

**Mémoire présenté devant l'ENSAE Paris
pour l'obtention du diplôme de la filière Actuariat
et l'admission à l'Institut des Actuaires
le 17/03/2022**

Par : **Archie Murcy**

Titre: **Etude d'un modèle de tarification de catastrophes
naturelles par simulations pour les traités CAT XL Vie**

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

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

Membres présents du jury de la filière

Entreprise : AXA Global Re

Nom :

Signature :

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

Directeur de mémoire en entreprise :

Nom : Jérôme CRETIEN

Signature :

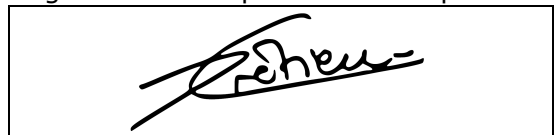


**Autorisation de publication et de
mise en ligne sur un site de**

diffusion de documents actuariels

*(après expiration de l'éventuel délai de
confidentialité)*

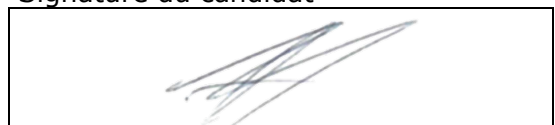
Signature du responsable entreprise



Secrétariat :

Signature du candidat

Bibliothèque :



Résumé

L'un des rôles de la réassurance est de permettre aux assureurs de mieux gérer le risque catastrophe notamment en recourant à des traités en excédent de sinistre par évènement. La tarification technique de ces traités est le résultat d'une analyse du risque CAT et du portefeuille d'assurance. Les modèles utilisés pour tarifier ce type de traité en réassurance vie et en réassurance non-vie se différencient pourtant fortement.

En réassurance non-vie, les modèles CAT sont de véritables modèles physiques agrégeant une quantité de données importantes permettant de simuler stochastiquement des évènements. En réassurance vie, le modèle par simulations fonctionne grâce à une base de données historiques en calibrant une sévérité et une fréquence.

Les limites du modèle actuellement utilisé en réassurance vie sont exposées et motivent cette étude dans la recherche d'un nouveau modèle.

Ce mémoire a ainsi pour objectif d'analyser la faisabilité d'un modèle de tarification en réassurance vie basé sur des simulations stochastiques de catastrophes naturelles. Cette faisabilité dépend des données disponibles pour un réassureur et des ajustements possibles sur les modèles existants en réassurance non-vie.

L'étude et la calibration d'un modèle de réassurance non-vie appliqué en réassurance vie sont justifiées par la dépendance qui existe entre le risque vie et non-vie liée aux catastrophes. Cette dépendance entre ces deux risques est mesurée et se révèle positive.

Dans le cadre de ce mémoire, le modèle physique en réassurance vie est calibré et appliqué au risque de tremblement de terre, ce qui permet d'évaluer sa performance et d'exposer la méthodologie mise en place.

Mots clés : Réassurance vie, Corrélation, Modèle physique, Tarification, Excédent de sinistre par évènement

Abstract

One of the roles of reinsurance is to enable insurers to better manage catastrophe risk, especially by using per event excess of loss treaties. The pricing of these treaties is the result of an analysis of the CAT risk and the insurance portfolio. However, the models used to price this type of treaties are very different between life and non-life reinsurance.

In non-life reinsurance, the CAT models are purely physical models that aggregate a large amount of data to stochastically simulate events. In life reinsurance, the simulation model works with a historical data base by calibrating a severity and a frequency.

The limits of the model currently used in life reinsurance are shown and motivate this study in the search for a new model.

The objective of this thesis is to analyse the feasibility of a pricing model for life reinsurance based on stochastic simulations of natural catastrophes. This feasibility depends on the data available to a reinsurer and the possible adjustments to existing non-life reinsurance models.

The study and calibration of a non-life reinsurance model applied in life reinsurance is justified by the dependence that exists between life and non-life catastrophe risk. The dependence between these two risks is measured and found to be positive.

In this thesis, the physical model in life reinsurance is calibrated and applied to earthquake risk, which allows us to assess its performance and to expose the implemented methodology.

Keywords: Life reinsurance, Correlation, Physical model, Pricing, Per event Excess of Loss

Note de synthèse

Contexte de l'étude et problématique

Face aux catastrophes, les assureurs sont exposés au risque de ne pas disposer de capitaux suffisants pour faire face aux pertes colossales engendrées par la survenance d'une catastrophe majeure, dans le contexte où l'occurrence de catastrophes naturelles devrait continuer à s'intensifier avec le changement climatique. Bien que les moyens de prévention face aux catastrophes naturelles aient progressé, le nombre de victimes engendré par ces événements reste important. La réassurance peut donc jouer le rôle de porter ce risque extrême non mutualisable pour un assureur. A cette fin, il est possible de mettre en place entre l'assureur et le réassureur un traité CAT XL Vie pour que la cédante bénéficie d'une couverture de réassurance face à ce risque.

Au sein du groupe AXA, l'entité AXA Global Re a pour mission de structurer et de placer des traités vie et non vie pour les autres filiales du groupe d'AXA. Ainsi, afin d'optimiser la structure de réassurance, une tarification technique des traités est effectuée par l'équipe Analytics & Pricing d'AXA Global Re en analysant les portefeuilles et les risques.

Les modèles de tarification vie et non-vie sont construits différemment. Les modèles de réassurance non-vie sont de véritables modèles de simulations stochastiques basées sur une approche physique demandant une quantité importante de données. Contrairement à la réassurance non-vie, les données recueillies auprès des cédantes pour les portefeuilles vie sont synthétiques. Les modèles de tarification en réassurance vie sont donc adaptés à cette contrainte et sont développés sur la base d'une approche sévérité-fréquence à l'aide d'un historique.

Face à ces différences d'approche, le mémoire vise à analyser la faisabilité d'un modèle physique de simulations de catastrophes naturelles pour la tarification des traités CAT XL Vie. L'enjeu est de savoir s'il est possible de recourir à une méthode utilisée actuellement en réassurance non-vie en faveur du domaine de la réassurance vie.

Cette faisabilité dépend des données disponibles pour un réassureur, des hypothèses à appliquer et des ajustements possibles sur les modèles de tarification existants. Pour justifier l'utilisation d'un modèle de réassurance non-vie en réassurance vie, la dépendance entre le risque vie et non-vie liée aux catastrophes naturelles a été analysée.

Le modèle de tarification en réassurance vie

Le modèle de tarification actuel d'AXA est basé sur un modèle de référence, celui de Ekhedden. Il s'agit d'un modèle de simulations stochastiques avec la calibration d'une sévérité et d'une fréquence basée sur un historique de données. La Figure 1 présente le fonctionnement du modèle.

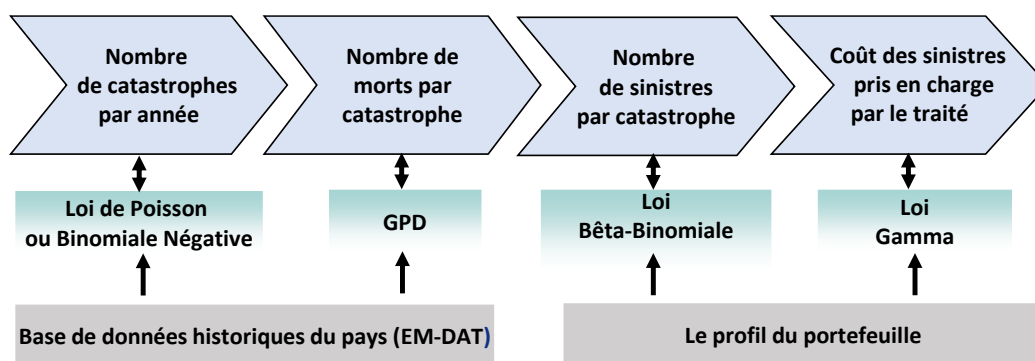


Figure 1 – Modèle de tarification actuel (Ekhedden)

Cette approche intègre la théorie des valeurs extrêmes en s’intéressant à la sévérité et à la fréquence des évènements au-delà d’un seuil, ce qui explique le choix d’une distribution de Pareto Généralisée (GPD) au sein du modèle.

Grâce à des hypothèses, ce modèle s’adapte bien aux données synthétiques recueillies auprès des cédantes pour appliquer le profil du portefeuille. Cependant, ce modèle pour tarifier les traités CAT XL Vie présente des limites importantes.

Une d’entre elles est que le faible nombre d’observations du passé ne permet pas de projeter l’ensemble des évènements dans le futur. Des évènements majeurs ont une période de retour très importante qui se compte en milliers d’année comme pour les grandes catastrophes. Il est ainsi probable qu’au sein d’un historique de 50 ans, que des observations de ces phénomènes ne se soient pas réalisées. Il y a donc une sous-estimation de catastrophes majeures au niveau des queues de distribution.

Pour contourner les limites posées par une base historique, il apparaît opportun d’engager en complément la simulation de catastrophes naturelles en ayant recours à un catalogue d’évènements stochastiques créé par des experts du risque CAT. La dépendance entre le risque CAT Vie et le risque CAT non-vie a été donc mesurée afin de justifier l’étude d’un modèle physique en réassurance vie pour l’analyse du risque catastrophe.

Dépendance entre le risque vie et non-vie liée aux catastrophes naturelles

La base internationale de catastrophes (EM-DAT) a été exploitée pour analyser la dépendance. La base EM-DAT recense le nombre de victimes et les pertes en assurance non-vie assurées pour des catastrophes naturelles survenues dans le monde.

Pour chaque évènement, il a été calculé des taux de destruction. Ne bénéficiant pas des expositions vie et non-vie dans tous les pays, il a été utilisé des données socio-économiques de la Banque Mondiale (taille de la population et PIB) pour calculer les taux de destruction.

$$\text{Taux de destruction}_{\text{vie}} = \frac{\text{Nombre de victimes}}{\text{Taille de la population du pays}}$$

$$\text{Taux de destruction}_{\text{Non-Vie}} = \frac{\text{Pertes assurances}}{\text{PIB du pays}}$$

L’utilisation du taux de destruction a permis de rendre comparable chaque évènement ayant eu lieu à des périodes différentes et dans des pays différents. Cette utilisation a aussi pour avantage d’éviter la

mise en « As If » qui signifie faire comme si les évènements étaient survenus aujourd’hui afin d’avoir un nombre de victimes et une perte assurée à aujourd’hui.

Il a été recouru une analyse avec les coefficients de corrélation linéaires (Pearson) et non-linéaires (Spearman). Les résultats démontrent l’existence d’une corrélation non-linéaire aux alentours de 50% entre les taux de destruction vie et non-vie avec une dépendance plus importante pour les évènements extrêmes. Les évènements ayant les taux de destruction vie les plus importants sont aussi ceux qui ont un taux de destruction les plus élevés en non-vie.

Cette dépendance est donc non-nulle contrairement à ce que suppose la formule standard de Solvabilité 2.

Les principes de la modélisation physique

En exploitant cette dépendance, il devient opportun d’étudier les possibilités l’application du modèle physique en réassurance vie.

Les modèles physiques sont d’un niveau de complexité plus élevé qu’en réassurance vie. Il s’agit de modèles physiques de simulations stochastiques d’évènements (ou de catastrophes). Il est aussi possible d’utiliser des modèles de marché développés par des sociétés spécialisées (RMS, AIR, ...). La structure du modèle physique est présentée par la Figure 2 ci-dessous.

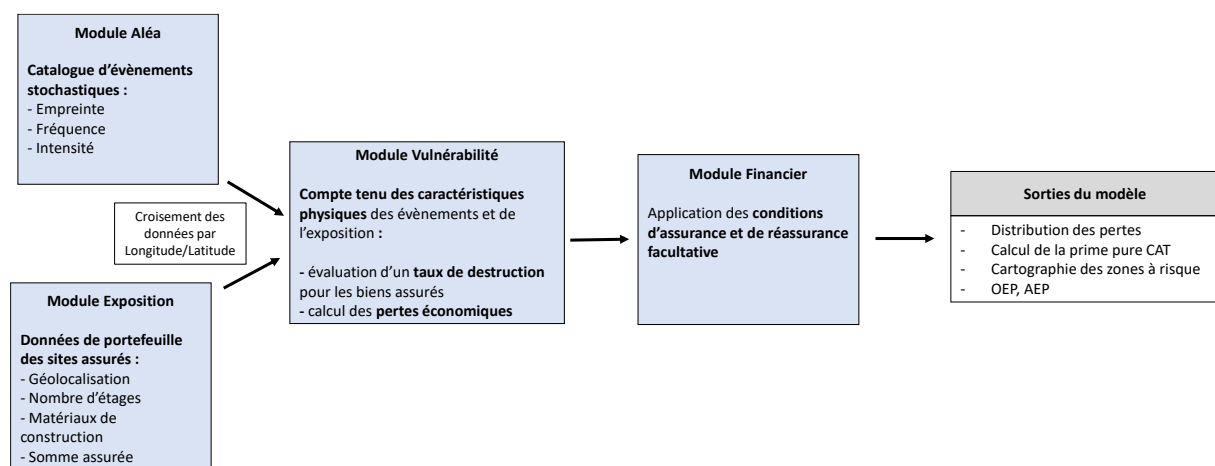


Figure 2 – Structure du modèle physique en réassurance non-vie

Le modèle physique est construit sur différents modules (aléa, exposition, vulnérabilité et financier), cela permet de prendre en compte de nombreux facteurs et d’analyser le plus finement le risque CAT. Ce modèle permet de simuler stochastiquement des évènements et d’analyser pour chacun des évènements leur impact sur un portefeuille d’assurance donné. L’efficacité d’une couverture de réassurance est testée car une distribution des pertes est dressée.

Calibration du modèle physique pour la réassurance-vie grâce à des backtests

En s’appuyant sur des backtests, l’objectif était de calibrer un modèle physique et d’évaluer sa capacité à simuler le nombre de morts engendrés pour une catastrophe donnée. La calibration s’est faite pour le risque tremblement de terre mais il est possible que ce modèle puisse être adapté à tous les risques CAT couverts par un traité de réassurance.

Les difficultés rencontrées pour la calibration sont liées à l’intégration d’une dimension humaine au sein du modèle. A l’opposé des bâtiments, les personnes sont mobiles. Selon, l’heure de la journée, le nombre de personnes exposées à une catastrophe sur une zone géographique donnée n’est pas la

même du fait de la mobilité. De plus, il n'est pas possible d'acquérir des données sur la localisation des individus car il s'agit de données personnelles.

La calibration du modèle a demandé ainsi de faire évoluer principalement deux modules du modèle physique de réassurance non-vie : le module d'exposition et le module de vulnérabilité. Il s'agissait d'inclure une base d'exposition dynamique et des courbes de vulnérabilité permettant de calculer une probabilité de décès.

Il a été inclus au sein de la base d'exposition le nombre d'occupants au sein des bâtiments (dimension humaine). De plus, cette base évolue selon l'heure de la journée, la répartition des occupants évoluent selon la structure socio-professionnelle du pays afin que l'exposition soit adaptée à l'heure de la catastrophe. Les probabilités de décès pour un niveau d'intensité physique donné à l'intérieur des bâtiments sont incluses à l'aide des courbes de vulnérabilité.

Il a été utilisé pour calibrer le modèle des données publiques comme la densité de population et autres informations disponibles auprès des instituts nationaux de statistiques. Des données de marché spécialisées sur le tremblement de terre issues du Global Earthquake Model ont été utilisées et adaptées au modèle.

Les résultats du backtest ont démontré une capacité du modèle à simuler un nombre de morts précis seulement où la densité de population est importante. Dans des zones rurales comme le centre de l'Italie, le modèle ne fonctionne pas. Nous avons remarqué que le modèle fonctionne comme sur le principe de la loi des grands nombres puisque les probabilités de décès appliquées avec les courbes de vulnérabilité sont faibles. Plus la population exposée aux ondes sismiques est importante, plus le nombre de morts simulés reflète la réalité.

Pour le Mexique où la densité de population est importante, le modèle simule les tremblements de terre avec un écart raisonnable de moins de 10% avec les données historiques. Cette différence est négligeable et encourageante au vu de la complexité de la modélisation.

Une étude de sensibilité sur le modèle a été réalisée. Cela a pu mettre en évidence que le modèle est capable de prendre en compte la surmortalité des catastrophes survenant la nuit. Nous retrouvons aussi des résultats observés en réassurance non-vie, les individus au sein du secteur résidentiel sont exposés à un risque de décès plus important du fait de la structure des bâtiments moins résistante aux ondes sismiques que les autres bâtiments du secteur commercial.

Cette étude de sensibilité a surtout prouvé l'importance sur la précision des données à entrer lors des simulations de catastrophes naturelles dans le modèle physique car les hypothèses sur la répartition et le nombre des occupants par bâtiment, les données sur les structures de bâtiment (le respect ou non des normes de construction pour le risque sismique), les courbes de vulnérabilité sont des facteurs très importants dont les résultats du modèle sont très sensibles.

Tarification du risque tremblement de terre dans un traité CAT XL Vie avec les données disponibles

Une fois avoir calibré le modèle, il a été proposé une première tarification du risque tremblement de terre d'un traité CAT XL Vie au Mexique avec la prise en compte d'un catalogue d'évènements stochastiques et des autres modules du modèle physique comme le montre la Figure 3. Cependant, les données à disposition pour un réassureur sur un portefeuille en assurance vie impliquent d'appliquer le profil du portefeuille sur les hypothèses du modèle de Ekhedden.

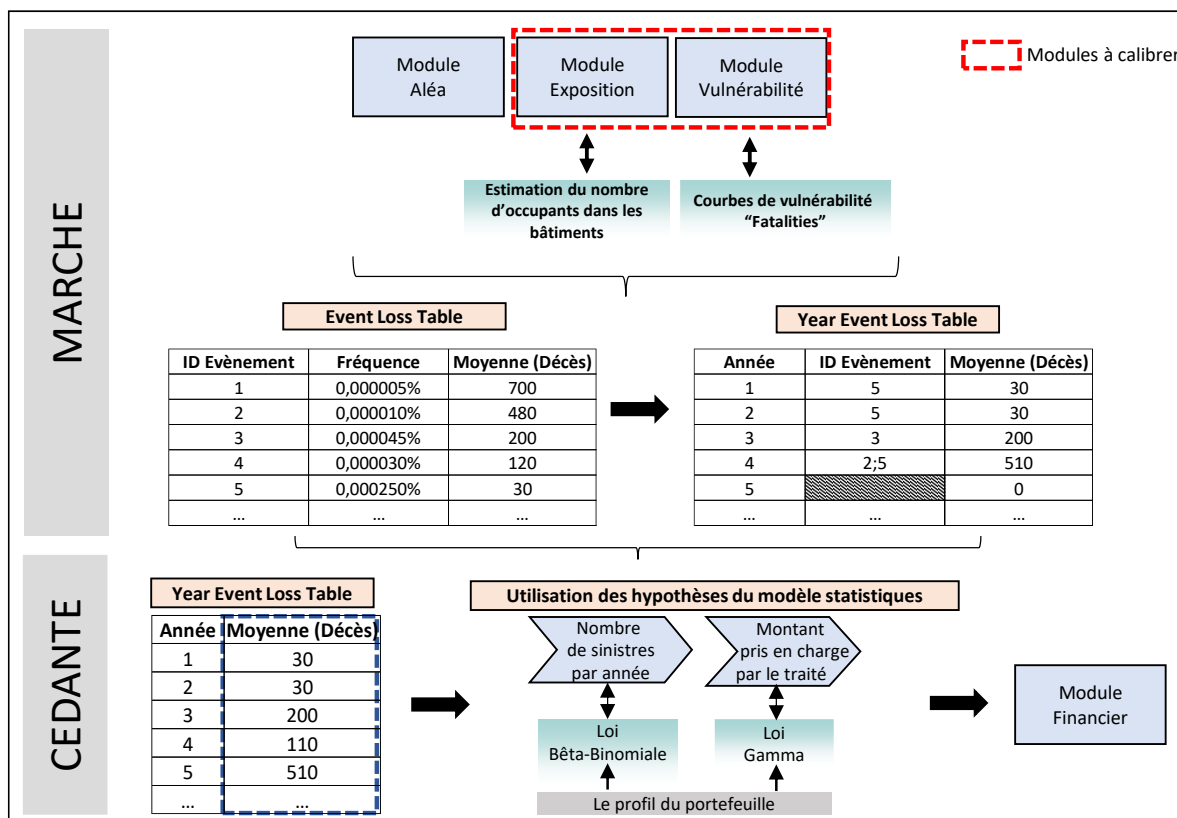


Figure 3 – Modèle de tarification proposé avec la prise en compte des modules physiques

Néanmoins, la modélisation est bien basée sur une simulation de catastrophes naturelles à partir d'un catalogue d'évènements stochastiques et le nombre de morts est simulé grâce aux modules physiques.

La simulation stochastique des évènements du catalogue permet de construire les Event Loss Table (ELT) répertoriant pour chaque évènement du catalogue une fréquence et une sévérité moyenne. A partir de ces ELT, il est possible de construire des scénarios par année (YELT) incluant la survenance et la sévérité des évènements de l'ELT. Cette méthodologie proposée permet de construire des scénarios par année répertoriant le nombre de décès pour chaque scénario.

Pour chaque scénario par année, il est utilisé une loi bêta-binomiale permettant de simuler le nombre d'assurés décédés au sein du nombre de victimes pour chaque scénario. Le montant de perte associé au décès de plusieurs assurés est estimé sur l'hypothèse de Ekheden avec l'utilisation d'une distribution gamma.

Avec le modèle de tarification étudié, le montant de la prime pure, c'est-à-dire la moyenne des récupérations sur tous les scénarios est supérieure de 41,18% à celle obtenue par le modèle standard de Ekheden. Cela s'explique principalement par une occurrence plus importante des sinistres dans le modèle physique.

Le modèle physique permet d'utiliser une base comprenant plus de 40 000 séismes alors que la base EM-DAT au sein du modèle de Ekheden répertorie seulement 32 séismes pour le Mexique. Le modèle physique présente un grand avantage par rapport au modèle statistique de mieux prendre en compte la sévérité des évènements catastrophes et de plus la fréquence est basée sur le catalogue développé par les experts.

Conclusions

Il s'avère ainsi possible d'envisager l'utilisation d'un modèle physique en réassurance-vie. L'intégration de la dimension humaine est la problématique majeure. Après calibration, le modèle est capable de simuler un nombre de morts avec des hypothèses simples mais fiables.

Il s'agit d'un modèle complexe, il est nécessaire de recourir à données de marché et surtout à des données précises sur l'exposition et les autres modules. Les données transmises actuellement par les cédantes sont synthétiques, ce qui aboutit actuellement à une prise en compte limitée du portefeuille d'assurance vie d'un assureur.

Le modèle physique présente de nombreux avantages, il permet de prendre en compte des évènements inconnus dans l'historique d'évènements disponible chez un assureur et apporte un degré de précision plus fin sur l'évaluation de la vulnérabilité et de l'exposition au risque. La réévaluation de la prime pure obtenue par ces calculs, d'environ 40 %, constitue également un point d'attention majeur, en apportant une nouvelle analyse sur le risque attaché au traité.

Ces différentes recherches constituent les prémises d'un éclairage nouveau permettant de mieux comprendre et évaluer le risque, en conservant à l'esprit les difficultés techniques de leur évaluation.

Il faudrait de ce fait continuer à approfondir cette démarche et recueillir notamment des données précises sur les portefeuilles en termes de nombres d'assurés ou de somme assurée par code postal par exemple. Cela permettrait d'implémenter un modèle complet pour effectuer des simulations en dépassant les hypothèses du modèle de référence (modèle de Ekhedén).

Executive summary

Context of the study and issues

In view of disasters, insurers are facing the risk of not having sufficient capital to cope with the colossal losses, especially in a context where the occurrence of natural disasters should continue to intensify with climate change. Although the means of prevention from natural disasters have improved, the number of victims generated by these events remains significant. Reinsurance can therefore play the role of carrying this extreme risk that cannot be mutualised for an insurer. To this end, it is possible to set up a Life CAT XL treaty between the insurer and the reinsurer so that the ceding company benefits from reinsurance cover against this risk.

Within the AXA Group, AXA Global Re is responsible for structuring and placing life and non-life treaties for other AXA Group subsidiaries. To optimise the reinsurance structure, a technical pricing of treaties is carried out by the Analytics & Pricing team of AXA Global Re by analysing the portfolios and risks.

The life and non-life pricing models are built differently. Non-life reinsurance models are true stochastic simulation models based on a physical approach requiring a large amount of data. Unlike non-life reinsurance, the data collected from cedants for life portfolios is synthetic. Life reinsurance pricing models are therefore adapted to this constraint and are developed based on a severity-frequency approach using a data history.

In view of these methodology differences, this paper aims to analyse the feasibility of a physical catastrophe simulation model for pricing Life CAT XL treaties. The issue at stake is whether it is possible to use the methodology currently used in non-life reinsurance for the life reinsurance domain.

This feasibility depends on the data available to a reinsurer, the assumptions to be applied and the possible adjustments to existing pricing models. To justify the use of a non-life reinsurance model in life reinsurance, the dependence between life and non-life risk related to natural catastrophes has been analysed.

The life reinsurance pricing model

AXA's current pricing model is based on a reference model, the Ekhedden model. It is a stochastic simulation model with calibration of a severity and frequency based on historical data. Figure 1 shows how the model works.

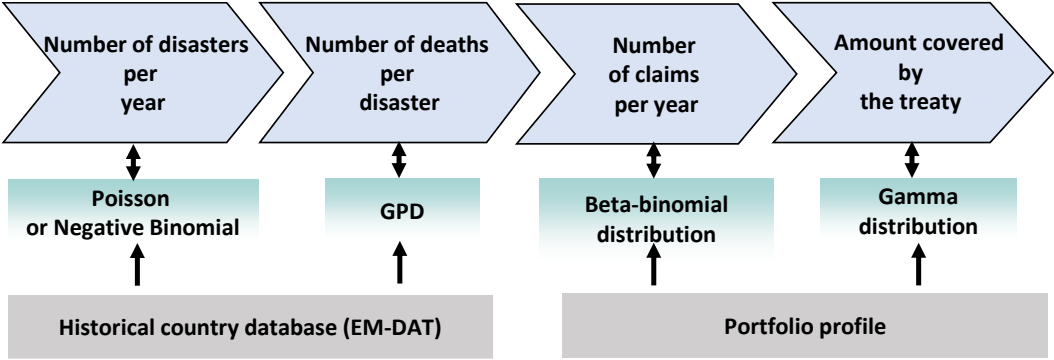


Figure 1 – The current pricing model (Ekhedden)

This approach incorporates extreme value theory by focusing on the severity and frequency of events above a threshold, which explains the choice of a Generalized Pareto Distribution (GPD) within the model.

Using the above assumptions, this model fits well with the synthetic data collected from cedants to apply the portfolio profile. However, this model for pricing CAT XL Life treaties has some important limitations.

One of these is that the small number of observations from the past does not allow the projection of all events into the future. Major events have a very high return period, which can be counted in thousands of years, as in the case of major catastrophes. It is therefore likely that within a 50-year history, some observations of these phenomena have not been made. There is therefore an underestimation of major disasters at the level of the distribution tails.

To overcome the limitations of a historical base, it seems appropriate to complement the simulation of natural catastrophes with a catalogue of stochastic events created by CAT risk experts. The dependency between the life CAT risk and the non-life CAT risk was therefore measured to justify the use of a physical approach to life reinsurance in order to analyse catastrophe risk.

Dependence between life and non-life risk from natural disasters

The International Catastrophe Database (EM-DAT) was used to analyse the dependency. The EM-DAT database records the number of victims and insured non-life losses for natural catastrophes worldwide.

For each event, destruction rates were calculated. As life and non-life exposures are not available for all countries, socio-economic data from the World Bank (population size and GDP) was used to calculate the destruction rates.

$$\text{Life rate destruction} = \frac{\text{Number of victims}}{\text{Size of the country's population}}$$

$$\text{Non - Life rate destruction} = \frac{\text{Insured Losses}}{\text{Country's GDP}}$$

The use of the destruction rate made the comparability of events that took place in different periods and in different countries possible. This use also has the advantage of avoiding the "As If" setting, that assumes that the events occurred today to have a number of victims and an assured loss as of today.

An analysis with linear (Pearson) and non-linear (Spearman) correlation coefficients was used. The results show that there is a non-linear correlation of around 50% between life and non-life destruction rates with a higher dependency on extreme events. The events with the highest life destruction rates are also those with the highest non-life destruction rates.

This dependency is therefore non-zero, opposed to what is assumed in the standard Solvency 2 formula.

The principles of physical modelling

By exploiting this dependency, it becomes appropriate to study the possibilities of applying the physical model in life reinsurance.

Physical models are of a higher level of complexity than in life reinsurance. They are physical models of stochastic simulations of events (or catastrophes). It is also possible to use market models developed

by specialised companies (RMS, AIR, etc.). The structure of the physical model is shown in Figure 2 below.

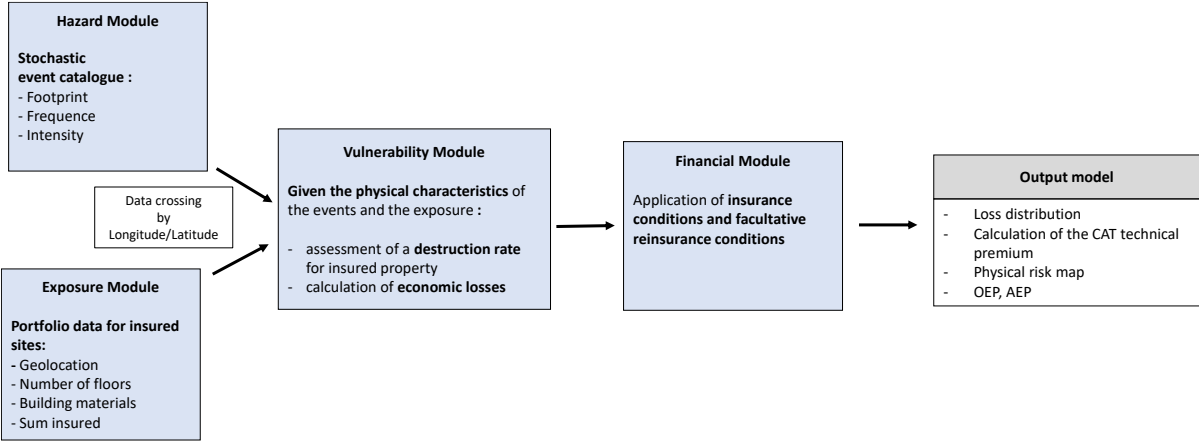


Figure 2 - The structure of the physical model in non-life reinsurance

The physical model is built on different modules (hazard, exposure, vulnerability and financial), which enables the use of many factors and hence to analyse the CAT risk in more detail. This model allows the stochastic simulation of events and the analysis of their impact on a given insurance portfolio for each event. The effectiveness of a reinsurance cover is tested because a loss distribution is drawn up.

Calibration of the physical model for life reinsurance through backtesting

Based on backtests, the objective was to calibrate a physical model and assess its ability to simulate the number of deaths produced for a given disaster. The calibration was done for earthquake risk but this model could be adapted to all the CAT risks covered by a reinsurance treaty.

The difficulties encountered in the calibration are related to the integration of a human dimension within the model. Unlike buildings, people are mobile. Depending on the time of day, the number of people exposed to a disaster in a given geographical area is not the same due to mobility. Furthermore, it is not possible to acquire data on the location of individuals as this is personal data.

The calibration of the model thus required the development of two main modules of the physical non-life reinsurance model: the exposure module and the vulnerability module. The aim was to include a dynamic exposure base and vulnerability curves allowing the calculation of a probability of death.

The number of occupants in the buildings (human dimension) was included in the exposure base. Moreover, this base evolves according to the time of day, the distribution of the occupants evolves according to the socio-professional structure of the country so that the exposure is adapted to the time of the disaster. The probabilities of death for a given level of physical intensity inside the buildings are included using vulnerability curves.

Public data such as population density and other information available from national statistical institutes were used to calibrate the model. Specialised earthquake market data from the Global Earthquake Model was used and adapted to the model.

The results of the backtest demonstrated the model's ability to simulate an accurate death toll only where population density is high. In rural areas like central Italy, the model does not work. We have noticed that the model works on the principle of the law of large numbers, since the probabilities of death applied with the vulnerability curves are low. The larger the population exposed to the seismic waves, the more accurate the simulated number of deaths.

For Mexico, where the population density is high, the model simulates earthquakes with a reasonable deviation of less than 10% from historical data. This difference is negligible and encouraging given the complexity of the modelling.

A sensitivity study on the model was performed. This showed that the model is able to take into account the excess mortality of disasters occurring at night. We can also see results observed in non-life reinsurance, individuals within the residential sector are exposed to a higher risk of death due to the less seismic resistant building structure than other buildings in the commercial sector.

This sensitivity study has shown above all the importance of the accuracy of the data to be included during the natural disaster simulations in the physical model, as the assumptions on the distribution and number of occupants per building, the data on building structures (compliance or not with the building standards for seismic risk), the vulnerability curves are very important factors to which the model results are very sensitive.

Pricing earthquake risk in a CAT XL Life treaty with available data

Once the model was calibrated, a first pricing of the earthquake risk of a CAT XL Life treaty in Mexico was proposed with the consideration of a catalogue of stochastic events and the other modules of the physical model as shown in Figure 3. However, the data available to a reinsurer on a life insurance portfolio implies to apply the portfolio profile on the assumptions of the Ekhedden model.

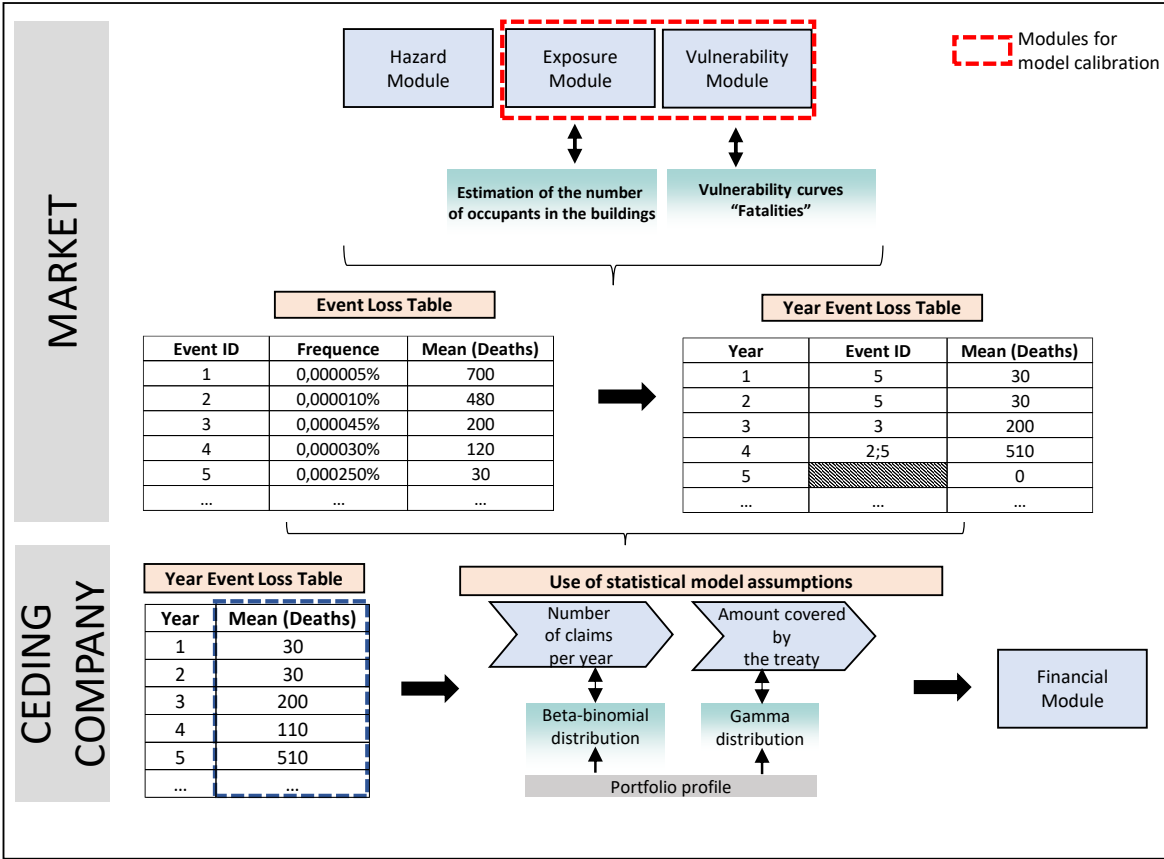


Figure 3 – The proposed pricing model with the inclusion of physical modules

Nevertheless, the modelling is based on a simulation of natural disasters from a catalogue of stochastic events and the number of deaths is simulated using physical modules.

The stochastic simulation of the events in the catalogue makes it possible to construct Event Loss Tables (ELT) listing the frequency and average severity of each event in the catalogue. From these ELTs, it is possible to build scenarios per year (YELT) including the occurrence and severity of the ELT events. This suggested methodology allows the construction of scenarios by year listing the number of deaths for each scenario.

For each scenario per year, a beta-binomial distribution is used to simulate the number of insured deaths within the number of victims for each scenario. The amount of loss associated with the death of several policyholders is estimated based on the Ekhedden assumption using a gamma distribution.

With the pricing model in question, the pure premium amount, i.e. the average recovery over all scenarios is 41.18% higher than that obtained with the standard Ekhedden model, which is mainly explained by a higher occurrence of claims in the physical model.

The physical model makes it possible to use a database of more than 40,000 earthquakes, whereas the EM-DAT database within the Ekhedden model lists only 32 earthquakes for Mexico. The physical model has a great advantage over the statistical model in that it better takes into account the severity of the disaster events. Besides, the frequency is based on the catalogue developed by the experts.

Conclusions

It is therefore possible to consider the use of a physical model in life reinsurance. The integration of the human dimension is the major problem. After calibration, the model can simulate a number of deaths with simple but reliable assumptions.

It is a complex model for it is necessary to use precise market data especially on exposure and other modules. The data currently provided by cedants is synthetic, which currently results in limited consideration of an insurer's life insurance portfolio.

The physical model has several advantages, allowing for unknown events in an insurer's available event history and providing a finer degree of precision in the assessment of vulnerability and risk exposure. The revaluation of the pure premium obtained by these calculations, by about 40%, is also a major point of attention, by providing a new analysis of the risk attached to the treaty.

These different research projects constitute the premises of a new light allowing to better understand and evaluate the risk, keeping in mind the technical difficulties in the assessment process.

It would therefore be necessary to continue to deepen this approach and to collect precise data on the portfolios in terms of numbers of insured or the sum insured per postal code, for example, in order to implement a complete model to carry out simulations to go beyond the hypotheses of the reference model (Ekhedden model).

Remerciements

Je voudrais adresser toute ma reconnaissance à mon tuteur d'entreprise, Jérôme CRETIEN, pour sa disponibilité et ses connaissances partagées. Son accompagnement et son suivi lors du stage ont été essentiels dans la progression de mon mémoire.

J'aimerais également remercier Aurélien BOISELET d'AXA Climate et Florient AUBRY du Risk Management d'AXA Global Re pour leurs conseils et expertises sur les problématiques traitées. Je voudrais exprimer aussi ma reconnaissance envers toute l'équipe d'Analytics & Pricing d'AXA Global Re qui m'a apporté son soutien intellectuel et moral au cours du stage.

Je porte aussi ma gratitude à mon tuteur académique, Christian Yann ROBERT, pour le suivi régulier et ses relectures qui m'ont permis d'améliorer mon mémoire.

Je remercie également toute l'équipe pédagogique de l'ENSAE en Actuariat pour leurs enseignements.

Enfin, je tiens à témoigner toute ma gratitude à ma famille pour les encouragements et son soutien.

Sommaire

RESUME	3
ABSTRACT	4
NOTE DE SYNTHÈSE	5
EXECUTIVE SUMMARY	11
REMERCIEMENTS	17
SOMMAIRE	19
INTRODUCTION	21
1 ASSURANCE ET CATASTROPHES NATURELLES.....	23
1.1 <i>Le risque Catastrophe</i>	23
1.1.1 Les évènements catastrophes et leurs sources.....	23
1.1.2 Exposition des assureurs au risque Catastrophe.....	25
1.1.3 Prise en compte de la dépendance risque Vie et Non-Vie au sein de Solvabilité 2	27
1.2 <i>Le risque tremblement de terre</i>	29
1.3 <i>Introduction à la réassurance.....</i>	31
1.3.1 Définitions	31
1.3.2 Les différentes natures de réassurance	33
1.3.3 Le rôle d'AXA Global Re	37
2 MODÈLE STATISTIQUE DE TARIFICATION DES EXCÉDENTS DE SINISTRE CATASTROPHE	39
2.1 <i>Présentation du modèle</i>	39
2.1.1 Méthodologie	39
2.1.2 Données disponibles pour la tarification des traités CAT XL.....	41
2.1.3 Limites du modèle	43
2.2 <i>Tarification des traités régionaux avec un simulateur d'évènements transfrontaliers.....</i>	44
2.2.1 Méthodologie	44
2.2.2 Analyse des évènements transfrontaliers	46
2.2.3 Application.....	48
2.3 <i>Etat de l'art des modèles de tarification proposés en CAT XL Vie</i>	50
3 ETUDE DE LA DÉPENDANCE ENTRE LES DOMMAGES NON-VIE ET LE NOMBRE DE MORTS LORS DES CATASTROPHES NATURELLES	53
3.1 <i>Enjeu et contexte</i>	53
3.2 <i>Mesure de la dépendance entre le risque Vie et Non-Vie.....</i>	54
3.2.1 Données disponibles	54
3.2.2 Utilité du taux de destruction et résultats	55
3.2.3 Impact sur le SCR d'un assureur	61

4 CALIBRATION D'UN MODÈLE PHYSIQUE POUR LA TARIFICATION DU RISQUE TREMBLEMENT DE TERRE	65
4.1 <i>Methodologie et principes de la modélisation</i>	65
4.1.1 Présentation du modèle physique	65
4.1.2 Le module aléa	68
4.1.3 Les courbes de vulnérabilité « Fatalities » pour estimer une probabilité de décès.....	70
4.2 <i>Calibration du modèle grâce à des backtests.....</i>	72
4.2.1 Description des événements et aléa physique.....	72
4.2.2 Calibration de l'exposition.....	76
4.2.3 Résultats et analyse de sensibilité.....	80
4.3 <i>Tarifification du risque tremblement de terre dans un traité CAT XL Vie avec les données disponibles.....</i>	86
4.3.1 Le module financier et hypothèses pour calculer une prime pure	86
4.3.2 Application du modèle pour un traité au Mexique.....	88
CONCLUSION.....	95
BIBLIOGRAPHIE	97
ANNEXES.....	99

Introduction

Au mois de juillet 2021, les inondations dans l'Ouest et le Centre de l'Europe notamment en Allemagne (190 morts) et en Belgique (38 morts) ont provoqué la mort de plus de 240 personnes et des pertes assurées de plus de 10 Mds\$. Il s'agit de la catastrophe naturelle la plus meurtrière de la Belgique et l'une des plus destructrices de l'année 2021. Avec le réchauffement climatique, l'occurrence de ce type d'évènements a augmenté et ils gagnent en intensité.

Les assureurs ne bénéficient pas de la mutualisation pour ce type de risque peu fréquent et d'une grande sévérité ni des fonds propres nécessaires pour absorber une perte engendrée par des évènements catastrophes. Les assureurs cherchent donc à se réassurer afin de mieux gérer le risque catastrophe. Il s'agit de l'une des fonctions de la réassurance.

Les assureurs vont se couvrir face au risque auprès du réassureur avec la mise en place d'un traité de réassurance du type excédent de sinistre par évènement. Il existe des traités pour le portefeuille vie et des traités pour le portefeuille non-vie. Le premier est la couverture de réassurance d'un portefeuille d'assurance vie (garantie décès et/ou invalidité) et le second d'un portefeuille d'assurance non-vie (biens et responsabilité civile).

Lorsque la cédante cherche à obtenir un traité de réassurance de ce type, elle va donc faire appel à un ou plusieurs réassureurs. Le traité doit donc être tarifé. Cette prime commerciale qui sera réglée par la cédante sera le résultat de l'addition d'une prime pure et d'un coefficient de chargement. Cette prime pure est calculée grâce à une analyse très fine du risque CAT et du portefeuille du réassureur.

L'état de l'art actuel des méthodes de tarification en réassurance vie et en réassurance non-vie sont différents. En réassurance-vie, un modèle par simulations de sévérité et fréquence reposant sur une base de données historiques est couramment utilisé. Alors qu'en réassurance non-vie, les modèles CAT sont développés par des scientifiques et actuaires permettant de générer stochastiquement des scénarios d'évènements catastrophes et d'analyser leur impact sur un portefeuille considéré. Ces modèles demandent une quantité de données importante. Les assureurs peuvent développer leur propre modèle CAT en interne et/ou requérir un modèle marché développé par des sociétés spécialisées (RMS, AIR,...) dans la modélisation des catastrophes naturelles et des catastrophes de cause humaine (man-made catastrophe).

L'analyse des risques CAT doit se faire avec des données limitées et un historique de sinistres très faible. En réassurance non-vie, la modélisation physique contourne très bien ce problème grâce à l'expertise et la connaissance de l'aléa physique (catalogue d'évènements stochastiques). A contrario, pour la réassurance vie, l'utilisation d'une base historique d'une cinquantaine d'années est souvent la seule source fiable pour effectuer des simulations et tarifier, malgré le faible nombre d'observations.

Le risque Vie et Non-Vie sont considérés comme indépendants par la réglementation. En effet, en formule standard, Solvabilité 2 fixe une corrélation nulle entre les risques Vie et Non-Vie. Les activités vie et non-Vie d'un assureur ne reposent pas sur le même business model, il s'agit de deux risques différents. Or, il est raisonnable de considérer une corrélation positive entre le risque CAT Vie et le risque CAT Non-Vie. Les catastrophes provoquant des dégâts matériels considérables engendrent en même temps un nombre de morts important.

En exploitant cette dépendance entre les risques CAT Vie et CAT Non-Vie, le mémoire consiste à engager la réflexion concernant la faisabilité d'un modèle physique de simulations de catastrophes naturelles pour la tarification des traités CAT XL Vie. Cette faisabilité dépend des données disponibles

pour un réassureur, des hypothèses à appliquer et des ajustements possibles sur les modèles de tarification existants.

Dans le cadre de ce mémoire, l'utilisation du modèle physique en réassurance-vie se fait sur le risque tremblement de terre mais il est possible de l'adapter à tous les types de catastrophes couverts par un traité en excédent de sinistre par évènement.

Pour ce faire, dans une première partie, il est explicité le risque Catastrophe, avec en mouvement de fond la recrudescence des catastrophes depuis ces 50 dernières années et leur impact sur l'activité des assureurs. Les caractéristiques et le vocabulaire couramment employés pour l'analyse du risque tremblement de terre sont décrits. Il est présenté à la suite le rôle de la réassurance pour mieux gérer ce risque CAT.

En seconde partie, le modèle de tarification actuel de simulations construit à l'aide d'une base historique est présenté. Ce modèle statistique permet de calibrer une sévérité et une fréquence selon l'historique de données. L'analyse porte alors sur l'examen des limites du modèle, qui constituent les sources de motivation pour proposer une nouvelle méthode de tarification des traités. Une amélioration du modèle existant est présentée pour la tarification des traités régionaux afin de prendre en compte des évènements transfrontaliers qui sont caractérisés le plus souvent par une sévérité plus importante que les évènements non transfrontaliers. Enfin, il est présenté l'état de l'art actuel des modèles de tarification des traités CAT XL Vie et les apports du modèle physique qui sera étudié.

Par la suite, la troisième partie s'attache à étudier la dépendance entre le risque Vie et le risque Non-Vie. Il est utilisé à ce titre la base internationale sur les catastrophes (EM-DAT) afin de mesurer cette dépendance. Les résultats sont présentés et démontrent une corrélation positive liée aux catastrophes naturelles. Cette corrélation incite à effectuer une analyse de sensibilité sur le besoin en capital d'un assureur lors de la prise en compte d'une corrélation. Cette analyse permet de déterminer la préconisation de capital supplémentaire dans la gestion du risque CAT.

Enfin en quatrième et dernière partie, à la suite de l'existence d'une corrélation, la description et la calibration d'un modèle physique sont explicitées pour la réassurance vie. Il s'agit d'exposer les changements et les adaptations nécessaires pour que le modèle soit appliqué sur un portefeuille en réassurance vie. Cela permettra d'évaluer sa faisabilité et sa performance sur la tarification du risque tremblement de terre au sein d'un traité CAT XL Vie par rapport au modèle statistique standard.

1 Assurance et Catastrophes Naturelles

Dans ce chapitre, la présentation du risque Catastrophe et l'exposition des assureurs face à ce risque sont décrites au point 1.1. Ensuite, une présentation de la réassurance et de la couverture possible face à ce risque est introduite au point 1.2.

1.1 Le risque Catastrophe

1.1.1 Les évènements catastrophes et leurs sources

Les évènements catastrophes sont considérés comme extrêmes puisqu'ils sont caractérisés par une très faible fréquence et une grande sévérité en termes de pertes humaines et dégâts matériels. Un effectif important est affecté et une accumulation de polices d'assurance sont touchées lors de ces évènements. Ces évènements peuvent fortement affaiblir le bilan d'une compagnie d'assurance. Il est usuel de distinguer deux grands types de catastrophes : les catastrophes « man made » et les catastrophes naturelles.

Les catastrophes « man made »

Les catastrophes « man made » qualifient les sinistres majeurs liés à des activités humaines, par exemple les accidents de transport (avion, train...), les accidents de mines et de carrières, le terrorisme et les catastrophes industriels. La fréquence de ces risques diffère selon le pays. Les accidents de transport et industriels ont diminué avec l'évolution des normes de sécurité dans ces secteurs surtout dans les pays développés. Comme le montre la Figure 1, selon la base de données mondiales sur le terrorisme (GTD), le risque terrorisme a fortement augmenté depuis le début des années 2000. Il est essentiel que tout modèle de tarification de réassurance vie et non-vie prenne en compte ce risque.

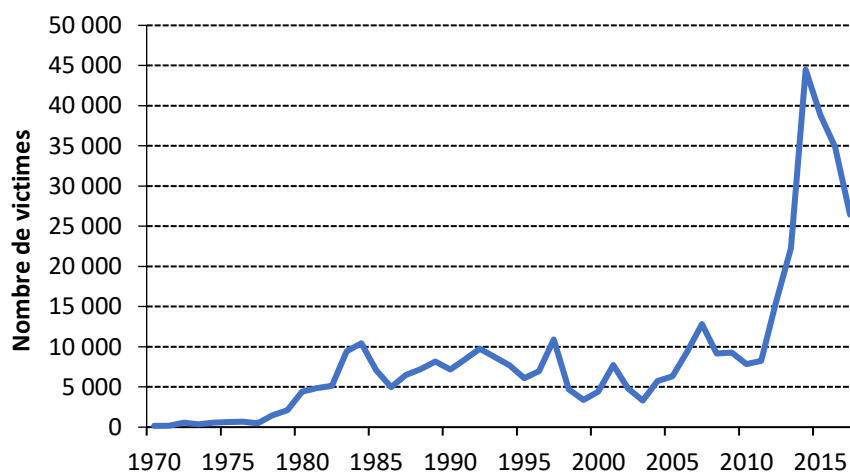


Figure 1 : Evolution du nombre de victimes d'actes terroristes dans le monde depuis 1970 à 2017

Les catastrophes naturelles

Les catastrophes naturelles peuvent avoir plusieurs sources et se distinguent selon leur origine : géologique (séisme, glissements de terrain...), hydrologique (inondations), météorologique (tempêtes, ouragans...) et climatologique (grêle, sécheresse...).

Selon les données de la base internationale sur les catastrophes (EM-DAT), il existe une forte augmentation des catastrophes naturelles depuis 1900. Ce phénomène est représenté sur la Figure 2.

Cette augmentation de la fréquence s'explique principalement pour 4 raisons :

- les moyens de mesure d'intensité et de traçage des phénomènes naturels sont récents, il y aurait donc un nombre significatif de catastrophes qui n'était pas comptabilisé au début de l'historique,
- un second facteur est la croissance de la population mondiale qui pousse les populations à s'installer sur des zones à risques,
- un troisième facteur lié au second est l'urbanisation qui entraîne une augmentation de la somme assurée par unité de surface (comptabilisation d'évènements CAT plus fréquente) ;
- le dérèglement climatique.

Le changement climatique devrait bouleverser la sévérité et la fréquence des catastrophes naturelles. D'après un rapport du CRED et UNDRR (2020), le changement climatique serait la principale cause du presque doublement des catastrophes naturelles de ces vingt dernières années. En effet, en se basant sur les données EM-DAT, entre 2000 et 2019, 7 348 désastres naturels ont été enregistrés dans le monde soit 74% de plus qu'entre 1980 et 1999 (4 212 catastrophes). Par exemple, le nombre d'inondations et de tempêtes a considérablement augmenté (+ 134% et +40% respectivement).

Cependant, il est important de souligner que le nombre de victimes engendré par les catastrophes est resté stable même avec la forte augmentation de la fréquence. Entre 1980 et 1999, 1.19M de personnes sont décédées à la suite de catastrophes contre 1.23M entre 2000 et 2019. Cela constitue une preuve des améliorations des moyens pour prévenir et protéger.

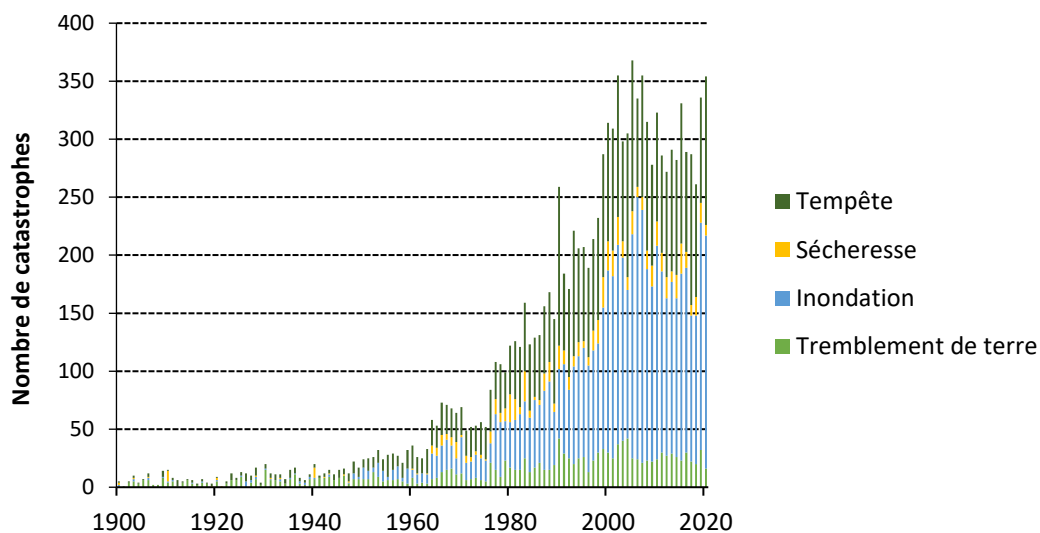


Figure 2 : Nombre de catastrophes dans le monde par année de 4 grands types de catastrophes naturelles

L'ampleur des dommages engendrés par une catastrophe naturelle ne dépend pas uniquement de l'intensité du phénomène naturel, mais aussi de nombreux facteurs humains comme la densité humaine, la structure des bâtiments et l'efficacité des moyens de protection mis en œuvre dans la région concernée.

Les pays développés (où la couverture d'assurance est élevée) ont pris des mesures de prévention et de secours. Ils les améliorent en permanence, afin d'éviter un grand nombre de victimes en cas de catastrophe.

Cela s'applique principalement aux catastrophes naturelles telles que les tremblements de terre (méthodes de construction), les tsunamis ou les tempêtes tropicales (systèmes d'alerte et abris) et les inondations (par exemple, les digues).

A l'heure actuelle, il n'est pas possible de prévoir le moment exact d'une catastrophe naturelle mais il est possible de connaître les zones géographiques les plus exposées. Par exemple, les zones exposées aux séismes sont les territoires proches d'une faille sismique. Le Japon illustre parfaitement une zone à risque, le pays est situé à la frontière de 4 grandes plaques : la plaque d'Okhotsk au nord, la plaque pacifique à l'est, la plaque philippine au sud et la plaque eurasiatique à l'ouest. Face à ce risque, les constructions au Japon répondent à des normes très strictes au niveau des séismes. Ainsi, un séisme de même magnitude causera probablement des dégâts plus forts aux Philippines qu'au Japon où les moyens de prévention ou de protection sont plus importants.

Les épidémies

Au sein du risque Catastrophe, nous avons aussi les épidémies qui demandent une indemnisation importante des assureurs et des réassureurs. Une épidémie se définit comme le développement et la propagation d'une maladie infectieuse. Elle occasionne des dommages importants en assurance non-vie, comme c'est le cas actuellement avec l'épidémie actuelle de COVID-19. Ainsi, les cessations d'activités avec les confinements ont engendré de nombreux sinistres. Ces épidémies causent en même temps des victimes, des pertes humaines touchant l'assurance-vie.

1.1.2 Exposition des assureurs au risque Catastrophe

L'augmentation de la fréquence des catastrophes durant ces 50 dernières années se traduit par une augmentation des pertes en valeur économique et en valeur assurée. Les dommages économiques englobent tous les dommages financiers (assurés ou non) engendrés par un évènement catastrophe. Il s'agit par exemple des dommages sur les infrastructures, les bâtiments, etc... La demande mondiale en assurance catastrophe est représentée sur la Figure 3. Le déficit de protection mondiale s'élève en 2018 à 80 milliards USD, avec des dommages assurés qui s'élevaient à 85 milliards USD en 2018 et des dommages économiques à 165 milliards USD. Le ratio de couverture qui est le rapport entre les dommages assurés et les dommages économiques s'élève donc à 50% au total en 2018.

En comparant les moyennes glissantes sur 10 ans, les dommages économiques ont un taux de croissance légèrement supérieur à celui des dommages assurés entre 1970 et 2018. Les dommages économiques ont progressé de 5 % entre 1992 et 2018, et les dommages assurés de 4,7 % au cours de la même période. La quasi-totalité des dommages assurés sont liées aux catastrophes météorologiques.

Cela démontre clairement que le changement climatique a un impact sur la soutenabilité du secteur

assurantiel avec une augmentation des dommages assurés (Ouragans, tempêtes...). La couverture de réassurance est plus importante dans les pays développés anciens (essentiellement en Europe, Amérique du Nord et en Océanie). Mais cette couverture reste plus faible en Asie, Afrique ou en Amérique du Sud.

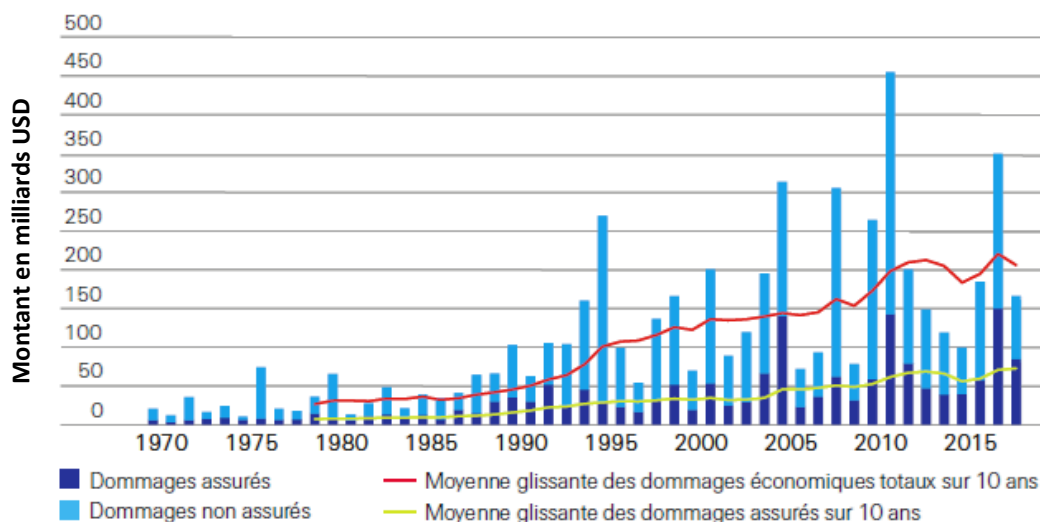


Figure 3 : Couverture d'assurance des catastrophes de 1970 à 2018 (en milliards USD, aux prix 2018)

Une majorité des groupes d'assurance ont une activité « Vie » et « Non Vie ». Les assureurs sont donc exposés à la fois par une augmentation des dommages en assurance IARD et d'une augmentation du risque de mortalité en assurance vie avec la survenance de catastrophes. Celles-ci nécessitent un besoin en capital important pour absorber les pertes extrêmes sur leurs bilans.

Les montants en euros des SCR Vie et Non-Vie sont représentés sur le Tableau 1 et le Tableau 2. Comme le montre les tableaux, les données mises à disposition par les assureurs ne permettent pas de définir précisément le capital de solvabilité requis pour faire face aux risques CAT (SCR CAT : *Solvency Capital Requirement for Catastrophe Risk*).

L'information n'est de plus pas disponible dans les rapports sur la solvabilité et la situation financière des compagnies d'assurance ou de réassurance (Rapport SFCR : *Solvency and Financial Conditions Report*). Sur 6 réassureurs et 6 assureurs, seulement deux d'entre eux publient des chiffres sur leur SCR CAT Vie et SCR CAT Non-Vie. Néanmoins, il n'est pas absurde de faire l'hypothèse que la part du risque CAT au sein des modules des SCR Vie et SCR Non-Vie soit un des risques prépondérants.

Pour tout assureur et réassureur, ayant des SCR Vie et Non-Vie importants comme la SCOR ou Allianz, une possible dépendance positive entre le risque Vie et Non-Vie viendrait augmenter leur besoin en capital.

Réassureurs	SCR Non-Vie (M€)	SCR Vie (M€)	SCR CAT Non-Vie (M€)	SCR CAT Vie (M€)
AXIS Re SE	264	34		
CCR	3 409	48	2 890	34
Hannover Re	4 353	3 140 (*)		
Munich Re	9 413	6 996 (*)		
Partner Reinsurance Europe SE	773	440		
SCOR	3 663	3 651 (*)		

(*) SCR Vie comprenant la Santé

Tableau 1 : Montant en euros des SCR de réassureurs en 2020 (Rapports SFCR)

Assureurs	SCR Non-Vie (M€)	SCR Vie (M€)	SCR CAT Non-Vie (M€)	SCR CAT Vie (M€)
AG2R	5	1 796		
Allianz	3 385	2 509		
CNP Assurances	600	4 600		
Crédit Mutuel	917	1 256	179	132
Generali Group	2 614	799		
Groupama	1 884	1 246		

Tableau 2: Montant en euros des SCR d'assureurs en 2020 (Rapports SFCR)

1.1.3 Prise en compte de la dépendance risque Vie et Non-Vie au sein de Solvabilité 2

Comme démontré précédemment, les catastrophes naturelles induisent un besoin en capital indispensable au bon fonctionnement de tout assureur. Depuis l'instauration de la directive Solvabilité 2 en Janvier 2016, une formule standard est utilisée majoritairement pour calculer le besoin en capital d'un assureur. La formule standard est conçue selon une architecture modulaire (risques de souscription, de marché, de crédit...). Cette approche modulaire, représenté par la Figure 4, permet l'agrégation d'un calcul de capital effectué individuellement pour chaque risque. Chaque module de risque est calibré pour que la probabilité de ruine à horizon 1 an soit égale à 0,5%. On obtient finalement un capital consolidé pour l'ensemble des risques auxquels la compagnie est exposée. Ce capital tient compte des phénomènes de dépendance entre les différents segments.

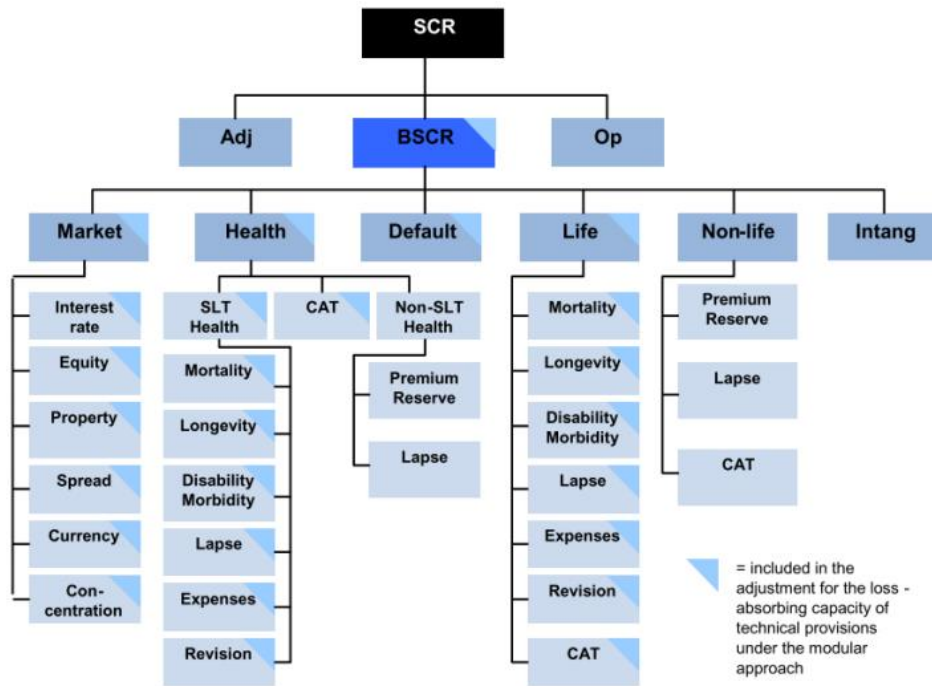


Figure 4 : Les différents modules du SCR

Le terme « agrégation des risques » a deux dimensions, agréger des capitaux économiques relatifs à des risques (longévité, mortalité...) d'un même module (corrélations intra-modulaires) mais aussi agréger des capitaux économiques de périmètres de modules différents (corrélations inter-modulaires). La prise en compte des effets de diversification dans l'architecture modulaire se fait à l'aide d'une matrice de corrélation. La matrice de corrélation de Solvabilité est représentée par le Tableau 3. Il s'agit d'un bénéfice de diversification lorsque le capital requis pour l'agrégat des modules considérés est inférieur à la somme des capitaux de chacun des modules.

CorrSCR	Market	Default	Life	Health	Non-Life
Market	100%	25%	25%	25%	25%
Default	25%	100%	25%	25%	50%
Life	25%	25%	100%	25%	0%
Health	25%	25%	25%	100%	0%
Non-Life	25%	50%	0%	0%	100%

Tableau 3 : Matrice de corrélation sous Solvabilité 2

La méthode d'agrégation standard se fait sous la forme suivante :

Soient :

- n risques : plusieurs modules de risque (marché, crédit, assurance non vie...).
- S un vecteur de SCR :

$$S = \begin{pmatrix} SCR_1 \\ \vdots \\ SCR_n \end{pmatrix} \in \mathbb{R}^n$$

- M la matrice de corrélation de ces risques : $M \in \mathcal{M}_n(\mathbb{R})$

Le SCR du risque total (somme des risques marginaux) est donné par la norme de S induite par la matrice M :

$$SCR_{\text{total}} = \sqrt{{}^tS \times M \times S} = \|S\|_M$$

Le choix de la racine carrée comme fonction pour agréger les modules implique que le gain de diversification est décroissant avec le taux de corrélation $Corr_{ij} \in [-1,1]$. Le gain de diversification est maximum pour un taux de corrélation de -100% et nulle pour un taux de corrélation de 100%.

$$\sqrt{SCR(X)^2 + SCR(Y)^2 + 2 * Corr_{XY} * SCR(X) * SCR(Y)} \leq SCR(X) + SCR(Y)$$

Solvabilité 2 considère un coefficient de corrélation de 0% entre le module Vie et Non-Vie. La principale raison est l'assurance vie est basée sur des branches d'activité longues alors que l'assurance non-vie est principalement basé sur des branches d'activité courtes ou moyennes. Il s'agit bien de deux modèles économiques différents dont les risques couverts sont différents.

En effet, la durée de l'engagement d'un assureur concernant le cycle économique complet, c'est-à-dire la souscription, la survenance du sinistre, puis la déclaration et enfin le règlement monétaire, permet de distinguer les branches courtes et les branches longues.

Les branches courtes sont caractéristiques de l'assurance non-vie : la déclaration des sinistres est faite rapidement une fois la survenance, l'estimation prend quelques jours et le règlement est immédiat, ainsi tout sinistre est indemnisé intégralement de façon rapide.

En revanche, avec les branches longues, la durée séparant la collecte de la prime et l'indemnisation complète d'un sinistre est très longue, supérieure à 10 à 20 ans. A noter qu'il existe en assurance non-vie des branches d'activité longues comme la responsabilité civile mais qui ne sont pas représentatives du modèle de l'assurance non-vie dans son intégralité.

Malgré ces différences, il est réducteur de considérer qu'il y n'ait aucune dépendance entre le risque Vie et Non-Vie. En effet, il faudrait au moins considérer qu'il existe une corrélation faible entre les sous-risques CAT de ces deux modules. Ces sous-modules CAT sont fortement impactés lorsque des catastrophes de grande ampleur éclatent.

Ainsi, une étude de la dépendance entre le risque vie et non-vie est justifiée, une dépendance non nulle et positive pourrait augmenter le SCR des assureurs et réassureurs exposés à leur bilan par un risque Vie et Non-Vie. Cette prise en compte pourrait permettre d'agréger de façon plus précise les modules et surtout limiter au mieux la probabilité de ruine pour les compagnies concernées.

1.2 Le risque tremblement de terre

Les tremblements de terre (y compris les tsunamis) sont le type de catastrophe naturelle le plus meurtrière des 20 dernières années, avec plus de 721 318 morts entre 2000 et 2019, soit 58% des victimes engendrées par des catastrophes naturelles selon les données EM-DAT. De plus, les mégapérils sont surtout des tremblements de terre (Séisme et tsunami de 2004 dans l'océan Indien ou Haïti en 2010) tuant des centaines de milliers de personne. Ce péril au sein du module CAT Vie a donc une place centrale pour les assureurs notamment dans des pays fortement exposés à ce risque comme le Japon, le Mexique ou l'Italie.

Origine des tremblements de terre

Les tremblements proviennent du mouvement des plaques tectoniques. Ces mouvements sont provoqués par la chaleur générée par la Terre qui remonte à la surface en créant des courants de convection. Sous l'effet de ces courants, les roches magmatiques rentrent en mouvement et elles entraînent avec elles les plaques tectoniques. Ces plaques peuvent coulisser l'une à côté de l'autre, s'écarter, se rapprocher et glisser l'une sous l'autre. Les tremblements de terre se produisent en bordure des plaques, c'est-à-dire aux failles.

Lorsqu'une plaque se glisse au-dessous de l'autre, il s'agit d'une zone de subduction. Au niveau de la zone de subduction, entre les deux plaques de l'énergie est stockée avec les frottements. Lorsque la roche se fissure sous la pression, l'énergie est libérée, ce qui fait trembler la terre. Des ondes sismiques sont provoquées, la terre tremble à la surface. Les zones de subduction sont les zones où les tremblements de terre les plus puissants se produisent.

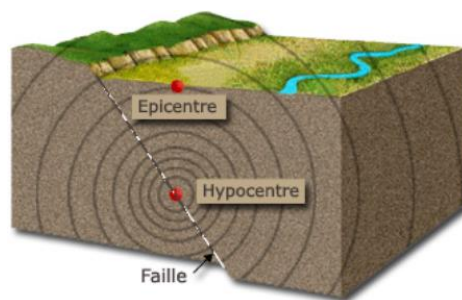


Figure 5 : Représentation d'un tremblement de terre

Comme le montre la Figure 5, le point où la roche se fissure sur le long de la faille est appelé hypocentre ou foyer. Les ondes sismiques partent du foyer et vont dans toutes les directions. L'épicentre est le point le plus proche de l'hypocentre à la surface terrestre, il se trouve à la verticale de l'hypocentre. Il s'agit de la zone où l'intensité perçue du tremblement sera la plus importante. La distance entre l'hypocentre et l'épicentre est la profondeur de l'hypocentre. Plus cette distance est faible, moins les ondes sismiques sont moins atténuées lorsqu'elles atteignent la surface terrestre (l'épicentre).

Force et intensité des tremblements de terre

Pour mesurer la force du séisme, il est courant d'utiliser la magnitude de Richter. Cette mesure estime l'énergie libérée par la rupture de la faille à l'origine des ondes sismiques. Elle se mesure à partir de l'amplitude du mouvement du sol enregistrée par des instruments sismologiques. Il s'agit d'une mesure logarithmique. Lorsque la magnitude est supérieure d'une unité, l'énergie est supérieure de 30 fois. L'échelle de Richter est utilisée communément par les médias pour relater chaque survenance de tremblements de terre notables. Cette échelle est représentée par le Tableau 4.

Magnitude	Catégorie	Observations
1,0 à 2,9	Micro	Non ressenti
2,0 à 2,9	Très mineur	Détecté par les instruments de mesure mais souvent non ressenti
3,0 à 3,9	Mineur	Généralement ressenti mais cause peu ou pas de dégâts
4,0 à 4,9	Léger	Secousses notables, chute d'objets et dégâts légers mais rarement importants
5,0 à 5,9	Modéré	Secousses importantes, dommages considérables sur les bâtiments les moins bien construits
6,0 à 6,9	Fort	Secousses violentes dans un rayon de moins de 200km de l'épicentre
7,0 à 7,9	Majeur	Secousses extrêmes, dommages modérés sur une zone plus vaste
8,0 à 8,9	Important	Secousses et dommages considérables sur des centaines de kilomètres
9,0 et plus	Dévastateur	Secousses dévastatrices sur des milliers de kilomètres

Tableau 4 : Magnitude de Richter

Pour mesurer l'intensité d'un séisme, l'intensité Mercalli notée MMI (Modified Mercalli Intensity) est utilisée, elle permet de représenter la nature et les dégâts provoqués par un séisme. Il ne s'agit pas d'un calcul mais d'une mesure basée sur le ressenti de la population. Le Tableau 5 résume la relation entre intensité et dégâts provoqués. Par définition, cette mesure est donc plus représentative de l'ampleur des sinistres. Elle permet d'avoir une idée de la destruction des bâtiments sur la zone du tremblement de terre et donc des victimes aussi. Un séisme de magnitude 7 peut par exemple provoquer une intensité maximum de 9 dans certaines zones géographiques, tout dépend des effets constatés.

Intensité (MMI)	Réaction des personnes	Effets sur les équipements d'une maison	Effets sur les constructions/bâtiments
I	Non ressenti	Aucun effet	Aucun effet
II	Peu de personnes peuvent ressentir les secousses	Les objets suspendus bougent très faiblement	Aucun effet
III	Plusieurs personnes ressentent les secousses	Les objets suspendus bougent beaucoup	Aucun effet
IV	La majorité des personnes ressentent les secousses	Les objets à l'intérieur vibrent	Les murs se fissurent
V	Les secousses sont ressenties par tout le monde; certaines personnes sont effrayées.	Les petits objets bougent	Le plâtre se fissure et les fenêtres s'entrechoquent aussi
VI	Plusieurs personnes sont effrayées et sont déplacées par les secousses	Les objets tombent et se déplacent	Les morceaux de plâtre peuvent céder et les fenêtres se cassent (dommages légers)
VII	Sentiment de panique; certaines personnes perdent l'équilibre	Les équipements se renversent	Dommages négligeables sur les bâtiments les plus solides et dégâts importants sur les bâtiments les moins bien construits
VIII	Des difficultés à se maintenir debout	Déplacements importants de tous les équipements	Dommages légers sur les bâtiments les plus solides (respectant les normes parasismiques) et dommages considérables sur les autres construction
IX	Certaines personnes sont jetées au sol par les secousses.		Dommages considérables sur les bâtiments les plus solides (respectant les normes parasismiques), certains bâtiments ordinaires sont quasiment effondrés
X			Destruction importante (les maisons ordinaires s'effondrent et les ponts cèdent), des fissures au sol, seulement quelques bâtiments peuvent rester debout
XI - XII			Destruction totale

Tableau 5 : Intensité de Mercalli

Pour se protéger des dégâts importants matériels et humains que provoquent les tremblements de terre et autres catastrophes, il est essentiel que les compagnies d'assurance se réassurent.

1.3 Introduction à la réassurance

Pour se prémunir contre le risque Catastrophe, il est indispensable qu'un assureur mette en place une couverture de réassurance. Le traité de réassurance généralement choisi est un excédent de sinistre par évènement. Il est donc essentiel de présenter le rôle d'un réassureur, le fonctionnement d'un traité et les différents types de couverture offertes à une cédante.

1.3.1 Définitions

Réassurance Vie et Réassurance Non-Vie

La réassurance est souvent définie comme l'assurance des assureurs. Elle permet aux assureurs de transférer une partie de leurs risques qu'ils se sont engagés à couvrir. Une partie ou l'intégralité des polices d'assurance du portefeuille de l'assureur (ou « cédante ») est cédée à un réassureur en échange d'une prime de réassurance. Le réassureur a la possibilité de transférer une partie des risques qu'il couvre via les polices d'assurances à un autre réassureur appelé « rétrocessionnaire ».

Les risques couverts en réassurance-vie par les traités peuvent être regroupés comme ci-dessous : la mortalité et morbidité ; l'incapacité ; les dépenses médicales ; les maladies graves ; la longévité et les autres risques. La réassurance non-vie ou en dénomination anglo-saxonne P&C : « *Property & Casualty Insurance* » couvre globalement les risques liés aux dommages aux biens (*property*) et à la responsabilité civile (*casualty*).

Il apparaît important de préciser des points au niveau des traités en réassurance-vie. Dans un traité de réassurance vie, il est courant d'avoir une liste d'exclusion (les risques exclus par les polices d'assurance, les suicides, la participation à un acte criminel, de s'occasionner des blessures volontaires...) entraînant la nullité du contrat si elle n'est pas respectée.

Il existe des « acceptations spéciales » qui doivent être acceptées par le réassureur si l'assureur veut l'inclure ou non au sein du périmètre du traité. On distingue deux acceptations spéciales :

- les « affaires à haut risque » : les polices d'assurance carte de crédit prévoyant la couverture automatique d'une assurance Individuelle Accident lors d'un voyage si le billet d'avion est réglé par cette carte etc..
- les « professions à risques élevés » : les pilotes d'avion, sportifs de haut niveau...

Les formes juridiques et fonctions de la réassurance

On distingue 3 formes juridiques qui fixent la relation de réassurance entre réassureur et cédante :

- Obligatoire : Tous les risques de portefeuilles couverts par le traité de réassurance doivent être obligatoirement présentés au réassureur par la cédante. Le réassureur s'engage à accepter tous les risques présentés par le réassureur.
- Facultative : Il s'agit d'une réassurance « police par police », la cédante a la liberté de présenter ou non le risque au réassureur et ce dernier est libre d'accepter ou de refuser le risque.
- Facultative-obligatoire : La cédante a la possibilité de présenter ou non le risque au réassureur mais le réassureur est obligé de l'accepter.

La réassurance a plusieurs fonctions et permet de remplir plusieurs objectifs. Globalement, elle permet de :

- transférer du risque entre cédant et réassureur, ce dernier mutualise et disperse les risques à l'échelle mondiale. Du fait de l'aversion aux risques des assureurs face aux risques extrêmes, l'assureur peut écarter les sinistres en partageant le risque avec le réassureur.
- limiter les effets de la sinistralité grâce au transfert de risque en stabilisant le résultat pour la cédante
- de diminuer les exigences réglementaires en capital sous Solvabilité 2 et d'augmenter en même temps la capacité de souscription de la cédante. Le réassureur joue un rôle « d'apporteur de fonds propres ».
- d'accéder aux conseils et à l'expertise des réassureurs
- d'apporter des liquidités en cas de sinistralité importante

La réassurance est donc très utile pour l'activité des assureurs notamment en termes de gestion des risques et du capital. Face au risque Catastrophe, la réassurance est indispensable, ce risque a comme caractéristique d'être un véritable risque de pointe pour un assureur. Un assureur ne dispose pas de la diversification et mutualisation à l'échelle internationale pour diluer le risque CAT. Sans réassurance,

le capital nécessaire pour absorber les pertes en cas d'un évènement CAT sur un portefeuille serait beaucoup trop importants.

Le risque « Catastrophe » au sein de la réassurance

Un évènement CAT en réassurance est défini comme la survenance d'un évènement engendré par un même fait générateur et causant un ensemble de sinistres dans une période déterminée. La clause horaire et le fait générateur sont les deux paramètres déterminants dans la définition d'un évènement. La clause horaire est fixe et varie selon le fait générateur. Cette clause horaire varie de quelques jours à plusieurs semaines. Il est donc indispensable qu'il n'existe aucune ambiguïté dans la définition d'un évènement CAT au sein des traités. Il est usuel dans les traités de réassurance non-vie et vie d'excédent de sinistre d'exclure certains faits générateurs comme les guerres, les pandémies...

En réassurance vie, un évènement CAT doit être un péril causant un nombre minimum de victimes dans une clause horaire fixée selon le traité de réassurance. Ce nombre minimum s'applique sur l'agrégation de toutes les polices d'assurance couvertes par le traité et non pas par police.

1.3.2 Les différentes natures de réassurance

Les traités proportionnels

Le réassureur s'engage à prendre en charge une proportion de risque de l'assureur, en échange de la même proportion de la prime perçue par l'assureur pour couvrir le risque. Il existe deux formes de traités proportionnels : les traités en quote-part et les traités en excédent de plein.

Traités en quote-part

L'assureur cède au réassureur une proportion fixée des risques réassurés et le même pourcentage de ses primes. En contrepartie, le réassureur s'engage à payer le même pourcentage de chaque sinistre et à payer des commissions correspondant à un pourcentage des frais.

Selon les conditions du traité et les clauses spécifiées sur lesquelles les deux parties se sont engagées, le réassureur peut être amené à reverser une commission supplémentaire en cas de bénéfices. A l'inverse, s'il fait des pertes importantes, l'assureur peut être amené à verser une prime supplémentaire.

Le résultat du réassureur peut s'écrire sous la forme :

$$R_{reassureur} = \alpha(P - S) - \varepsilon * F$$

Où :

- $R_{reassureur}$: le résultat du réassureur
- α : la proportion de risque fixée de la cédante et transférée au réassureur
- P : le montant de prime reçu par l'assureur
- S : le montant des sinistres
- $\varepsilon * F$: le montant total des commissions

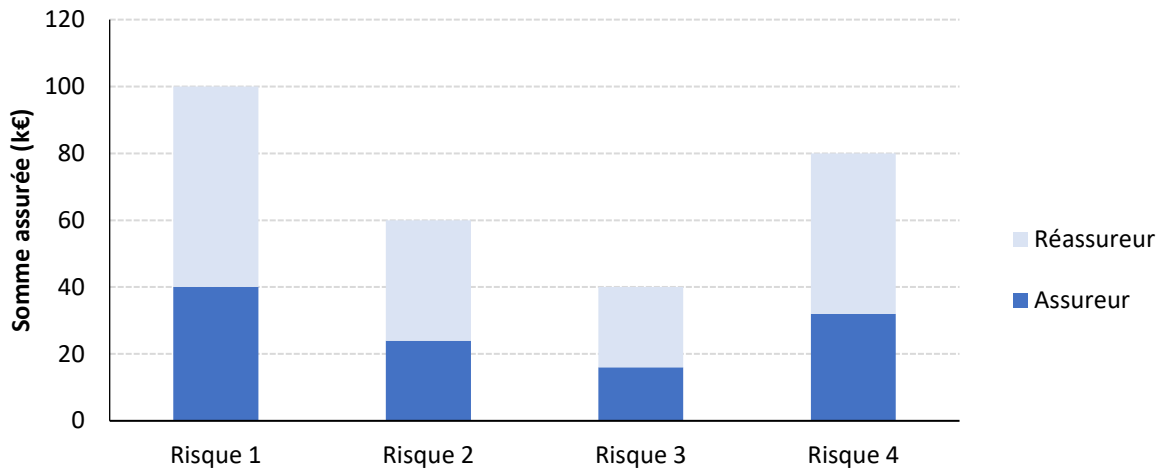


Figure 6 : Exemple d'un traité en quote-part avec un taux de rétention à 40% et un taux de cession à 60%

La Figure 6 représente l'engagement d'un assureur et d'un réassureur dans le cadre d'un traité en quote-part avec un taux de cession de 60%. Par exemple, pour le risque 2 dont la somme assurée totale s'élève à 60 000€, l'assureur a une rétention de 24 000€ et cède au réassureur un montant de 36 000€.

Il s'agit du traité le plus simple, il ne permet pas à l'assureur de modifier sa structure mais il réduit son engagement. Il est souvent utilisé par les assureurs qui cherchent à se développer sur de nouvelles branches aux résultats incertains ou tout simplement d'augmenter leurs capacités de souscription.

Traités en excédent de plein

L'assureur conserve sur chaque risque un montant identique d'exposition ou de plein de conservation (PC) et cède la partie la partie qui dépasse ce plein. Un plein de souscription (PS) est fixé pour chaque risque, ce qui permet de limiter la garantie à chaque risque. Ce plein de souscription se définit comme la somme pleine du plein de conservation et la capacité (limite de l'engagement du réassureur).

Ainsi, la part de cotisation (PC) versée au réassureur est limitée à la différence entre le plein de souscription et le plein de cotisation. La part de cotisation peut s'écrire sous la forme :

$$P'_{reassurance} = \min[\max[(P - PC), 0], (PS - PC)]$$

Où :

- $P'_{reassurance}$: le résultat du réassureur
- P : la prime reçue par l'assureur

En cas de sinistre, le réassureur prend en charge un montant proportionnel au taux de cession de la cédante :

$$C_{reassurance} = S * \frac{P'_{reassurance}}{P}$$

Où :

- $S'_{reassurance}$: la charge du réassureur
- S : le montant total du sinistre

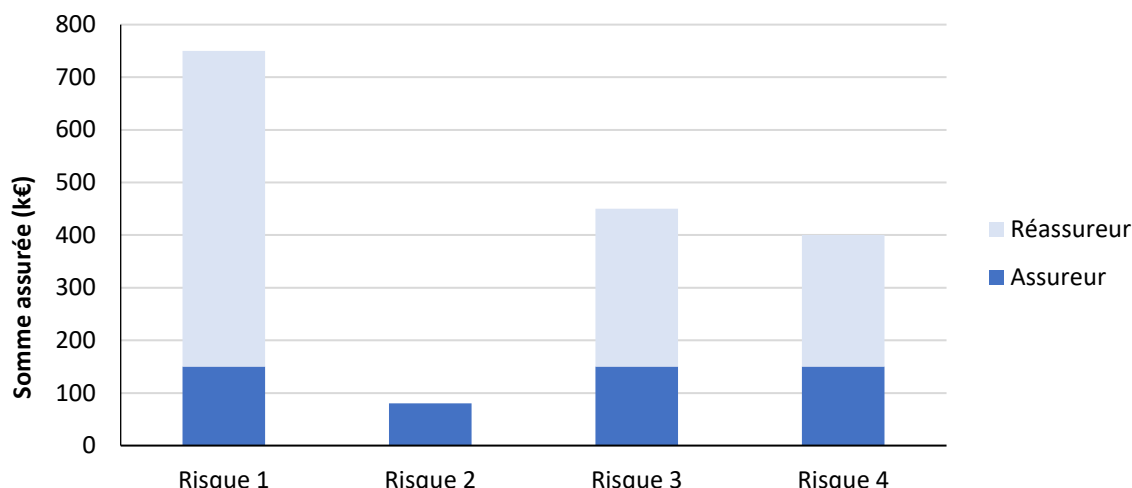


Figure 7 : Exemple de traité en excédent de plein : plein de rétention à 150k€

La Figure 7 représente l'engagement d'un assureur et d'un réassureur dans le cadre d'un traité en excédent de plein. Par exemple pour le risque 2, il n'y a aucune cession, la somme assurée totale du risque est de 80 000€ donc inférieure au plein de rétention (150 000€). L'assureur reste engagé à 100% sur ce risque.

Ce traité permet à l'assureur de bénéficier d'une homogénéisation de son portefeuille et d'augmenter sa capacité de souscription. Pour chaque risque, la cédante obtient une suppression des pointes d'exposition et une conservation des petits risques.

Les traités non proportionnels

Dans ces traités, l'engagement de la cédante et du réassureur ne sont pas proportionnels. Cette différence est induite par des effets de seuil. Il est indispensable de définir 2 notions :

- La priorité est le seuil au-dessous duquel la cédante conserve engagement
- La portée est le montant maximum de l'engagement du réassureur

En sommant priorité et portée, la capacité du traité est obtenue, il s'agit du montant au-dessus duquel la cédante n'est plus couverte. Il est possible de segmenter la couverture de réassurance en plusieurs traités, en additionnant un traité dont sa capacité coïncide avec la priorité d'un autre traité par exemple. Cela permet à la cédante de mieux négocier chaque tranche de réassurance et possiblement avec plusieurs réassureurs ou contreparties. Il s'agit d'une logique purement commerciale.

Il existe 3 formes de traités non proportionnels : l'excédent de sinistre par risque, l'excédent de sinistre par évènement et l'excédent de perte annuelle (ou *Stop Loss*).

Traité en Excédent de Sinistre par risque

Un traité en Excédent de Sinistre (XS ou XL : *Excedent of Loss* en anglais) peut se définir par risque ou par évènement, il se note *Portée XS Priorité*.

Pour les polices couvertes par le traité, le réassureur s'engage à couvrir le montant de chaque sinistre dépassant la priorité, le montant indemnisé pour chaque sinistre sera inférieur ou égal à la portée.

En cas de sinistre, le montant pris en charge par le réassureur peut s'écrire sous la forme suivante :

$$C_{reassureur} = \text{Min}(\text{Max}(S - \text{Priorité}, 0), \text{Portée})$$

Ce type de traité permet à la cédante d'avoir une protection contre les sinistres importants et de lisser la sinistralité.

Au sein de ce type traité, il est fréquent de retrouver un AAL (*Annual Aggregate Limit*) venant limiter l'engagement du réassureur sur le montant annuel agrégé des sinistres couverts. Il s'agit du montant total maximum des sinistres réassurés.

Le traité peut aussi inclure un AAD (*Annual Aggregate Deductible*) qui se définit comme le montant de sinistre agrégé au-dessous duquel le réassureur n'interviendra pas.

Le traité spécifie aussi souvent une clause de reconstitution exprimant un nombre de reconstitution. Il s'agit d'une reconstitution de la portée. La portée est dite « consommée », une fois que la portée du traité est dépassée par le montant global des sinistres réassurés. Dans ce cas, la cédante peut à nouveau bénéficier d'une couverture gratuitement (reconstitution gratuite) ou en payant une prime supplémentaire (reconstitution payante). Il est possible que certains traités de réassurance aient une portée illimitée notamment en responsabilité civile en France mais cela reste un cas exceptionnel.

Traité en Excédent de Sinistre par évènement

Un traité *CAT XL* couvre l'ensemble des sinistres générés par un même fait générateur sur une période déterminée. Il permet à la cédante de se protéger contre une accumulation de sinistres de plusieurs polices causés par un même évènement. Ce traité peut être mis en place en plus d'un traité d'excédent sinistre par risque, en complément. Comme présenté ci-dessus, cet évènement est défini contractuellement, dans sa nature, espace et le temps. C'est ce type de traité qui sera analysé au sein du mémoire notamment pour les catastrophes naturelles.

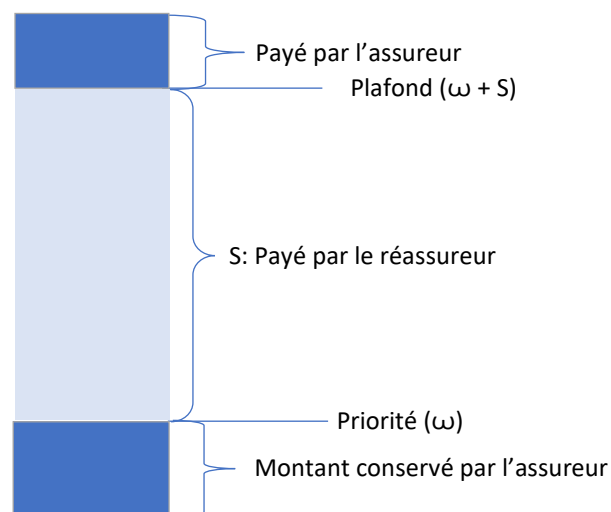


Figure 8 : Répartition du coût d'un sinistre entre assureur et réassureur dans le cadre d'un excédent de sinistre par évènement

La Figure 8 représente la répartition du coût d'un sinistre entre assureur et réassureur dans le cadre d'un excédent de sinistre par évènement. Par exemple dans le cadre d'un traité 40 XS 20, lors d'un sinistre s'évaluant à 70M€, 40M€ seront à la charge du réassureur et 30M€ seront à la charge de l'assureur. Les 30M€ que paye l'assureur lors du sinistre s'explique par la priorité de 20M€ du traité et du montant du sinistre dépassant le plafond du traité pour un montant de 10M€.

Traité en Excédent de perte annuelle (Stop Loss)

L'objectif d'un traité en excédent de perte annuelle permet à la cédante de se garantir d'un résultat. Il ne s'exprime pas en montant mais en ratio de sinistralité (Sinistres/Primes). Il fonctionne comme un traité non proportionnel avec une priorité et une portée sur le résultat comptable d'un exercice donné d'un portefeuille. Ce traité permet de lisser le résultat de la cédante. La charge du réassureur s'exprime sous la formule suivante :

$$C_{reassureur} = P * \text{Min} \left(\text{Max} \left(\frac{S}{P} - \text{Priorité}, 0 \right), \text{Portée} \right)$$

1.3.3 Le rôle d'AXA Global Re

AXA Global Re centralise l'ensemble des activités de réassurance du groupe AXA. Lorsqu'une entité veut souscrire ou renouveler un traité de réassurance, elle doit passer par AXA Global Re. Elle joue le rôle de réassureur pour les entités d'AXA en nouant des traités avec les différentes entités du groupe dans le monde.

AXA Global Re gère seulement la réassurance interne, et agit comme une plateforme de réassurance. En jouant ce rôle, elle a comme premier avantage de détenir un pouvoir de négociation conséquent vis-à-vis des courtiers. Ces courtiers en réassurance (Aon, Guy Carpenter, ...) auront pour mission de placer les traités de réassurance au près des réassureurs. AXA bénéficie aussi d'une diversification et mutualisation à l'échelle mondiale indispensable dans la gestion des risques.

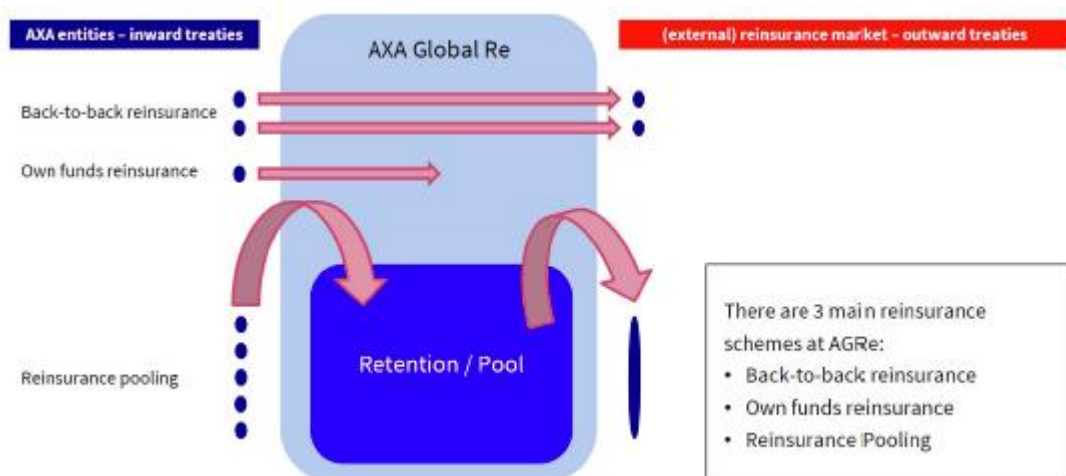


Figure 9 : Schéma de l'activité d'AXA Global Re

Le fonctionnement d'AXA Global Re est décrit sur la Figure 9. La gestion des traités est opérée de 3 façons :

- Les traités peuvent être rétrocédés totalement et identiquement à un autre réassureur (« *back-to-back reinsurance* »). La rétrocession est dite spécifique, AXA Global Re joue le rôle de courtier.
- Une deuxième possibilité est qu'AXA Global Re ne rétrocède aucune partie du traité et garde l'intégralité du risque au sein de son bilan (« *Own funds reinsurance* »).

- Une dernière possibilité est qu'AXA Global Re ne réassure qu'une seule partie du traité et rétrocède l'autre partie. Il s'agit de placer les traités au sein du pool de réassurance. Les primes et les sinistres sont mutualisés par toutes les cédantes. Les primes sont envoyées au pool, qui paie tous les sinistres. Chaque assureur récupère une partie du résultat proportionnelle à sa part de primes versées.

AXA Global Re peut ainsi mutualiser l'achat de protection de réassurance (couvertures "groupe") en mutualisant les portefeuilles de plusieurs entités, et décider de garder une partie du risque en rétention selon les opportunités de marché.

Pour le risque Catastrophe Vie, AXA Global Re est exposée à l'échelle mondiale, avec des traités en Excédent de Sinistre par événement concernant près de 17 entités différentes. Il existe aussi des traités régionaux regroupant plusieurs entités au sein du même traité.

2 Modèle statistique de tarification des excédents de sinistre catastrophe

Dans cette partie, il s'agit de présenter en premier (point 2.1) le modèle de tarification par simulation actuellement utilisé pour les traités CAT XL et ses limites. Ensuite (point 2.2), il est décrit une des améliorations du modèle avec la création d'un générateur d'évènements transfrontaliers pour tarifer les traités régionaux. Enfin (point 2.3), il est présenté l'état de l'art actuel des modèles de tarification proposés pour les traités CAT XL Vie.

2.1 Présentation du modèle

2.1.1 Méthodologie

Les traités CAT XL Vie couvrent sur une année donnée le montant des sinistres engendrés par un évènement ayant causé le décès d'au moins d assurés dans une période déterminée (3 jours). Le montant couvert par le traité est le montant excédant la priorité dans la limite de la portée.

Dans la suite du mémoire, il s'agira d'étudier des modèles de tarification appliqués aux risques de catastrophes naturelles (les tremblements de terre, tempêtes, inondations, tsunamis, feux de forêt, activités volcaniques). Néanmoins, ces modèles peuvent s'appliquer au risque terrorisme et autres évènements cat-man made (accidents de transport, accidents industriels...). Le modèle statistique ci-dessous peut ainsi inclure un module terrorisme en plus du module catastrophe naturelle pour tarifer entièrement le traité.

Le modèle statistique considéré a été développé dans l'article E. EKHEDEEN (2008). Ce modèle est basé sur une approche sévérité-fréquence. Il peut se découper en plusieurs parties comme le montre la Figure 10. La fréquence et la sévérité des évènements CAT sont calibrés grâce une base de données historiques.

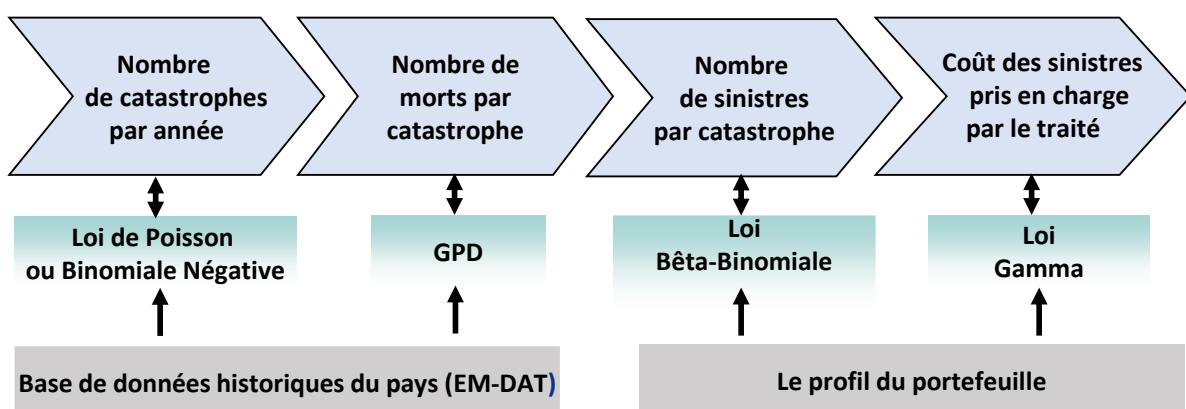


Figure 10 : Schéma récapitulatif du modèle par simulation

Calibration de la fréquence

Soit K_d le nombre d'évènements causant au moins d morts sur la zone considérée. Pour déterminer la fréquence du modèle, l'hypothèse suivante est appliquée : si $E(K_d) \geq V(K_d)$ alors K_d suit une distribution de Poisson sinon si $E(K_d) < V(K_d)$ alors K_d suit une distribution Binomiale Négative. En général, la loi de Poisson sera utilisée, il est fréquent d'observer dans les différentes régions du monde une sur-dispersion c'est-à-dire que la variance est supérieure à l'espérance. Les paramètres de la loi sont estimés par la méthode des moments.

Calibration de la sévérité

Pour déterminer le nombre de morts Z généré par les catastrophes sur la zone considérée, il est pertinent d'utiliser une distribution des valeurs extrêmes. L'utilisation d'une distribution des valeurs extrêmes comme en assurance non-vie a comme caractéristique d'avoir une queue de distribution épaisse. Par définition, les catastrophes sont des évènements majeurs provoquant un nombre de morts important.

Il s'agit donc de calibrer une GPD (Generalized Pareto Distribution) discrétisée sur les catastrophes provoquant la mort d'au moins d morts pour déterminer Z :

$$Z \sim GPD(d, \sigma, \varepsilon)$$

Les paramètres sont fixés par maximum de vraisemblance.

Après avoir estimé Z le nombre de victimes lors des catastrophes, il faut déterminer le nombre X d'assurés décédés parmi les victimes. Pour déterminer le nombre d'assurés conditionnellement au nombre de victimes lors des catastrophes, la loi beta-binomiale est utilisée

$$X|Z \sim Binomiale(Z, p)$$

$$\text{avec } p|Z \sim Beta(\delta \ln(Z)q, \delta \ln(Z)(1 - q))$$

Où q est la proportion d'assurés dans la population totale (taux de pénétration) sur la zone considérée, on a :

$$q = \frac{\text{nombre d'assurés}}{\text{population totale de la zone considérée}}$$

La probabilité q est fixée avec un paramètre δ prenant en compte la dépendance entre le nombre de morts et le nombre d'assurés parmi les victimes lors catastrophes. En effet, pour certaines catastrophes (accidents industriels ou dans les mines par exemple) ou pour les polices de groupe c'est-à-dire une police d'assurance couvrant collectivement un même groupe d'individus (salariés d'une même entreprise par exemple), il est possible que le nombre de victimes et le nombre d'assurés parmi victimes soient dépendants. Le paramètre δ est fixé par jugement d'expert.

Si $\delta \ln(Z) \rightarrow \infty$ alors : $X|Z \sim Binomiale(Z, p) \rightarrow$ Indépendance si δ est grand

Si $\delta \ln(Z) \rightarrow 0$ alors : $P(X = 0 | Z) = 1 - q$ et $P(X = Z | Z) = q \rightarrow$ Dépendance si δ est petit

Le montant des sinistres à la charge du réassureur

Il faut évaluer le coût économique des sinistres. Or, le réassureur ne dispose pas de toute la distribution de la répartition de la somme assurée de chaque assuré. La somme assurée peut être différente pour chaque police d'assurance couverte par le traité de réassurance. L'hypothèse utilisée pour estimer la perte assurée est celle du modèle de Strickler (cf. Annexes).

Soient :

- Y la variable aléatoire représentant le coût des sinistres pour un évènement CAT
- \overline{SAR} est la somme assurée moyenne couverte par le traité de réassurance
- $Exp(j)$ est la distribution exponentielle de paramètre j
- $\Gamma(k, \theta)$ est la distribution gamma de paramètre k et θ

Le montant du sinistre Y lors d'une catastrophe provoquant la mort d'une personne est : $Y \sim \overline{SAR} * Exp(1)$. Pour une catastrophe causant la mort de n morts, le coût des sinistres est : $Y \sim \overline{SAR} * \Gamma(n, 1)$. Ce résultat peut être démontré par récurrence. Le montant des sinistres pour chaque victime lors d'un évènement CAT est supposé comme indépendant et identiquement distribué.

Le montant d'un sinistre après application du traité s'écrit :

$$Y^{RE} = \text{Min}(\text{Max}(Y - \text{Priorité}; 0); \text{Portée}) \mathbb{1}_{X \geq d}$$

Le montant annuel des sinistres à la charge du réassureur est défini par :

$$S = \text{Min} \left(\sum_{i=1}^{K_d} Y_i^{RE}; AAL \right)$$

La prime pure du traité de réassurance est la moyenne des montants annuels des sinistres à la charge du réassureur obtenus lors des simulations.

2.1.2 Données disponibles pour la tarification des traités CAT XL

Informations sur les traités disponibles

Pour chaque traité CAT XL souscrit avec le réassureur, la cédante lui fournit des informations essentielles pour l'analyse des risques et la tarification du traité. Il s'agit de données de portefeuille permettant de connaître l'exposition du portefeuille sur la zone géographique considérée. Les éléments contractuels fixés entre la cédante et le réassureur sont disponibles.

Pour raisons de confidentialité, les chiffres présentés ci-dessous ne sont pas les données réelles du portefeuille d'AXA Global Re.

Pour le Mexique, les informations ci-dessous sur le Tableau 6 sont disponibles :

Contrat groupe			Contrat individuel		
Somme assurée (SAR)	Nombre d'assurés	Somme assurée moyenne	Somme assurée (SAR)	Nombre d'assurés	Somme assurée moyenne
900 000 000	5 000	180 000	400 000 000	4 000	100 000

Traité	Min. victimes	Priorité	Portée	Nb de reconstitutions	AAL
CAT-MEX	3	2 000 000	20 000 000	1	40 000 000

Tableau 6 : Données sur les traités

En plus, de ces informations pour certaines cédantes, les réassureurs disposent d'informations sur des zones à risque (bâtiments, etc...) où il existe un risque de concentration. Il s'agit d'un bâtiment par

exemple où plusieurs personnes sont assurées par un même assureur, salariés d'une même entreprise couverte par un contrat groupe par exemple. Ainsi un accident provoquant l'effondrement de ce bâtiment présente un risque de mortalité prépondérant pour l'assureur, il s'agit d'un risque de concentration.

La base de données EM-DAT

La base de données EM-DAT (*Emergency Events Database*) recense les catastrophes naturelles et technologiques depuis 1900. Il s'agit d'une base qui a été créée en 1988 par le Centre de recherche sur l'épidémiologie des catastrophes (CRED : *Centre for Research on the Epidemiology of Disasters*). Il s'agit d'une base objective visant à promouvoir la prise de décision pour se préparer au mieux face aux catastrophes et à améliorer l'efficacité de l'action humanitaire. Les données proviennent de plusieurs sources, notamment les compagnies d'assurance, les agences des Nations Unies, des organisations non gouvernementales, des instituts de recherche... Ces données sont utilisées pour évaluer la vulnérabilité et définir les priorités en matière de gestion de catastrophes.

Pour des raisons de fiabilité des données dans le cadre d'une tarification, il est indispensable d'utiliser les données seulement depuis 1970, ce qui correspond à un historique total de 50 ans. En effet, les moyens de mesure d'intensité et de traçage des phénomènes naturels sont récents. La comptabilisation d'évènements très anciens ne pourrait en effet, être complètement exact et juste. Par exemple pour les catastrophes du début du XXème siècle les évènements ne seraient probablement pas recensés de façon exhaustive, ou le nombre de victimes répertorié serait estimé, mais vraisemblablement inexact.

De plus, le choix de se limiter à une période récente permet de travailler avec un historique plus représentatif de la réalité actuelle (infrastructures etc...).

Cette base comptabilise les évènements s'il respecte au moins l'un des critères suivants :

- Cause au moins la mort de 10 personnes
- Affecte au moins 100 personnes (affectées/blessées/sans-domiciles)
- Affecte un pays ayant fait la déclaration de l'état d'urgence et/ou appel à l'aide internationale.

Sur cette base répertoriant diverses informations sur les catastrophes, ne sont présentées ci-dessous uniquement les colonnes utilisées pour la modélisation CAT en assurance vie :

- Numéro du désastre : un numéro à 8 chiffres composé de l'année et d'un numéro séquentiel à 4 chiffres propre à l'évènement
- La date de l'évènement
- Le groupe de désastre : Naturel, technologique ou désastre complexe (plusieurs facteurs)
- Le sous-groupe de désastre : biologique, géophysique, climatologique, hydrologique, météorologique...
- Le type de catastrophe : tremblement de terre, tempête, températures extrêmes, épidémies...
- Le sous-type de catastrophe : tremblement de terre, tempête tropicale, tempête extratropicale, etc...
- La catastrophe associée : si une catastrophe est associée à la catastrophe, la catastrophe associée est mentionnée (Tremblement de terre → Tsunami)
- La localisation (Villes, Continent, Pays, coordonnées GPS)
- Un indicateur d'intensité de l'évènement (Magnitude pour les tremblements de terre, Vitesse du vent pour les tempêtes, la surface inondée en km²...)
- Le nombre de victimes

- Le coût total des dommages assurés (en milliers de dollars US)

De 1970 à 2020, la base est composée de 22 877 lignes avec majoritairement des événements de l'Asie du Sud, de l'Est et du Sud-Est. Comme le montre le Tableau 7, les types de péril les plus représentés dans la base sont : les accidents de transport, les inondations et les tempêtes dans la base. Ce sont les mêmes périls qui sont les plus représentés annuellement. Les tremblements de terre sont le type de catastrophes entraînant en moyenne le plus de décès. Le Tableau 8 et la Figure 11 représentent la répartition géographique des événements. L'Asie compte le plus grand nombre d'événements CAT au sein de la base.

Type d'événements	Effectif	Freq. annuelle moyenne	Maximum de victimes engendré par un seul événement	Moyenne de victimes	Mediane (nbr de victimes)
Tremblement de terre	1 118	22	242 000	1 178	3
Sécherresse	675	13	300 000	1 038	0
Température extrême	582	11	55 736	324	18
Tempête	3 922	77	300 000	208	5
Epidémie	1 431	28	8 000	180	29
Activité volcanique	214	4	21 800	127	0
Inondation	5 112	100	30 000	64	7
Glissement de terrain	703	14	2 800	58	23
Accident de transport	5 683	111	4 000	40	22
Divers accidents (feu, explosion..)	1 403	28	2 236	36	17
Accident industriel	1 449	28	2 500	33	18
Feu de forêt	436	9	240	7	0

Tableau 7 : Récapitulatif des événements, EM-DAT, 1970 – 2020

Region	Repartition
Asie du Sud	15%
Asie de l'Est	12%
Asie du Sud-Est	11%
Amerique du Sud	8%
Afrique de l'Est	7%
Afrique de l'Ouest	6%
Amerique du Nord	6%
Amerique Centrale	4%
Europe de l'Est	4%
Europe du Sud	4%
Asie de l'Ouest	4%
Afrique du Nord	4%
Afrique Central	3%
Europe de l'Ouest	3%
Les Caraibes	3%

Tableau 8 : Régions les plus représentées, EM-DAT, 1970 - 2020

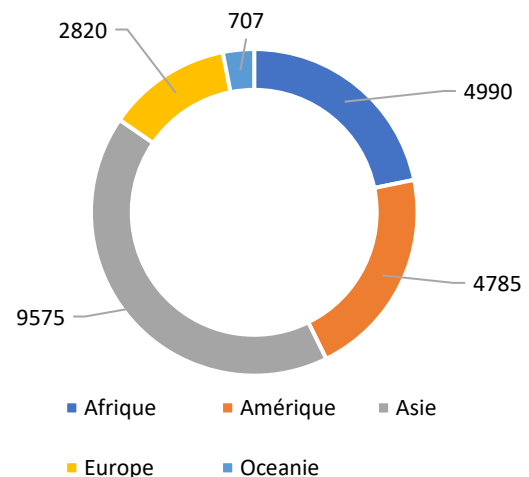


Figure 11 : Nombre d'événements par continent, EM-DAT, 1970-2020

2.1.3 Limites du modèle

Ce modèle a pour avantage d'être facile à mettre en place. Le modèle est basé sur un historique riche de 50 ans d'événements bien documentés permettant de calibrer une sévérité et une fréquence des

catastrophes. Les données d'assurance sur le portefeuille nécessaires pour ce modèle sont facilement accessibles (somme assurée moyenne et nombre d'assurés). Il a notamment la caractéristique d'intégrer la dépendance entre le nombre de victimes et les assurés parmi les victimes selon l'exposition aux polices individuelles et aux polices de groupe.

Le modèle présente néanmoins certaines limites :

- 1) Les simulations sont basées uniquement sur une base de données historique, le modèle repose sur la phrase suivante : « les événements passés dans l'historique sont représentatifs des événements futurs ». Or, la période de la base de données est limitée à 50 ans alors que des événements comme les grands tremblements de terre ont des périodes de retour qui se comptent en milliers d'année. Il existe un risque de sous-estimation d'événements les plus extrêmes et de non-comptabilisation de l'ensemble des catastrophes possibles. Le fait que le pays n'a pas subi un sinistre ou un type d'événement ne veut pas dire qu'il ne le connaîtra jamais. De plus, par les critères de comptabilisation de la base EM-DAT, la fréquence des événements est potentiellement sous-estimée notamment des événements CAT d'ampleur faible ou moyenne tuant moins de 10 personnes.
- 2) L'utilisation de la GPD n'est pas justifiée quantitativement, il pourrait donc être approprié d'utiliser d'autres distributions.
- 3) Le modèle ne prend pas en compte la localisation des assurés (spécificités géographiques, risque de concentration). Le risque CAT est parfois plus important dans un ou plusieurs territoires d'une même zone géographique, l'exposition varie. Il existe ainsi un risque de concentration et/ou des spécificités à prendre en considération grâce certaines informations disponibles (bâtiments spécifiques...) lors de la tarification.
- 4) Hypothèse forte sur la distribution du montant des sinistres réassurés (Loi Gamma)
- 5) Aucune spécificité selon le type de péril, le modèle ne calibrant qu'une seule loi de sévérité et fréquence pour l'ensemble des catastrophes.
- 6) Le modèle doit être adapté afin de prendre en considération les événements transfrontaliers pour tarifier au mieux les traités régionaux ou les couvertures groupes.

Dans la suite du mémoire, pour répondre à ces limites, il sera proposé une méthode de tarification des traités régionaux en vérifiant la pertinence de la GPD par rapport à d'autres distributions pour la sévérité (limites n°4 et n°6). Et dans un deuxième temps, il est proposé en complément un modèle physique pouvant mieux prendre compte le risque physique, l'exposition et les événements extrêmes.

2.2 Tarification des traités régionaux avec un simulateur d'événements transfrontaliers

2.2.1 Méthodologie

Pour tarifier un traité régional ou une couverture groupe, il est plus réaliste de prendre en considération les événements transfrontaliers. Ces événements ont la particularité de toucher deux pays ou plus en même temps. Dans le cas d'une tarification d'un traité de réassurance CAT XL couvrant plusieurs cédante, le fait de calibrer une sévérité et fréquence au niveau des deux pays comme une même zone géographique sans prise en compte des événements cross-border sous-estime potentiellement la prime pure. Il est donc indispensable pour AXA Global Re de disposer d'un modèle statistique de simulations pouvant comptabiliser correctement les événements transfrontaliers.

Par exemple, dans le cas où un même évènement comme une tempête touche deux pays à la fois la Belgique et l'Allemagne. En faisant, l'hypothèse que ce même évènement cause des morts en Belgique et en Allemagne. Le fait de comptabiliser cette tempête, comme deux évènements distincts touchant une même zone géographique, cela va impacter la fréquence (deux évènements sont comptabilisés au lieu d'un) et la sévérité (sous-estimation que cet évènement passe le seuil du traité de d morts). Des simulations basées uniquement sur des catastrophes non transfrontalières ne sont pas optimales.

Pour remédier à ce problème, pour la tarification des traités régionaux ou de la couverture groupe, il est présenté une nouvelle méthode basée sur le modèle de simulation ci-dessus. Il est tout d'abord indispensable de vérifier si les évènements transfrontaliers et les évènements non-transfrontaliers suivent les mêmes distributions. Pour cela, il sera utilisé un test de Wilcoxon-Mann-Whitney.

Il faut par la suite calibrer une sévérité et fréquence pour calculer la prime pure pour les évènements transfrontaliers et une prime pure pour les évènements non-transfrontaliers. La tarification du traité CAT XL sera le résultat de l'addition de ces deux primes pures.

Pour calculer ces primes pures, au niveau de la fréquence, le choix se fera entre une binomiale négative ou une loi de Poisson. Cependant pour la sévérité, le choix de la Pareto Généralisée (GPD) sera remis en question afin que la distribution choisie s'adapte le mieux aux données. Pour cela, il sera utilisé des tests et outils statistiques permettant de savoir si l'échantillon peut être approximé par une distribution précise (théorique) et vérifiant la qualité de l'ajustement. Les tests utilisés seront le test de Kolmogorov-Smirnov, le test d'Anderson Darling et le critère d'information AIC.

Le test d'Anderson-Darling est plus sensible aux queues de distribution. Alors que le test de Kolmogorov-Smirnov est plus au sensible au centre de la distribution. Pour l'évaluation d'une sévérité dans le cas de sinistres extrêmes (CAT), le test d'Anderson-Darling est très utile. Il pourra être associé à un qq-plot (représentation graphique des quantiles) afin de choisir la meilleure distribution qui s'adapte le mieux au niveau de la queue de distribution.

Le test de Kolmogorov-Smirnov :

Il s'agit de tester si un échantillon suit loi de probabilité donnée.

Soit (X_1, \dots, X_N) un échantillon i.i.d de même loi que X admettant F comme fonction de répartition. On teste l'hypothèse :

$$\begin{cases} H_0 : F = F_0 \\ H_1 : F \neq F_0 \end{cases}$$

Où F_0 est la fonction de répartition d'une loi connue.

Le test se base une statistique de test notée D_n calculant la distance maximale entre la fonction de répartition empirique de l'échantillon et la fonction de répartition théorique de la loi de probabilité connue. La statistique de test s'écrit :

$$D_n = \sup_{x \in \mathbb{R}} |F_{emp}(x) - F_0(x)|$$

Où :

$$F_{emp}(x) = \frac{1}{n} \sum_{i=1}^n 1_{X_i \leq x}$$

Le test d'Anderson-Darling :

L'hypothèse que l'échantillon suive une distribution spécifique ou non sera déterminé par une statistique de test précise, notée :

$$AD = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 [F(x)(1 - F(x))] dF(x)$$

Critère d'information Akaike (AIC) :

Ce critère permet d'évaluer la qualité d'ajustement d'un modèle par rapport à d'autres modèles. Un choix entre plusieurs modèles conduira à choisir celui qui minimisera ce critère.

Le critère AIC se note :

$$AIC = 2k - 2 \ln(\tilde{L})$$

L'AIC est calculé à partir du nombre de paramètres k utilisés pour construire le modèle et le maximum de la fonction de vraisemblance du modèle notée \tilde{L} . La vraisemblance explique dans quelle mesure le modèle reproduit les données. Il s'agit d'un compromis entre la qualité de l'ajustement (la vraisemblance) et la complexité du modèle (le nombre de paramètres).

2.2.2 Analyse des évènements transfrontaliers

Empiriquement, au sein de la base de données historiques EM-DAT, de 1970 à 2020, il a été référencé 628 évènements transfrontaliers (hors évènements exclus des traités de réassurance vie : sécheresse, épidémie, températures extrêmes etc...). Ces catastrophes transfrontalières ont lieu sur tous les continents. En moyenne, il y a 12 catastrophes transfrontalières par an. Ces catastrophes transfrontalières sont moins fréquentes mais plus sévères que les catastrophes non transfrontalières comme le montre le Tableau 9. Cela s'explique notamment par le fait que les catastrophes les plus meurtrières (les méga-périls : tsunamis, tempêtes) touchent souvent plusieurs pays en même temps et sont donc comptabilisés comme des évènements transfrontaliers. Il s'agit principalement de tempêtes et inondations, la répartition des évènements est présentée sur le Tableau 10.

	Effectif	Moyenne	Q1	Q2	Q3	Max
Evènements transfrontaliers	628	913	6	17	62	226 408
Evènements non-transfrontaliers	18 227	125	3	15	33	300 000

Tableau 9 : Nombre de morts provoqués par les évènements transfrontaliers et non-transfrontaliers

Type d'évènements	Effectif
Tremblement de terre	47
Inondation	206
Accident industriel	2
Glissement de terrain	6
Tempête	349
Activité volcanique	4
Feu de forêt	14

Tableau 10 : Types d'évènements transfrontaliers, EM-DAT, 1970-2020

Les évènements transfrontaliers sont donc moins nombreux, il est évident que la fréquence annuelle des catastrophes transfrontalières et non-transfrontalières sont donc différentes. Au niveau de la sévérité, les données statistiques précédentes la mettent en avant, mais reste relative et à démontrer. Il est nécessaire d'utiliser un test statistique non paramétrique pour confirmer le résultat. Le test de Wilcoxon-Mann-Whitney est adapté pour cette étude.

Le test de Wilcoxon-Mann-Whitney

Il s'agit d'un test non-paramétrique basé sur les rangs permettant de tester l'hypothèse selon laquelle deux échantillons suivent la même distribution.

Hypothèse :

- H_0 : les deux échantillons appartiennent à la même population
- H_1 : les deux échantillons sont issus de deux populations différentes

La procédure du test consiste à :

- Classer toutes les observations par ordre croissant de deux échantillons a et b (respectivement de taille n_a et n_b)
- Affecter le rang à chaque observation
- Calcul de W la Somme des rangs d'un échantillon de petite taille où $W = \sum_{i=1}^N r_{X,i}$ telle que $r_{X,i}$ le rang de la $i^{\text{ème}}$ observation de X
- Calcul de la statistique : $U = \min(U_a, U_b)$

$$\text{On a : } U_a = W_a - \frac{n_a(n_a+1)}{2} \text{ et } U_b = W_b - \frac{n_b(n_b+1)}{2}$$

Sous l'hypothèse H_0 vraie, l'espérance et la variance U s'écrivent :

$$E(U) = \frac{n_a n_b}{2} \text{ et } V(U) = \frac{(n_a + n_b + 1)n_a n_b}{12}$$

Le test est basé sur la comparaison de la valeur observée à cette moyenne et écart-type, en calculant une p-value. Le rejet de l'hypothèse H_0 se base sur le calcul d'une p-value.

En appliquant, le test de Wilcoxon-Mann-Whitney aux nombres de victimes des évènements transfrontaliers et aux nombres de victimes des évènements non-transfrontaliers, la p-value est égale à : 1.191e-07, la p-value est inférieure à 5%. Les deux échantillons sont donc significativement différents. Il est donc important de calibrer une sévérité propre aux évènements transfrontaliers, en soulignant l'existence d'une queue de distribution plus importante, sur des évènements extrêmes.

2.2.3 Application

Il s'agit d'appliquer le modèle de tarification à un traité régional couvrant des entités asiatiques :

- La priorité est de 10M et la portée est de 50M (CAT 50M XL 10M)
- Le seuil minimum de victimes parmi les assurés est d'au moins 5 morts. Il s'agit du seuil minimum pour que le traité de réassurance soit actif.
- Entités couvertes : Malaisie, Philippines, Hong-Kong, Thaïlande, Indonésie

Comme données de portefeuille, les hypothèses sont les suivantes :

Entités	Police groupe			Police individuelle		
	Somme assurée	Nombre d'assurés	Somme assurée moyenne	Somme assurée	Nombre d'assurés	Somme assurée moyenne
Malaisie	1 344 000 000	168 000	8 000	2 822 400 000	252 000	11 200
Philippines	495 000 000	99 000	5 000	567 000 000	81 000	7 000
Hong Kong	3 900 000 000	60 000	65 000	3 640 000 000	40 000	91 000
Thaïlande	1 920 000 000	320 000	6 000	4 032 000 000	480 000	8 400
Indonésie	10 000 000 000	400 000	25 000	126 000 000 000	3 600 000	35 000

Entite	retention	limite	nb.reconstitution	seuil min de victimes
Regional - Asiatique	10 000 000	50 000 000	1 (Gratuite)	5

Tableau 11 : Informations du traité régional asiatique

Au niveau des données EM-DAT, depuis 1970, pour ces pays, il a été comptabilisé 1265 évènements non-transfrontaliers et 36 évènements transfrontaliers provoquant le décès d'au moins 5 personnes. Il y a donc 25 catastrophes par an en moyenne touchant au moins un pays du traité régional et causant 5 victimes.

	Min	Q1	Q2	Q3	Max
Evenements transfrontaliers	10	38	102	243	174 204
Evenements non-transfrontaliers	5	13	25	57	7 354

Tableau 12 : Sévérité des évènements au sein de la base historique pour le traité régional

Au niveau de la sévérité, il y a bien une différence entre le nombre de morts provoqué par les évènements transfrontaliers et non-transfrontaliers. Notamment, au niveau de la médiane du nombre de victimes pour les évènements transfrontaliers qui est plus de 4 fois supérieure à celui des évènements non-transfrontaliers. De plus, la présence d'un évènement extrême causant le décès de plus de 174 204 et d'un 3^{ème} quantile plus important a pour conséquence une queue de distribution beaucoup plus épaisse pour les évènements transfrontaliers que non-transfrontaliers.

Pour tarifier ce traité régional, il faut donc :

- 1) Effectuer les simulations pour chaque pays (Malaisie, Philippines, Hong-Kong, Thaïlande, Indonésie) en prenant en considération uniquement les évènements non-transfrontaliers, il s'agit du risque CAT spécifique à chaque pays. Sur 500 000 simulations, il faut calculer les pertes brutes annuelles en retenant uniquement les évènements non transfrontaliers causant un nombre de victimes supérieur ou égal à 5 assurés et des pertes supérieures à la rétention.
- 2) Avant d'appliquer les conditions de réassurance aux pertes calculées ci-dessus, pour chaque année de simulation, il faut ajouter les pertes brutes associées aux évènements transfrontaliers pour tous les pays au sein du traité régional :

Pour la fréquence, les paramètres estimés sont les suivants :

Espérance	Variance	Loi de Poisson : P(λ)		Loi Binomiale Négative : BN(n,p)	
		λ	n	p	
0,71	1,41	0,71	0,41	0,71	

La variance est inférieure à l'espérance, dans ce cas la distribution négative est utilisée lors de la modélisation de la fréquence avec les paramètres n et p trouvés ci-dessus dans le tableau.

Pour la sévérité, il faut utiliser les critères d'ajustement afin de déterminer la loi de sévérité qui s'ajuste au mieux aux données pour cela, il est utilisé le test de Kolmogorov-Smirnov (*KS test*), d'Anderson-Darling (*AD test*) et le critère d'information AIC. Puis il faut bien analyser, que la distribution choisie puisse reproduire au mieux la queue de distribution des données empiriques.

Pour cela, il faut tout d'abord estimer les paramètres de la GPD et des autres distributions (Log-Normal, Weibull, Log-gamma). Il s'agit de distributions continues puisque la distribution du nombre de morts ne révèle aucune répétition sur les 36 événements non-transfrontaliers, les lois de probabilités discrètes ne convient pas dans ce cas afin d'approximer au mieux la sévérité. Les distributions à queue de distribution épaisses notamment à droite sont favorisées.

Ainsi, pour la sévérité, le tableau les résultats sont ci-dessous :

Distributions	Weibull	Log-normal	Log-gamma	GPD
<i>KS test</i>	0,2605	0,1866	0,1380	0,1381
<i>AD test</i>	3,5299	1,3633	0,5364	0,7773
<i>AIC</i>	524,1602	500,9544	490,3510	493,8546

En utilisant, uniquement les différents tests de Kolmogorov-Smirnov, la loi Log-gamme et la GPD sont les distributions les plus adaptées pour approximer la sévérité. En effet les statistiques de test sont moins élevées pour ces deux distributions et équivalentes. Or, le test d'Anderson-Darling en étant plus sensible aux valeurs extrêmes de la distribution (queues de distributions) permet de révéler la loi Log-gamma s'ajuste mieux que la loi GPD. Le critère d'information AIC est moins élevée pour la loi Log-gamme, ce qui vient confirmer le choix de cette loi.

Pour visualiser, les deux graphiques ci-dessous représentés sur la figure peuvent être utilisés. Le premier permet de s'apercevoir de la distance entre la fonction empirique de l'échantillon utilisée et les distributions théoriques (log-gamme et GPD) et le second compare les quantiles empiriques et théoriques. Le choix n'est pas simple visuellement, les lois GPD et log-gamma sont très proches et s'ajustent bien à l'échantillon empirique. L'utilisation du critère d'information AIC et du test d'Anderson-Darling étaient indispensables.

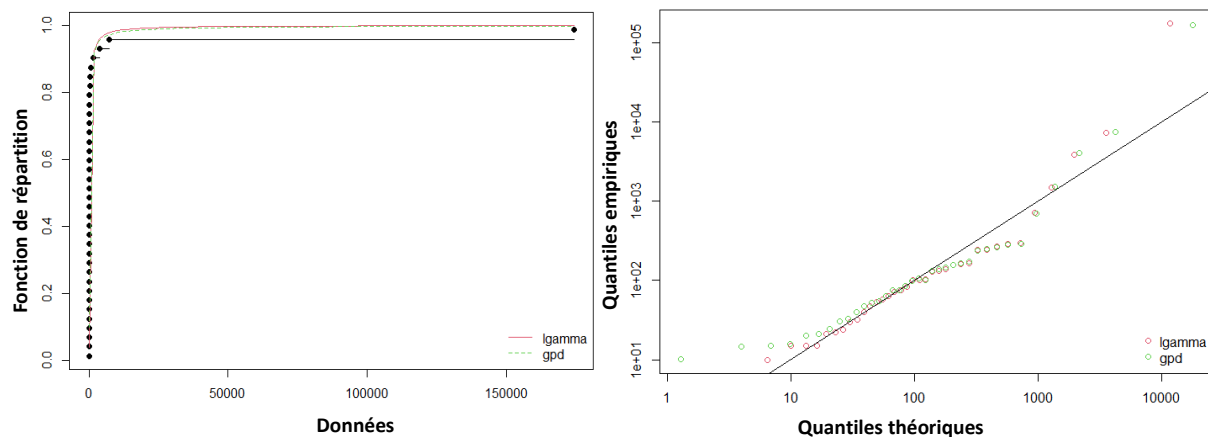


Figure 12 : Fonction de répartition empirique et qq-plot

Pour la sévérité, il est donc retenu une distribution log-gamma de paramètres : $\alpha = 7,64$ et $\beta = 1,94$.

Résultats : Sur 500 000 scénarios, 3 488 passent la priorité de 10M€ du traité, la prime pure s'élève à 62 405€. Il s'agit de la moyenne des récupérations sur les 500 000 scénarios. La prime pure obtenue est supérieure de 21% à celle obtenue avec un modèle comptabilisant les évènements transfrontaliers en plusieurs évènements non transfrontaliers (51 615€).

2.3 Etat de l'art des modèles de tarification proposés en CAT XL Vie

Il est couramment utilisé pour tarifer les traités CAT XL Vie, deux modèles : le modèle de Strickler (cf. Annexes) et le modèle de Ekhedden (modèle actuellement utilisé par AXA Global Re). Le modèle de Strickler, développé dans les années 1960 est un modèle déterministe considéré anciennement comme une référence en réassurance vie. Ce modèle présente de nombreuses limites (cf. Annexes) et tend à être remplacé par le modèle de Ekhedden. Le modèle Ekhedden a été construit pour répondre aux limites du modèle de Strickler mais reprend une des hypothèses du modèle de Strickler avec l'utilisation d'une distribution exponentielle pour représenter le montant d'un sinistre à la suite du décès d'un assuré lors d'un évènement.

Les deux modèles Strickler et Ekhedden sont cités et utilisés dans l'ensemble des mémoires traitant de la tarification des traités CAT XL Vie (N. LAMARI MFITH (2010) , J. MESSIAS (2017), A. PELLERIN (2019)).

A noter que le mémoire de N. LAMARI MFITH (2010) et J. MESSIAS (2017) utilisent et adaptent ces deux modèles de référence pour deux portefeuilles d'assurance. Leurs conclusions démontrent que le modèle de Ekhedden est le seul modèle apte pour tarifer les traités puisque le modèle de Strickler n'est plus adapté aux données actuelles.

Ces auteurs répertorient les mêmes limites évoquées précédemment (point 2.1.3) sur le modèle de Ekhedden. D'une façon générale, les données disponibles en réassurance sont très synthétiques, ce qui rend difficile la mise en place d'un modèle de tarification en réassurance vie sans hypothèses fortes. Il faut retenir notamment que l'hypothèse d'une distribution exponentielle pour représenter la distribution des sommes assurées par tête se justifie par sa simplicité et reste une représentation très plausible des sommes assurées par tête au sein d'un portefeuille en assurance-vie. Cependant, cette hypothèse instaurée par Strickler a l'inconvénient de sous-estimer la probabilité de pertes monétaires très élevées (quantiles élevées).

Le mémoire de A. PELLERIN (2019) consiste à améliorer un modèle de tarification basé sur le modèle de Strickler à l'aide d'un algorithme de Panjer pour mieux prendre en compte le risque de concentration et l'hétérogénéité des risques. Une des améliorations du modèle est l'utilisation d'une loi bêta-binomiale, il s'agit de la même hypothèse que le modèle de Ekhedden. La loi bêta-binomiale permet de mieux représenter le nombre d'assurés décédés à la suite d'une catastrophe en prenant en compte le fait qu'il existe des polices individuelles et groupes. La conclusion de l'auteur est que l'algorithme de Panjer permet à partir des quantiles de la loi bêta de donner la distribution de toutes les possibilités du nombre de personnes touchées lors d'une catastrophe. Pour une catastrophe donnée, mieux comptabiliser le nombre d'assurés décédés permet de mieux estimer le montant engagé d'un assureur.

De plus, le modèle proposé par A. PELLERIN fonctionne sur la simulation de scénarios pour construire un Event Loss Table associant une probabilité à chaque scénario de montant engagé par la compagnie d'assurance. Sur ce point, la construction du modèle se rapproche de notre modèle étudié, à la grande

différence que le modèle de A.PELLERIN utilise des scénarios basés sur la base historique EM-DAT avec des périodes de retour de 300 ou 500 ans. Il s'agit d'évènements historiques où le nombre de morts est déterminé par une mise en As-if des données historiques. Le modèle de A.PELLERIN sans approche physique, apparaît dans son apport plus limité alors que notre modèle étudié fonctionne à l'aide d'un catalogue d'évènements stochastiques, construit sur l'expertise de scientifiques.

Le nombre d'évènements simulés grâce au catalogue est très important, permettant de palier le risque de non prise en compte d'évènements majeurs ayant une période de retour très longue et par là, de pallier les défauts d'une base historique.

En outre, les nombres de victimes engendrés par les évènements stochastiques de notre modèle sont générés à l'aide de simulations prenant en compte l'intensité physique des évènements, l'exposition et la vulnérabilité. Les spécificités physiques et démographiques d'une zone géographique donnée sont prises en considération au sein du modèle physique.

Ainsi, le modèle cet auteur se différencie fortement de notre modèle étudié et proposé.

3 Etude de la dépendance entre les dommages Non-Vie et le nombre de morts lors des catastrophes naturelles

Dans cette partie, la dépendance entre les sous-modules CAT Non-Vie et CAT Vie est exposée et quantifiée. L'existence d'une dépendance entre les deux sous-modules permettra de justifier l'utilisation d'un modèle physique d'assurance non-vie en assurance vie. Au point 3.1, il sera présenté l'enjeu et le contexte de cette étude sur la dépendance et au point 3.2, les résultats de l'étude seront proposés ainsi qu'une analyse de cette dépendance sur le SCR d'un assureur.

3.1 Enjeu et contexte

Afin d'analyser et tarifier au mieux les traités en réassurance vie, l'application d'un modèle de tarification non-vie en réassurance vie permettrait à AXA Global Re de mieux appréhender les risques catastrophes et de pouvoir avoir une méthode alternative de tarification pour les traités. L'objet et la présentation de la mise en place de ce modèle se trouve en dernière partie de ce mémoire.

L'existence d'une dépendance entre le risque CAT Vie et CAT Non-Vie permettrait de justifier l'utilisation d'un modèle de tarification non-vie en réassurance vie. Or, Solvabilité 2 suppose une corrélation nulle entre le module Vie et Non-Vie puisqu'il s'agit de deux modèles de business différent (*point 1.1.3*).

Cependant, il est logique de considérer qu'il existe une dépendance entre le risque Vie et Non-Vie liée aux catastrophes naturelles. Les catastrophes de grande ampleur (man-made ou naturelles) impactent de façon considérable les activités Vie et Non-Vie de l'assurance ou de la réassurance.

Plus généralement, au niveau du secteur de l'assurance, il est possible de quantifier cette dépendance entre le risque Vie et Non-Vie. Le nombre de victimes est corrélé aux pertes en assurance non-vie lors des catastrophes. Par exemple dans le cas d'un accident très important, le crash d'un avion, il est probable que des passagers soient assurés par un assureur et par la même occasion le bâtiment détruit soit assuré par le même assureur.

Il est évident que cette dépendance existe seulement pour un réassureur si et seulement s'il regroupe des activités Non-Vie et Vie sur une même zone géographique. L'exposition compte beaucoup dans cette dépendance, notamment le taux de pénétration de l'assurance sur la zone géographique. Le niveau de vie au sein de la zone géographique doit être aussi pris en compte. En effet les pays développés ont mis en place des moyens de prévention efficaces pour faire face aux catastrophes. Le risque de subir des catastrophes causant de nombreuses victimes est donc plus faible dans ces zones où la couverture d'assurance est importante. Pour les pays en développement, c'est le contraire, la probabilité de constater un nombre de victimes important est forte mais la couverture de réassurance est plus faible.

La mesure de cette dépendance peut se faire en étudiant une base de données « marché » regroupant les pertes assurance P&C et le nombre de victimes (assurés décédés) lors des catastrophes dans le monde. Le nombre de victimes sera donc utilisé comme proxy de la mortalité d'assurés. Utiliser cette hypothèse a pour conséquence de surestimer le risque de mortalité de victimes couvertes en assurance vie mais reste la seule manière d'étudier le risque. Il n'est pas possible d'acquérir une base de données donnant des chiffres sur la mortalité des assurés au niveau marché ou même au niveau

d'AXA. AXA Global Re n'a jamais connu de récupération sur le risque Vie, aucune catastrophe en réassurance n'a dépassé la priorité d'un traité de réassurance. Il est donc essentiel de réaliser l'étude sur une base d'assurance « marché ».

En plus du volet tarification qu'implique cette dépendance, l'existence d'une corrélation entre les deux sous modules CAT Vie et CAT Non-Vie impacte le besoin en capital d'un assureur. Il est approprié de réaliser une analyse de sensibilité sur le besoin en capital supplémentaire qu'implique la prise en compte d'une dépendance entre ces modules. En cas de corrélation non nulle mais positive, le SCR d'un assureur augmente puisqu'il bénéficie moins de la diversification.

3.2 Mesure de la dépendance entre le risque Vie et Non-Vie

3.2.1 Données disponibles

Base de données publique EM-DAT :

La mobilisation de la base de données EM-DAT apparaît adéquate dans le cadre de cette étude. Pour certaines catastrophes, elle répertorie à la fois le coût des pertes assurées et le nombre de victimes. Par définition, les dégâts assurés répertoriés dans la base sont la valeur en dollars des dégâts causés aux biens, dont les cultures agricoles. Ces dégâts assurés correspondent aux pertes des assureurs en assurance non-vie.

Au sein de la base de données, pour les événements datés depuis 1970, seulement 5% des événements renseignés sont enregistrés avec les pertes P&C assurance et le nombre de morts. Il s'agit tout de même de 1038 événements au total, ce qui est un échantillon important pour réaliser une étude de dépendance. Ces événements sont répartis dans le monde. Il est important de relever qu'une fois la base filtrée sur uniquement les événements ayant des informations complètes, il n'y a plus aucun événement CAT Man-Made, il s'agit alors uniquement de catastrophes naturelles.

La base de données utilisée est représentée par le Tableau 13. Cette base permettra d'évaluer la dépendance entre le CAT Vie et le CAT Non-Vie. Il y a une proportion plus importante de tempêtes et d'inondations.

Le pays le plus représenté est les USA notamment avec 457 événements dont 366 tempêtes.

Type d'évènements	Effectif	Proportion au sein de la base	Nombre total de victimes	Total des pertes assurées (USD\$)
Tremblement de terre	90	9%	601 155	100 596 741 000
Tempête	654	63%	189 450	595 189 226 000
Inondation	223	21%	62 481	81 296 666 000
Température extrême	15	1%	9 752	5 841 200 000
Glissement de terrain	6	1%	2 094	443 900 000
Feu de forêt	43	4%	664	52 536 000 000
Sécheresse	7	1%	0	20 954 000 000

Tableau 13 : Présentation des événements utilisés au sein de l'étude

Pays (ISO)	Effectif	Proportion au sein de la base
USA	457	44%
JPN	62	6%
AUS	47	5%
CHN	41	4%
DEU	38	4%
GBR	33	3%
FRA	31	3%
MEX	24	2%
IND	21	2%
CHE	18	2%

Tableau 14 : Les 10 pays les plus représentés au sein de la base

Bases de données de la Banque Mondiale :

Par exemple, pour récupérer des données démographiques (taille de la population, densité...), socio-économiques (PIB, PIB/habitant etc...) à l'échelle nationale, la Banque Mondiale met à disposition des bases de données riches et variées sur l'ensemble des pays. Les données disponibles permettent d'avoir des séries de données complètes de 1970 à nos jours pour de nombreux pays. Grâce au code « ISO » (3 lettres uniques pour chaque pays du monde entier) et les années, cette base est facilement manipulable notamment pour effectuer des jointures avec la Base EM-DAT en cas de besoin.

Cette base est utilisée dans l'étude afin de normaliser les données brutes (victimes et pertes assurances USD) en taux de destruction et rendre les différents événements comparables entre eux. L'utilité d'employer des taux de destruction sera évoquée ci-dessous.

3.2.2 Utilité du taux de destruction et résultats

Taux de destruction

Comme les événements ne sont pas survenus ni à la même année et ni dans le même pays, il est trompeur de réaliser l'étude avec les données brutes. En effet, l'exposition de l'assurance n'est pas la même dans les différents pays, ne disposant pas d'information sur les expositions Vie et Non Vie de 1970 à 2020 sur tous les pays constituant la base EM-DAT. Les meilleurs proxys pouvant être utilisés comme exposition Vie et Non-Vie à l'échelle d'un pays sont respectivement la taille de la population et le PIB (en USD\$). Ces informations sont disponibles pour chaque année dans les bases de la Banque Mondiale et après la jointure entre la base EM-DAT et la base de la banque mondiale, il est possible de calculer les taux de destruction.

Pour chaque événement, les taux de destructions se définissent par :

$$\text{Taux de destruction}_{Vie} = \frac{\text{Nombre de victimes}}{\text{Taille de la population du pays}}$$

$$\text{Taux de destruction}_{Non-Vie} = \frac{\text{Pertes assurances}}{\text{PIB du pays}}$$

L'utilisation du taux de destruction a pour avantage d'éviter la mise en « *As If* » qui signifie faire comme si les événements étaient survenus aujourd'hui afin d'avoir un nombre de victimes et une perte

assurée à aujourd'hui. Avec les taux de destruction, il ne faut donc pas prendre en compte l'inflation. Cependant, travailler avec des taux de destruction a comme inconvénient de manipuler des chiffres très petits, de l'ordre de 10^{-6} par exemple.

Les outils mathématiques utilisés pour évaluer cette dépendance sont : les coefficients de corrélation (Pearson, Spearman) en plus de l'analyse des diagrammes de dispersion.

Les coefficients de corrélation et méthodologie

Coefficient de corrélation de Pearson :

Ce coefficient mesure une corrélation linéaire entre deux variables X et Y. C'est une méthode dite paramétrique car elle dépend de la distribution des données.

Pour calculer le coefficient, la formule suivante peut être utilisée :

$$\rho = \frac{COV(X, Y)}{\sqrt{V(X) * V(Y)}}$$

$$\rho = \frac{\sum[(x - \mu_x)(y - \mu_y)]}{\sqrt{\sum(x - \mu_x)^2 \sum(y - \mu_y)^2}}$$

$\rho \in [-1; 1]$; si $\rho = 0$ les deux variables ne sont pas corrélées sinon elle est positive ($\rho > 0$) ou négative ($\rho < 0$).

En principe, il est recommandé d'utiliser cette méthode lorsque les données X et Y doivent suivre une distribution normale, ce qui n'est pas souvent le cas.

Ainsi, pour déterminer s'il existe ou non une corrélation entre deux distributions, il est possible d'utiliser un test de corrélation non-paramétrique de Spearman par exemple.

Coefficient de corrélation de Spearman :

Il s'agit d'une mesure non paramétrique puisqu'elle est basée sur les rangs. Cette corrélation évalue la relation monotone entre deux variables X et Y.

Elle se calcule avec la formule suivante

$$\rho_{Spearman} = \frac{\sum \left(r_X - \frac{N+1}{2} \right) \left(r_Y - \frac{N+1}{2} \right)}{\sqrt{\sum \left(r_X - \frac{N+1}{2} \right)^2 \sum \left(r_Y - \frac{N+1}{2} \right)^2}}$$

Où :

- r_X et r_Y sont respectivement le rang dans l'échantillon X et Y d'une observation
- N est la taille de l'échantillon

Comme pour la corrélation de Pearson, le coefficient : $\rho \in [-1; 1]$.

Choix du coefficient de corrélation

Comme le coefficient de corrélation de Pearson est très sensible aux valeurs extrêmes. Il n'est donc pas très révélateur de la dépendance dans plusieurs cas.

	Tempête - USA	
	Avec l'évènement 'Katrina'	Sans l'évènement 'Katrina'
Coefficient de corrélation Pearson	78,09%	34,01%
Coefficient de corrélation Spearman	47,61%	47,22%

Tableau 15 : Coefficients de corrélation et sensibilité aux valeurs extrêmes

En effet, dans la base, l'évènement « Tempête » aux USA est l'évènement le plus représenté. Il y a 366 tempêtes ayant touchées les USA qui sont référencées dans la base, ce qui représente 35% de la base au total. Lorsque l'on calcule le coefficient de corrélation linéaire avec les 366 tempêtes dont la tempête la plus sévère en termes de morts et dégâts assurés (la tempête « Katrina » de 2005 tuant 1833 personnes et provoquant des pertes assurance à hauteur de 41 Mds de dollars USD), le coefficient de corrélation de Pearson s'élève à 78%. Or, dès le retrait de l'évènement « Katrina » du calcul, le coefficient de Pearson chute à 34%. Il convient par ailleurs de maintenir dans l'échantillon cet évènement extrême, révélateur de la sévérité des risques catastrophes.

Néanmoins, comme le montre le Tableau 15, dans ce cas, le coefficient de Spearman reste stable.

Il faut donc être prudent sur l'utilisation du coefficient de corrélation linéaire. Un écart important entre le coefficient de corrélation Spearman et de Pearson dans certains doivent demander une analyse plus précise notamment graphique à l'aide des diagrammes de dispersion.

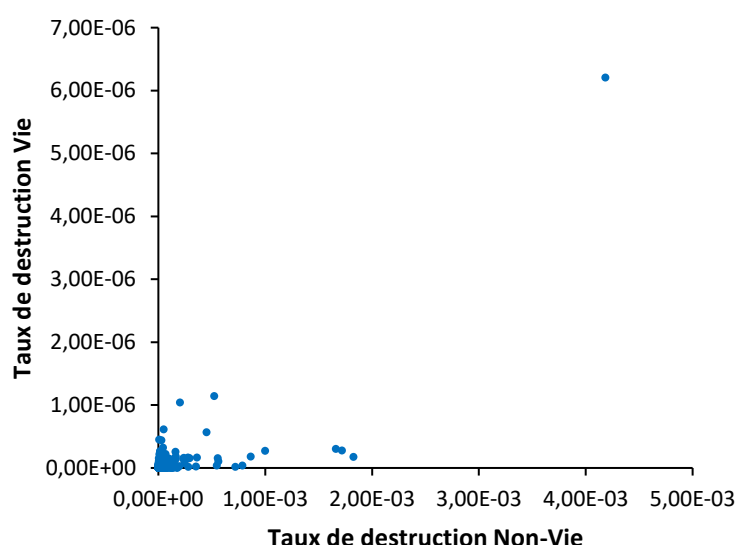


Figure 13 : Nuage de points des taux de destruction pour chaque tempête aux USA

Comme démontré sur la Figure 13, les taux de destruction fluctuent fortement selon les évènements. Il est impossible de visualiser et justifier une corrélation linéaire de l'ordre de 78% en visualisant le graphique.

Un moyen pour visualiser les données notamment par rang est de visualiser les données sous forme d'une copule empirique. Il faut calculer les rangs des taux de destruction puis les normaliser entre 0 et 1. Pour un échantillon U de taille n, $\{u_1, \dots, u_n\}$, les étapes de calcul sont présentées ci-dessous :

- Ranger l'échantillon dans l'ordre croissant
- Calculer les rangs r_j puis remplacer les valeurs u_j par leur rang

- Diviser chaque rang par le rang maximal noté ($r_n = \max(r_j)$)
- $\{u'_1, \dots, u'_n\}$ est donc compris entre 0 et 1.

Il s'agit d'une méthode qui s'apparente à celle utilisée pour les copules empiriques.

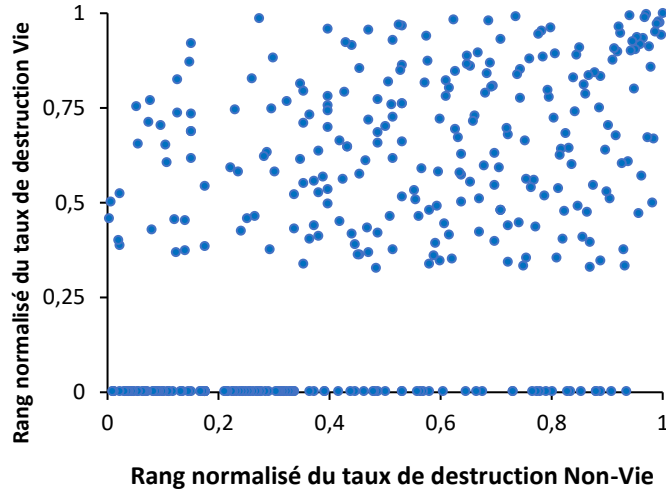


Figure 14 : Diagramme des rangs des taux de destruction pour le risque tempête aux USA

Appliquer cette méthode aux tempêtes aux USA permet d'obtenir la figure 14. Les événements les plus extrêmes sont regroupés avec les autres mais il reste difficile de mettre en évidence une corrélation par rangs de 47%. Les points sur l'axe des abscisses sont les événements n'ayant causé aucune victime. Les événements les plus mortels et ayant causé aussi des dégâts P&C importants se trouvent en haut à droite du graphique.

Analyse des résultats

Sur l'ensemble des données, le coefficient de corrélation de Pearson s'élève 6,42% alors que celui de Spearman s'élève à 46,49%. La relation n'est pas linéaire, l'analyse par rangs est plus adaptée à notre cas de dépendance.

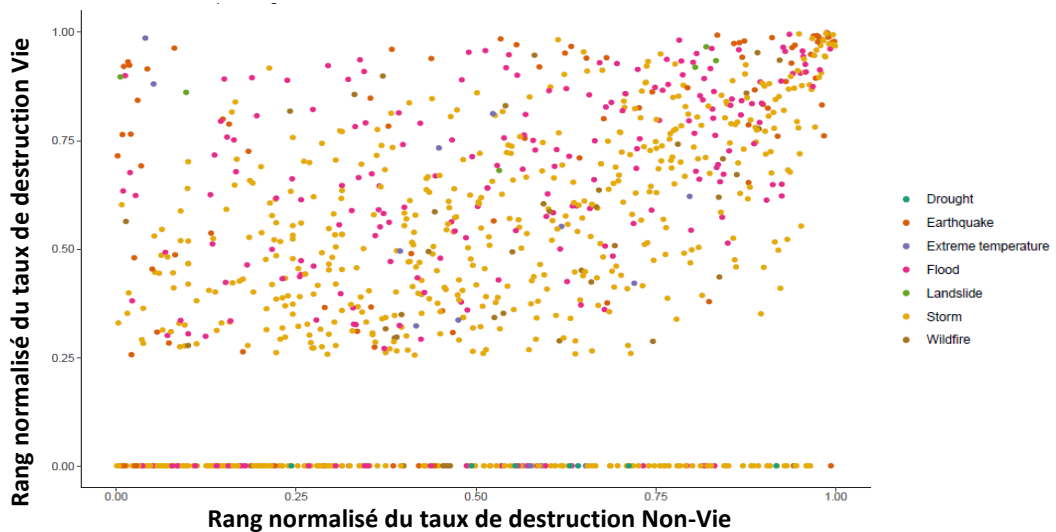


Figure 15 : Diagramme des rangs sur l'ensemble des événements

La Figure 15 présente le nuage de points obtenus pour une analyse par rangs. Il n'est pas évident de se rendre compte d'une dépendance à un niveau proche de 50% mais le graphique démontre bien une dépendance notamment au niveau des événements les plus extrêmes (en haut à droite du graphique). Ce graphique est très proche du graphique précédent puisque le risque tempête aux USA est le risque le plus représenté au sein de la base.

Analyse par type d'évènements

A partir des niveaux de corrélations présentés sur le Tableau 16 et la Figure 16, l'analyse pour les 3 événements CAT les plus représentés (Tempête, Inondation, Tremblement de terre) démontrent des coefficients de corrélation de l'ordre de 50% par rangs mais en linéaire très faibles.

Les tremblements de terre sont le type d'évènement présentant la plus forte dépendance avec un coefficient de corrélation de Spearman de 56%. En effet, les tremblements de terre sont les catastrophes les plus meurtrières, les personnes décèdent lorsque des dégâts matériels sont importants (effondrement d'un bâtiment par exemple).

Types d'évènement	Effectif	Proportion dans la base	Corr Pearson	Corr Spearman
Tempête	654	63,01%	8,46%	48,11%
Inondation	223	21,48%	4,48%	50,66%
Tremblement de terre	90	8,67%	16,80%	56,70%

Tableau 16 : Dépendance par type d'évènements

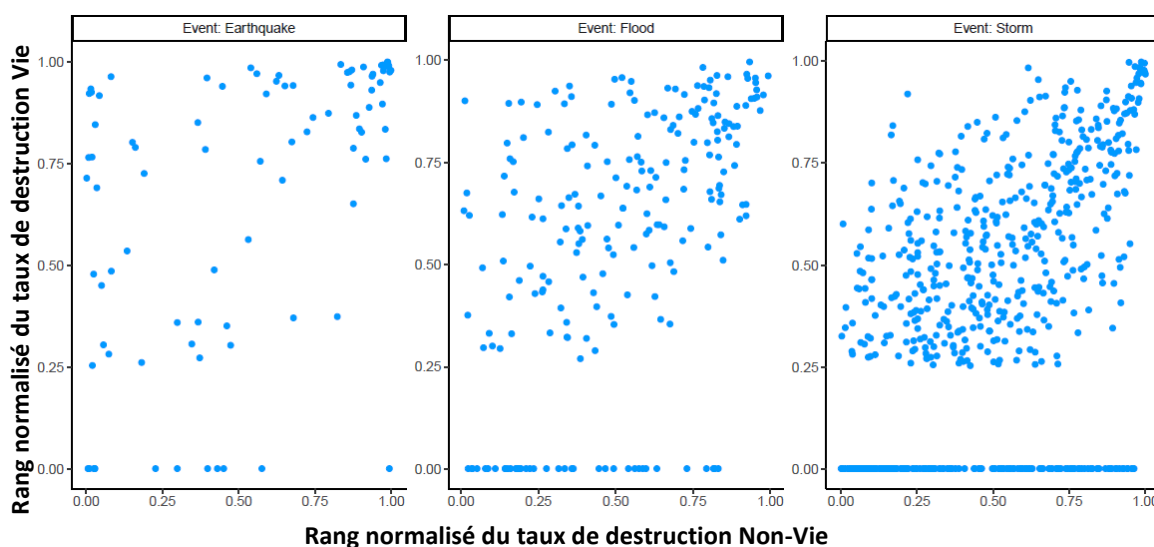


Figure 16 : Diagramme de dispersion de la dépendance par rangs pour les 3 grands types d'évènements

Mises à part les évènements n'ayant causé aucune mortalité, les 3 courbes font apparaitre une copule similaire, seule le nombre d'observation d'évènements engendre une densité différenciée pour chacun des graphes.

Analyse par pays

Il est intéressant d'étudier la dépendance selon les pays, un assureur pouvant être fortement plus exposé à une zone géographique qu'une autre ; Axa est par exemple plus présent en France.

ISO	Effectif	Proportion dans la base	Corr Spearman
CHE	18	1,73%	69,94%
DEU	38	3,66%	47,14%
FRA	31	2,99%	65,41%
GBR	33	3,18%	50,59%
JPN	62	5,97%	39,41%
USA	457	44,03%	47,59%

Tableau 17 : Corrélation de Spearman pour 6 pays majeurs

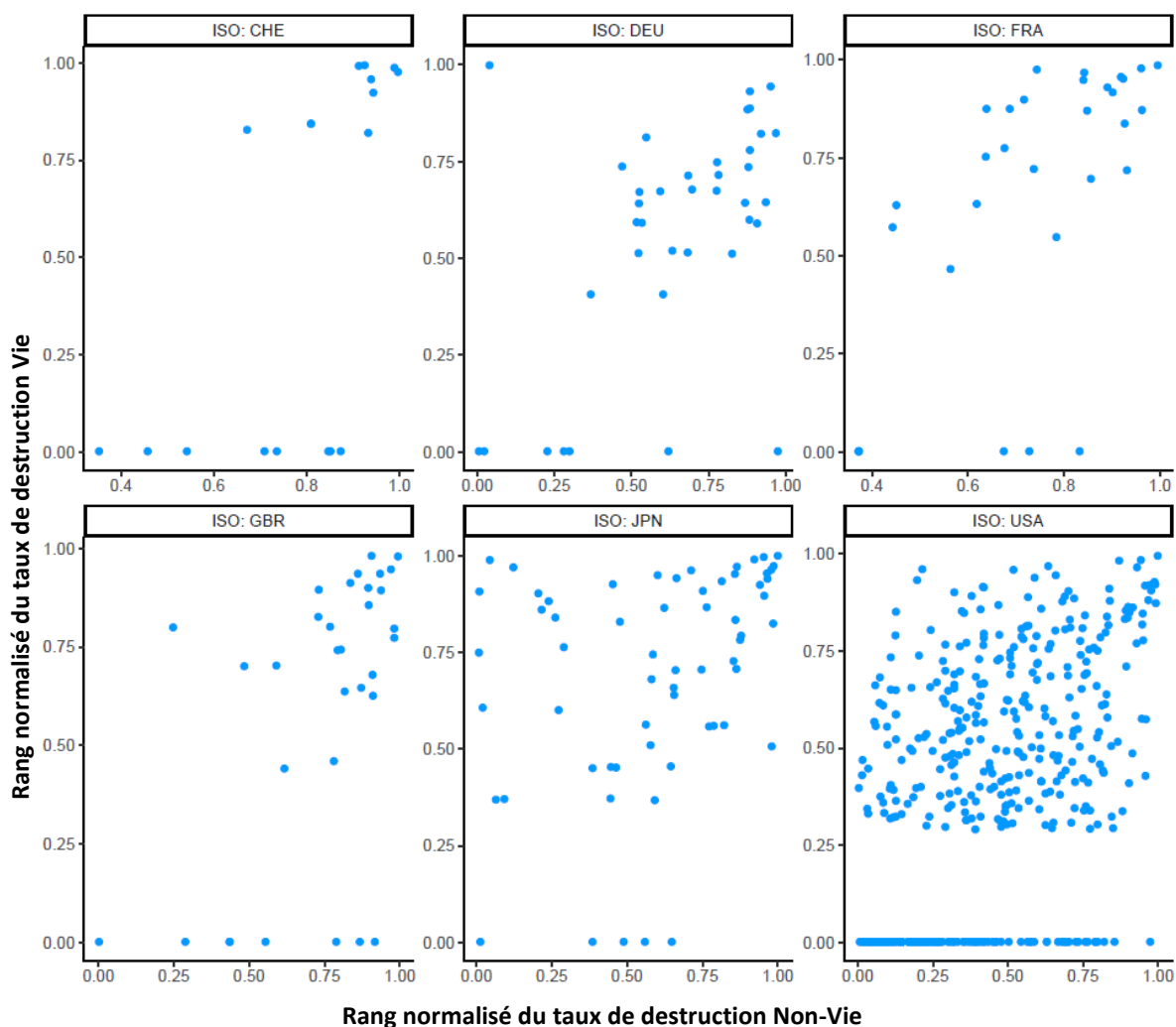


Figure 17 : Dépendance par rangs de 6 pays majeurs

Comme le montre le Tableau 17 et Figure 17, les corrélations sont moins marquées dans certains pays comme le Japon ou l'Allemagne, contrairement à la France et à la Suisse.

Pour le Japon, le nuage de points sur la Figure 17 révèle des catastrophes très meurtrières mais causant peu de dégâts P&C. Pour l'Allemagne, il y a aussi une catastrophe extrême du même type que ceux observés au Japon. Ces catastrophes diminuent la corrélation entre les risques CAT Vie et CAT Non-Vie.

Il faut noter que la dépendance est élevée en France et en Suisse. Ces pays se caractérisent par un nombre important de tempêtes au sein de l'observation. Par exemple, la France compte 31 observations dont 21 tempêtes ayant une dépendance forte de 74,71% (cf. 100).

Le risque tremblement de terre aux USA a un coefficient de corrélation de 73% sur 10 tremblements de terre.

Ainsi, évaluer la corrélation dans une zone géographique délimitée, c'est-à-dire à l'échelle d'un pays pour un risque spécifique (tremblement, etc...) permet de capter une dépendance plus élevée.

Il est finalement constaté une corrélation élevée entre les modules CAT Vie et Non-Vie, qui s'établit dans un intervalle large de 25% à 50%, avec dans la plupart des situations, une dépendance plus forte sur les évènements les plus extrêmes.

3.2.3 Impact sur le SCR d'un assureur

À la suite de la conclusion de l'existence d'une corrélation entre le risque CAT Vie et CAT Non-Vie, une analyse de sensibilité sur l'évolution du besoin en capital d'un assureur avec la prise en compte d'un niveau de corrélation positif est pertinente.

Pour tout assureur ou réassureur, il est essentiel que dans tout modèle interne prenne en considération la dépendance entre les risques. Lors de l'agrégation des deux modules Vie et Non-Vie, une corrélation positive liée aux catastrophes naturelles a pour effet une diminution de la diversification c'est-à-dire une augmentation du besoin en capital (SCR).

Il est pertinent de mesurer le capital supplémentaire exigé lors d'une dépendance positive entre les deux modules. Pour cela, il est possible avec la formule Standard de Solvabilité 2 d'appliquer des hypothèses de corrélation entre le module CAT Vie et CAT Non-Vie en créant une matrice d'agrégation sur la base de celle proposée par le régulateur.

Méthodologie

En formule Standard, la corrélation est égale à 0%, la matrice est la suivante :

Standard 0%	Life CAT	Mortality	Longevity	Disability	lapse	Expenses	Revision	Premium & Reserve	CAT
Life CAT	100%	0%	0%	0%	0%	0%	0%	0%	0%
Mortality	0%	100%	-25%	50%	0%	25%	0%	0%	0%
Longevity	0%	-25%	100%	0	25%	25%	25%	0%	0%
Disability	0%	50%	0	100%	0%	50%	0%	0%	0%
Lapse	0%	0%	25%	0%	100%	50%	0%	0%	0%
Expenses	0%	25%	25%	50%	50%	100%	25%	0%	0%
Revision	0%	0%	25%	0%	0%	25%	100%	0%	0%
Premium & Reserve	0%	0%	0%	0%	0%	0%	0%	100%	25%
CAT	0%	0%	0%	0%	0%	0%	0%	25%	100%

Tableau 18 : Matrice de corrélation des sous-modules (Formule Standard)

Par exemple, pour une hypothèse de 50% de corrélation, la matrice devient :

Corr de 50%	Life CAT	Mortality	Longevity	Disability	lapse	Expenses	Revision	Premium & Reserve	CAT
Life CAT	100%	0%	0%	0%	0%	0%	0%	0%	50%
Mortality	0%	100%	-25%	50%	0%	25%	0%	0%	0%
Longevity	0%	-25%	100%	0	25%	25%	25%	0%	0%
Disability	0%	50%	0	100%	0%	50%	0%	0%	0%
Lapse	0%	0%	25%	0%	100%	50%	0%	0%	0%
Expenses	0%	25%	25%	50%	50%	100%	25%	0%	0%
Revision	0%	0%	25%	0%	0%	25%	100%	0%	0%
Premium & Reserve	0%	0%	0%	0%	0%	0%	0%	100%	25%
CAT	50%	0%	0%	0%	0%	0%	0%	25%	100%

Tableau 19 : Matrice de corrélation des sous-modules (hypothèse de 50% de corrélation entre CAT Vie et CAT Non-Vie)

En appliquant différentes hypothèses de corrélation (25%, 50%), l'objectif est de mesurer l'écart de capital absolu et relatif par rapport à la formule standard dont l'hypothèse est une corrélation de 0%. De plus, il s'agira d'évaluer la corrélation implicite entre le module Vie et Non-Vie générée par la dépendance entre les deux sous-modules CAT.

La corrélation implicite est calculée avec la formule suivante :

$$Corr_{implicite} = \frac{SCR_{Total}^2 - SCR_{Vie}^2 - SCR_{Non-Vie}^2}{2 * SCR_{Vie} * SCR_{Non-Vie}}$$

Démonstration :

$$SCR_{Total} = \sqrt{SCR_{Vie}^2 + SCR_{Non-Vie}^2 + 2 * Corr * SCR_{Vie} * SCR_{Non-Vie}}$$

$$\Leftrightarrow SCR_{Vie}^2 + SCR_{Vie}^2 + SCR_{Non-Vie}^2 = 2 * Corr * SCR_{Vie} * SCR_{Non-Vie}$$

$$\Leftrightarrow Corr_{implicite} = \frac{SCR_{Total}^2 - SCR_{Vie}^2 - SCR_{Non-Vie}^2}{2 * SCR_{Vie} * SCR_{Non-Vie}}$$

Hypothèses et résultats

- 1^{er} exemple fictif :

Il s'agit d'une première structure de SCR proche de celle du Crédit Mutuel selon son rapport sur la solvabilité et situation financière (SFCR). La structure est présentée dans le tableau 17 ci-dessous.

Module	Sub-Module	Value
Life	Life CAT	132 000 000
	Mortality	478 000 000
	Longevity	428 000 000
	Disability	7 000 000
	Lapse	809 000 000
	Expenses	329 000 000
	Revision	3 000 000
P&C	Premium & Reserve	878 000 000
	Non-Life CAT	179 000 000

Life SCR	1 297 955 508			
P&C SCR	938 885 510			
P&C and Life SCR				
0% btw CAT Life and CAT P&C	1 601 934 612	Difference	Relative difference	Implicit Corr btw Life & Non Life SCR
25% btw CAT Life and CAT P&C	1 605 617 794	3 683 182	0,23%	0,48%
50% btw CAT Life and CAT P&C	1 609 292 546	7 357 935	0,46%	0,97%

Tableau 20 : 1^{er} exemple fictif, impact d'une corrélation entre sous-modules CAT sur le SCR

- 2^{ème} exemple fictif :

Pour ce second exemple, il s'appuie sur le 1^{er} exemple mais il s'agit d'accroître avant la prise en compte de toute diversification la part du risque Catastrophe en Vie et en Non-Vie dans le calcul SCR. Dans cet exemple, la part du SCR CAT Vie représente 20% du module Vie et la part du SCR CAT Non-Vie représente 35% du module Non-Vie avant toute diversification.

Module	Sub-Module	Value
Life	Life CAT	515 000 000
	Mortality	478 000 000
	Longevity	428 000 000
	Disability	7 000 000
	Lapse	809 000 000
	Expenses	329 000 000
	Revision	3 000 000
P&C	Premium & Reserve	878 000 000
	Non-Life CAT	474 000 000

Life SCR	1 390 140 101			
P&C SCR	1 097 108 017			
P&C and Life SCR				
0% btw CAT Life and CAT P&C	1 770 913 747	Difference	Relative difference	Implicit Corr btw Life & Non Life SCR
25% btw CAT Life and CAT P&C	1 805 045 844	34 132 097	1,93%	4,00%
50% btw CAT Life and CAT P&C	1 838 544 397	67 630 650	3,82%	8,00%

Tableau 21 : 1er exemple fictif, impact d'une corrélation entre sous-modules CAT sur le SCR

Ces deux exemples démontrent que dans une structure d'assurance où l'exposition au risque CAT-Vie et CAT Non-Vie ne sont pas prépondérantes (le 1^{er} exemple), une corrélation de 50% a un impact faible sur le SCR Total d'une compagnie d'assurance. Cela fait augmenter le SCR de 0,46%, ce qui est négligeable et induit une corrélation implicite des deux modules Vie et Non-Vie très faible (0,97%).

Mais dès lors que l'exposition CAT est prépondérante, le besoin en capital augmente considérablement, une corrélation de 50% entre les deux sous-modules augmente de 3,8% le SCR. Le niveau de corrélation implicite entre le risque Vie et Non-Vie induite par le risque CAT s'élève à 8%. Il faut donc du capital supplémentaire afin de mieux gérer les risques.

Ainsi, il est fondamental de pouvoir quantifier au mieux la dépendance entre les deux risques pour un portefeuille d'assurance donné même si le besoin en capital pourrait globalement augmenter de façon marginale pour un assureur dont le risque CAT n'est pas prépondérant.

4 Calibration d'un modèle physique pour la tarification du risque tremblement de terre

Cette dernière partie porte sur le risque tremblement de terre mais le modèle présenté pourra être répliqué sur tous les risques CAT (catastrophes naturelles ou man-made) pour tarifier l'ensemble des risques couverts par un traité CAT XL Vie. Le choix du péril « tremblement de terre » se justifie par la quantité de données disponibles lors de l'étude, mais également, en opportunité, de l'accompagnement par un expert du risque sismique d'AXA Climate lors de sa réalisation.

Par la dépendance qui existe entre les risques Vie et Non-Vie au niveau des catastrophes naturelles, il peut s'avérer pertinent d'appliquer une méthode de tarification utilisée en réassurance Non-Vie pour tarifier les traités CAT XL Vie. Il s'agit de mettre en place un modèle physique permettant de répondre aux limites du modèle de tarification actuel d'AXA Global Re (point 2.1.3).

L'étude a pour but de vérifier la possibilité de calibrer un modèle pouvant simuler le nombre de personnes décédées lors d'un tremblement de terre et de proposer une première tarification d'un traité CAT XL Vie pour le risque tremblement de terre. Cela permettrait de savoir si la mise en place d'un modèle physique en réassurance vie est possible ou non avec les informations et données que disposent AXA Global Re actuellement.

Une présentation de la méthodologie et des principes de la modélisation sera effectuée au point 4.1. Par la suite, la calibration du modèle grâce à des backtests en Italie et au Mexique est présentée au point 4.2. Enfin, au point 4.3, une première application à une tarification d'un traité est proposée et les limites du modèle sont exposées.

4.1 Méthodologie et principes de la modélisation

4.1.1 Présentation du modèle physique

Principe et modules de la modélisation physique

Tout d'abord, il est important d'expliquer le principe d'un modèle probabiliste ou physique. Cette modélisation a pour origine première de résoudre le problème du faible nombre de données historiques de sinistres liés aux phénomènes CAT. L'historique très faible de moins d'une cinquantaine d'années ne permet pas refléter le risque réel du portefeuille assuré. Il faut ainsi enrichir les données par des événements fictifs mais probables appelés scénarios afin d'obtenir une distribution des pertes liées au portefeuille. Ce modèle probabiliste consiste à simuler des événements possibles sur une période importante, des milliers d'années. Cette modélisation permet d'obtenir un catalogue d'événements avec une probabilité de survenance et une charge moyenne pour chaque événement.

Ces modèles sont développés par des spécialistes du risque climatique et sont souvent gérés par d'importants acteurs sur le marché comme : AIR (Applied Insurance Research), RMS (Risk Management Solutions) ... Pour évaluer le risque tremblement de terre, le GEM (Global Earthquake Model) a vu le jour en 2009 lors du Forum mondial de la science de l'OCDE en tant que fondation à but non lucratif financée par un partenariat public et privé. Le GEM s'impose comme l'une des sources les plus complètes et une véritable norme mondialement reconnue pour l'évaluation des risques sismiques.

Ce modèle est construit sur 4 modules qui ont des fonctions différentes :

- **Module Aléa** : Ce module permet en appliquant souvent des considérations physiques et statistiques de créer des événements fictifs probables qui peuvent toucher le portefeuille. Il permet d'obtenir un catalogue d'évènements (empreintes d'une inondation, épiceutre et profondeur d'un tremblement de terre, trajectoire d'une tempête...).
- **Module Exposition** : Ce module permet d'intégrer au sein de la modélisation les informations concernant les biens assurés (matériaux, géolocalisation, nombre de bâtiments etc...), leur nature et leur valeur (somme assurée...).
- **Module Vulnérabilité** : il permet d'associer une perte économique (ou monétaire) subie par les biens assurés et endommagés en fonction de l'intensité de l'évènement et leurs caractéristiques physiques. Il s'agit de l'impact physique des catastrophes sur les biens assurés (taux de destruction, dégâts...), il faut ainsi que les caractéristiques physiques renseignées sur les biens assurés dans le module d'exposition soient correctement renseignées.
- **Module Financier** : Ce dernier module applique les conditions d'assurance aux biens assurés comme la franchise, les limites, la coassurance... Ce module demande des données complètes et précises sur la couverture des biens assurés.

Le modèle permet d'obtenir des indicateurs, de cartographier et de tarifier le risque CAT. Il permet d'évaluer une structure de réassurance et de piloter au mieux cette couverture. Ces différents modules seront développés et appliqués au contexte des tremblements de terre dans les parties suivantes.

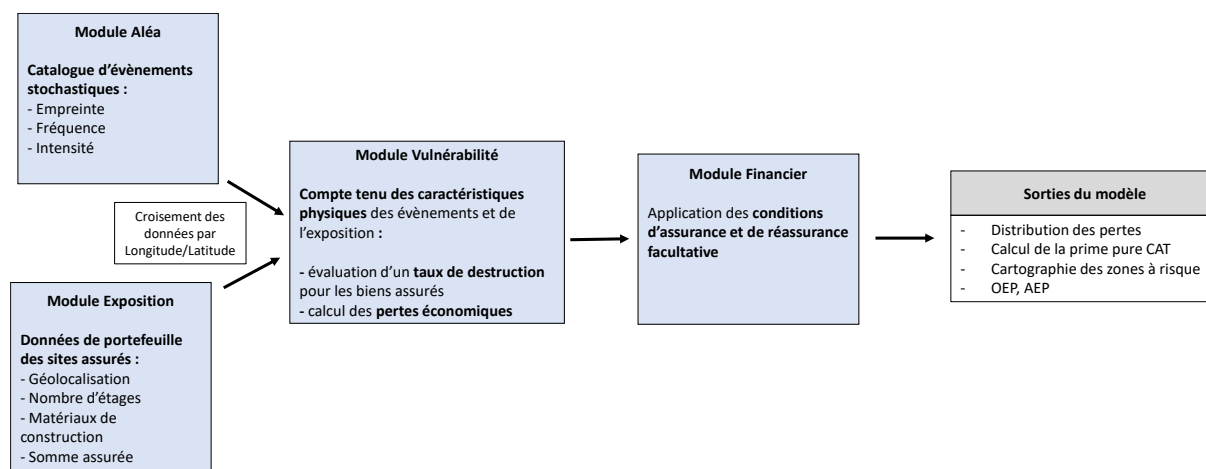


Figure 18 : Schéma récapitulatif du modèle physique

Intérêts de l'application en réassurance vie

Ce modèle est devenu la norme en modélisation des risques catastrophes naturelles en réassurance non-vie pour la tarification des traités. Il permet de simuler dans un monde virtuel, les sinistres causés par des catastrophes naturelles dans la réalité. Appliquer un modèle de ce type en réassurance vie permettrait de mieux modéliser les risques CAT et la prime pure.

Cette approche physique présente de nombreux avantages, elle permet de prendre en compte des événements inconnus dans l'historique d'évènements disponibles chez un assureur et apporte de la précision sur l'évaluation de la vulnérabilité et de l'exposition. Cependant, ce modèle est complexe dans sa mise en place et demande une qualité et une quantité de données importantes.

Il est évident que cette méthode pourrait pallier le manque de données historiques dans les modèles statistiques en tarification CAT XL Vie. Il s'agirait de simuler grâce au catalogue d'évènements un nombre de morts au sein des assurés couverts dans le portefeuille sur un territoire donné. Les avantages du modèle reposent surtout sur l'expertise et la connaissance de l'aléa physique. Au sein de la tarification, un plus grand nombre d'évènements et notamment les évènements les plus extrêmes non présents dans les bases historiques seront comptabilisés. L'hétérogénéité des risques et la vulnérabilité seront de nouveaux paramètres pour évaluer la prime pure.

En termes d'exposition grâce à l'analyse de l'aléa, il est facile de démontrer que certaines zones sur un même territoire sont plus exposées à un péril précis selon l'historique et l'analyse d'experts. S'il est possible de connaître les différents départements où résident les assurés, il serait possible de montrer qu'une partie du portefeuille d'un assureur par sa localisation est plus exposée que d'autres à une catastrophe naturelle (étude de cartographie et du risque de concentration).

Par exemple, pour le risque tremblement à l'échelle d'un pays, il existe des départements où la probabilité de connaître un séisme destructeur est importante (proximité avec une faille, type de sol particulier...). Ces caractéristiques géographiques ou géologiques d'une zone sont intégrées au sein du module aléa pour construire le catalogue d'évènements fictifs. Ainsi, en spécifiant les zones touchées, un nombre de logements bien précis (exposition) sera donc touché lors d'un évènement du catalogue. Les assurés se trouvant sur ces zones à risque ont donc potentiellement une probabilité plus importante de décéder s'ils n'ont aucun moyen de protection et de prévention. La tarification est donc plus juste en simulant le nombre de morts sur une partie du portefeuille réellement exposée au risque.

De plus, l'exposition au sein d'un modèle physique est beaucoup plus fine, les logements ou bureaux où résident et travaillent la population sont détaillés au sein du modèle afin de détenir le plus précisément la vulnérabilité de chaque personne ou assuré. Cette vulnérabilité en réassurance-vie est une probabilité de décès d'un occupant (assuré) au sein d'un bâtiment lors d'un tremblement de terre.

Méthodologie de l'application en réassurance-vie

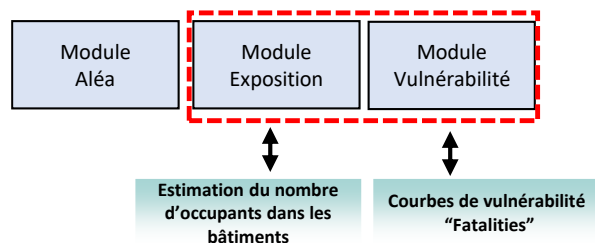


Figure 19 : Modules à calibrer pour le modèle physique

L'application du modèle physique en réassurance-vie demande principalement à faire évoluer deux modules du modèle physique : le module d'exposition et le module vulnérabilité.

Il est important de signaler qu'aucune catastrophe naturelle couverte par les traités CAT XL en réassurance vie à AXA Global Re n'a déjà provoqué des pertes. L'étude du modèle s'est donc faite sur le nombre de morts à l'échelle du pays non pas à l'échelle d'AXA puisque le nombre d'assurés décédés est inconnu. De plus, ce nombre est potentiellement très faible puisqu'il n'a pas dépassé le minimum de morts de la clause des traités CAT XL.

Dans le module d'exposition, en plus de bien modéliser les bâtiments affectés lors des catastrophes naturelles, il faut pouvoir modéliser le nombre d'occupants dans les bâtiments. Pour calibrer correctement le modèle, il faudrait connaître la localisation précise à une heure donnée des personnes.

A l'heure actuelle, il est malheureusement impossible d'obtenir une localisation des assurés couverts dans un portefeuille.

En effet, une localisation géocodée étant une donnée personnelle, il est donc impossible de récupérer et utiliser l'information. De plus, une position unique d'un assuré n'est pas suffisante, il faut pouvoir modéliser sa mobilité et avoir un traçage précis durant la journée des assurés. Il est donc très compliqué de modéliser une mobilité sans des données sur la mobilité, le style de vie ou les flux des personnes de façon très précise. Par exemple, dans l'idéal, il faudrait savoir le déroulement exact de la vie des personnes : « A quelle heure une personne active quitte-elle la maison pour aller travailler ? Combien de temps passe-t-elle dans les transports ? Où déjeune-t-elle le midi ? Change-t-elle de zone pour aller travailler ? etc.. ». Il s'agit de données très précises même trop précises. Ces problématiques sont équivalentes à l'assurance automobile, où les voitures changent de lieu durant la journée. Néanmoins, l'enjeu n'est pas le même, un dégât automobile n'a pas la même valeur économique ou la même importance qu'un décès.

Face à ces contraintes, pour calibrer le modèle avec un degré de finesse acceptable, la densité de population a donc été fixée selon les statistiques nationales, les seules hypothèses faites se font sur l'occupation des bâtiments sur deux plages horaires, la journée de 7h à 18h lorsque les personnes sont au travail et de 18h à 7h du matin lorsque les personnes sont dans leur logement. Ces hypothèses font fluctuer l'occupation dans les différents secteurs de l'exposition : résidentiel, commercial et industriel.

Dans le module vulnérabilité, pour modéliser les décès et non pas les pertes économiques sur les bâtiments, il faut que le module vulnérabilité inclut des courbes de vulnérabilité « fatalités » permettant de représenter la probabilité de décès moyenne des occupants en fonction de la structure des bâtiments et de l'intensité de l'évènement. Le GEM les modélise, l'enjeu est ainsi de savoir si ces courbes coïncident avec la réalité.

4.1.2 Le module aléa

Dans le cadre d'une approche stochastique, le module aléa permet de créer un catalogue stochastique d'évènements. Ce catalogue est composé d'empreintes d'évènements modélisés. Ainsi, pour chaque évènement du catalogue, l'empreinte permet de générer une carte représentant les effets du séisme. Ces catalogues de tremblement de terre sont créés par des géologues. Lors de l'application du modèle physique dans le cadre d'un modèle de tarification en réassurance vie, ce module n'a pas besoin d'être changé.

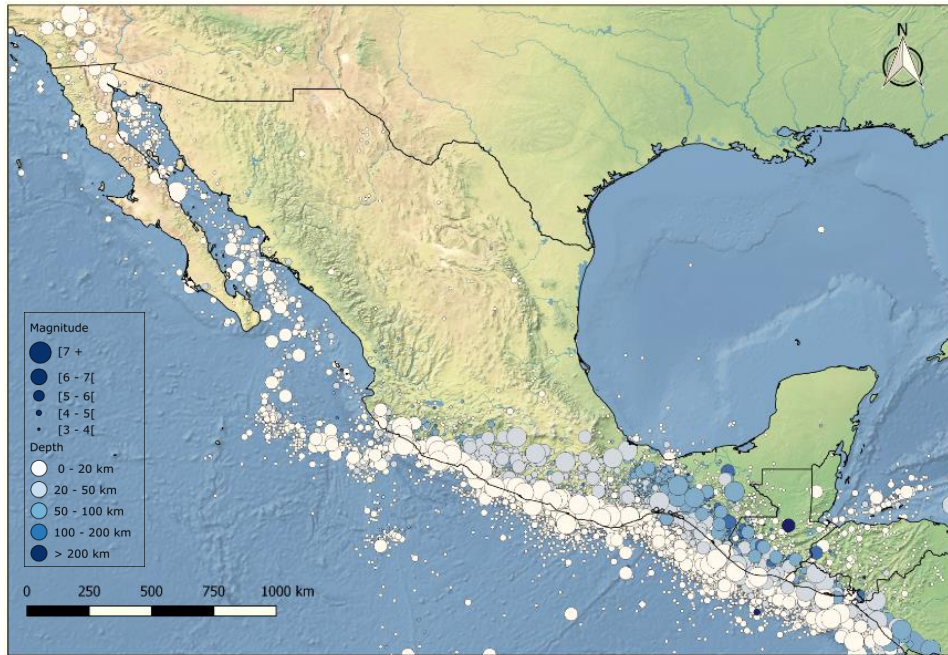


Figure 20 : Catalogue d'évènement Mexique – Tremblements de terre

La Figure 20 représente les tremblements de terre qui constituent le catalogue d'évènements pour le Mexique. Ce catalogue a été développé par les experts d'AXA. Tout au long des plaques tectoniques, des séismes de magnitude différents sont simulés avec une profondeur de l'hypocentre différente. Il s'agit du résultat du module d'aléa pour le Mexique.

Ce catalogue doit être complet et homogène pour qu'il soit le plus représentatif du risque et manipulable. Pour constituer ce catalogue, il est indispensable d'analyser les failles, les spécificités géologiques et l'historique des tremblements de terre de la zone. Il faut réaliser une analyse de la magnitude des tremblements de terre et de leur fréquence. Pour chaque magnitude, il est indispensable d'attribuer une fréquence, la période de retour de puissants séismes se compte en dizaine de milliers d'années par exemple.

Pour calculer, ces probabilités d'occurrence, il est possible de mettre en place une approche statistique en utilisant l'historique des séismes sur une zone délimitée. Une fois analysée les différentes failles et la fréquence des séismes sur la zone, la répartition des évènements dans le catalogue doit respecter cette loi d'occurrence avec les paramètres spécifiques de la zone.

Chaque évènement dans le catalogue se distingue par sa magnitude, sa profondeur de l'hypocentre, la localisation de son épicentre, le type de sol touché... L'intensité physique des ondes du tremblement de terre dépendra de ces paramètres. Pour chaque évènement du catalogue, il sera possible de tracer une carte de l'aléa, c'est-à-dire une carte géocodée de PGA (*peak ground acceleration*). Il s'agit de la mesure d'accélération au sol la plus utilisée pour les études en actuariat.

Le PGA est une mesure d'accélération du sol induite par un tremblement de terre, elle s'exprime généralement en m/s^2 ou en g (multiple de l'accélération de la pesanteur). Elle donne une indication précise sur l'intensité des secousses sur une zone géographique lors d'un tremblement de terre.

Cette mesure n'indique pas l'ampleur globale d'un séisme mais elle donne une indication de l'intensité pour différentes zones géographiques touchées par un même tremblement de terre. Contrairement à l'intensité Mercalli, il s'agit d'une grandeur physique, mesurée par des instruments spécifiques non

pas une mesure construite sur le ressenti de la population lors du tremblement de terre. Toutefois, le PGA est corrélé avec l'échelle de Mercalli.

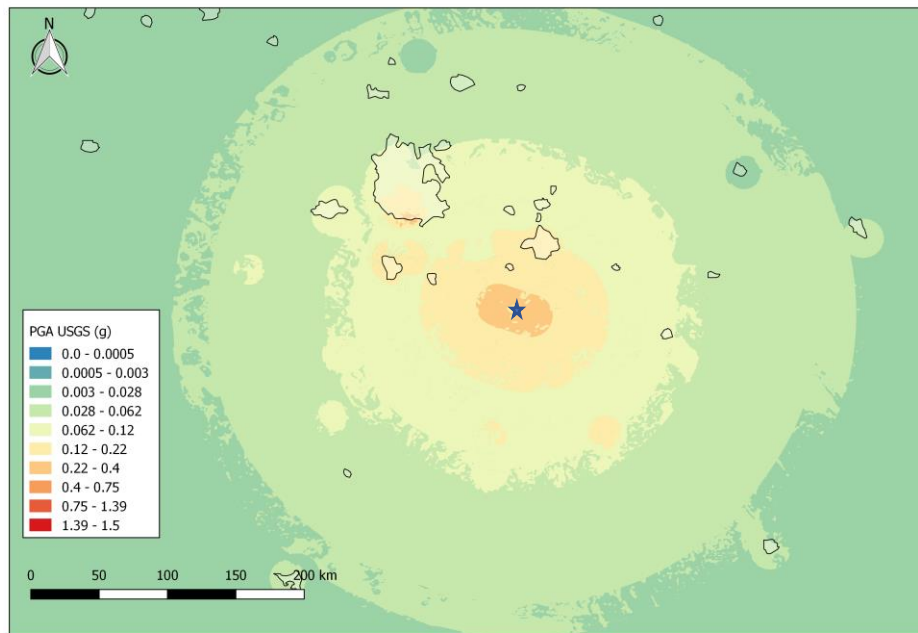


Figure 21 : Empreinte du séisme de Puebla 09/2017 – Carte PGA – Mexique

La Figure 21 représente l'empreinte du tremblement de terre de Puebla en Septembre 2017. Comme le montre la figure, le PGA globalement est élevé près de l'épicentre puis diminue fortement plus on s'éloigne de l'épicentre. Mais, cette relation entre PGA et distance de l'épicentre n'est pas fixe, il est possible de trouver un PGA plus important dans une zone à 80 km de l'épicentre qu'à 20 km de l'épicentre. Cela s'explique par la structure des sols, certains types de sol comme le sable peuvent amplifier les ondes sismiques alors que d'autres types de sol peuvent amortir les ondes sismiques.

Ainsi à l'aide de ce module, pour chaque tremblement de terre du catalogue d'évènements, il sera possible de tracer une représentation du PGA, représentant l'intensité des ondes sismiques qu'il faudra croiser avec l'exposition. En croisant avec l'exposition, cela permettra de repérer les bâtiments et donc les occupants exposés aux effets du séisme.

Les empreintes d'évènements historiques sont disponibles sur le site de l'USGS ((United States Geological Survey). Pour les deux tremblements de terre étudiés dans le cadre du backtest, les empreintes des tremblements de terre proviennent de l'USGS. Ces empreintes sont reconnues par les experts du monde entier. Elles permettent à ceux qui modélisent les tremblements de terre de vérifier la robustesse de leur modélisation en les comparant avec les cartes fournies par l'USGS.

4.1.3 Les courbes de vulnérabilité « Fatalities » pour estimer une probabilité de décès

Au sein du module vulnérabilité, il faudrait affecter pour chaque structure de bâtiment (béton, maçonnerie) une probabilité de décès des occupants. La fonction de vulnérabilité associe une probabilité de décès à un niveau d'intensité physique (module aléa) et pour un type de bâtiment donné (module exposition).

Pour déterminer les courbes de vulnérabilité, c'est-à-dire les taux de létalité pour un niveau d'accélération donné pour une structure, il faut généralement combiner deux approches : une approche expérimentale et une approche empirique.

La première approche consiste à faire des expériences dans un hangar (laboratoire) afin d'évaluer le taux de destruction d'un bâtiment pour un niveau d'intensité précis. A l'aide de maquettes avec de véritables murs, des expériences sont réalisées, il s'agit d'observer et de mesurer la déformation d'une structure (ex : bâtiment en béton) lorsque celle-ci est soumise à des vibrations.

Cette approche demande de beaucoup de moyens en termes d'outils et d'expertise. Elle est effectuée par des spécialistes en génie civil. Lorsque de telles expérimentations sont effectuées, les résultats sont précis et permet d'avoir un constat réel de la destruction de la structure. Cette approche constitue la première façon d'obtenir des courbes de fragilité, c'est-à-dire d'évaluer la probabilité de taux de dommage (ou d'effondrement) d'un bâtiment pour une mesure d'intensité précise du tremblement de terre. Ces taux de dommage ou taux d'effondrement seront combinés à des taux de létalité. Ce qui permettra de relier cette courbe de fragilité à une courbe de vulnérabilité qui mesure la probabilité de décès d'un occupant. Il s'agit d'une méthode indirecte pour évaluer la probabilité de décès d'un occupant.

La deuxième approche est une approche empirique qui repose sur des données afin d'évaluer ces fonctions de létalité. Cette approche demande un historique de données sur les morts et les blessés lors des tremblements de terre. Il s'agit de recueillir des questionnaires et des enquêtes à la suite des tremblements de terre.

Ces questionnaires regroupent des informations sur les entrées aux hôpitaux des enquêtes et questionnaires des gouvernements ou des ONG et bien-sûr des rapports d'expertise qui sont recueillis à la suite des tremblements de terre. Le recours au jugement d'expert vient compléter les données. La qualité et quantité des données sont déterminantes pour cette approche.

Après avoir compilé l'ensemble des données sur une multitude de tremblements de terre, il est possible d'utiliser des modèles de régression afin de modéliser la vulnérabilité. Les courbes de vulnérabilité générées seront calibrées selon plusieurs variables : le type de structure (les matériaux de construction du bâtiment), l'intensité physique appliquée au bâtiment, le nombre d'étages, le code de construction (normes parasismiques) ...

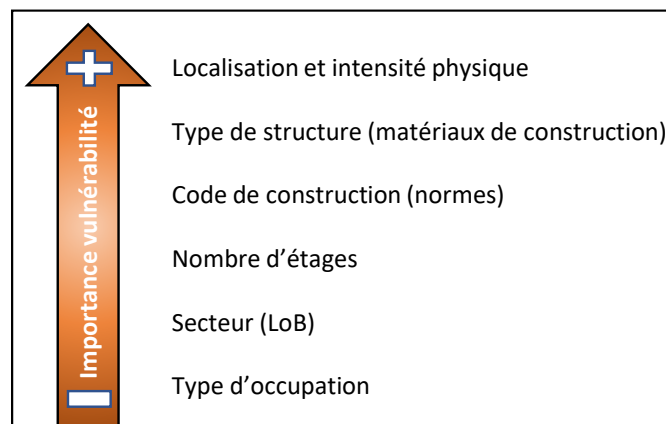


Figure 22 : Variable impactant la vulnérabilité

Le facteur qui influe le plus sur le taux de vulnérabilité est l'intensité physique enregistrée à la localisation du bâtiment. Plus l'intensité physique est importante (PGA), plus la probabilité de décès

est importante. Il peut être considéré qu'un PGA au-delà de 25%g, les secousses sont importantes, les dégâts peuvent commencer à être importants d'où une probabilité de décès des occupants qui commencent à être non nuls.

Les courbes utilisées dans le cadre de cette étude sont issues du Global Earthquake Model, elles sont réalisées par des experts. Pour l'Italie, les courbes correspondent à celles utilisées pour tous les pays européens alors que pour le Mexique, il s'agit de courbes modélisées uniquement pour le Mexique.

Grâce à la géolocalisation des bâtiments et de la taxonomie des bâtiments (informations sur le type de structure, nombre d'étages...) de ces courbes, il est facile de croiser l'exposition et ces courbes. Pour chaque pays, le nombre de courbes est important et se distingue selon les caractéristiques de l'exposition.

Globalement, les probabilités de décès sont très faibles de l'ordre 10^{-6} pour un indicateur d'accélération entre 25%g et 30%g par exemple. Ainsi, pour obtenir un décès en utilisant ce modèle, il faut qu'au moins 10^6 occupants soient exposés à ce niveau de PGA.

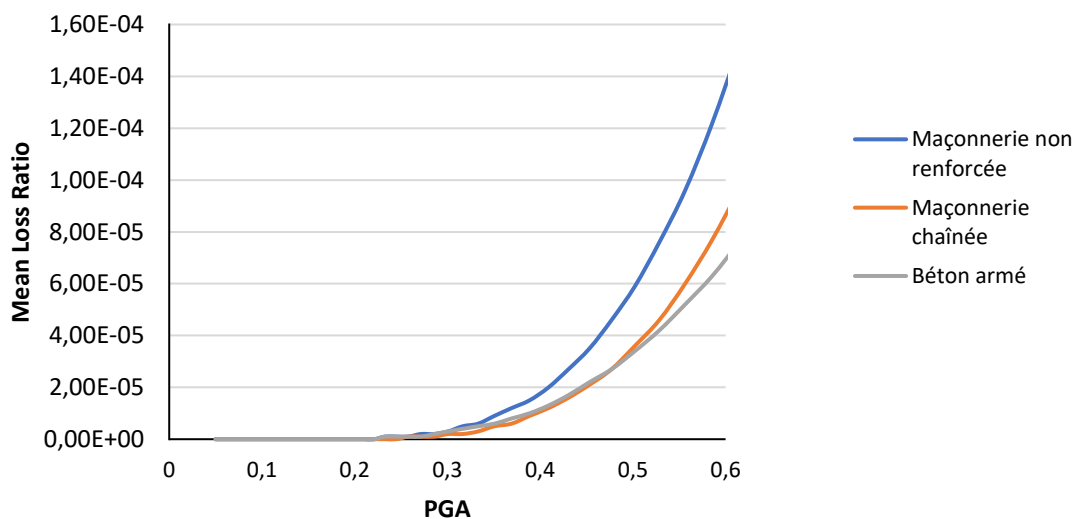


Figure 23 : Exemple courbe de vulnérabilité sur les niveaux de PGA globalement observés – Mexique

4.2 Calibration du modèle grâce à des backtests

Afin d'appliquer un modèle physique en tarification vie, il faut bien veiller que la structure du modèle convient à une approche dite de « fatalité ». Les backtests permettent de calibrer le modèle pour une zone géographique donnée et de vérifier la capacité du modèle à répliquer la mortalité lors de différents séismes. Les backtests sont effectués pour l'Italie et le Mexique. Deux backtests sont détaillés : « Amatrice » en Italie, 08/2016 (296 morts) et « Puebla » au Mexique, 09/2017 (369 morts). Le résultat de quatre autres backtests sont donnés.

4.2.1 Description des évènements et aléa physique

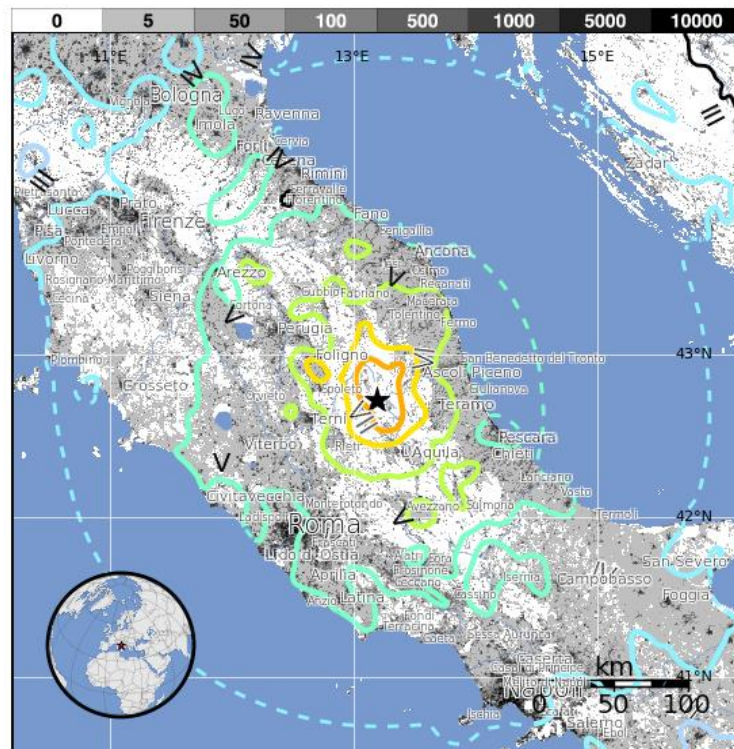
L'étude s'est faite sur des tremblements de terre les plus meurtriers de ces dix dernières années au Mexique et en Italie. Deux seront exposés intégralement au sein du mémoire, celui d'Amatrice en Italie, qui a causé de nombreuses pertes humaines dans des zones rurales (villages dans des montagnes) où la densité de population est faible et un second séisme, à Puebla, au Mexique, où les villes touchées sont caractérisées par une forte densité urbaine. Les caractéristiques principales des tremblements étudiés sont résumées dans le Tableau 22.

Séismes	Italie Province du Rieti - 24/08/2016	Mexique Etat de Puebla - 19/08/2017
Heure locale	03h36	13h14
Magnitude	6.2	7.1
Profondeur	4 km	51 km
Nombre de morts	296 morts	369 morts
Zone la plus touchée (morts, distance par rapport à l'épicentre)	Amatrice (234 morts, à 40 km de l'épicentre)	Mexico (228 morts, à 120 km de l'épicentre)
Densité de la population et nombre d'habitants de la zone la plus touchée	14 habitants par km ² avec 2 400 habitants	6 200 habitants par km ² avec 9 209 000 habitants
Pertes P&C	122M USD	1 200M USD

Tableau 22 : Description des tremblements de terre étudiés

Le 1^{er} tremblement de terre en Italie s'est produit dans le centre de l'Italie, il a eu lieu le 24 août 2016 à 03h36 en heure locale, il s'agissait d'un séisme de magnitude 6.2 dont l'épicentre est situé dans la province de Rieti à 4 km de profondeur.

De nombreuses fortes secousses ont suivi la secousse principale. Ce tremblement de terre a causé des dégâts dans plusieurs communes, Accumoli, Arquata del Tronto, Amatrice et Pescara del Tronto. Des secousses ont même été ressenties à Rome qui se situe à environ 150 km de l'épicentre comme le montre la Figure 24. Sur l'échelle Mercalli, l'intensité ressentie est mesurée à un niveau de VII près de l'épicentre, c'est-à-dire que les dommages sur les bâtiments les moins bien construits et les moins solides sont importants.



Population per ~1 sq. km. from LandScan

Figure 24 : Estimation de la population exposée aux secousses et intensité ressentie (MMI) lors du tremblement de terre d’Amatrice (2016)

La province qui a été le plus sévèrement notamment en pertes humaines est celle d’Amatrice qui est située à environ 8.5 km de l’épicentre. Amatrice a quasiment été rasée par le tremblement de terre avec plus de trois-quarts du village qui a été détruit (cf. 100).

Le bilan humain total du séisme s’élève à 296 morts, dont plus de 234 morts dans le seul village d’Amatrice. Ce bilan humain très lourd s’explique principalement par 3 facteurs : la structure vulnérable des bâtiments exposés au tremblement de terre qui ne répondaient pas aux normes de construction antisismiques ; la magnitude de l’évènement et la survenance du tremblement de terre pendant la nuit lors de la période estivale.

En effet, ce séisme a touché des petites villes et villages, des zones historiques composés majoritairement de logements anciens et construits avec des briques et pierres arrondies. Les habitations étaient construites en maçonnerie traditionnelle italienne non armée qui ne correspondait plus aux nouvelles normes parasismiques, le risque d’effondrement des bâtiments était important.

De plus, ces facteurs ont été amplifiés par la survenance du séisme pendant la nuit et la densité de population plus importante durant la période estivale au mois d’août. Pendant la nuit, tous les habitants sont à l’intérieur des bâtiments donc exposés aux risques d’effondrement. La densité de population dans les différents villages était très importante avec la saison estivale et la forte influence touristique. En effet, la commune d’Amatrice est une commune comptant environ 2400 résidents (14 habitants par km²) mais durant cette période de vacances, la population avoisine les 30 000 à 40 000 personnes. Le soir du tremblement, il s’agissait en plus d’une nuit précédant une fête qui attire de nombreux touristes et italiens dans ce petit village le dernier week-end d’août. Le nombre d’occupants des bâtiments aux secousses sismiques était donc plus important.

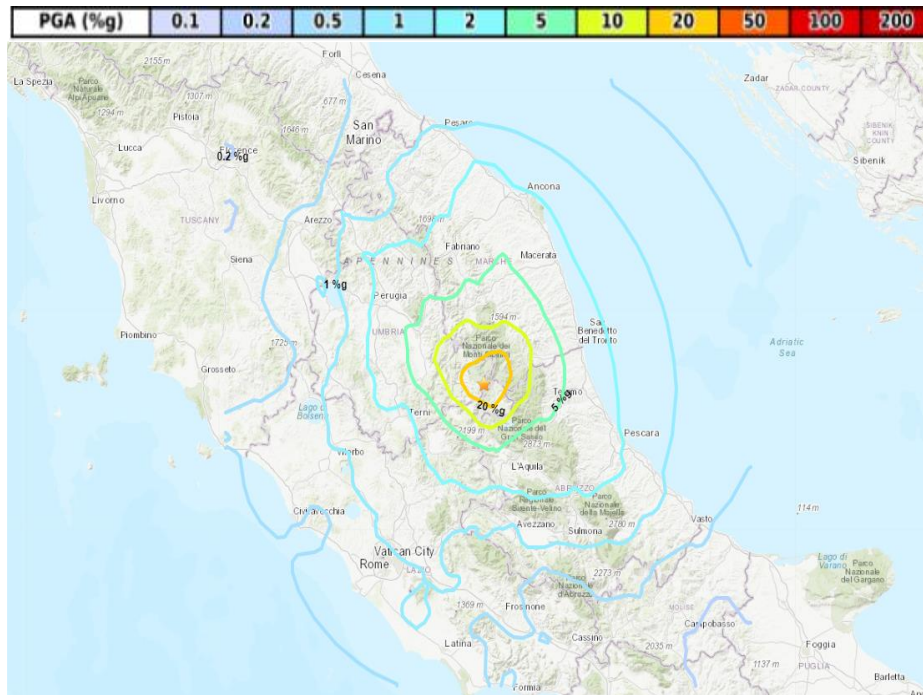


Figure 25 : Intensité physique enregistrée (PGA) lors du tremblement de terre d'Amatrice (2016)

Sur la Figure 25, le PGA qui est l'indicateur physique qui permet de mesurer l'intensité des secousses atteint un niveau d'au moins 20%g près de l'épicentre sur une zone faiblement peuplée où les constructions sont anciennes et peu solides au centre de l'Italie. Il faut bien noter que les appareils ont enregistré les secousses à des centaines de kilomètres autour de l'épicentre. L'empreinte s'étend quasiment sur toute l'Italie.

Le 2^{ème} tremblement de terre étudié est celui dans l'Etat de Puebla au Mexique. Ce séisme s'est produit le 19 Septembre 2017 à 13h14 (heure locale) avec une magnitude 7.1 et une profondeur de l'épicentre à 51 km. La mégalopole mexicaine, la ville de Mexico à 120 km de l'épicentre, a été la plus touchée.

Les pertes assurés P&C s'évaluent à près de 1,2 milliards USD en P&C. Plus d'une quarantaine d'édifices se sont effondrés et des milliers de logements sont inhabitables du fait des dommages irréremédiables. La structure des bâtiments et leur résistance sont représentatives des fractures sociales économiques au sein du pays, les quartiers les plus aisés ont été épargnés par rapport aux quartiers les plus modestes. Au total 369 personnes sont mortes dont 228 dans la seule ville de Mexico. Le séisme est survenu 32 ans jour pour jour après le séisme de 1985 de Mexico ayant provoqué la mort de plus 10 000 personnes.

La ville de Mexico est d'une géographie particulière, il s'agit d'une zone lacustre, placée sur un ancien lac qui s'est asséché au fil du temps. La structure du sol est donc constituée d'argiles molles qui amplifient les oscillations sismiques. Démonstré par le PGA et l'intensité de Mercalli sur les Figure 26 et Figure 27, malgré la distance avec l'épicentre, la ville de Mexico est fortement impactée par les secousses. Le PGA est élevé sur la mégalopole mexicaine et l'intensité est aussi importante que les zones les plus proches de l'épicentre.

Des normes de construction antisismiques ont normalement été mises en place pour les bâtiments construits après 1985 dans la ville de Mexico, mais il s'est révélé des défaillances avec l'effondrement d'écoles et de collèges lors de ce tremblement de terre provoquant de nombreuses victimes.

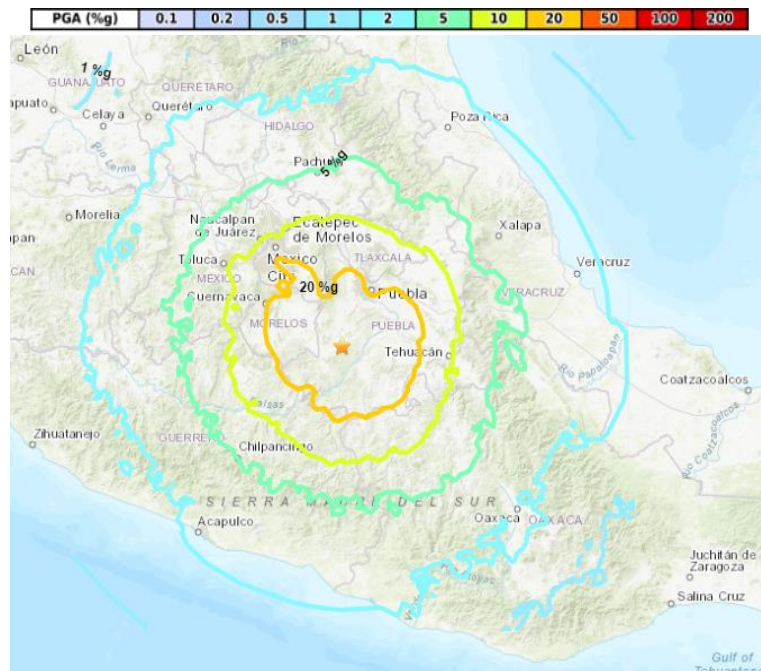


Figure 26 : Intensité physique enregistrée (PGA) lors du tremblement de terre de Puebla (2017)

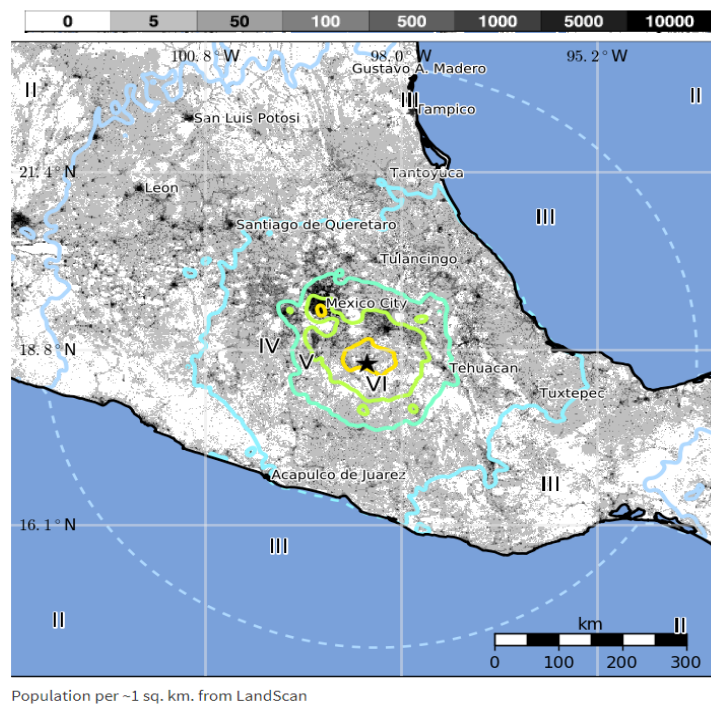


Figure 27 : Estimation de la population exposée aux secousses et intensité ressentie (MMI) lors du tremblement de terre de Puebla (2017)

4.2.2 Calibration de l'exposition

Description de la base de données

Les données d'exposition concernant l'Italie et le Mexique proviennent comme les courbes de vulnérabilité, du GEM (Global Earthquake Model). Ces données d'exposition sont géocodées, la localisation des bâtiments est exprimée en longitude et latitude. Ces données se présentent sous la forme de deux bases de données permettant d'avoir des informations sur les bâtiments présents en Italie et au Mexique. Sur ces deux bases, nous trouvons au total, les variables suivantes :

- *ID_1* : Identifiant de la région administrative (1^{er} niveau administratif)
- *NAME_1* : Nom de la région administrative
- *ID_2* : Identifiant de la commune (2^{ème} niveau administratif)
- *NAME_2* : Nom de la commune
- *LONGITUDE, LATITUDE* : Coordonnées géographiques attribuées à l'actif immobilier. Selon la base, ces coordonnées peuvent représenter des emplacements individuels de l'actif ou peuvent être géocodés selon la densité de la population dans une commune précise.
- *IMPLANTATION* : rurale ou urbaine
- *TAXONOMIE* : Taxonomie du bâtiment utilisée pour appliquer le module de vulnérabilité sur l'actif. (Structure externe, structure interne et nombre d'étages)
- *ETABLISSEMENTS* : Le nombre total de maisons ou de logements dans le secteur résidentiel et le nombre total d'établissements dans les secteurs commerciaux et industriels.
- *BÂTIMENTS* : Le nombre total de bâtiments composant l'actif.
- *COÛT_DE_REMPLACEMENT_USD* : Le coût de remplacement total de l'actif (en 2018 USD, y compris les composants structurels et non structurels et le contenu du bâtiment).
- *COÛT_STRUCTUREL_USD* : Le coût de remplacement des composants structurels de l'actif (en 2018 USD).
- *COÛT_NON_STRUCTUREL_USD* : Le coût de remplacement des composants non structurels de l'actif (en USD).
- *COUT_DES_CONTENUS_USD* : Le coût de remplacement des contenus de l'actif (en dollars US).
- *SECTEUR* : Le secteur principal de l'actif (RES : résidentiel /COM : commercial /IND : industriel).
- *OCCUPANTS_PAR_ACTIF* : Le nombre moyen d'occupants la nuit dans tous les bâtiments composant l'actif, ce nombre est disponible seulement pour le secteur résidentiel.

Ces deux bases de données sont néanmoins différentes. Sur la base de données de l'Italie, il existe trois variables supplémentaires qui peuvent permettre d'estimer plus précisément le nombre d'occupants dans un bâtiment :

- *TYPE_D'OCCUPATION* : Le type d'établissement pour un chaque actif du secteur commercial (commerce, hôtel ou bureau)
- *SUPERFICIE_EN_M²_DES_ETABLISSEMENTS* : La surface moyenne d'une maison ou d'un logement dans le secteur résidentiel ; la surface moyenne d'un établissement dans les secteurs industriels et commerciaux. Ce chiffre évolue selon le type d'occupation.
- *COÛT_PAR_SUPERFICIE_USD* : Le coût moyen de remplacement par unité de surface (en 2018 US\$/m², incluant les éléments structurels et non structurels mais excluant le contenu du bâtiment).

Hypothèses et exposition en Italie

Italie - Base de données	
Nombre d'actifs	1 216 118
Nombre de bâtiments	12 010 784
Nombre d'occupants	56 593 516

Secteur	Répartition
Résidentiel	93,5%
Commercial	4,1%
Industriel	2,4%

Type de structure	Répartition
Maçonnerie non renforcée	70,0%
Béton armé	27,3%
Autres	2,7%

Hauteur des bâtiments	Répartition
Faible (1-5)	99,0%
Moyenne (6-10)	1,0%

Normes de construction	Répartition
Aucune norme	89,7%
Faibles	10,1%
Elevées	0,2%

Tableau 23 : Description de l'exposition en Italie

Comme le montre les différents tableaux, les bâtiments en Italie sont surtout des bâtiments construits de façon traditionnelle avec de la maçonnerie non renforcée. La hauteur des bâtiments est faible. Les bâtiments ne respectent pas des normes parasismiques. Ils sont donc vulnérables aux puissants séismes. Au total, au sein de la base, l'Italie est composée de 12 000 000 de bâtiments.

Pour la modélisation, ce sont les bâtiments localisés qui se situent au sein de l'empreinte du tremblement qui sont cruciaux puisqu'ils sont exposés aux ondes sismiques du tremblement de terre. Ces bâtiments selon leur structure se verront appliqués un taux de mortalité moyen des occupants. Il est donc primordial d'estimer au mieux les occupants des bâtiments à l'heure donnée du tremblement de terre.

Dans le cas de l'Italie, le séisme a eu lieu à 3h du matin, il est donc raisonnable de faire l'hypothèse que les personnes au sein de l'empreinte dorment chez eux ou à l'hôtel (répartition des occupants à 100% dans le secteur résidentiel). Grâce à la base de données, le nombre d'occupants dans le secteur résidentiel (logements) est déjà déterminé.

Afin d'être plus précis pour le backtest, il est nécessaire de faire une hypothèse sur les occupants dans les hôtels. La capacité des hôtels à 3h du matin a été définie en multipliant le nombre moyen de chambres par hôtel au taux d'occupation moyen et au nombre de clients moyen par chambre.

Le séisme a eu lieu durant le mois d'août, selon les statistiques nationales d'Italie le taux d'occupation des hôtels était à son plus haut niveau durant ce mois avec d'un taux d'occupation de 79% des hôtels (ISTAT (2017)). Ce taux est donc utilisé comme taux d'occupation pour les hôtels au sein de l'empreinte. Selon le rapport écrit par le cabinet de conseil Horwath HTL (2018), leader dans le secteur du conseil en Hôtellerie, Tourisme, Loisirs, en moyenne, les hôtels en Italie sont composés de 33 chambres contre 36 en moyenne en Europe. En moyenne, 1.5 personnes par chambre viennent dormir sur une année donnée.

Hypothèses et exposition au Mexique

Mexique - Base de données	
Nombre d'actifs	425 204
Nombre de bâtiments	31 455 379
Nombre d'occupants	119 536 680

Secteur	Répartition
Résidentiel	85,1%
Commercial	13,0%
Industriel	1,9%

Type de structure	Répartition
Maçonnerie non renforcée	47,3%
Maçonnerie chaînée	25,0%
Béton armé	13,6%
Maçonnerie renforcée	7,7%
Bois	4,5%
Autres	1,9%

Hauteur des bâtiments	Répartition
Faible (1-5)	95,8%
Moyenne (6-10)	4,2%

Normes de construction	Répartition
Aucune norme	72,2%
Faibles	22,8%
Elevées	5,0%

Tableau 24 : Description de l'exposition au Mexique

Au Mexique, le nombre total de bâtiments au sein de la base s'élève à 31 455 379. Ces bâtiments sont construits principalement en maçonnerie non renforcée, chaînée et béton armé. La présence d'une part importante de maçonnerie chaînée, renforcée et de béton armé témoignent du besoin à se protéger face aux risques sismiques. Le risque sismique est fréquent et redouté par la population et le gouvernement. Au total plus de 20% des bâtiments respectent à des normes parasismiques dont 11,3% avec des normes élevées. Pour ces bâtiments, le risque de décès n'existe que pour les tremblements avec une très forte intensité.

Pour le Mexique, au vu de la base des données et de l'heure du tremblement de terre (13h en heure locale), il faut donc poser d'avantage d'hypothèses. A 13h, il est raisonnable de supposer qu'à cette heure la population mexicaine travaille (dans les écoles, bureaux, commerces, industries...) et donc le taux d'occupation du secteur résidentiel est faible. Une hypothèse réaliste pour le backtest est de répartir la population de chaque département, en plaçant 5% dans le secteur résidentiel et 95% dans le commercial et de l'industrie.

Tout d'abord, il est important de noter que le secteur commercial au sein des données d'exposition inclut les établissements commerciaux et les services. Grâce aux statistiques nationales de l'institut national de statistique et de géographie du Mexique (INEGI), il est possible de déterminer la capacité des actifs. Le nombre total d'établissements dans le secteur de l'industrie et du secteur commercial de la base de données du GEM (Global Earthquake Model) et les chiffres nationaux sur les unités économiques de ces différents secteurs sont quasiment similaires (Tableau 25).

Il est possible de déterminer la capacité d'un actif du secteur commercial en multipliant le nombre de bâtiments composant l'actif et le nombre d'employés moyen dans le secteur. Par exemple pour le secteur de l'industrie, selon les données nationales du Mexique en 2018, sur le Tableau 25, le nombre d'employés moyen dans ce secteur s'élève à 11,2. En faisant des hypothèses avec le nombre d'employés moyens, il est possible de remplir la base intégralement et de faire des simulations.

Activité économique	Nombre d'unités économiques		Nombre de personnes employées	
	Absolu	%	Absolu	%
Services	1 899 812	39,6%	10 844 080	40,0%
Commerce	2 294 932	47,8%	8 712 491	32,1%
Industrie	605 413	12,6%	7 576 306	27,9%
Total	4 800 157	100%	27 132 877	100%

Tableau 25 : Unités économiques au Mexique, statistiques nationales

4.2.3 Résultats et analyse de sensibilité

Résultats pour le séisme d'Amatrice (ITA) et pour d'autres séismes en Italie

Au sein du module aléa, le PGA maximal enregistré dans le village d'Amatrice est de 46%g, ce qui est important. Au niveau de l'exposition, 368 566 actifs immobiliers composés de 4 118 456 bâtiments sont à l'intérieur d'une zone où il a été enregistré des ondes sismiques. Pour seulement 11 605 bâtiments, un PGA au moins égale de 20%g a été enregistré. La probabilité décès la plus élevée appliquée aux occupants d'un actif immobilier s'élève $9,09 \cdot 10^{-3}$.

	Nombre	% par rapport au nombre total dans l'empreinte
Actifs avec un taux de décès moyen supérieur à 0	4 775	1,30%
Bâtiments avec un taux de décès moyen supérieur à 0	111 615	2,71%
Occupants exposés à un risque de décès	366 152	1,59%

Min	Q1	Mediane	Moyenne	Q3	Max
1,93E-06	5,00E-06	1,01E-05	1,22E-04	3,49E-05	2,27E-03

Tableau 26 : Résultats des taux de vulnérabilité appliqués - Amatrice

Le backtest n'est pas concluant pour le séisme, le résultat est de 79 morts dont 18 morts à Amatrice. Ce qui est très différent du nombre de victimes réel enregistré après le séisme qui est de 296 morts au total dont 234 morts à Amatrice.

Le modèle ne fonctionne pas pour ce séisme. La mortalité importante de ce séisme à Amatrice s'explique par un contexte particulier : la saison estivale (population multipliée par 10) et la localisation de principalement de bâtiments anciens sur la zone sismique.

Deux raisons qui sont reliées peuvent expliquer la faible capacité du modèle à reproduire le nombre de décès :

- Ne disposant d'aucune densité de population lors des vacances, il y a un risque de sous-estimation de l'exposition c'est-à-dire du nombre de personnes exposées sur l'empreinte aux secousses sismiques lors du backtest.
- Les courbes de vulnérabilité ne semblent pas assez sévères pour ce tremblement de terre. Par exemple, Amatrice est un petit village de 2400 habitants en temps normal, ainsi 234 morts représentent un taux de mortalité d'environ 10% des résidents de ce village. Or, aucune courbe de vulnérabilité n'atteint une probabilité de décès de 10% quel que soit le niveau de PGA.

Ces deux raisons s'amplifient : une exposition plus faible couplée à une sévérité moins élevée que la réalité se traduit automatiquement par un nombre de décès moins élevé.

Dans une zone où la densité de population est faible comme le centre de l'Italie, le modèle ne permet pas de simuler un nombre de décès adéquat. En effet, lorsque la densité de population est faible (nombre d'occupants et de bâtiments faibles sur une zone géographique), l'application de taux de vulnérabilité faibles ne permet pas en moyenne de simuler un nombre de décès important.

Sur le Tableau 27, afin de garder une exposition et des hypothèses réalistes, deux autres séismes qui se sont produits à des années proches de celui d'Amatrice (2016) sont simulés à l'aide du modèle physique. Il s'agit de deux séismes qui ont lieu dans des zones rurales mais le bilan de victimes est beaucoup plus faible que celui d'Amatrice. Le modèle n'est pas performant, la différence entre le nombre de victimes simulés et les chiffres réels est importante.

Tremblements de terre	Emilia Romagna 29/05/2012	Cittareale 18/01/2017
Caractéristiques	Magnitude : 5,5 Profondeur : 6,8 km Heure : 10:55	Magnitude : 5,7 Profondeur : 7 Heure : 10:14
Densité de population	199 hab/km ²	8,2 hab./km ²
Bilan officiel	17	10
Nombre de morts simulés	12	3
Différence (%)	-29,41%	-70,00%

Tableau 27 : Résultats du backtest de deux autres séismes - Italie

Malheureusement pour l'Italie, les courbes utilisées ne sont pas des courbes développées uniquement pour ce pays. Ce sont les courbes utilisées pour les simulations de tous les pays européens. Il faudrait pouvoir pour chaque simulation revoir les courbes de vulnérabilité, les déplacer en les rendant plus ou moins sévères selon la zone géographique. Ce qui est très compliqué dans le cas de milliers de simulations. La densité de population et les lieux où les séismes se sont produits ne facilitent pas la calibration du modèle.

Résultats pour le séisme de Puebla (MEX) et pour d'autres séismes au Mexique

Au sein du module aléa, sur l'empreinte, le PGA maximal enregistré est de 48%g près de l'épicentre et notamment près de la ville de Mexico. Au niveau de l'exposition, 166 539 actifs immobiliers composés de 13 865 960 bâtiments sont à l'intérieur d'une zone où il a été enregistré des ondes sismiques. Pour 15,56% de ces actifs ont été calculés un taux de décès moyen supérieur à 0. La probabilité décès la plus élevée appliquée aux occupants d'un actif immobilier s'élève $3,06 \cdot 10^{-3}$.

Pour le backtest, avec une hypothèse de 95% des occupants répartis dans le secteur commercial et industriel et 5% dans le secteur résidentiel lors du tremblement de terre à 13h en heure locale, le modèle permet de simuler 397 décès. Le nombre simulé est supérieur de n'est supérieur que de très peu (+7,59%) par rapport au bilan officiel de 369 morts pour ce séisme. Cette différence est négligeable et encourageante au vu de la complexité de la modélisation.

	Nombre	% par rapport au nombre total dans l’empreinte
Actifs avec un taux de décès moyen supérieur à 0	25 921	15,56%
Bâtiments avec un taux de décès moyen supérieur à 0	3 933 054	28,36%
Occupants exposés à un risque de décès	12 568 779	23,25%

Min	Q1	Mediane	Moyenne	Q3	Max
4,11E-08	2,00E-06	9,11E-06	7,72E-05	4,30E-05	3,06E-03

Tableau 28 : Résultat des taux de vulnérabilité appliqués – Puebla

Secteur	Nombre d'occupants exposé à un taux de décès non nul	Nombre de décès
Commercial	8 108 043	253
Residentiel	3 319 353	125
Industriel	1 141 382	19
Total	12 568 778	397

Tableau 29 : Résultats du backtest par secteur – Puebla

En analysant de plus près les résultats, les résultats trouvés sont proches de l’assurance non-vie, le secteur résidentiel semble plus vulnérable que le secteur commercial et industriel. Cela peut s’expliquer par la structure des bâtiments, le nombre de bâtiments non renforcés est plus important au sein du secteur résidentiel que celui de l’industriel ou du commercial qui comptent un nombre plus important de bâtiment en béton renforcé par exemple.

Villes (distance par rapport à l’épicentre)	Bilan officiel de victimes	Nbre de victimes simulées
Ciudad de Mexico (119 km)	228	106
Morelos (66 km)	74	97
Puebla (62 km)	45	162
Estado de México (60 km)	15	28
Guerrero (137 km)	6	1
Oaxaca (249 km)	1	1
Tlaxcala (91 km)	0	2
Total	369	397

Tableau 30 : Résultats du backtest par rapport au bilan officiel

Le modèle fonctionne comme sur le principe la loi des grands nombres (LGN), du fait des probabilités de décès très faibles : plus le nombre d’occupants touchés par les ondes sismiques est important, plus le modèle permettra, en moyenne, de simuler un nombre moyen proche de la réalité. Ainsi, en regardant, le nombre de décès à l’échelle d’un quartier ou d’une ville, le modèle sera peu précis. A l’échelle d’un pays, il réplique les grandes tendances.

Pour le Mexique, deux autres tremblements de terre ont été simulés avec le modèle physique. Il s'agit de deux tremblements qui se sont produits à des années proches de celui de Puebla (2017).

Tremblements de terre	Oaxaca 08/09/2017	Oaxaca 23/06/2020
Caractéristiques	Magnitude : 8,2 Profondeur : 47,4 km Heure : 04:49	Magnitude : 7,4 Profondeur : 20 Heure : 15:29
Densité de population	44,1/km ²	44,1/km ²
Bilan officiel	98	10
Nombre de morts simulés	106	11
Différence (%)	8,16%	10,00%

Tableau 31 : Résultats du backtest de deux autres séismes - Mexique

Il est impossible d'obtenir des expositions pour toutes les années afin de réaliser une multitude de backtests. Les résultats obtenus sur les 3 backtests démontrent une réelle capacité du modèle à répliquer la sévérité des séismes avec une différence de moins de 10% pour tous les séismes. Le modèle semble fonctionner pour le Mexique.

Une analyse de sensibilité permettra de mesurer l'impact d'un tremblement de terre la nuit sur le nombre de décès simulé et de l'effet des autres paramètres sur les résultats en sortie du modèle.

Analyse de sensibilité du modèle au Mexique (séisme de Puebla)

- Changement de l'heure du tremblement (séisme nocturne)

Le but de cette analyse de sensibilité est de savoir si le modèle est capable de capter la surmortalité des tremblements lors d'un séisme nocturne, les personnes dorment dans leurs logements (100% secteur résidentiel). Pour cela, il faut changer l'hypothèse de la répartition de la population au sein des différents secteurs. Il faut répartir toute la population dans le secteur résidentiel et dans les hôtels.

Le nombre d'occupants dans les logements la nuit est déjà une variable présente dans la base d'exposition, aucune hypothèse ne doit être faite. Pour être plus précis, il faut estimer le nombre d'hôtels et d'occupants dans ces établissements. Pour déterminer les hôtels dans le secteur commercial, une approche par valeur économique peut être utilisée.

Comme les hôtels ne sont pas référencés dans notre base, il faut dans un premier temps rechercher les hôtels dans les actifs du secteur commercial de la base d'exposition.

Pour déterminer le nombre d'hôtels dans la base d'exposition, il faut pour cela utiliser une base de données sur laquelle il est possible de retrouver des informations complètes sur le type de bâtiments au Mexique. Les autres bases à disposition chez AXA sont la base RMS ou le portefeuille client d'AXA au Mexique (P&C). La base RMS ne donne aucune information sur le type d'établissement, la base cliente d'AXA géocodée est donc retenue. Il est important de souligner que cette proposition repose sur l'hypothèse que le portefeuille d'AXA est un bon proxy du portefeuille marché du GEM.

En effet, la base cliente d'AXA géocodée est une base de données confidentielle présentant des informations intéressantes comme la localisation des bâtiments (longitude et latitude), la structure des bâtiments, le nombre d'étages, le secteur (RES/COM/IND), la valeur économique du bien, la valeur assurée, le type d'établissement (hôtels, parking, école, usine d'assemblage de matériaux lourds...).

Pour estimer le type d'établissement par la valeur économique des bâtiments, il faut calculer la valeur économique moyenne d'un étage pour chaque type d'établissement de la base d'AXA ; puis tenter d'approximer le type d'établissement en comparant ces valeurs moyennes aux valeurs économiques par étage de chaque bâtiment de la base d'exposition marché.

Sur la base d'AXA, les hôtels sont le type d'établissement commercial le plus cher avec une valeur économique de plus d'un million d'euros par étage. Cela s'explique par la structure des hôtels qui sont composés d'un nombre important de chambres et d'équipements pour satisfaire au mieux les clients (par exemple pour un hôtel de luxe : spa, restaurant, etc...), ce qui fait augmenter la valeur économique du bien assuré.

Au final, cette technique permet de placer 22 846 hôtels dans la base marché (contre 22 560 au sein des statistiques nationales). Elle fonctionne bien, les hôtels d'une part par leur structure externe et équipements à l'intérieur (contents) sont facilement repérables, leurs valeurs économiques sont importantes.

Après avoir repéré les hôtels dans l'empreinte du séisme, il s'agit de faire une hypothèse sur le nombre d'occupants dans les hôtels grâce aux statistiques nationales mexicaines. Selon les résultats de l'activité du tourisme au Mexique (Horwath HTL (2020)), le taux d'occupation moyen au Mexique des hôtels sur l'année est de 60.9% et le nombre moyen de chambres en moyenne au Mexique retenue est de 35.

	Nombre de morts simulés
Séisme (13h)	397
Séisme Nocturne	710
Différence (%)	78,84%

Tableau 32 : Analyse de sensibilité – séisme nocturne Puebla - Mexique

Le séisme nocturne provoque un nombre de décès supérieur de 78% par rapport à un séisme durant la journée. Cela s'explique par les structures différentes des bâtiments du secteur commercial, industriel et résidentiel.

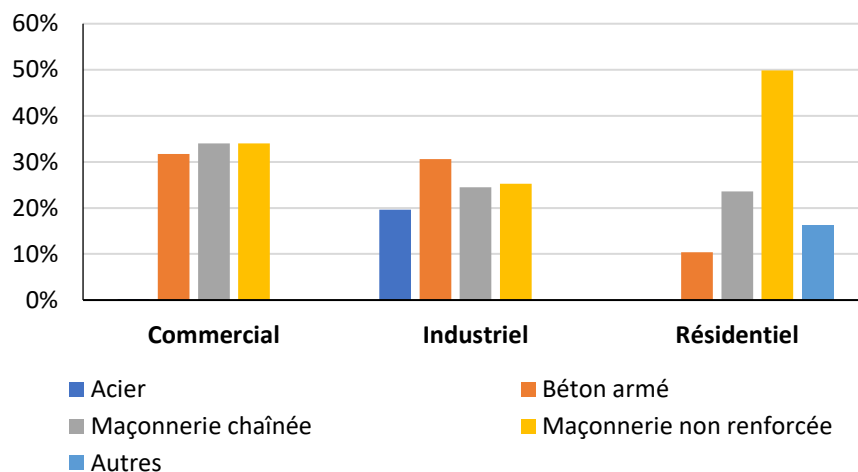


Figure 28 : Répartition par secteur des matériaux de construction des bâtiments

- Impact de l'hypothèse de répartition des occupants entre les différents secteurs

Résidentiel	100%	80%	60%	40%	20%	0%
Commercial et Industriel	0%	20%	40%	60%	80%	100%
Nombre de victimes simulé	710	688	656	607	521	329

Tableau 33 : Analyse de sensibilité Puebla – Répartition des occupants entre les différents secteurs

Comme le montre le Tableau 33, il est extrêmement important de prendre en considération la répartition des occupants. Puisque les bâtiments des différents secteurs (commercial, industriel, résidentiel) ne sont pas construits de la même façon. Plus généralement, plus le niveau de vie de la population est faible, plus les matériaux de construction des habitations sont basiques et moins solides face aux tremblements de terre.

En effet, du fait du niveau de vie peu élevé au Mexique, une bonne partie des habitants n’ont pas les moyens d’habiter dans des logements respectant des normes sismiques élevées ou vivre dans des bâtiments dont les matériaux de construction sont très résistants aux ondes. Selon la Figure 28, 50% des bâtiments du secteur résidentiel sont construits en maçonnerie traditionnelle. Le secteur industriel et commercial sont composés à 30% de bâtiments en béton armé alors que seulement 10% de bâtiments du secteur résidentiel sont en béton armé. Les occupants sont donc plus en sécurité sur leur lieu de travail ou dans les commerces. Finalement comme en assurance non-vie, le secteur résidentiel est plus vulnérable que les autres secteurs.

- Impact des normes de construction sur la sévérité des tremblements de terre

Normes de construction	Situation initiale	1ère hypothèse	2ème hypothèse
Aucune norme	72,2%	100%	0%
Faibles	22,8%	0%	100%
Elevées	5,0%	0%	0%
Nombre de victimes simulé	397	439	319

Tableau 34 : Analyse de sensibilité Puebla – Normes de construction

Les résultats obtenus sur le Tableau 34 sont cohérents, plus les normes de constructions sont mises en place, plus le nombre de décès lors des catastrophes diminuent.

L’analyse de sensibilité démontre que les données prélevées pour construire l’exposition notamment les normes de construction doivent être précises afin d’obtenir une sévérité proche de la réalité

Conclusions sur le modèle

Le modèle physique permet donc d’inclure dans la simulation en plus de l’aléa (événements physiques) des facteurs d’exposition très précis (structure des bâtiments, etc..) venant impacter considérablement la vulnérabilité, la sévérité de chaque événement. Plus les hypothèses et les données seront précises sur les assurés ou occupants des bâtiments, plus le modèle sera performant.

Ce modèle fonctionne comme sur la loi des grands nombres puisque les probabilités de décès exercées sont très faibles. La densité de population et le nombre d’occupants à l’intérieur de l’empreinte font varier la sévérité. Le modèle a tendance à moins bien fonctionner à l’échelle d’un quartier ou d’un petit village contrairement à de grandes villes ou à un pays densément peuplé.

Avec les données actuelles, l’analyse et les backtests au Mexique prouvent que le modèle est pertinent au Mexique par la construction des courbes de vulnérabilité qui semblent convenir très bien à la réalité. L’analyse de sensibilité a mis en évidence des mécanismes rencontrés en assurance non-vie et l’importance de la prise en considération de tous les facteurs physiques lors des simulations.

Ce modèle peut être considéré comme plus performant que les modèles statistiques par sa construction et ses avantages. A l'avenir, plus les données seront enrichies et précises sur les assurés, leurs lieux de travail et leurs habitations, plus le modèle sera robuste et répliquera au mieux la réalité. AXA dispose des données nécessaires pour effectuer une première simulation de tarification.

4.3 Tarification du risque tremblement de terre dans un traité CAT XL Vie avec les données disponibles

À la suite des conclusions satisfaisantes de la partie précédente pour le Mexique, il est pertinent de calculer une première prime pure à l'aide du modèle physique. Pour cela, il est présenté une méthodologie adaptée aux données accessibles à AXA Global Re sur le portefeuille au point 4.3.1. Puis la prime pure est calculée et comparée à celle obtenue lors de l'utilisation du modèle statistique pour tarifier le risque tremblement de terre au Mexique.

4.3.1 Le module financier et hypothèses pour calculer une prime pure

Méthodologie

Ne disposant d'aucune information sur la localisation des assurés. Le modèle physique notamment la partie exposition fonctionne en vision marché. Il est donc indispensable d'utiliser les hypothèses faites pour le modèle statistique afin d'obtenir les résultats au niveau d'AXA (cédante).

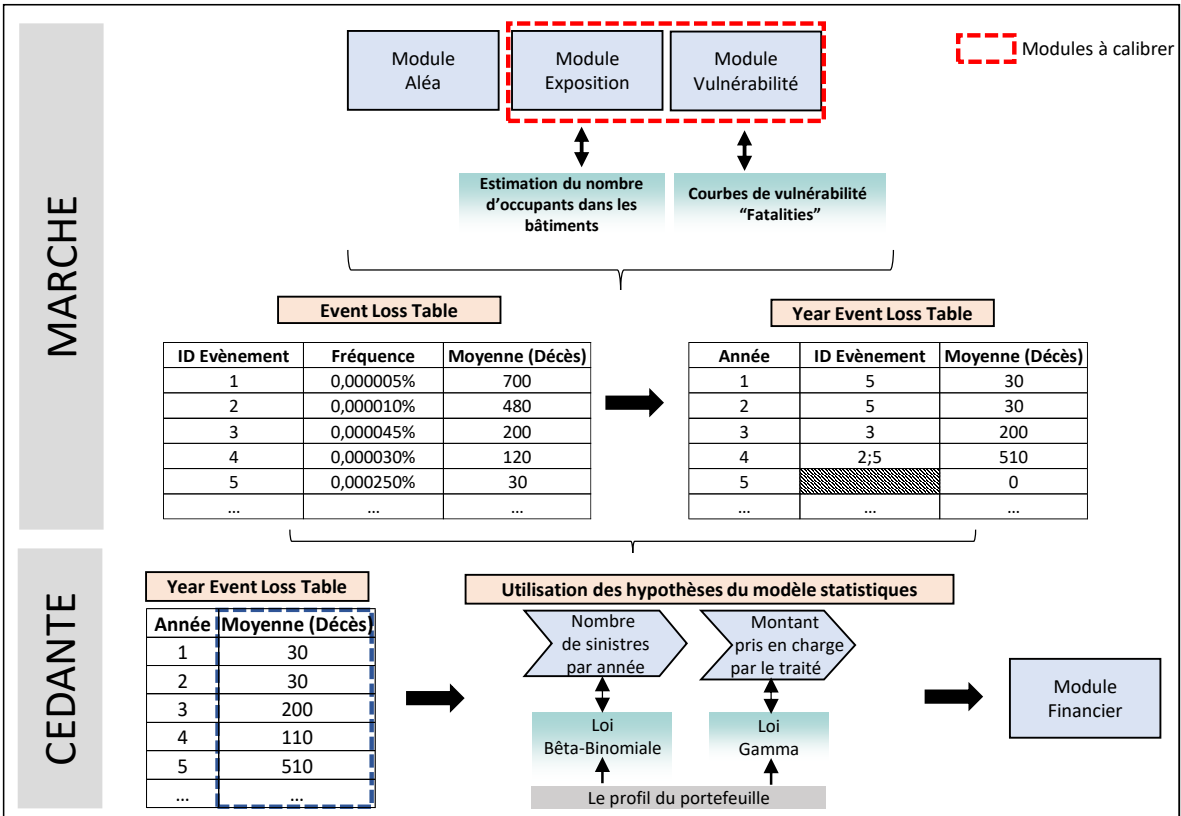


Figure 29 : Méthodologie pour la tarification d'un traité CAT XL par une approche physique

En effet, en vision marché, une fois avoir calibré les différents modules (Aléa, Exposition, Vulnérabilité), il faut évaluer le montant assuré à la charge du réassureur (vision cédante). Disposant de la puissance de calculs nécessaires et de l'expertise, AXA Climate effectue les simulations avec le catalogue d'évènements. L'output des simulations permet de construire un ELT (Event Loss Table).

L'ELT répertorie pour chaque évènement du catalogue le nombre de morts engendré au niveau du pays. L'ELT est aussi composé de la fréquence de chaque évènement. Avec l'ELT, il faut par la suite construire un YELT (Year Event Loss Table), le nombre de décès pour une année donnée.

L'ELT et YELT obtenus, ici correspondent à des tables au niveau marché, il s'agit du nombre de morts à l'échelle du pays et non le nombre de victimes assurées par AXA lors des tremblements de terre.

Il faut alors ajouter des hypothèses notamment le nombre d'assurés décédés sur une année donnée et la somme assurée associée pour ces décès avant d'appliquer les conditions de réassurance.

A l'heure actuelle, il a été choisi de reprendre les hypothèses du modèle statistique en utilisant une loi bêta-binomiale pour évaluer le nombre de sinistres (assurés décédés) et une distribution exponentielle pour approximer les pertes assurance associées aux différents sinistres.

Comme le montre la Figure 29, la méthodologie mise en place inclut la partie Aléa (le catalogue d'évènements) qui est constituée d'un nombre d'évènements très important, la partie Exposition (base d'exposition incluant une dimension humaine avec le nombre d'occupants au sein des bâtiments) et la partie Vulnérabilité (courbes de vulnérabilité). Les simulations d'évènements ne sont plus réalisées par la calibration d'une fréquence et sévérité issues d'une base de données historiques mais par le modèle physique, les parties conservées du modèle statistique sont celles permettant d'insérer le profil du portefeuille et de calculer le montant des sinistres.

Passage d'ELT à YELT

La forme des ELT (Event Loss Table) utilisés sont basés sur celles du modèle RMS. Ces tables indiquent pour chaque évènement numéroté par un identifiant, la fréquence, la moyenne du nombre de morts, l'écart type du nombre simulé et l'exposition (la population au sein de l'empreinte).

ID Evènement	Fréquence	Moyenne	Ecart-type	Exposition (Pop. dans l'empreinte)
1	0,000005%	700	110	2 800 000
2	0,000010%	480	80	17 000 000
3	0,000045%	200	98	18 000 000
4	0,000030%	120	40	1 000 000
5	0,000250%	30	8	800 000
...

Tableau 35 : ELT fictif

Il faut transformer cette table en YELT (Year Event Loss Table) permettant d'avoir une vision sur une année donnée du nombre d'évènements réalisés (provenant de l'ELT) et du nombre de décès en moyenne engendré par les catastrophes. Dans le cadre d'une tarification, c'est cette vision par année qui est intéressante puisqu'un traité de réassurance couvre les pertes sur une année donnée.

Année	ID Evènement	Moyenne (Décès)
1	5	30
2	5	30
3	3	200
4	2;5	510
5		0
...

Tableau 36 : YELT fictif

Comme écrit sur le Tableau 36, il est possible que pour une année donnée, aucun évènement se produise. Il est aussi possible au contraire que plusieurs catastrophes se produisent la même année.

Afin de passer d'une ELT à une YELT, il faut dans un premier temps évaluer le nombre d'évènements pour une année donnée (fréquence).

Soient : $\lambda = \sum Freq_i$ et N le nombre d'évènements pour une année donnée.

$N \sim Poisson(\lambda)$, N suit une loi de Poisson de paramètre λ .

Dans un deuxième temps, il faut déterminer lequel des évènements de l'ELT ou laquelle des lignes de l'ELT se réalisera. La probabilité de réalisation d'un évènement de l'ELT dépend de son taux de fréquence $Freq_i$.

Soient : C une variable aléatoire représentant la réalisation ou non d'un évènement au sein de l'ELT et $U \sim Uniforme(0,1)$ la loi uniforme.

Ainsi :

$$C = \min \left\{ c : U \leq \sum_{i=1}^r \frac{Freq_i}{\lambda} \right\}$$

Il s'agit d'un tirage aléatoire réalisé grâce une loi uniforme.

Puis, il s'agit d'agréger la moyenne des décès des évènements réalisés pour obtenir l'YELT désiré.

4.3.2 Application du modèle pour un traité au Mexique

Le Mexique est un pays où le risque de tremblement de terre est important. Il est intéressant d'évaluer la faisabilité et l'utilité du modèle physique pour ce pays. Dans le cadre de cet exemple, il s'agit de déterminer si le modèle physique permet de mieux analyser le risque et pourrait justifier le coût important de sa mise en place.

Indicateurs essentiels de l'analyse d'un traité

Pour un traité donné, l'équipe « Analytics & Pricing » d'AXA Global Re a pour mission d'effectuer une analyse du risque CAT associé au traité. Cette analyse devrait faciliter la prise de décision de souscrire à un traité ou de le céder à un autre réassureur. Généralement, la souscription à un traité dépend fortement de sa rentabilité.

Pour évaluer cette rentabilité, il est utilisé couramment des indicateurs essentiels en réassurance qui sont obtenus grâce aux simulations : la prime pure (ou l'espérance des récupérations), la volatilité des récupérations, le RoL (*Rate on Line*) et la période de retour.

Le RoL pur se définit comme le rapport entre la prime pure et la portée du traité. Le RoL reflète le coût d'achat de la couverture.

La période de retour se définit comme le rapport entre le nombre de récupérations et le nombre de simulations effectuées. Il s'agit de la durée moyenne pour qu'un évènement CAT entraîne une récupération.

Chargements, prime commerciale et coût du capital

Il est possible d'estimer une prime commerciale pour le traité. Pour calculer une prime commerciale, il faut ajouter à la prime pure des chargements : un chargement de sécurité prenant en compte la volatilité des récupérations, la rémunération du capital en plus des frais de gestion.

Au sein du modèle interne d'AXA Global Re, le SCR CAT est égal à la variation de fonds propres liée à une déviation de la sinistralité à 200 ans moins la moyenne des pertes.

Pour prendre en compte le coût du capital lié à un traité, au sein de la tarification (prime commerciale), il faut par exemple utiliser une hypothèse de coût du capital (CoC) de 8% et calculer le SCR en faisant la différence entre le quantile à 99,5% sur la distribution des pertes simulées et la moyenne des pertes.

RoL commercial et RoL minimum

Avec une prime commerciale, il est tout à fait possible aussi de calculer un RoL commercial se définissant comme le rapport entre la prime commerciale et la portée du traité.

Lorsque le RoL commercial obtenu est faible, il est utilisé généralement un RoL minimum qui se définit comme le Rate on Line au-dessous duquel le réassureur ne serait pas prêt à souscrire le traité. Ce RoL minimum est habituellement compris entre 0,5% et 1% en réassurance vie.

Estimation du coefficient de chargement et formule de tarification

Les formules de tarification des réassureurs sont confidentielles. Le principe de prime par écart-type est une mesure de risque simple souvent utilisée pour calculer la prime commerciale. Ce principe de prime se définit par la formule suivante : $P(X) = E[X] + \beta * \sigma[X]$

En réassurance avec la prise en compte des frais, cette formule peut s'écrire de la façon suivante :

$$Prime\ commerciale = \frac{Prime\ Pure + Coeff.\ Chargement * Volatilité}{(1 - Taux\ de\ frais\ total)}$$

La prime pure est définie comme la moyenne des récupérations et la volatilité est usuellement définie en réassurance comme l'écart-type des récupérations. Le taux de frais total inclut les frais de gestion du réassureur et les frais de courtage. Les frais de gestion sont généralement exprimés en pourcentage de la prime totale et permettent au réassureur de payer ses salariés, couvrir les frais de l'entreprise, etc.. Les frais de courtage permettent de rémunérer les courtiers pour le placement des traités.

Au sein de AXA Global Re, grâce aux études en interne, il est considéré que les frais de gestion d'un réassureur sont en moyenne de 4% et les frais de courtage de 10%.

Il reste ainsi à définir le coefficient de chargement pour pouvoir calculer la prime commerciale. En réassurance, il est communément utilisé un seul coefficient de chargement qui englobe l'ensemble des chargements restants (notamment le chargement de sécurité et le coût du capital). La problématique dans la détermination du coefficient de chargement est souvent délicate.

En effet, chaque réassureur a sa propre vision du risque mais aussi des frais, des marges commerciales et des besoins en capital différents. Le coefficient de chargement est donc fixé de façon souveraine par chaque réassureur selon la branche d'assurance et la cédante concernée.

Pour estimer le coefficient de chargement, des cotations des réassureurs pour les traités CAT XL Vie en Amérique Latine ont été utilisées. Une cédante comme AXA peut estimer le coefficient de chargement qui est appliqué par un réassureur pour un traité dont la prime commerciale est connue en utilisant la formule ci-dessous :

$$\text{Coeff. Chargement} = \frac{\text{Prime commerciale} * (1 - \text{Taux de frais total}) - \text{Prime Pure}}{\text{Volatilité}}$$

Grâce à un traité benchmark (traité ayant des caractéristiques similaires au traité de l'application), il a été considéré comme raisonnable d'appliquer un coefficient de chargement de 10% au sein de ce mémoire.

Ce coefficient de chargement provenant d'une cotation ne peut pas être décomposé, il englobe des chargements. L'utilisation du principe de la prime par écart-type et du coefficient de chargement au sein de l'application permet simplement d'estimer la prime commerciale du traité en gardant les mêmes hypothèses de tarification d'un réassureur qui a tarifé le traité benchmark.

Hypothèses sur le traité CAT au Mexique

Après avoir présenté, les indicateurs nécessaires au comité de souscription d'AXA Global Re, il s'agit de comparer les résultats obtenus avec le modèle statistique et le modèle physique pour un traité au Mexique. Les hypothèses retenues pour ce traité sont les suivantes :

Contrat groupe			Contrat individuel		
Somme assurée (SAR)	Nombre d'assurés	Somme assurée moyenne	Somme assurée (SAR)	Nombre d'assurés	Somme assurée moyenne
900 000 000	5 000	180 000	400 000 000	4 000	100 000

Traité	Min. victimes	Priorité	Portée	Nb de reconstitution	AAL
CAT-MEX	3	5 000 000	50 000 000	1	100 000 000

Tableau 37 : Hypothèses pour la tarification du traité CAT XL Mexique

Analyse du traité avec le modèle statistique

Afin d'utiliser le modèle statistique pour tarifier le risque tremblement de terre au Mexique, il faut uniquement s'intéresser aux séismes dans la base EM-DAT (la base historique mondiale des catastrophes). Entre 1970 et 2020, au sein de la base EM-DAT, au Mexique, il y a 32 tremblements dont 19 provoquant le décès d'au moins 3 personnes. La calibration de la sévérité et de la fréquence se font donc sur ces 19 événements.

Au niveau, de la fréquence, l'espérance s'élève à 0,37 et la variance à 0,44. Comme la variance est supérieure à l'espérance, la loi Binomiale Négative est retenue, les paramètres de la distribution sont : $n = 1,58$ et $p = 0,37$.

Pour la sévérité, le choix d'une GPD de paramètres : $\xi = 1,99$ et $\beta = 9,14$ est pertinente selon les critères d'information et les tests (Kolmogorov-Smirnov). Ici, ξ est supérieur à 1 donc la GPD n'admet pas d'espérance. Cependant, la portée du traité de 50 000 000€ vient borner le risque, la prime pure (moyenne des récupérations) admet donc une espérance.

Comme le montre la Figure 30, après 1 million de simulations, la prime pure reste stable à 1 190 euros.

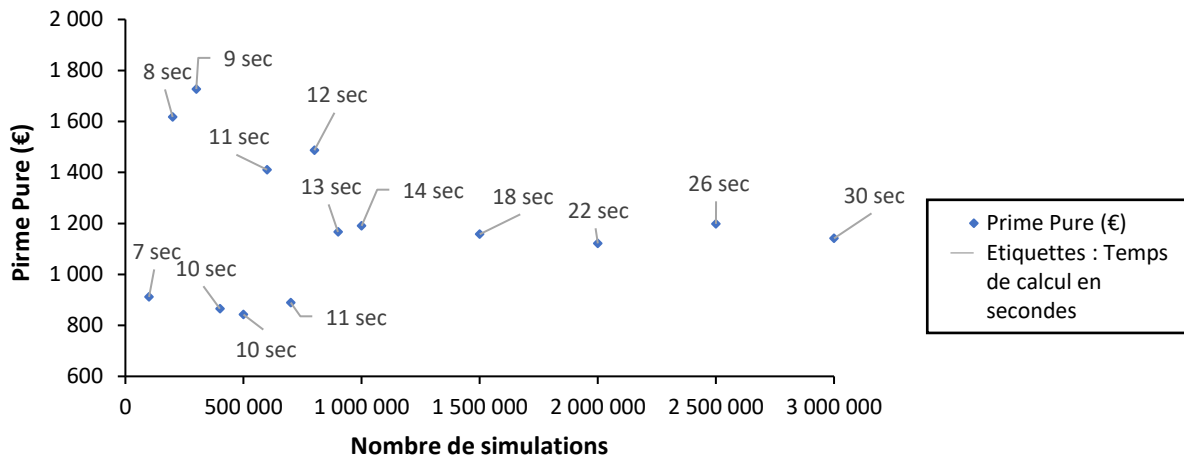


Figure 30 : Evolution de la prime pure et du temps de calcul avec le nombre de simulations

En réassurance, la volatilité est généralement définie comme l'écart-type des récupérations. Après 1 million de simulations, elle s'élève à 183 300€.

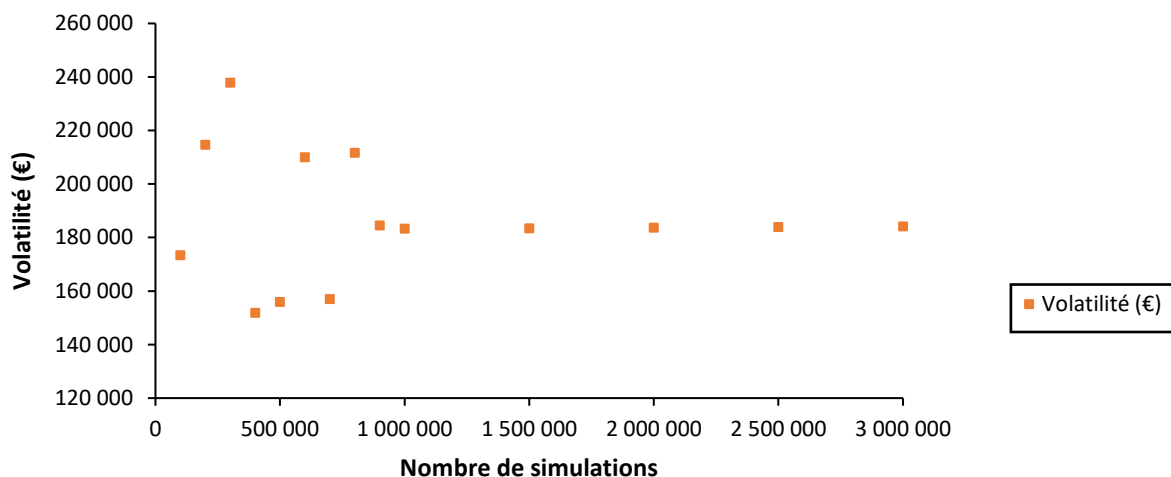


Figure 31 : Evolution de la volatilité avec le nombre de simulations

Seulement 86 scénarios sur 1M de simulations dépassent la priorité du traité. La période de retour s'élève à 11 628 années.

Vision Cédante	Nombre de simulations provoquant une récupération	Période de retour	Prime pure
	86	11 628	1 190

Tableau 38 : Tarification du risque Tremblement de terre (MEX) avec le modèle statistique

Une tarification commerciale incluant les chargements donne le résultat suivant :

$$\begin{aligned}
 \text{Prime commerciale} &= \frac{\text{Prime Pure} + \text{Coeff. Chargement} * \text{Volatilité}}{(1 - \text{Taux de frais total})} \\
 &= \frac{1\,190 + 10\% * 183\,300}{(1 - 14\%)} = 22\,698\text{€}
 \end{aligned}$$

Dans le cas présent, la volatilité est élevée par rapport à la prime pure donc le choix du coefficient de chargement détermine quasiment la valeur de la prime commerciale.

Ainsi :

$$\begin{aligned}
 \text{RoL commercial} &= \frac{\text{Prime commerciale}}{\text{Portée}} \\
 &= \frac{22\,698}{50\,000\,000} = 0,045\%
 \end{aligned}$$

Le RoL commercial est faible et inférieur à 0,5% (le RoL minimum), ainsi la prime commerciale retenue s'élève finalement à 250 000€.

$$\text{Prime commerciale} = \text{Portée} * \text{RoL minimum} = 0,5\% * 50\,000\,000 = 250\,000\text{€}$$

Analyse du traité avec le modèle physique

Le modèle est basé sur le catalogue d'évènements développé par AXA comportant plus de 40 000 tremblements de terre de magnitude supérieure ou égale à 3. Les simulations sont effectuées par un expert d'AXA Climate, pour des raisons de confidentialité, seulement quelques statistiques descriptives sont présentées.

Tout d'abord, il s'agit d'estimer le nombre de décès que pourrait entraîner chaque tremblement de terre du catalogue avec leur fréquence.

Vision Marché	Event Loss Table	
	Nombre d'évènements	Nombre d'évènements causant le décès d'au moins 3 personnes
>40 000	>7 829	

Après avoir obtenu les résultats (fréquence, moyenne, écart-type) composant l'Event Loss Table. Il s'agit de simuler 1 millions d'années dont se produit 0 ou plusieurs tremblements pour constituer l'YELT (Year Event Loss Table).

Vision Marché	Year Event Loss Table	
	Nombre d'années simulées	Nombre d'années provoquant le décès d'au moins 3 personnes
1 000 000	147 586	

Après avoir construit l'YELT, il nous faut estimer le nombre de sinistres touchant le traité d'AXA, en appliquant les caractéristiques du profil du portefeuille.

Les hypothèses du traité sont les suivantes :

Nombre d'assurés	Pop. Mexique	Taux de pénétration AXA
9 000	128 900 000	0,0069822%

Pour cela, comme dans le modèle statistique, il faut utiliser la loi bêta-binomiale afin d'estimer le nombre d'assurés décédés chaque année et une loi gamma pour estimer le montant des rétentions après application des conditions de réassurance.

Vision Cédante	Nombre de simulations provoquant une récupération	Période de retour	Prime pure
	127	7 874	1 680

Tableau 39 : Tarification du risque Tremblement de terre (MEX) avec le modèle physique

Ainsi, 127 simulations provoquent des pertes couvertes par le réassureur. La prime pure s'élève à 1680€ et la période de retour à 7 874 ans.

La volatilité obtenue lors de modélisation s'élève à 224 800, une tarification commerciale incluant les chargements donne le résultat suivant :

$$\text{Prime commerciale} = \frac{1\,680 + 10\% * 224\,800}{(1 - 14\%)} = 28\,093\text{€}$$

Ainsi :

$$\text{RoL commerciale} = \frac{28\,093}{50\,000\,000} = 0,056\%$$

Le RoL commercial est faible et inférieur à 0,5% (le RoL minimum), ainsi la prime commerciale retenue s'élève comme précédemment à : 250 000€.

Conclusions sur l'application

Autour de cet exemple sur le péril tremblement de terre au Mexique, l'utilisation du modèle a mis en évidence qu'il est possible de mettre en place le modèle physique en réassurance vie. Le catalogue physique permet d'utiliser une base comprenant plus de 40 000 séismes alors que la base EM-DAT répertorie seulement 32 séismes pour le Mexique.

L'augmentation de la prime pure avec le modèle physique reste néanmoins très faible en montant (+490€) pour ce traité. Cette augmentation de 41% de la prime pure s'explique principalement par une occurrence plus importante des sinistres dans le modèle physique.

La prime commerciale estimée reste la même avec les deux modèles. Le choix d'AXA Global Re de souscrire ou de céder le traité n'est pas influencé par une augmentation de la prime pure de 490€. Le modèle n'est pas utile à AXA Global Re pour ce cas particulier.

Il convient de constater que la simulation des scénarios possibles dans le futur mais inexistant dans une base historique avec un catalogue d'événements permet de calculer une prime pure plus juste, avec la réserve qu'il s'agit d'un modèle complexe à mettre en place.

Le modèle physique exige des moyens techniques, humains et financiers importants. Actuellement, en prenant en considération ces éléments, ce modèle physique pourrait avoir un intérêt pour les périls et zones géographiques dont les observations historiques sont très faibles (inférieur à 10 par exemple) ne permettant pas de calibrer une loi statistique pour la fréquence et la sévérité dans le modèle actuel d'AXA Global Re.

Conclusion

Pour mettre en place le modèle de simulations de catastrophes naturelles, il est possible de s'appuyer sur des données marché et des données de l'assurance non-vie. Cependant, les données actuelles récupérées par un réassureur en réassurance vie sont trop synthétiques. Le présent mémoire s'est ainsi attaché à poser des hypothèses afin d'obtenir une méthodologie de tarification complète basée sur des simulations de catastrophes naturelles à l'aide d'un modèle physique.

Pour justifier l'utilisation de ce modèle en réassurance vie, la dépendance entre le risque vie et le risque non-vie liée aux catastrophes naturelles a été mesurée. La corrélation entre les deux modules CAT s'élève à près de 50% et est non linéaire. Il s'agit d'une corrélation élevée et marquée au niveau des événements les plus extrêmes entraînant les dégâts matériels et le nombre de victimes les plus importants. Les tremblements de terre présentent la plus forte dépendance.

Ne disposant pas des données sur les taux de pénétration entre 1970 et 2020 pour tous les pays présents dans la base de catastrophes internationales (EM-DAT), la corrélation mesurée au sein de cette étude ne prend pas en compte ce facteur. Généralement, le taux de pénétration de l'assurance est généralement plus important dans les pays où le niveau de vie est élevé contrairement aux pays en développement où le taux de pénétration est plus faible. Il faudrait par exemple à l'avenir pouvoir récupérer les sommes assurées du marché de l'assurance vie et de l'assurance non-vie de tous les pays de la base EM-DAT pour mesurer plus précisément la dépendance.

Le modèle physique, en s'appuyant sur un nombre beaucoup plus large de simulations, apparaît constituer le seul moyen de capturer correctement cette corrélation en prenant en compte des facteurs comme la densité de population, l'intensité physique de l'évènement ou le niveau de vie à travers la structure des bâtiments (matériaux et normes de construction). La simulation des catastrophes naturelles à l'aide du modèle physique permet de plus de ne pas sous-estimer la survenance de catastrophes majeures au niveau des queues de distribution.

Les résultats du backtest du modèle physique sur le risque tremblement de terre en Italie et au Mexique ont démontré que le modèle s'adapte mieux aux pays où la densité de population est importante sur les zones touchées. En effet, par la structure des courbes de vulnérabilité, le modèle fonctionne comme sur le principe de la loi des grands nombres. Les probabilités de décès appliquées sont très faibles en général et pour ce motif, il faut un nombre important de personnes exposées aux ondes sismiques pour obtenir une estimation du nombre de victimes proche de la réalité.

Pour le risque tremblement de terre, une première tarification avec le modèle sur un traité au Mexique a été effectuée. L'expertise et la connaissance des scientifiques sur l'aléa physique permettent de générer un catalogue d'évènements stochastiques générant des dizaines de milliers d'évènements. Le résultat de la tarification démontre ainsi une capacité du modèle à tarifier le traité en prenant en compte un nombre beaucoup plus important de catastrophes que sur une base de données historiques. Les événements les plus extrêmes sont comptabilisés et augmentent la prime pure par rapport au modèle statistique classique.

Le modèle physique pour ces raisons présente de nombreux avantages, il permet de prendre en compte des événements inconnus dans l'historique d'évènements disponible chez un assureur et apporte un degré de précision plus fin sur l'évaluation de la vulnérabilité et de l'exposition au risque. Pour un réassureur comme AXA Global Re, il est nécessaire de disposer de la meilleure évaluation possible du risque catastrophe attaché à un traité avant de faire le choix de souscrire ou non à un traité.

Cependant, le modèle physique présente certaines limites. Il est surtout complexe à mettre en place et demande des données précises pour fonctionner (catalogue d'évènements, courbes de vulnérabilité, base d'exposition dynamique avec une dimension humaine...). Pour chaque pays et pour chaque péril, il faudrait construire et calibrer le modèle, ce qui demande des moyens importants pour un réassureur couvrant de nombreuses cédantes dans le monde comme AXA Global Re.

Il faudrait de ce fait continuer à approfondir cette démarche et recueillir notamment des données précises sur les portefeuilles. Par exemple, acquérir des données en termes de nombres d'assurés ou de somme assurée par code postal permettrait ainsi d'implémenter un modèle complet pour effectuer des simulations dépassant en précision les hypothèses du modèle de référence (modèle de Ekhedden).

En définitive, ce modèle apparaît comme une réelle alternative au modèle statistique de calibration de sévérité et fréquence recourant à une base historique pour la tarification. Ce modèle de tarification par sa structure pourrait être de plus appliqué en Risk Management et notamment au sein d'un modèle interne par exemple pour évaluer le SCR CAT Vie d'un assureur. L'une des caractéristiques du modèle physique est d'obtenir une distribution des pertes. En se basant, sur les scénarios construits, il est donc possible d'avoir une estimation de la surmortalité provoquée par un scénario atypique ou du risque engendré par un scénario bicentenaire.

Les différentes recherches du présent mémoire constituent les prémisses d'un éclairage nouveau permettant de mieux comprendre et évaluer le risque, en conservant à l'esprit les difficultés techniques de leur évaluation. La mise en place d'un modèle physique en réassurance vie est faisable, cependant, les apports de ce modèle à AXA Global Re à l'heure actuelle restent insuffisants pour justifier les moyens nécessaires pour sa mise en place.

Bibliographie

CREDIT MUTUEL (2020), *Rapport sur la solvabilité et situation financière*

CCR (2016), *Séisme du 24 août 2016 en Italie, Service R&D modélisation*

CEIOPS (2010), *Advice for Level 2 Implementing Measures on Solvency II: SCR STANDARD FORMULA*, Article 111(d) Correlations

CRED et UNDRR (2020), *Human cost of disasters, an overview of the last 20 years 2000-2019*

Christian Y. ROBERT (2020), *Théorie du risque*, ENSAE

E. EKHEDEEN (2008), *The Pricing of Catastrophe Cover in Life Reinsurance*, Mathematical Statistics, Stockholm University

GEM (2013), *Glossary for the GEM Building Taxonomy*, Technical report

GEM (2015), *Existing Empirical Fragility and Vulnerability Relationships: Compendium and Guide for Selection*, Vulnerability and loss modelling

Horwath HTL (2018), *Hotels & Chains in Italy 2018*, The Report

Horwath HTL (2020), *Market report, Mexico*

INEGI (2019), *Personal ocupado subcontratado en las unidades económicas*, Censos económicos

ISTAT (2017), *Hotels and similar establishments, net occupancy rate of bed-places*, Database

Jérôme CRÉTIEN (2019), *Enterprise Risk Management*, Cours

K. JAIWAL & D. WALD (2020), *An Empirical Model for Global Earthquake Fatality Estimation*, USGS, Research Article

N. LAMARI MFITH (2010), *Le risque Evénement en Réassurance Vie*

J. MESSIAS (2017), *Etude de la couverture de réassurance du pool catastrophe du Bureau Commun des Assurances Collectives (BCAC)*

D. NGUYEN (2017), *Reinsurance management in a big insurance group: Risk-costing and optimization*

A. PELLERIN (2019), *Une modélisation du risque sous-jacent aux traités CAT XL par évènement en réassurance vie*

USGS, *Earthquake Hazard programs, ShakeMap*

SWISS RE INSTITUTE (2019), *Catastrophes naturelles et techniques en 2018: les périls « secondaires » en première ligne*

Annexes

1- Le modèle de Strickler

Le modèle repose sur deux caractéristiques : la fréquence des évènements causant le décès d'au moins M morts et la distribution du montant des sinistres assurés.

Pour calibrer le modèle, Strickler a utilisé des données issues du *Statistical Bulletin of the Metropolitan Life Insurance Company* à New-York. Ces données correspondent à des accidents avec cinq ou plus de victimes décédées entre 1946 et 1950.

Strickler définit la fonction ci-dessous :

$$A(n) = a * 100^{\frac{1}{n}} * x^{-b}$$

$A(n)$: pour une population d'un million d'individus, il s'agit du nombre de personnes décédant annuellement dans un évènement entraînant le décès de n ou plus de personnes.

Par exemple : $A(1)$ représente le nombre total de personnes décédées dans une catastrophe pour une année donnée.

Le paramètre a représente le taux de décès dans un évènement CAT de la population étudiée et le paramètre b décrit la queue de distribution de $A(n)$. Plus b est petit, plus le nombre de décès à la suite des évènements est grand.

Avec les données publiées du *Statistical Bulletin of the Metropolitan Life Insurance Company* à New-York, les paramètres retenus par Strickler sont : $a = 8$ et $b = 1/3$. Cette fonction permet d'estimer le nombre de victimes.

Le nombre annuel d'évènements causant un nombre exact de décès de n personnes pour une population d'un million d'habitants peut s'exprimer en fonction de $A(y)$:

$$H(n) = \frac{A(n) - A(n + 1)}{n}$$

D'où le nombre total d'évènements peut s'écrire :

$$\sum_{z=1}^{\infty} H(z)$$

Ainsi, la probabilité qu'un évènement cause le décès de n personnes exactement s'écrit :

$$h(n) = \frac{H(n)}{\sum_{z=1}^{\infty} H(z)}$$

Il faut noter que Strickler fait l'hypothèse conservatrice suivante : quand $n \rightarrow \infty$, $H(n) \rightarrow 0$ et $h(n) \rightarrow 0$; notamment dès que $n > 1500$ alors $H(n) \rightarrow 0$.

De plus, Strickler fait l'hypothèse que le coût d'un sinistre lié au décès d'un assuré suit une loi exponentielle de paramètre 1 multipliée par la somme assurée moyenne du portefeuille. Ce qui peut s'écrire avec les formules suivantes :

Soient :

- Y le montant de sinistre lié au décès d'un assuré
- \overline{SAR} la somme assurée moyenne au sein du portefeuille

$$Y \sim \overline{SAR} * Exp(1)$$

Ainsi, le décès de n assurés provoque un montant de pertes assurées qui suit une loi Gamma de paramètres : $\Gamma(n, 1)$. On a :

$$Y \sim \overline{SAR} * \Gamma(n, 1)$$

La densité de la variable aléatoire Y peut s'écrire :

$$w_n(y) = e^{-y} \frac{y^{n-1}}{(n-1)!}$$

Ainsi, toute la distribution des montants de sinistres engendrés par des événements causant le décès d'au moins M victimes sur lesquels le réassureur va prendre en charge une partie se note :

$$w(y) = \sum_{n=M}^{\infty} h(n)w_n(y)$$

Le modèle a pour l'avantage d'être simple et d'être le premier modèle de tarification pour le risque Vie en réassurance. Cependant, ce modèle n'est plus adapté aux données récentes, la fonction $A(n)$ a été calibrée et définie sur une analyse de données des USA datant des années 1950. Il n'existe aucune méthode pour mettre à jour cette fonction $A(n)$. L'hypothèse que le nombre de morts n ne peut excéder 1500 est beaucoup trop conservatrice.

2- Tableau - Corrélation par pays et par événements

Types d'évènement	ISO	Effectif (sup ou égale à 10)	Proportion dans la base	Corr Pearson	Corr Spearman
Tempête	USA	366	35,26%	78,09%	46,80%
	JPN	38	3,66%	46,45%	42,23%
	DEU	28	2,70%	18,85%	50,63%
	FRA	20	1,93%	83,98%	74,71%
	GBR	19	1,83%	69,51%	67,19%
	CHE	10	0,96%	97,87%	55,69%
Inondation	USA	52	5,01%	42,66%	62,17%
	GBR	12	1,16%	60,16%	27,16%
	FRA	10	0,96%	6,46%	15,15%
Tremblement de terre	JPN	17	1,64%	97,45%	26,18%
	USA	10	0,96%	52,35%	73,16%

3- Images - Destruction du village d'Amatrice par le tremblement de terre (24/08/2016)



Centre-ville d'Amatrice (Italie) avant et après le tremblement de terre (Source : CCR)