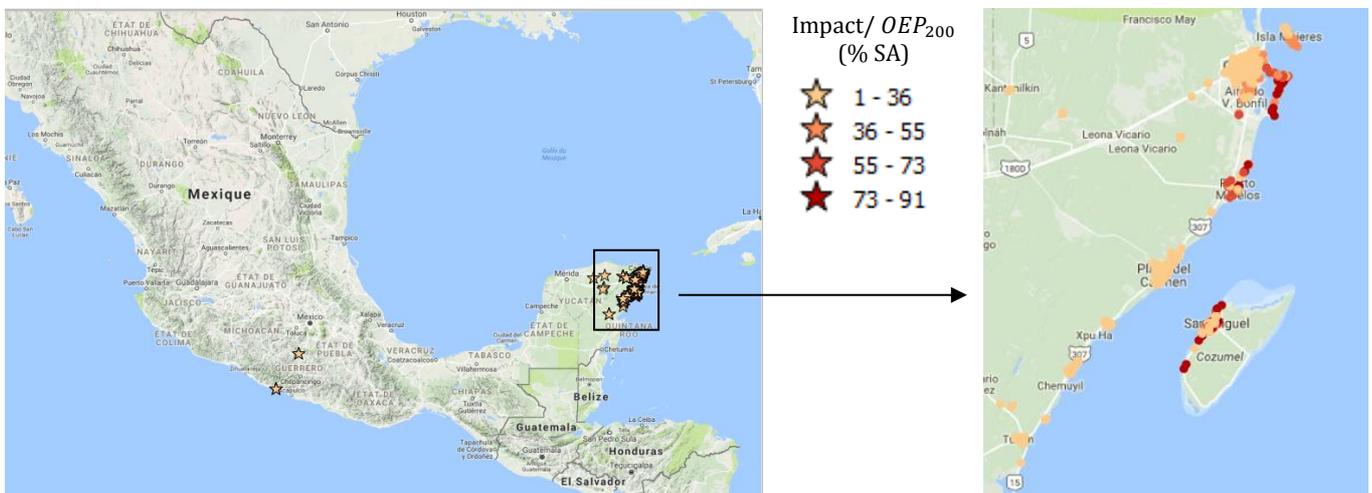
Une première méthode consiste à essayer de prédire l'impact d'un nouveau site sur l'événement deux-centenaire, et donc sur la portée du traité, en observant le comportement de l' $OEP_{200}$  avec le retrait de chacun des sites du portefeuille.

On constate que seuls les sites situés à l'Est du pays, dans l'état de Quintanaroo, ont un impact non négligeable sur l' $OEP_{200}$ , à l'exception de deux sites situés dans l'état de Guerrero, identifié dans les points précédents comme le second état en termes de chargement de réassurance. Ces sites se concentrent de plus sur la côte.

En partant de ce portefeuille, l'ajout d'un petit nombre de sites n'aurait donc a priori d'impact que s'ils sont situés dans ces deux états.

On pourrait effectuer un calcul d'algorithme de prédiction tel que présenté dans le b. de cette partie, mais cette prédiction serait limitée. En effet, il suffirait que l'événement correspondant à l' $OEP_{200}$  change légèrement pour observer rapidement des résultats différents. Il convient

donc d'être prudent, surtout dans le cas où l'on n'est pas capable de mettre à jour régulièrement l'événement set pour ajuster la calibration des algorithmes de prédiction au fil de l'évolution du portefeuille.



**Figure 40 :** Carte des sites ayant un impact sur l' $OEP_{200}$  de plus de 1% de leur somme assurée. Les teintes de couleurs correspondent à l'impact, en pourcentage de somme assurée, du site sur l' $OEP_{200}$ . On observe que l'événement de l' $OEP_{200}$  se situe clairement sur la zone côtière de l'état de Quintanaroo.

Au Mexique, contrairement à d'autres pays, la réglementation est plutôt plus flexible, notamment en termes de résiliation. Les assureurs font donc face à une concurrence dynamique et leur portefeuille peut évoluer rapidement.

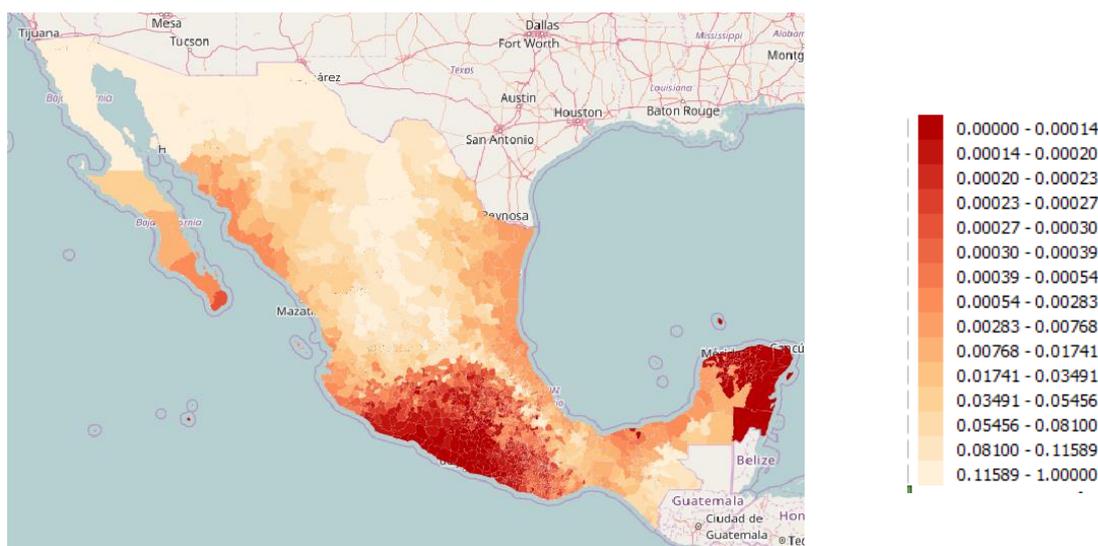
Une étude similaire a été effectuée par municipalité afin de mutualiser à l'échelle de chaque municipalité l'effet sur l' $OEP_{200}$ . Néanmoins, cette étude n'est pas suffisante car seules les municipalités de l'état de Quintanaroo ont un impact sur l' $OEP_{200}$ .

### iii. Capacité restante par municipalité

Au lieu de tenter d'estimer l'impact du retrait d'un site ou d'une municipalité sur l' $OEP_{200}$ , on effectue un calcul opposé. On effectue le calcul de la somme assurée que l'on peut souscrire sans impacter l' $OEP_{200}$ .

Ce calcul avait été effectué par un membre de l'équipe en dehors de ce mémoire. Les étapes sont les suivantes :

- Hypothèse : pour x% de somme assurée ajoutée dans la municipalité, les pertes par municipalité, par année, par événement augmentent de x%.
- Pour chaque municipalité, on incrémente la somme assurée de x%. On en déduit une augmentation des pertes par municipalité, par année, par événement, de x%.
- On calcule l' $OEP_{200}$  :
  - o Si la valeur est la même que l' $OEP_{200}$  d'origine, on réitère une augmentation de somme assurée.
  - o Sinon, on retient la valeur de x% sur cette municipalité. Il s'agit alors du x% maximal qui conserve l' $OEP_{200}$ . A cette valeur en pourcentage, on associe une somme assurée au-delà de laquelle toute souscription supplémentaire se traduit par une augmentation de l' $OEP_{200}$ .
- On réitère l'exercice sur l'ensemble des municipalités.



**Figure 41** : carte de la capacité (en somme assurée) pouvant être souscrite sans variation de  $OEP_{200}$  modélisée. Les résultats sont indiqués en valeur relative à la capacité à souscrire maximale. (Source : étude interne AGPC)

Une fois ce calcul de capacité effectué, on peut établir un indicateur de répartition qui permette de :

- Charger la souscription dans les municipalités dans lesquelles il n'est plus possible de souscrire sans augmenter l' $OEP_{200}$ .
- Ne pas charger les autres municipalités tant que cette capacité n'est pas atteinte.

A défaut de proposer une solution définitive sur le calcul d'augmentation de la limite du traité de réassurance, cette étude permet de filtrer l'immense majorité des souscriptions qui n'auront pas d'impact sur le traité. Elle permet aussi d'effectuer un suivi des seuils de somme assurée à souscrire par municipalité à partir desquels il convient d'augmenter la capacité du traité, à condition de la mettre régulièrement à jour pour prendre en compte les accumulations inter-municipalités.

En conclusion de cette partie sur l'augmentation potentielle de la limite du traité de réassurance pour pouvoir couvrir l' $OEP_{200}$ , ce mémoire propose de combiner :

- Un chargement par municipalité de l'état de Quintanaroo.
- Un suivi de la capacité de souscription restante par municipalité sans effectuer d'étude de chargement lié à une extension de la capacité du traité de réassurance.
- Une mise à jour régulière de l'étude pour prendre en compte les évolutions du portefeuille et les interactions entre municipalités.

## **IV- EXTENSION AU COUT DU CAPITAL**

La partie précédente se concentrait sur l'allocation du chargement de réassurance associé à un traité non proportionnel. Cette allocation vise à segmenter davantage le portefeuille en ajoutant la contribution des différents sites assurés au chargement de réassurance.

Dans un cadre Solvabilité 2, les assureurs sont tenus de conserver un niveau de capital suffisant pour supporter un événement survenant en moyenne une fois tous les 200 ans. De la même manière que pour l'allocation du chargement de réassurance, il est donc intéressant de prendre aussi en compte le coût du capital réglementaire dans la tarification des catastrophes naturelles.

### **a. Notion de coût du capital**

#### **i. Le capital réglementaire dans l'assurance**

La notion de capital réglementaire, aussi appelé marge de solvabilité, existait déjà avec la directive européenne Solvabilité 1. Avec l'introduction de la directive européenne Solvabilité 2, entrée en vigueur au 1<sup>er</sup> Janvier 2016, cette notion est modifiée en profondeur.

L'idée de la marge de solvabilité est de rendre obligatoire la constitution d'un « matelas » de capital permettant à une entreprise d'assurance ou de réassurance d'encaisser un scénario de résultat extrême, ce qui signifie concrètement d'être capable de répondre à ses engagements de sinistres sans faire faillite lors d'un scénario de très faible fréquence mais de forte sévérité.

Avec Solvabilité 1, l'exigence de marge faisait partie du pilier 1 et se calculait, selon la branche, par des pourcentages de primes acquises, de variations de PSAP (provisions pour sinistres), de provisions mathématique, d'actifs sous gestion, etc. La méthode était relativement simple, bien que critiquée pour son manque de prise en compte du profil de risque de l'entreprise. Solvabilité 2 a tenté de répondre en partie à ces limitations.

La marge de solvabilité fait partie du le pilier 3 de la directive Solvabilité 2. Elle prend la forme d'un SCR (Solvency Capital Requirement) qui doit être respecté par l'assureur. Cette exigence de solvabilité permet de couvrir cinq grands types de risques auxquels s'ajoute le risque opérationnel :

- Le risque de marché : taux de change, actions, immobilier, spread, taux d'intérêt, accumulation.
- Le risque vie : mortalité, longévité, maladie, dépenses, catastrophe, etc.
- Le risque santé : mortalité, longévité, maladie, dépenses, etc.
- Le risque non-vie : prime, réserve, catastrophe.
- Le risque de contrepartie : en cas de défaut d'un fournisseur.

Ce capital réglementaire doit permettre de couvrir « la valeur en risque des fonds propre de base de l'entreprise d'assurance ou de réassurance avec un niveau de confiance de 99,5% à l'horizon d'un an » (article R352-2 du code des assurances). Autrement dit, l'entreprise doit être capable d'encaisser le pire scénario de résultat annuel survenant en moyenne une fois tous les 200 ans.

Le SCR peut être calculé via une formule standard ou à l'aide d'un modèle interne qui peut être partiel ou complet et doit être préalablement validé par le régulateur. On calcule généralement plusieurs sous-SCR par module des 5 risques. On agrège les résultats avec des matrices de corrélation dont les coefficients sont fixés réglementairement.

La réassurance est un des moyens de réduire l'exigence de capital en cédant une partie des pertes des scénarios extrêmes à un réassureur. En effet, cela permet de réduire les pertes associées à la queue de la distribution des scénarios de l'entreprise étudiée et donc de réduire a priori le quantile à 99,5% de cette distribution.

Dans le contexte de l'étude dont il est question, on s'intéresse au coût du capital de catastrophes naturelles de type ouragan. Ce coût du capital est considéré net de réassurance puisque la réassurance permet de réduire ce coût et a par ailleurs déjà été étudié dans la seconde partie de ce mémoire. Ce coût du capital est considéré pour le capital diversifié, c'est à dire après prise en compte de l'effet de sous additivité de la corrélation.

Le coefficient de diversification (noté  $d$ ) a été calculé dans une étude interne comme égal à 80%. Ceci signifie qu'une hausse de 100€ de capital requis pour les catastrophes naturelles (SCR Cat) n'entraîne qu'une hausse de 20€ sur le capital requis au total sur l'ensemble des risques (SCR Total).

## ii. Le coût du capital

Le coût du capital est un indicateur de finance d'entreprise qui vise à quantifier le taux de rentabilité économique attendu par les actionnaires et créanciers d'une entreprise. D'un point de vue de l'entreprise, c'est donc le coût que représente l'immobilisation du capital pour la gestion de l'activité de l'entreprise.

On distingue dans le capital deux grandes catégories : les fonds propres et les dettes. La méthode la plus classique (Vernimmen, 2017) pour calculer le coût du capital est de calculer le coût moyen pondéré du capital sur les dettes et les fonds propres :

$$WACC = CoCP * P_{CP} + CoD * (1 - TxIS) * P_D$$

Où :

- WACC : coût moyen pondéré du capital (weighted average cost of capital)
- CoCP : coût des capitaux propres (rentabilité demandée par les investisseurs)
- CoD : coût de la dette (rentabilité des créanciers)
- TxIS : taux d'imposition sur les sociétés
- $P_{CP}$  et  $P_D$  : proportion de capital représenté par les capitaux propres et la dette (somme unitaire)

On constate que le taux d'imposition rentre en compte dans le calcul car les intérêts des dettes sont déductibles de l'impôt sur les sociétés, ce qui est reflété directement ici.

Pour une entreprise d'assurance, le capital se réduit généralement aux fonds propres, en raison du cycle inversé de production qui entraîne une trésorerie positive et un besoin de provisionnement pour être capable d'honorer les engagements de sinistres futurs. Ces fonds propres sont plutôt un « coussin règlementaire » qu'un réel outil de financement à long terme, contrairement à des entreprises au cycle de production plus classique.

Par ailleurs, dans une note de l'EIOPA (ex-CEIOPS, 2009), le coût du capital à prendre en compte est de 6%. Ceci se justifie par le coût représenté par un financement d'une entreprise d'assurance ou de réassurance notée BBB.

Malgré des discussions soulevées par ce chiffre (notamment au CRO forum de 2008), ce chiffre apparaît comme une référence marché fermement soutenue par le régulateur et est donc retenu pour l'étude.

Par conséquent, le montant que l'on cherche à allouer site par site est le suivant :

$$CoC = WACC * (1 - d) * SCR(Cat)$$

Où :

- WACC : coût du capital pour les assureurs, égal à 6%
- (1-d) : coefficient de transmission du SCR Cat dans le SCR total de l'entreprise d'assurance, égal à 20%.

En ce qui concerne le calcul de SCR Cat, il est nécessaire de se baser sur un modèle interne de façon à prendre en compte l'information site par site produite par le modèle physique de catastrophe naturelle. Dans cette étude, on prend donc en considération le SCR Cat généré par les ouragans au Mexique sur le portefeuille assuré.

Pour simplifier, on définira le SCR Cat comme le quantile à 200 ans des pertes modélisées sur la totalité du portefeuille, pour l'ensemble des périls significatifs, en négligeant les primes d'assurance mais en prenant en compte la réassurance.

## **b. Application de l'allocation du CoC dans la tarification Cat**

### **i. Choix d'un indicateur**

Pour allouer le coût du capital, l'indicateur qui est utilisé doit refléter au mieux le quantile à 99,5% de la distribution du résultat net de réassurance. Sur le même principe que pour l'allocation du chargement de réassurance, on peut définir :

- Un « coût individuel du capital » : on prend le quantile de perte nette de réassurance à 200 ans modélisé pour chaque site pris isolément.
- Un « coût marginal du capital » : on prend l'impact du site sur le quantile de perte nette de réassurance à 200 ans du portefeuille pour chaque site.

Néanmoins, pour ces deux indicateurs, l'effet capté ne prendra pas en compte le potentiel réordonnement des événements susceptibles de se situer sur le quantile à 200 ans des pertes. Autrement dit, lorsque le portefeuille évolue, l'événement qui était à l'origine de la perte à 200 ans peut devenir celui à l'origine de la perte à 201 ans. Un autre événement, probablement localisé à un autre endroit du pays modélisé prendrait alors sa place.

De façon à éviter que seuls les sites de l'événement constitutif du deux-centenaire au moment de l'étude soient impactés pour le coût du capital, on prend un indicateur qui intègre plus largement les événements de la queue de distribution de pertes du portefeuille.

Un indicateur répondant à ce critère peut être la valeur moyenne des valeurs en risque supérieures au quantile 200 ans, ou Tail-VaR. La TVaR se définit comme :

$$TVaR_{\alpha}(X) = E(X|X \geq VaR_{\alpha}(X))$$

Où  $\alpha$  est le quantile correspondant à la TVaR recherchée. Il sera pris à 99,5%, ce qui correspond au quantile à 200 ans.

De cette manière, on capte plus généralement les événements susceptibles de constituer l'événement deux-centenaire dans le cas où le portefeuille évolue.

### **ii. Allocation du CoC**

Avec un coefficient de Gini à 99,31%, le coût du capital est réparti sur un petit nombre de site. En effet, 100 sites sur un portefeuille de plus de 100 000 assurés représentent 76% du chargement de réassurance alloué avec la méthode de la TVaR pour 8% de la somme assurée du portefeuille. Ceci confirme l'emploi d'une méthode qui tient compte de la distribution au-delà d'un quantile et non d'un quantile uniquement.

Néanmoins, l'incertitude reste importante sur les différentes étapes du calcul et notamment sur la stabilité au portefeuille. En effet, ce petit nombre de sites représente une part importante de la somme assurée du portefeuille (8%) pour un faible pourcentage en nombre (0,1%). Deux solutions sont alors possibles ici :

- Allouer le coût du capital sur les gros sites uniquement.
- Répartir le coût du capital globalement sur le portefeuille par zone géographique, par exemple à la municipalité.

L'avantage de la première méthode est qu'elle est plus juste théoriquement, mais en pratique pose de réels problèmes de stabilité du coût alloué. En effet, il suffirait que les deux premiers sites changent d'assureur pour que 6% du coût du capital ne soit plus alloué. A l'inverse, une poignée de nouvelles souscriptions suffirait à allouer un coût supérieur au coût du capital.

La seconde méthode paraît donc la plus raisonnable d'un point de vue de la souscription et de la stabilité du coût alloué. L'indice de Gini de cette méthode descend d'ailleurs légèrement (99,26%) par rapport à celui de l'allocation site par site (99,31%) mais reste très élevé.

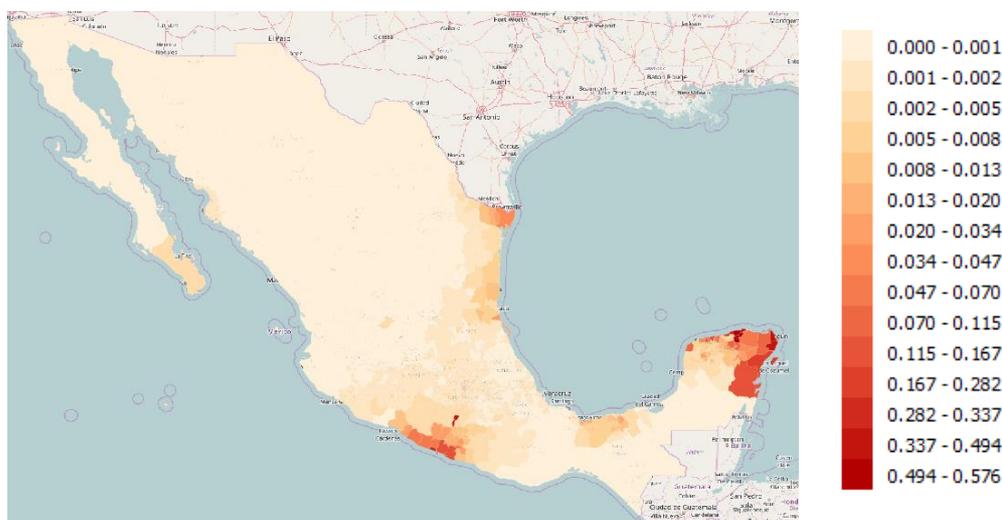
En observant le graphe de répartition du coût du capital avec cette méthode, on observe que quelques municipalités regroupent la majorité du coût du capital (cf. Figure 42). Néanmoins, cette observation est biaisée par le fait que ces municipalités contiennent des sites à forte valeur assurée.



**Figure 42 :** Carte de la répartition du coût du capital en vision modèle, par municipalité. On constate qu'un petit nombre de municipalités présentent un coût non négligeable. Echelle en dollars.

En traçant la carte des taux de coût du capital sur la somme assurée des sites, on distingue cette fois deux ou trois zones qui semblent avoir un rôle important sur les événements les plus extrêmes et susceptibles de jouer le rôle de deux-centenaire (cf. Figure 43).

En effet, les zones situées dans les états de Guerrero, Quintana Roo et Yucatan présentent des taux de coût du capital supérieurs à 0,28%. Dans les autres états, ce taux descend rapidement en dessous de 0,05%.



**Figure 43 :** Carte du coût du capital proposé pour la tarification, en taux de somme assurée. Les deux zones les plus sensibles correspondent aux zones de plus forte sévérité de l'aléa des ouragans. Légende en % somme assurée.

Avec cette partie de l'étude, on propose ainsi une solution pour allouer le coût du capital représenté par le risque ouragan sur la tarification locale. Cette solution consiste à ajouter au tarif un coût du capital qui dépend uniquement de la municipalité et de la somme assurée.

## **V- LIMITES ET PISTES D'AMELIORATION DE L'ETUDE**

### **a. Tarification Cat**

#### **i. Les limites de la segmentation**

Avec cette étude, on cherche à construire une segmentation du prix et du chargement de réassurance des catastrophes naturelles sur le portefeuille, en fonction de variables endogènes (structure, ligne de métier, nombre d'étages, etc.) ou exogène (zones de risques, localisation géographique).

La segmentation permet en théorie à un assureur d'être plus compétitif qu'un autre en faisant payer à chaque assuré le prix le plus adapté aux caractéristiques du risque qu'il constitue. Un assureur qui segmente ses risques de façon adaptée va proposer à chaque classe de risque un prix cohérent quand un assureur sans segmentation proposera par exemple le même prix à chaque assuré. Sous l'hypothèse de rationalité des assurés, les assurés les moins risqués iront à l'assureur qui lui propose le meilleur prix, en l'occurrence celui qui segmente, quand les plus risqués iront aussi chercher le meilleur prix, cette fois-ci chez l'assureur qui ne segmente pas.

En construisant cette segmentation, on permet ainsi à l'assureur mexicain de se créer un avantage compétitif par rapport aux autres assureurs. Les autres assureurs devraient en théorie être anti-sélectionnés par les risques les plus sensibles aux ouragans.

Pourtant, dans le cas des catastrophes naturelles, il n'est pas évident que segmenter au maximum soit la meilleure démarche à adopter pour les assureurs. En effet, la segmentation fonctionne bien si les risques considérés sont relativement peu volatiles et que l'on peut les étudier avec des outils statistiques robustes.

Dans le cas des catastrophes naturelles, les résultats de distributions de valeurs extrêmes obtenues sont généralement très sensibles aux maximas les plus élevés, et une incertitude relativement importante existe à chaque étape de la construction d'un modèle Cat. Les résultats finaux sont donc plus dépendants des hypothèses de modélisation et des données de calibration que pour la plupart des modèles de modélisation de pertes assurantielles.

Les modèles de tarification classiques d'assurance sont généralement calibrés sur des bases de données de plusieurs milliers de sinistres chaque année. Dans le cas des modèles Cat, il n'est pas rare de n'avoir qu'une poignée d'événements sur les 15-20 dernières années dont les bases de sinistres se réduisent à quelques milliers, voire centaines de sites touchés. Calibrer une segmentation qui cherche à reproduire une moyenne d'événements pouvant se produire avec une fréquence inférieure à 1 fois tous les 100 ans est donc un exercice ambitieux, sinon dangereux pour l'assureur qui l'applique.

La segmentation issue de ce type de modèle est donc plus volatile que pour celles issues de modèles plus classiques de tarification, à la fois en termes de précision et de temps.

En termes de précision, les zones géographiques et autres paramètres sélectionnés pour segmenter peuvent être erronés lorsque l'on souhaite avoir une granularité de segmentation trop fine. Dans l'exemple présenté dans la partie III, il est peut-être plus judicieux d'avoir une segmentation à l'échelle de la municipalité, voire du CRESTA, que de chercher absolument à construire un modèle d'allocation au niveau coordonnée. Les erreurs liées à la modélisation se mutualisent ainsi un peu plus et évitent à l'assureur de prendre de mauvaises décisions en refusant des risques qui seraient sur-tarifés ou inversement en acceptant des risques qui ont été sous-tarifés.

En termes de temps, le problème d'une segmentation trop précise de catastrophes naturelles est lié à leur très faible fréquence. Effectuer une segmentation précise pour des tremblements de terre reviendrait quasiment à se priver d'une partie des assurés en ayant une vision moyenne modélisée qui ne tiendrait pas compte de la volatilité très élevée des événements naturels.

Ainsi, un assureur qui segmenterait de façon trop précise sur ce type de risques de catastrophes naturelles risquerait de perdre des parts de marché au profit d'autres assureurs dans certaines zones, sans que celles-ci soient nécessairement sinistrées dans le futur.

#### ii. La dépendance à une unique vision du risque

En termes de vulnérabilités relatives, les résultats du modèle ont été acceptés sans vérification. Pourtant, il aurait pu être intéressant de valider les résultats des primes pures Cat site par site à l'aide d'une base de données de sinistres historiques survenus lors de gros ouragans passés. En effet, cela aurait permis de gagner en confort sur les choix de modélisation effectués, notamment sur la hiérarchisation des variables et l'impact des modalités de celles-ci sur les résultats.

Une autre piste d'amélioration de l'étude peut être d'affiner la vision du risque en prenant en compte plusieurs modèles d'ouragan. En effet, plusieurs entreprises de modèle vendeurs proposent un modèle d'ouragan pour le Mexique et il aurait été intéressant d'avoir les moyens de comparer les résultats obtenus avec les mêmes méthodes, mais à partir d'*event set* produits par différents modèles vendeurs.

La méthode employée pour prendre en compte ces différents modèles resterait à définir, mais on pourrait penser à plusieurs éléments :

- Un simple mélange : On peut moyenniser les résultats obtenus sur chacun des modèles, éventuellement avec des poids différents selon le degré de confiance que l'on a dans chaque vision du risque proposée.
- Une segmentation géographique : après une étude spécifique, on pourrait imaginer que certaines zones soient mieux modélisées par un modèle que par un autre, et inversement sur d'autres zones. On utiliserait alors tantôt les résultats issus d'un modèle, tantôt ceux d'un autre modèle, selon que l'on se trouve dans une zone géographique ou une autre.
- Un *event set* par sévérité : on pourrait imaginer avoir plus de confiance pour des événements de forte sévérité pour un modèle et plus de confiance pour les événements d'impact moyen pour un autre modèle. On construirait alors son propre *event set* à partir des événements de chaque modèle vendeur. On pourrait même y inclure des événements historiques.

#### iii. Limitation du périmètre de péril

La présente étude s'est concentrée sur le péril ouragan uniquement. Il aurait été intéressant d'intégrer, dans une vision globale, d'autres périls naturels comme les séismes. Ce dernier péril est très présent dans l'histoire du Mexique. Ainsi, on distingue plus de 20 séismes de magnitude supérieure à 7 survenus au cours du XX<sup>ème</sup> siècle (cf Tableau 5).

Intégrer une approche unifiée pour le chargement de réassurance permet de bien croiser les différentes sources d'aléa sur le portefeuille.

Dans l'idéal, on aurait pu construire un *event set* multi-péril pour effectuer les calculs de récupérations de réassurance et ainsi prendre bien en compte les années pour lesquelles on aurait par exemple épuisé le traité de réassurance.

Nom	Date	Magnitude	Nombre de victimes
<u>1907 Guerrero earthquake</u>	1907-04-15	7.7	
Guerrero	1909-07-30	7.6	
"Temblor maderista" <u>1911 Michoacán earthquake</u>	1911-06-07	7.6	45
<u>1911 Guerrero earthquake</u>	1911-12-16	7.6	
Acambay	1912-11-19	7.0	
<u>1931 Oaxaca earthquake</u>	1931-01-15	7.8	114
<u>1932 Jalisco earthquake</u>	1932-06-03	8.1	400
<u>1932 Colima earthquake</u>	1932-06-18	7.8	
<u>1932 Manzanillo earthquake</u>	1932-06-22	7.0	
<u>1937 Orizaba earthquake</u>	1937-07-26	7.3	34
<u>1941 Colima earthquake</u>	1941-04-15	7.6	90
<u>1957 Guerrero earthquake</u>	1957-07-28	7.9	68
<u>1964 Guerrero earthquake</u>	1964-07-06	7.4	40
<u>1965 Oaxaca earthquake</u>	1965-08-23	7.5	6
Guerrero, Oaxaca	1968-08-02	7.3	
<u>1973 Veracruz earthquake</u>	1973-08-28	7.0	600–1,200
<u>1979 Imperial Valley earthquake</u>	1979-10-15	6.4	N/A
<u>1980 Oaxaca earthquake</u>	1980-10-24	7.0	300
<u>1981 Playa Azul earthquake</u>	1981-10-25	7.2	
<u>1985 Mexico City earthquake</u>	1985-09-19	8.0	10,000–65,000
<u>1995 Guerrero earthquake</u>	1995-09-14	7.4	3
<u>1995 Colima–Jalisco earthquake</u>	1995-10-09	8.0	49
<u>1995 Chiapas earthquake</u>	1995-10-21	7.1	
<u>1999 Tehuacán earthquake</u>	1999-06-15	7.0	15
<u>1999 Oaxaca earthquake</u>	1999-09-30	7.4	35

**Tableau 5 :** Liste des séismes survenus au Mexique au cours du XXème siècle. On constate que le nombre de victimes n'a dépassé le millier que dans deux cas, en 1973 et lors de l'événement de 1985. (source : wikipedia)

On aurait aussi pu penser à d'autres périls comme le volcanisme ou les inondations, qui sont néanmoins des périls plus réduits en termes de pertes assurancielles au cours de dernières décennies.

## **b. Allocation de la réassurance et coût du capital**

Globalement, les limites et pistes d'amélioration de l'étude peuvent être portées à la fois sur l'allocation du chargement de réassurance et du coût du capital. En effet, les méthodes étudiées sont les mêmes et les difficultés rencontrées sont proches.

### **i. Traités spécifiques et réassurance non traditionnelle**

Au niveau des types de traités étudiés, certains présentent des caractéristiques de fonctionnement intéressantes qui ne peuvent être prises en compte en l'état actuel de l'étude. Il s'agit notamment des couvertures de fréquence avec les AAL/AAD (*annual aggregate limit/deductible*).

L'intérêt de ces spécificités expliquées au II.c s'observe particulièrement dans le cas de fréquence élevée pour des tranches de réassurance dites « travaillantes » (qui sont touchées régulièrement par des événements Cat). Etant donné que l'étude ne portait que sur un péril relativement peu fréquent, l'introduction des AAL et AAD n'a pas été prise en compte. Cependant, il conviendrait d'envisager cette possibilité dans la perspective d'une modélisation multi-péril.

La réassurance non traditionnelle regroupe des formes de transfert de risque différentes de celles présentées dans l'étude. Cela représente un marché faible de la réassurance totale mais en croissance régulière. Parmi les produits proposés dans ce cadre, on distingue par exemple les Cat bonds qui consistent à déclencher les récupérations pour l'assureur si un paramètre (un indice météo par exemple) dépasse un certain seuil. L'idée pour l'assureur est de diversifier ses fournisseurs de réassurance en permettant à des acteurs de marchés purement financiers de profiter de ce type de produit. Selon les cas, cela peut aussi permettre à l'assureur de réduire son chargement de réassurance.

### **ii. Calculs de chargement de réassurance global et extension de capacité**

Le chargement de réassurance a été considéré comme acquis dans l'étude. La méthode de calcul varie cependant fortement d'un acteur du marché de la réassurance à un autre. Deux réassureurs n'ont en effet pas la même structure de frais généraux et n'utilisent pas nécessaire la même métrique pour prendre en compte la volatilité d'un traité (écart type ici). De plus, les facteurs de chargement peuvent aussi varier.

Pour un unique réassureur, on peut même observer des chargements variables selon le temps de souscription. En effet, si un réassureur prend en compte les accumulations de risques sur une même localisation, il aura tendance à charger de plus en plus les traités sur une même zone au fur et à mesure qu'il les souscrira. Il prend alors en compte le coût marginal lié à l'ajout d'un nouveau traité sur ses accumulations.

Un second point sur les calculs de chargement de réassurance tient à ce que, en période de marché fortement compétitif (du fait de taux d'intérêt bas et d'une abondance de liquidités sur les marchés financiers), il est possible de constater des primes commerciales inférieures à ce que donnent les modèles Cat en prime pure. Dans ce cas-là, le réassureur préfère prendre le risque de sous-tarifier plutôt que de laisser passer des affaires de façon à conserver sa part de marché et sa relation client avec les assureurs.

Au sujet de l'extension de la portée d'un traité qui serait fixée à un quantile de perte, l'estimation du chargement de réassurance généré par cette extension n'est pas évidente. Dans l'idéal, il faudrait en effet être capable de calculer le chargement de réassurance de l'extension uniquement. Une manière de se rapprocher de ce coût de l'extension serait de considérer que le

réassureur utilise la même méthode de chargement que présentée au III.f et calculer les variations de prime pure et de volatilité liée à l'ajout d'un site dans le portefeuille.

### iii. Un approfondissement des méthodes d'allocation et de prédiction

Les méthodes qui donneraient potentiellement les résultats les plus « justes », en prenant en compte l'ensemble des possibilités d'entrée dans le jeu de chaque site, sont la méthode de Shapley et celle du Nucléole. Ces méthodes, particulièrement gourmandes en calcul, sont néanmoins impossibles à mettre en œuvre avec les outils à disposition aujourd'hui sur un portefeuille dépassant quelques milliers de sites.

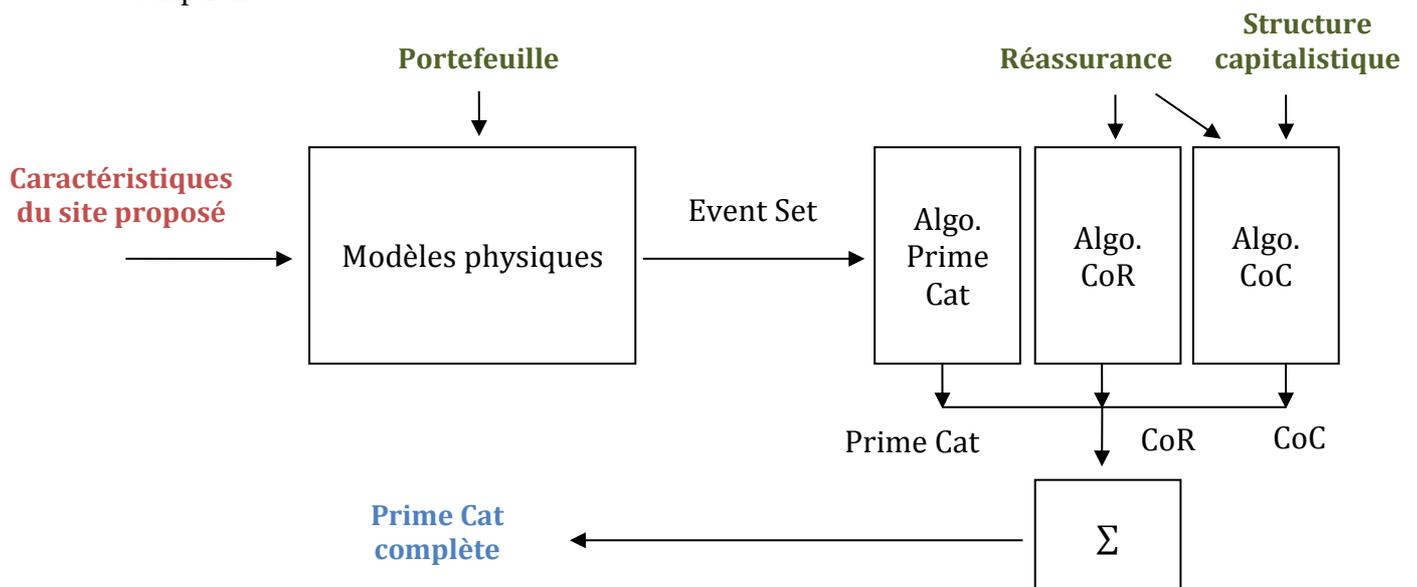
Au niveau des méthodes de prédiction statistiques des résultats produits par le modèle stochastique, il aurait été possible de tester d'autres alternatives pour améliorer encore la qualité de la prédiction. En effet, certaines méthodes de machine learning aurait pu permettre d'accroître la qualité des prédictions. Par exemple, utiliser un gradient boosting model (arbre de décision sur lequel on applique du boosting) ou un réseau de neurones auraient pu être des alternatives intéressantes au modèle de forêt aléatoire. Le risque de ces méthodes est néanmoins de devenir moins interprétable que les modèles classiques.

L'ensemble des méthodes d'allocation utilisées ne présente que l'utilisation des récupérations de réassurance sans inclure la volatilité de celles-ci, facteur qui entre aussi en compte dans le chargement de réassurance. Inclure cette volatilité permettrait de raffiner l'étude.

### iv. Vers un véritable outil de machine learning

Globalement, on a cherché à calculer le chargement de réassurance (ou du capital) à partir de résultats site par site d'un modèle. Le calcul a été assez artisanal (via des algorithmes spécifiques) et s'appuie sur des résultats de modèles eux aussi spécifiques.

Dans l'idéal, il pourrait être intéressant d'avoir un processus d'allocation complètement intégré à la tarification site par site, avec un cadre de calcul intégré qui s'adapterait à chaque nouvelle souscription.



**Figure 44 :** Schéma d'un modèle idéal de calcul intégré de prime Cat intégrant à la fois la prime Cat, le coût de réassurance et le coût du capital. En rouge, l'input de départ serait entré par un souscripteur qui recevrait en bleu l'output. Ce dernier serait calculé à partir des paramètres du modèle en vert.

Au sujet du renouvellement des polices, on pourrait aussi imaginer un système qui mettrait à jour chaque année la prime en prenant en compte les nouvelles conditions de réassurance négociées pendant la période de renouvellement ainsi que les nouveaux sites souscrits depuis l'entrée du site re-tarifé dans le portefeuille. Ainsi, la prime évoluerait à chaque reconduction du contrat en prenant en compte ces divers éléments. Le chargement de réassurance et du capital serait alors intégré quasiment en temps réel dans le système tarifaire de l'assureur.

Enfin, le processus opérationnel pourrait garder en mémoire le chargement de réassurance qui a déjà été alloué au portefeuille et ajuster le chargement des futures affaires en fonction des choix de souscription. Si, par exemple, un souscripteur choisit de ne pas allouer le chargement de réassurance estimé pour une nouvelle affaire, ce « manque à allouer » sera reporté sur les affaires suivantes de façon à inclure un pilotage de cette allocation en fonction de l'intérêt commercial des affaires.

Ainsi, pour un assureur qui déciderait de démarrer dès demain l'allocation de son chargement de réassurance sur son portefeuille, la démarche pourrait être la suivante :

- Calcul de l'allocation des chargements de réassurance et du capital sur les sites existants selon la méthode choisie selon le degré de segmentation souhaité.
- Ajout des chargements de réassurance et du coût du capital à la prime commerciale au moment du renouvellement de la police (l'allocation prendrait alors au moins un an) en intégrant bien sûr les autres éléments de la tarification (prime Cat, prime incendie, frais de gestion, etc.).
- Pour les nouvelles souscriptions, estimation du chargement de réassurance et du coût du capital et proposition d'un tarif au souscripteur (ou à l'assuré pour un portefeuille de risques « retail »).
- En risques entreprise (« commercial »), ajustement de la prime selon la vision du souscripteur du site.
- En fonction des ajustements précédents, le chargement de réassurance serait « répercuté » sur les souscriptions suivantes.

De cette manière, on aurait une intégration complète des chargements de réassurance et du capital dans la chaîne de valeur de l'assureur, ce qui permettrait un pilotage très fin de la souscription.

## **VI- CONCLUSION**

Les travaux de ce mémoire avaient pour objectif de proposer une approche conceptuelle puis pratique de l'allocation du chargement de réassurance et du coût du capital dans la tarification du risque ouragan au Mexique. A cet effet, différentes méthodes de théorie des jeux coopératifs ont été étudiées au plan théorique puis appliquées à un exemple simple avec pour objectif de proposer une approche d'allocation du chargement de réassurance au niveau des assurés. Il a été conclu que les méthodes les plus satisfaisantes au niveau théorique (méthode de Shapley, méthode du Nucléole) sont aussi les plus difficiles à mettre en œuvre en termes de calcul à grande échelle. Les méthodes sélectionnées ont donc été la méthode des coûts proportionnels et celle des coûts marginaux appliquées aux récupérations de réassurance générées par le modèle sur les sites assurés.

D'un point de vue théorique, chaque méthode présente des avantages et des inconvénients en termes de justice de l'allocation entre les assurés.

D'un point de vue pratique, il convient aussi de prendre en compte l'objectif de segmentation de l'assureur. En effet, contrairement à un tarif d'assurance classique qui vise à segmenter chaque assuré selon ses qualités intrinsèques, l'allocation proposée dépend en grande partie des sites déjà en portefeuille. Le choix d'une méthode d'allocation par rapport à une autre repose donc aussi sur le niveau de segmentation du chargement de réassurance souhaité par l'assureur. En effet, si l'assureur alloue le chargement de réassurance sur un trop petit nombre de sites, il risque de se trouver hors marché au moment de la souscription et de perdre des parts de marché.

De plus, étant donné les incertitudes liées à la modélisation et au choix de la méthode d'allocation, l'assureur peut avoir intérêt à rester pragmatique en appliquant un chargement global (ou mutualisé à un niveau CRESTA par exemple) pour un portefeuille de risques de faible somme assurée (particuliers) et en segmentant davantage pour les risques de forte valeur ajoutée (entreprises).

Une approche pour prendre en compte l'effet de la souscription d'un nouveau site sur la capacité de réassurance lorsque ce site se situe dans une zone fortement exposée aux catastrophes naturelles a été aussi proposée. Dans cette approche, on calcule la capacité qu'il est possible de souscrire dans chaque municipalité sans impacter l'événement deux-centenaire (supposée pris comme référence pour l'achat de capacité).

Enfin, de la même manière que le chargement de réassurance a été alloué dans la première partie du mémoire, une approche spécifique a été proposée pour l'allocation du coût du capital au sens Solvabilité 2. Cette approche se base sur la queue de distribution après application de la réassurance. Ici encore, les réflexions sur le choix d'un niveau de segmentation s'appliquent.

Finalement, s'il n'a pas la prétention de proposer une approche définitive en matière d'allocation des coûts du capital et de réassurance, ce mémoire suggère un référentiel de réflexion et une méthode applicable à un portefeuille d'assurance, à condition d'avoir accès aux données individuelles site par site d'un modèle interne.

## **VII- BIBLIOGRAPHIE**

*Home Shapes And Roofs That Hold Up Best In Hurricanes*, 21 Juin 2007, New Jersey Institute of Technology, <https://www.sciencedaily.com/releases/2007/06/070619155735.htm>

*Final CEIOPS' Advice for Level 2 Implementing Measures on Solvency II: Technical Provisions – Article 86 (d) Calculation of the Risk Margin*, Oct 2009, CEIOPS (EIOPA).

*Describing inequality in plant size or fecundity*, 2000, C. Damgard & J. Weiner, *Ecology* , 81(4), 2000, pp. 1139–1142.

*Insurance risk management and insurance*, January 2016, G.Gorge & al.

*Coherent allocation of risk capital*, January 2001, Michel Denault, HEC Montréal.

*Etudes internes sur l'inondation*, 2014-2015, AGPC.

*Etude interne de capacité de souscription restante pour les ouragans*, 2016, AGPC.

*Introduction à la théorie des jeux (2) : les jeux coopératifs*, Dominique LEPELLEY, Michel PAUL et Hatem SMAO, écoflash n°283, Décembre 2013.

*Méthodes d'allocation du capital*, 2008, V. Monin, AXA.

*Probabilités, analyses de données et statistiques*, 2011, Gilbert Saporta, Technip.

*Théorie des valeurs extrêmes*, 2002, Thierry Roncalli, Note de cours de l'université Paris Dauphine.

CATNET, SwissRe.

*La modélisation des catastrophes naturelles, les innovations face au développement des catastrophes non-vie*, Journée d'étude IARD, 2013, EQECAT.

*RMS® North Atlantic Hurricane Models*, 2016, RMS.

*The AIR tropical cyclone model for Mexico*, 2016, AIR Worldwide.

*A guide to catastrophe modelling*, 2008, RMS, The Review.

*About Catastrophe Models*, 2012, AIR Worldwide.

*Catastrophes naturelles et réassurance*, 2003, Swiss Re.

*Finance d'entreprise*, 2017, Pierre Vernimmen, Dalloz.

*Code des assurances (n.d.)*.

*Directive 2009/138/CE*, 2009, Parlement européen et Conseil de l'Union européenne, Journal officiel de l'Union européenne.

## VIII- LISTE DES FIGURES

Figure 1 : Carte de la tarification Cat obtenue avec le modèle.....	5
Figure 2 : Graphes des chargements de réassurance prédits en fonction des coûts à prédire.....	8
Figure 3 : Classement des variables explicatives par ordre d'importance dans les arbres de régression	8
Figure 4 : Carte du coût du capital proposé pour la tarification, en taux de somme assurée .....	9
Figure 5: Map of the Cat pricing produced with the stochastic model result.....	10
Figure 6 : graphs of predicted reinsurance costs over costs to predict.....	13
Figure 7 : Explanatory variables ranking in order of importance for the regression trees .....	13
Figure 8 : Map of the proposed cost of capital pricing, in sum insured rate .....	14
Figure 9 : carte des sites géo localisés à l'aide de leurs coordonnées. ....	23
Figure 10 : comparaison de la répartition du type de structure entre le portefeuille utilisé pour la modélisation (gauche) et un portefeuille marché (droite, source : AIR, 2016) .....	23
Figure 11 : exemple de carte du risque tempête d'un point de vue de l'aléa selon un modèle marché (EQE). ....	24
Figure 12 : Fit d'une distribution de valeurs extrêmes sur les maxima d'une série temporelle de hauteur d'eau pour une station de mesure hydrologique.....	24
Figure 13 : Propagation de l'ouragan Patricia sur la période du 20 au 24 Octobre 2015.....	25
Figure 14 : Exemple de courbe de vulnérabilité.....	26
Figure 15 : Exemple de vulnérabilité relative en fonction de la ligne de métier.....	26
Figure 16 : Application des conditions financières à une perte brute d'assurance.....	27
Figure 17 : application des conditions financières à une perte brute sur une police multi-site.....	27
Figure 18 : Comparaison des pertes modélisées aux pertes historiques lors d'événements passés.....	28
Figure 19 : Courbes d'OEP et d'AEP pour un modèle d'ouragan mexicain.....	29
Figure 20 : Comparaison des OEP de trois modèles vendeurs pour un même portefeuille d'assurés, pour le péril tempête mexicaine. ....	31
Figure 21 : Comparaison des sites les plus détruits par le modèle avec la carte des vitesses de vent... ..	34
Figure 22 : carte de la tarification Cat obtenue avec le modèle. ....	35
Figure 23 : Comparaison du fonctionnement d'un traité par risque par rapport à un par événement ..	37
Figure 24 : Calcul de la valeur de Shapley pour les trois sites de l'exemple .....	51
Figure 25 : Cartes de la répartition des pertes individuelles par municipalité mexicaine (% du total). ..	56
Figure 26 : Cartes des taux moyen de pertes individuelles par municipalité mexicaine.....	57

Figure 27 : Cartes de la répartition des récupérations de réassurance par municipalité mexicaine .....	60
Figure 28 : Cartes des taux moyen de récupérations de réassurance par municipalité mexicaine .....	60
Figure 29 : Illustration du fonctionnement d'une forêt aléatoire.....	64
Figure 30 : Evolution de l'erreur de prédiction au sens de la somme du carré des résidus en fonction du nombre d'arbres de régression choisi .....	65
Figure 31 : Classement des variables explicatives dans les arbres de régression.....	66
Figure 32 : Graphes des taux de réassurance prédits en fonction des taux à prédire. ....	66
Figure 33 : Graphes des chargements de réassurance prédits en fonction des coûts à prédire.....	67
Figure 34 : Graphes des chargements de réassurance prédits en fonction des coûts à prédire.....	67
Figure 35 : Cartes de la répartition du chargement de réassurance par état mexicain (% du total).....	70
Figure 36 : Carte des différences en points entre le pourcentage de chargement de réassurance de l'état avec la méthode des récupérations marginales et celui avec la méthode des récupérations individuelles. ....	70
Figure 37 : Illustration de la courbe de Lorentz. ....	72
Figure 38 : Schéma du procédé de calcul de la valeur de Shapley et de la méthode du nucléole. ....	77
Figure 39 : Comparaison des résultats d'allocation de récupération par méthode en pourcentage de somme assurée.....	78
Figure 40 : Carte des sites ayant un impact sur l'OEP200 de plus de 1% de leur somme assurée. ....	80
Figure 41 : carte de la capacité pouvant être souscrite sans variation de OEP200 modélisée. ....	81
Figure 42 : Carte de la répartition du coût du capital en vision modèle, par municipalité.....	86
Figure 43 : Carte du coût du capital proposé pour la tarification, en taux de somme assurée. ....	86
Figure 44 : Schéma d'un modèle idéal de calcul intégré de prime Cat intégrant à la fois la prime Cat, le chargement de réassurance et le coût du capital. ....	92
Figure 45 : Arbre de régression du taux de destruction moyen des sites sur le portefeuille ouragan de l'étude.....	102

## **IX- ANNEXES**

### **a. Rappels sur la théorie des valeurs extrêmes**

La théorie des valeurs extrêmes est basée sur le théorème de Fisher-Tippett-Gnedenko qui donne une distribution des maxima d'une suite de variables aléatoires indépendantes.

#### i. Théorème

Soit  $(X_1, \dots, X_n)$  une suite de variables aléatoires indépendantes et identiquement distribuées, et  $(M_1, \dots, M_n)$  la suite des maxima associée.

S'il existe une suite de couples  $(a_n, b_n)$  dans  $R_+^* \times R$  tel que

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} < x\right) = F(x)$$

alors  $F$  est une distribution généralisée des valeurs extrêmes (GEV).

On dit alors que la loi des  $X_i$  appartient au max-domaine d'attraction de  $F$ .

#### ii. Caractéristiques

La distribution de valeurs extrêmes se caractérise par trois paramètres et sa fonction de répartition est :

$$G_{(\xi, \mu, \sigma)}(x) = e^{-\left(1 + \xi \left(\frac{x - \mu}{\sigma}\right)\right)^{-1/\xi}} \quad \text{si } 1 + \xi \left(\frac{x - \mu}{\sigma}\right) > 0 \text{ et } \xi \neq 0$$

Elle est prolongée par continuité en  $\xi = 0$  par :

$$G_{(0, \mu, \sigma)}(x) = e^{-e^{-\left(\frac{x - \mu}{\sigma}\right)}}$$

La méthode la plus couramment utilisée en pratique pour construire un échantillon de maxima de suite de variables aléatoires indépendantes consiste à découper une série temporelle en blocs de même durée, puis à en extraire les maxima sur lesquels on cherchera ensuite à calibrer une distribution de valeurs extrêmes.

Ces distributions de valeurs extrêmes ont un intérêt particulier par rapport aux distributions plus classiques (normales, poisson, etc.) en ce qu'elles présentent une queue de distribution plus épaisse. Autrement dit, elles permettent de mieux prendre en compte la génération d'événements rares, ce qui est l'objectif recherché en réassurance.

Cette épaisseur de queue de distribution est caractérisée par le paramètre  $\xi$  et plus sa valeur absolue est importante, plus l'épaisseur de la distribution sera importante aussi.

Enfin, selon la valeur du paramètre  $\xi$ , la distribution se simplifie en loi de Fréchet ( $\xi > 0$ ), loi de Weibull ( $\xi < 0$ ) ou loi de Gumbel ( $\xi = 0$ ).

## **b. Démonstration de la propriété sur la méthode des bénéfices résiduels**

### i. Propriété

Si la fonction de coût vérifie les deux conditions ci-après, alors la méthode des bénéfices résiduels vérifie la rationalité collective.

$$(1) \forall i, c(i) > CM(i)$$

$$(2) \sum_{k=1}^n CM(k) \leq c(N)$$

### ii. Démonstration

$$\forall i, x_i \geq CM(i) \Leftrightarrow c(i) - CM(i) \geq \frac{c(i) - CM(i)}{\sum_{j=1}^n (c(j) - CM(j))} \left[ \sum_{k=1}^n c(k) - c(N) \right]$$

Sachant que  $c(i) - CM(i) > 0$  et  $\sum_{k=1}^n c(k) - c(N)$ , il vient :  $\forall i$ ,

$$x_i \geq CM(i) \Leftrightarrow \sum_{j=1}^n (c(j) - CM(j)) \geq \sum_{k=1}^n c(k) - c(N)$$

En développant, on aboutit à :

$$x_i \geq CM(i) \Leftrightarrow \sum_{j=1}^n c(j) - \sum_{j=1}^n CM(j) \geq \sum_{k=1}^n c(k) - c(N)$$

Ce qui donne, après simplification et multiplication par -1 :  $\forall i$ ,

$$x_i \geq CM(i) \Leftrightarrow \sum_{j=1}^n CM(j) \leq c(N)$$

### c. Démonstration du nombre de calculs pour la méthode de Shapley

Pour calculer la valeur de Shapley des sites d'un ensemble de  $N+1$  sites, il convient de calculer, pour chaque site  $i$ , le chargement de réassurance correspondant à son arrivée au rang  $k+1$  ( $k$  appartenant à  $[0 ; N]$ ), dans toutes les configurations possibles des autres sites du sous-portefeuille de  $N$  sites  $\{N+1 \setminus i\}$ .

Il y a donc, pour chaque site  $i$ , pour chaque rang d'arrivée  $k+1$ ,  $k$  parmi  $N$  possibilités pour les autres sites. Autrement dit :

$$Nb\ Calcul = \sum_{i=1}^{N+1} \sum_{k=0}^N \binom{N}{k} = (N+1) \cdot \sum_{k=0}^N \binom{N}{k}$$

Or, la somme des coefficients binomiaux présente la propriété suivante (que l'on démontre grâce au binôme de Newton, lui-même démontré par récurrence) :

$$\sum_{k=0}^N \binom{N}{k} = \sum_{k=0}^N \binom{N}{k} \cdot 1^k \cdot 1^{N-k} = (1+1)^N = 2^N$$

Par suite, on a donc :

$$Nb\ Calcul = (N+1) \cdot 2^N$$

Ce qui démontre que le calcul de la valeur de Shapley sur le portefeuille de  $N+1$  sites nécessite un nombre de calculs égal à  $(N+1) \cdot 2^N$ .

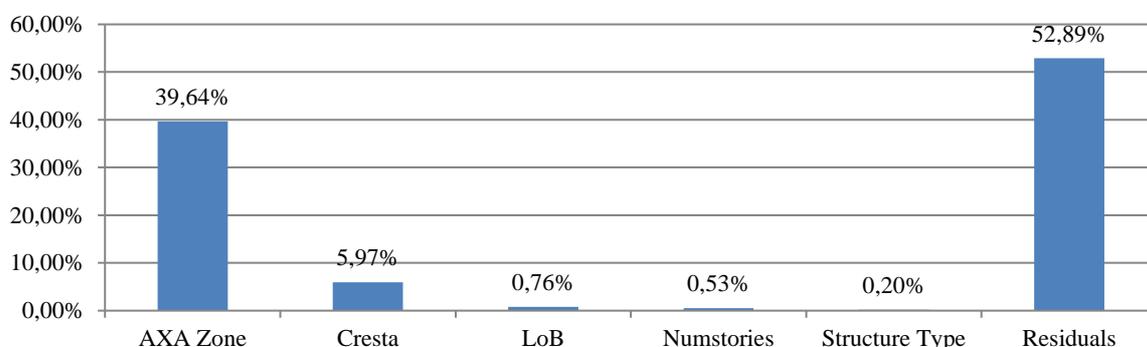
#### d. Résultats de la construction des modèles de prédiction (GLM)

Le tableau ci-dessous présente les résultats d'une analyse de variance sur un modèle linéaire simple calibré sur les résultats de chargement de réassurance au niveau site

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Variance expliquée
factor(AXA_zone)	6	1,14E-01	1,90E-02	13530	0,00E+00	39,64%
Cresta	31	1,72E-02	5,55E-04	394	0,00E+00	5,97%
LOB	3	2,20E-03	7,33E-04	521	0,00E+00	0,76%
factor(NumStories)	56	1,52E-03	2,72E-05	19	7,17E-189	0,53%
StructureType	4	5,90E-04	1,48E-04	105	2,40E-89	0,20%
Residuals	108299	1,52E-01	1,41E-06	NA	NA	52,89%

On constate que la variable AXA\_Zone est la plus importante et contient à elle seule 84% de la variance expliquée par le modèle.

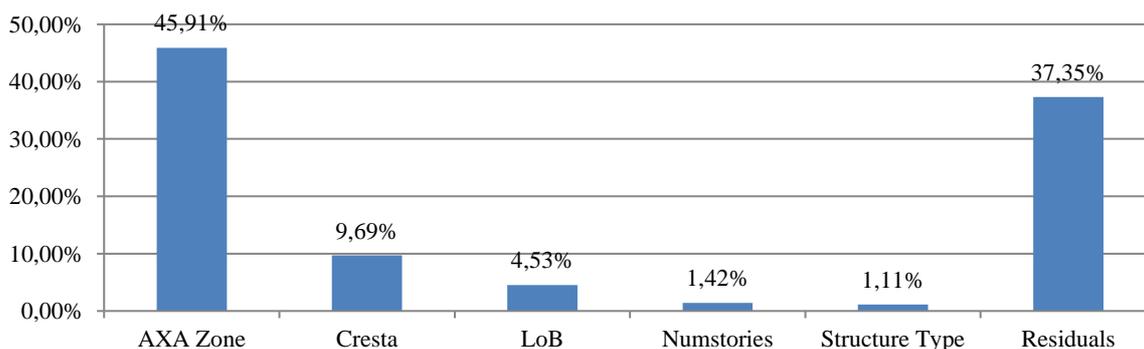
**Variance expliquée par le modèle linéaire**



Dans le tableau ci-dessous, on a restreint la prédiction aux sites présentant une perte moyenne de plus de 100€, ce qui permet de conserver 97% du chargement de réassurance et d'améliorer la qualité de la prédiction par modèle linéaire simple de 15 points de variance expliquée.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Variance expliquée
factor(AXA_zone)	6	9,5E-02	1,6E-02	2960	0,0E+00	45,91%
Cresta	31	2,0E-02	6,5E-04	121	0,0E+00	9,69%
LOB	3	9,4E-03	3,1E-03	584	0,0E+00	4,53%
factor(NumStories)	44	2,9E-03	6,7E-05	12	1,8E-86	1,42%
StructureType	4	2,3E-03	5,7E-04	107	6,0E-90	1,11%
Residuals	14445	7,7E-02	5,3E-06	NA	NA	37,35%

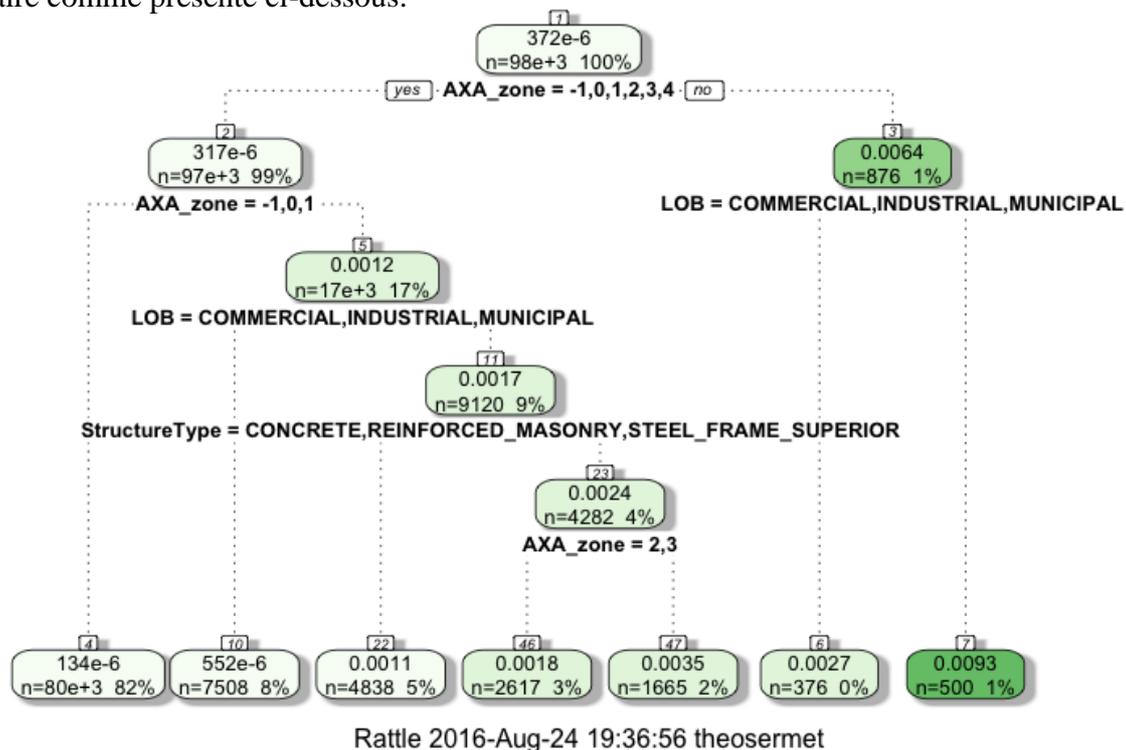
**Variance expliquée par le modèle linéaire**



### e. Rappels sur les arbres de régression et de décision

Cette méthode fait partie des outils de prédiction de data mining. Elle permet de prédire une variable quantitative ou qualitative. Dans le premier cas, on utilisera un arbre de régression, dans le second, un arbre de décision (ou discrimination).

L'idée est de diviser l'échantillon d'apprentissage en deux sous-ensembles à l'aide d'un des prédicteurs. On recommence ensuite avec chaque sous-ensemble. Cela construit un arbre binaire comme présenté ci-dessous.



**Figure 45 :** Arbre de régression du taux de destruction moyen des sites sur le portefeuille ouragan de l'étude. On constate clairement que la zone de risque AXA\_zone est discriminante quant au taux de destruction.

Le critère de séparation d'une branche (nœud) dépend de la variable à prédire.

Si la variable est quantitative, on cherchera à minimiser la variance interclasse ou à maximiser la variance intraclasse, ce qui est équivalent (théorème de Koenig Huygens appliqué à une classification).

Si elle est qualitative, on s'intéressera alors plutôt à l'entropie ou à l'indice de diversité de Gini.

Une fois l'arbre construit, il suffit de suivre les règles associées à chaque nœud selon les valeurs des prédicteurs de l'individu étudié pour obtenir sa valeur prédite.

La qualité de la prédiction est alors mesurée à l'aide d'indicateurs assez classiques comme par exemple le  $R^2$  pour une variable continue et le tableau de classement pour une variable qualitative.