

**Mémoire présenté devant le CNAM pour l'obtention du Diplôme
du Master Droit Economie Gestion mention Actuariat et
l'admission à l'Institut des Actuaires
le 24 janvier 2022**

Par : Stella JOVET

Titre: Optimisation d'une politique de résiliations ciblées à l'aide d'un algorithme
génétique : application à un portefeuille ouragan mexicain

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

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

Présidente du Jury

Sandrine LEMERY

signatures

Entreprise:

Nom : AXA Global Re

Membres présents du jury de la filière

Nathanaël ABECERA

Olivier DESMETTRE

François WEISS

Directeur de mémoire en entreprise :

Nom : Jérôme CRÉTIEN

Signature :

Invité :

Nom :

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

Amina ABED

Emmanuel DUBREUIL

Alexandre HASSLER

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

Mots-clés : catastrophe naturelle, ouragan, réassurance, AEP, OEP, ratio combiné économique, résiliation, algorithme génétique.

Avec un ratio combiné économique de 113%, le portefeuille de polices ouragan mexicain étudié dans ce mémoire est déficitaire et les coûts de réassurance et du capital pèsent lourdement sur sa rentabilité. Dans ce contexte, il a été décidé d'effectuer une politique de *pruning*, c'est-à-dire des résiliations ciblées, pour optimiser la rentabilité globale du portefeuille.

Le portefeuille est modélisé à l'aide d'un modèle CAT ouragan, considéré comme complexe et consommateur en temps de calcul. L'utilisation actuelle de ce modèle est adaptée aux besoins du *Risk Management* – pour quantifier les besoins en capital – mais pas de la souscription : en raison des temps de calcul, il est impossible de lancer un grand nombre de fois le modèle. Sans outil pour pouvoir estimer l'impact des actions de *pruning*, l'entité n'est pas en mesure d'identifier les combinaisons de polices dont la résiliation permettrait de maximiser sa rentabilité.

Tout d'abord, ce mémoire développe une méthode permettant de quantifier l'impact d'un ensemble de résiliations sur le ratio combiné économique du portefeuille. Cette méthode permet de quantifier les effets non-proportionnels des coûts de réassurance et du capital sans avoir recours au modèle CAT.

Grâce à l'implémentation d'un algorithme génétique, ce mémoire permet la mise en place de solutions d'aide aux décisions de *pruning*. Cet algorithme, adapté aux problèmes d'optimisation combinatoire multi-objectif, permet d'identifier les scénarios de résiliation qui optimisent la rentabilité du portefeuille tout en minimisant le pourcentage de primes résiliées et en s'adaptant aux différentes contraintes opérationnelles.

Abstract

Keywords: natural catastrophes, hurricane, reinsurance, AEP, OEP, economic combined ratio, pruning, genetic algorithm.

The portfolio studied in this thesis is composed of Mexican property policies covering hurricane risk. The portfolio has an economic combined ratio of 113% and reinsurance and capital costs weigh heavily on its profitability.

In this context, it was decided to carry out pruning actions, i.e. targeted policy terminations. The overarching goal is to optimize the overall profitability of the portfolio.

The portfolio is modeled using a CAT hurricane model, which is considered complex and time-consuming. The current use of this model is adapted to Risk Management needs - to quantify the capital requirements - but not to underwriting needs: due to its computation time, the model can't be run numerous times. Without tools to estimate the impact of pruning actions, the entity is not able to identify the policies whose termination would maximize the portfolio profitability.

First, this thesis develops a method that allows quantifying the impact of pruning actions on the portfolio economic combined ratio. This method enables quantifying the non-proportional effects of reinsurance and capital costs without launching the CAT model.

This thesis develops solutions to support pruning decisions via a genetic algorithm. This algorithm, suited to multi-objective discrete optimization problems, helps to identify policy terminations that would optimize the profitability of the portfolio while minimizing terminated premiums and adapting to the different operational constraints.

Note de synthèse

Les méthodes d'estimation du ratio combiné économique et d'optimisation à l'aide d'un algorithme génétique ont été développées dans le cadre de ce mémoire.

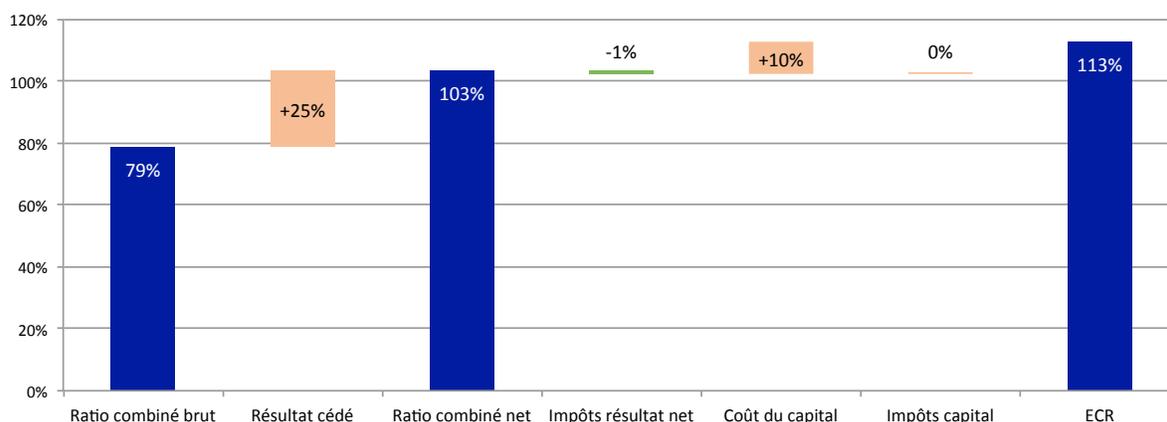
Contexte et problématique de l'étude

1.1. Le portefeuille étudié

Notre étude porte sur un portefeuille d'assurance dommage composé de polices couvrant le risque ouragan mexicain¹. On suppose que ce portefeuille a été souscrit par une entité du groupe AXA qui a par ailleurs des activités en assurance IARD et Vie.

Avec un ratio combiné économique de 113%, le portefeuille étudié n'est pas rentable : après rémunération des actionnaires, chaque euro de prime encaissé entraîne une détérioration des fonds propres de 13 centimes.

Le ratio combiné économique, aussi appelé ECR (*Economic Combined Ratio*), permet de quantifier la performance de la souscription et de piloter le résultat technique. Cet indicateur estime la charge de l'assureur pour un euro de prime souscrite : on souhaite donc qu'il soit le plus petit possible. Schématiquement, l'ECR se calcule comme la somme des sinistres bruts, coûts de réassurance, coût du capital, impôts et frais divisés par les primes brutes. Le graphique ci-dessous présente la décomposition de l'ECR du portefeuille étudié, modélisé avec un modèle catastrophe *inhouse*.



Pour un euro de prime encaissée, l'entité doit faire face aux coûts suivants :

- Le ratio combiné brut, c'est-à-dire les sinistres bruts et les frais d'exploitation, représente la plus grosse charge : 79 centimes.
- Le résultat cédé entraîne un coût de 25 centimes : le portefeuille est protégé à la hauteur de son événement bicentenaire par un traité de réassurance par événement.
- Le ratio combiné net est supérieur à 100% : l'entité bénéficie donc d'un crédit d'impôt de 1 centime.
- Le coût du capital représente une charge de 10 centimes. Il s'agit des coûts associés au rendement des fonds propres attendu par les actionnaires de l'entreprise.
- Il n'y a pas d'impôt sur les rendements du capital immobilisé car ce dernier est supposé placé sur des actifs peu risqués ayant une rentabilité nulle.

¹ Les polices d'assurance mexicaines ont la spécificité de couvrir le risque CAT à l'aide de polices dédiées. De ce fait, on dispose d'informations détaillées sur les primes ouragan du portefeuille.

² Solvabilité II est une directive européenne et ne s'applique donc pas au Mexique. Elle est utilisée dans ce mémoire. De ce fait, on dispose d'informations détaillées sur les primes ouragan du portefeuille.

1.2. Problématique

Avec un ratio combiné économique de 113%, le portefeuille étudié dans ce mémoire est déficitaire et les coûts de réassurance et du capital pèsent lourdement sur sa rentabilité. Dans ce contexte, il a été décidé d'effectuer une politique de *pruning*, c'est-à-dire des résiliations ciblées, pour optimiser la rentabilité globale du portefeuille.

La distribution des sinistres du portefeuille a été obtenue à l'aide d'un modèle CAT ouragan, considéré comme complexe et consommateur en temps de calcul. L'utilisation actuelle de ce modèle est adaptée aux besoins du *Risk Management* – pour quantifier les besoins en capital – mais pas de la souscription : en raison des temps de calcul, il est impossible de lancer un grand nombre de fois le modèle. Sans outil pour pouvoir estimer l'impact des actions de *pruning*, l'entité n'est pas en mesure d'identifier les combinaisons de polices dont la résiliation permettrait de maximiser sa rentabilité.

Tout d'abord, ce mémoire développe une méthode permettant de quantifier l'impact d'un ensemble de résiliations sur le ratio combiné économique du portefeuille. Une attention particulière a été accordée aux effets non-proportionnels des coûts de réassurance et du capital.

Grâce à l'implémentation d'un algorithme génétique, ce mémoire permet la mise en place de solutions d'aide aux décisions de *pruning*. Cet algorithme, adapté aux problèmes d'optimisation combinatoire multi-objectif, permet d'identifier les résiliations qui optimisent la rentabilité du portefeuille tout en s'adaptant aux différentes contraintes opérationnelles.

Notre problème doit s'adapter à plusieurs types de contraintes :

- Tout d'abord, il est nécessaire de prendre en compte certaines contraintes métier. Par exemple, il n'est pas possible de résilier partiellement une police : cela nécessiterait d'avoir recours à la réassurance facultative, jugée trop coûteuse en temps de gestion. Par ailleurs, toutes les polices ne sont pas considérées comme résiliables car certaines sont souscrites en bouquet avec d'autres polices rentables.
- Ensuite, l'entité souhaite résilier un volume de polices et de primes limité pour ne pas envoyer de mauvais signal au marché. L'entité souhaite résilier au maximum 4% de ses primes brutes.
- Enfin, notre étude risque de demander des temps de calcul importants. Même en restreignant l'étude aux cent polices les moins rentables du portefeuille, cela représente $1,3 \times 10^{30}$ scénarios de résiliation possibles à évaluer (2^{100}). Dès lors, il est nécessaire d'utiliser des algorithmes adaptés pour limiter les temps de calcul.

1.3. Données disponibles

On dispose de la base de données d'exposition du portefeuille. Pour chaque police, des informations sont disponibles sur les caractéristiques physiques (localisation, activité, matériau de construction, etc.), sur les conditions financières (franchise, limite, etc.) ainsi que sur les primes payées au titre du risque ouragan.

On dispose également de la distribution des sinistres du portefeuille obtenus à l'aide d'un modèle CAT ouragan mexicain. Il s'agit d'un modèle *inhouse* développé au sein d'AXA. Les résultats ont la forme d'une table contenant les pertes par site, par année simulée et par événement.

2. Démarche de résolution

2.1. Cadrage du problème

L'entité souhaite résilier un ensemble de polices pour minimiser à la fois le ratio combiné économique du portefeuille et le pourcentage de primes résiliées.

Le problème étant bi-objectif, une méthode a posteriori sera utilisée : on souhaite résoudre le problème en déterminant une frontière optimale, c'est-à-dire un ensemble de solutions pour lesquelles il n'est pas possible d'améliorer un critère sans en dégrader un autre.

Notre problème a été encodé en binaire car il n'est pas possible de résilier une partie d'une police :

- Si une police est 100% résiliée, elle est encodée 1.
- Si une police est conservée (0% résiliée), elle est encodée 0.

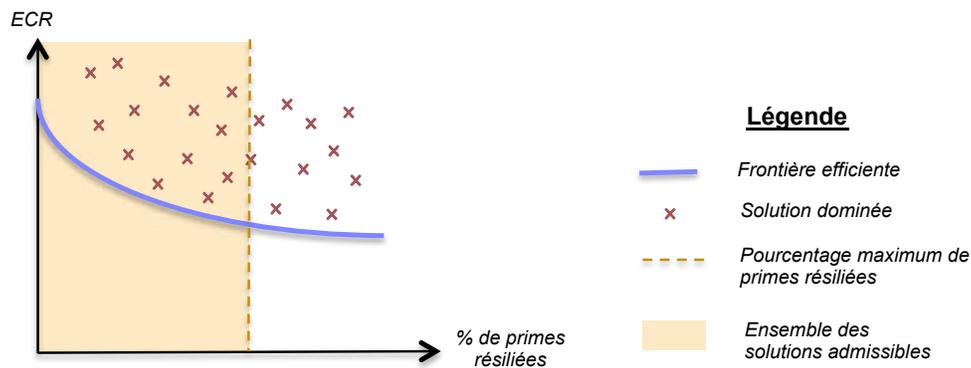
Soit n le nombre total de polices. Notre problème consiste à déterminer un ensemble de n -uplets $b = (b_1, \dots, b_n)$, avec $b_i \in \{0; 1\}$ qui optimise le système suivant :

$$\begin{cases} \min ECR \\ \min \text{pourcentage de primes résiliées} \\ \text{avec pourcentage de primes résiliées} \leq 4\% \end{cases}$$

Pour chaque n -uplets b , nous sommes capable de :

- Calculer le pourcentage de primes résiliées car le risque ouragan est couvert à l'aide de polices dédiées : pour chaque police, on dispose donc des informations de prime.
- Estimer l'ECR du portefeuille après résiliation grâce à une méthode développée dans ce mémoire. Cette méthode permet de quantifier les effets non-proportionnels (traité de réassurance par événement, coût du capital).

Comme on le voit sur la figure ci-dessous, aucune contrainte n'est appliquée sur les primes. Les solutions admissibles sont déterminées a posteriori. Cette méthode permet de disposer d'un niveau de flexibilité important.



Notre problème appartient à la classe de complexité NP : sa résolution exacte requiert un temps de calcul exponentiel à la taille du problème. De ce fait, ce type de problème est traditionnellement résolu à l'aide d'algorithmes heuristiques. Une heuristique est une stratégie pour se déplacer intelligemment dans l'espace des solutions, afin d'obtenir une solution approchée la meilleure possible, dans un temps de calcul raisonnable. L'avantage de cette classe d'algorithmes est qu'il n'y a pas besoin que le problème vérifie certaines propriétés (convexité, etc.). Il est simplement nécessaire de pouvoir évaluer les solutions candidates.

2.2. Algorithme génétique

L'optimisation s'appuie sur un algorithme heuristique et plus particulièrement sur un algorithme génétique. Les algorithmes génétiques reposent sur une population de chromosomes, eux-mêmes composés de gènes. Chaque gène correspond à une police (résiliée ou non) et chaque chromosome à un scénario de résiliation.

À chaque itération de l'algorithme, une nouvelle population de chromosomes est générée : il s'agit d'une nouvelle « génération ». Chaque génération est composée du même nombre de chromosomes et subit une sorte de sélection naturelle : seuls les chromosomes les plus adaptés à leur environnement survivent. Au fil des générations les chromosomes tendent ainsi vers l'optimum de la fonction sélective.

La création d'une nouvelle population de chromosomes s'appuie sur cinq principes :

- **La sélection** : cette étape consiste à sélectionner des chromosomes « parents » qui seront par la suite combinés pour produire des chromosomes « enfants ».
- **La reproduction** : à partir des chromosomes sélectionnés (chromosomes « parents »), des chromosomes « enfants » sont générés. Un parallèle est fait ici avec l'héritage génétique en biologie : un enfant possède à la fois des gènes de sa mère et de son père. L'objectif sous-jacent est de faire se reproduire les meilleurs chromosomes pour les perfectionner au fil des générations.
- **La mutation** : la mutation consiste à modifier un gène d'un chromosome. Ce chromosome est habituellement choisi aléatoirement. Ce phénomène a pour but de sortir des minimums locaux.
- **L'évaluation** : les individus sont évalués au regard de certains critères définis par l'utilisateur. Pour survivre, les chromosomes doivent être adaptés ou s'adapter à ces critères.
- **L'élitisme** : au cours de cette opération, l'algorithme sélectionne les chromosomes qui répondent le mieux au problème d'optimisation. Un parallèle est fait ici avec la théorie darwinienne de l'évolution : seuls les meilleurs individus survivent.

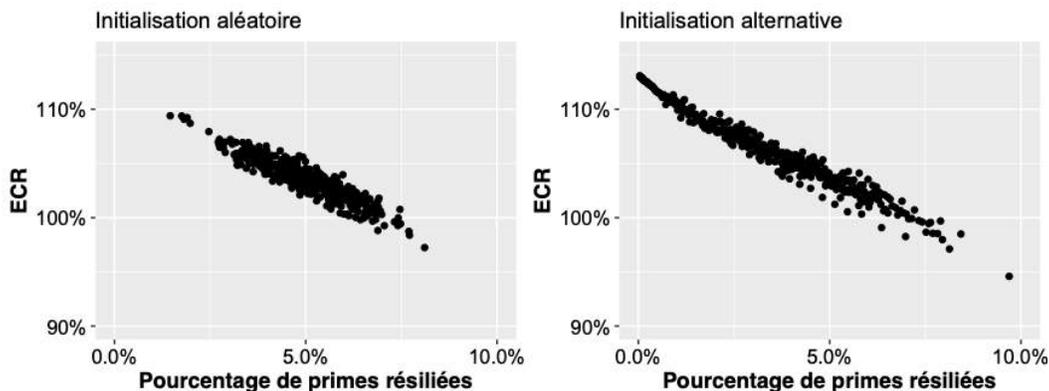
Pour chaque étape décrite ci-dessus, des paramètres adaptés à notre problème ont été choisis. Des études de sensibilité ont ensuite été réalisées pour tester la stabilité des résultats obtenus.

2.3. Population initiale

Une méthode a été développée pour identifier les polices qui pénalisent le plus la rentabilité du portefeuille. Les équipes de souscription ont ensuite étudié ces polices et en ont retenu 100 considérées comme résiliables. On suppose que les autres polices sont renouvelées à l'identique.

Dans ce mémoire, une attention particulière a été accordée à la génération de la population initiale. Habituellement, la population initiale est générée de manière aléatoire : chaque gène a une probabilité de 50% d'être tiré. Une méthode « alternative » a été développée dans le but de générer une population initiale ayant un bon niveau de diversité génétique, permettant de couvrir uniformément l'espace de recherche et favorisant la vitesse de convergence de l'algorithme.

Les nuages de points ci-dessous présentent la répartition de deux populations de 400 chromosomes : la première a été générée aléatoirement et la seconde avec la méthode dite alternative.

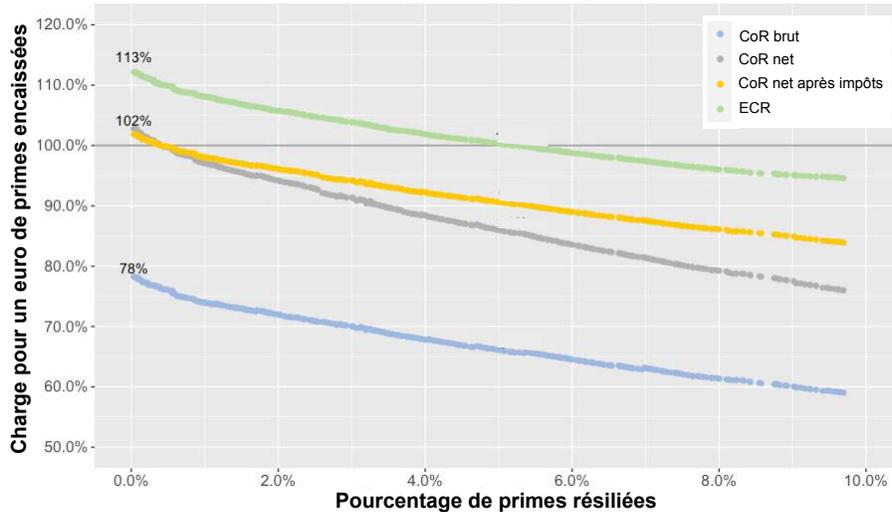


Les chromosomes de la population aléatoire sont concentrés entre 2,5% et 7,5% de primes résiliées alors que la population alternative permet d'explorer l'espace des primes entre 0% et 7,5%. L'initialisation alternative permet donc de mieux explorer l'espace.

3. Résultats et applications

3.1. Résultats

Le graphe ci-dessous présente les résultats de la minimisation de l'ECR et du pourcentage de primes résiliées. La courbe verte représente l'ECR : chaque point correspond à un scénario de résiliation optimal. Pour chaque scénario, les autres courbes présentent la décomposition de l'ECR.



Pour revenir à l'équilibre, c'est à dire à un ECR de 100%, il serait nécessaire de résilier au minimum 5% des primes. Avec un objectif de résiliation de maximum 4% des primes, l'entité peut espérer réduire son ECR d'environ 10 points : pour un euro de prime encaissé, la détérioration de fonds propres serait alors réduite à 3 centimes au lieu de 13.

3.2. Enrichissement de la méthode

Une fois les premiers résultats communiqués aux équipes de souscription, on suppose que l'entité souhaite affiner sa politique de résiliation avec une nouvelle contrainte sur les primes puis un objectif supplémentaire.

Ajout d'une contrainte sur le pourcentage de primes résiliées

Dans un premier temps, l'entité souhaite ajouter une contrainte sur les primes : on souhaite résilier entre 2% et 4% des primes brutes. Pour ce faire, on ajoute une pénalisation dans la fonction d'évaluation : l'ECR des scénarios de résiliation qui ne respectent pas la contrainte de prime est pénalisé.

L'ajout d'une contrainte sur les primes permet de concentrer les chromosomes entre 2% et 4% de primes résiliées sans impacter négativement les résultats sur l'intervalle de primes étudié. Cela offre davantage de souplesse aux équipes de souscription qui peuvent faire leur choix parmi un plus grand nombre de scénarios de résiliation respectant la contrainte.

Ajout d'un troisième objectif : minimisation du nombre de polices résiliées

Dans un second temps, l'entité souhaite minimiser le nombre de polices résiliées pour limiter les temps de gestion des résiliations. Dès lors, on souhaite ajouter un troisième objectif à notre problème : minimiser le nombre de polices résiliées.

Les contraintes sur les primes ont été relâchées et on cherche à optimiser le système suivant :

$$\left\{ \begin{array}{l} \min ECR \\ \min \text{pourcentage de primes résiliées} \\ \min \text{nombre de polices résiliées} \end{array} \right.$$

L'algorithme génétique utilisé dans ce mémoire est multi-objectif et il s'adapte facilement à l'ajout d'un objectif supplémentaire : il suffit d'ajouter le troisième objectif dans la fonction d'évaluation.

Conclusion

Le **premier objectif** de ce mémoire était de développer une méthode permettant de quantifier l'impact de n'importe quel scénario de résiliation sur le ratio combiné économique du portefeuille. Des méthodes d'estimation du résultat cédé et du coût du capital ont été développées pour prendre en compte les effets non-proportionnels de l'ECR. Il est désormais possible de quantifier en moins d'une seconde l'impact des résiliations sur la rentabilité du portefeuille. Des résultats similaires auraient été obtenus en plusieurs heures avec un modèle CAT. Cette méthode a permis de renforcer la connaissance du portefeuille : elle a permis d'identifier les polices qui pèsent le plus sur les coûts de réassurance et du capital. Cette méthode permet également de piloter la politique de résiliation de risques et d'anticiper ses impacts sur les coûts de réassurance et du capital.

Le **second objectif** de notre étude consistait à proposer une optimisation des combinaisons de polices à résilier pour minimiser le ratio combiné économique tout en minimisant le pourcentage de primes résiliées. Pour résoudre ce problème d'optimisation combinatoire multi-objectif, notre étude s'est appuyée sur un algorithme génétique. Les résultats se présentent sous la forme d'une frontière d'efficacité composée d'un ensemble de scénarios de résiliation répondant au problème. L'algorithme est flexible et stable. Il permet aussi de prendre en compte les différentes contraintes métier.

Ce mémoire présente cependant certaines limites. Tout d'abord, seules les polices couvrant le risque ouragan ont été considérées : les polices tremblement de terre et incendie n'ont pas été prises en compte car nous ne disposons pas de leur modélisation. Le lecteur souhaitant mettre en pratique les méthodes développées ici devra considérer toutes les polices du portefeuille dommage.

Par ailleurs, on suppose dans notre étude que les polices non résiliées sont renouvelées à l'identique et qu'il n'y a pas de nouvelles souscriptions. Or en pratique, toutes les polices ne seront pas renouvelées et de nouvelles polices seront souscrites. Pour remédier à ce problème, il est possible de développer une méthode pour quantifier l'impact des nouvelles souscriptions sur la rentabilité du portefeuille en restructurant le catalogue d'événement stochastique du modèle CAT et en utilisant un système de gestion de bases de données (*El Alj, 2019*).

Enfin, ce mémoire a l'intérêt de présenter des méthodes qui peuvent être utilisées dans d'autres contextes. Les méthodes pourraient par exemple être appliquées pour optimiser un portefeuille de rétrocession ou pour anticiper l'impact d'un changement de franchise ou de limite au sein d'un portefeuille.

Executive summary

The methods for estimating the economic combined ratio and optimizing the portfolio using a genetic algorithm were developed as part of this thesis.

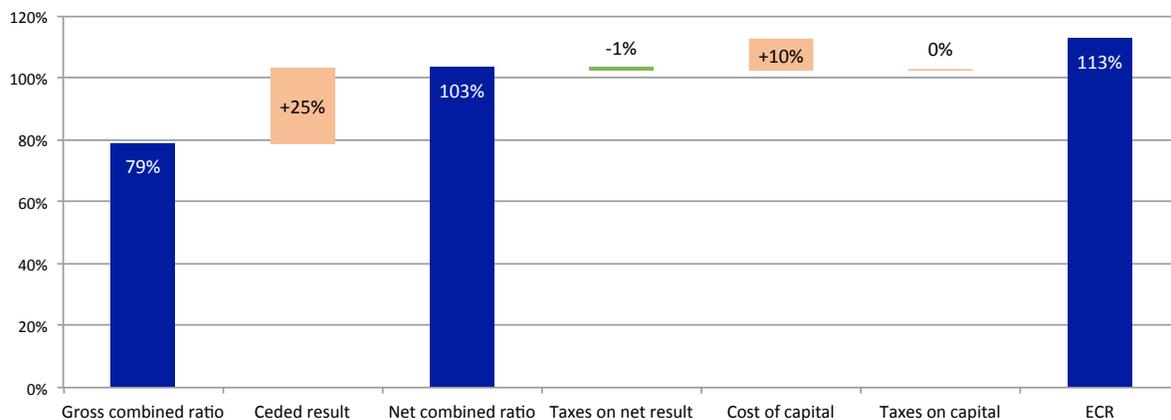
1. Context and objective

1.1. Portfolio studied

This thesis focusses on a property insurance portfolio composed of policies located in Mexico and covering hurricane risk. This portfolio was underwritten by an AXA entity, which operates in the Property & Casualty and Life & Health segments.

With an economic combined ratio of 113%, the portfolio studied is not profitable: each euro of premium collected results in a 13 cents decrease of the own funds.

The economic combined ratio, also known as ECR, quantifies the underwriting performance. This metric estimates the insurer's cost per euro of premium written: we therefore want it to be as low as possible. Schematically, the ECR is calculated as the sum of the gross claims, reinsurance costs, cost of capital, taxes and expenses divided by the gross premium. The graph below shows the decomposition of the ECR of the studied portfolio, modeled with an in-house CAT model.



For one euro of premium collected, the costs faced by the entity are as follows:

- The gross combined ratio, i.e. gross claims and operating expenses, represents the largest cost: 79 cents.
- The ceded result results in a cost of 25 cents: the portfolio is protected by a per event reinsurance treaty.
- The net combined ratio is above 100%: the entity therefore benefits from a tax credit of 1 cent.
- The cost of capital represents a charge of 10 cents: this amount corresponds to the return on capital expected by the shareholders.
- There is no tax on the returns of the capital tied up because the capital is assumed to be invested in low-risk assets with a zero return.

1.2. Objective

The portfolio studied in this thesis is composed of property policies located in Mexico and covering hurricane risk. With an economic combined ratio of 113%, the portfolio generates losses and reinsurance and capital costs weigh heavily on its profitability.

In this context, it was decided to carry out pruning actions, i.e. targeted policy terminations. The overarching goal is to optimize the overall profitability of the portfolio.

The distribution of the portfolio's claims was obtained using a CAT hurricane model, which is considered complex and time-consuming. The current use of this model is adapted to the needs of Risk Management - to quantify the capital requirements - but not to the needs of underwriting: due to the computation time, it is impossible to run the model a large number of times. Without tools to estimate the impact of pruning actions, the entity is not able to identify the policies whose termination would maximize its profitability.

First, this thesis develops a method to quantify the impact of policy terminations on the portfolio economic combined ratio. Particular attention has been paid to non-proportional effects such as the impact on reinsurance costs and capital.

Thanks to the implementation of a genetic algorithm, this thesis allows the implementation of solutions to assist in pruning decisions. This algorithm, adapted to multi-objective discrete optimization problems, allows us to identify the policy terminations that optimize the profitability of the portfolio while taking into account the different operational constraints.

Our problem must adapt to several types of constraints:

- First, it is necessary to take into account certain business constraints. For example, it is not possible to partially terminate a policy: this would require the use of facultative reinsurance, which is considered too costly in terms of management time. In addition, not all policies are considered terminable because some are written in combination with other profitable policies.
- Secondly, the entity wishes to terminate a limited number of policies and premium in order not to send a wrong signal to the market. The entity wishes to terminate no more than 4% of its gross premium.
- Finally, our study may require significant computation time. Even restricting the study to the 100 least profitable policies in the portfolio, this represents 1.3×10^{30} possible termination scenarios to be evaluated (2^{100}). Therefore, it is necessary to use suitable algorithms to limit the computation time.

1.3. Available data

The portfolio exposure database is available. For each policy, information is available on physical characteristics (location, activity, construction material, etc.), on financial conditions (deductible, limit, etc.) and on premium paid for the hurricane risk.

We also have the distribution of the portfolio's losses obtained using a Mexican hurricane CAT model. This is an in-house model developed at AXA. The results are in the form of a table containing the losses per site, per simulated year and per event.

2. Problem solving

2.1. Study framework

The entity wishes to terminate a set of policies in order to minimize the economic combined ratio of the portfolio while minimizing the percentage of pruned premium.

The problem being bi-objective, an a posteriori method will be used. As a matter of fact, we wish to solve the problem by determining an optimal frontier, i.e. a set of solutions for which it is not possible to improve one criterion without degrading another.

Our problem has been encoded in binary because it is not possible to terminate only a part of a policy:

- If a policy is 100% terminated, it is encoded 1.
- If a policy is renewed (0% terminated), it is encoded 0.

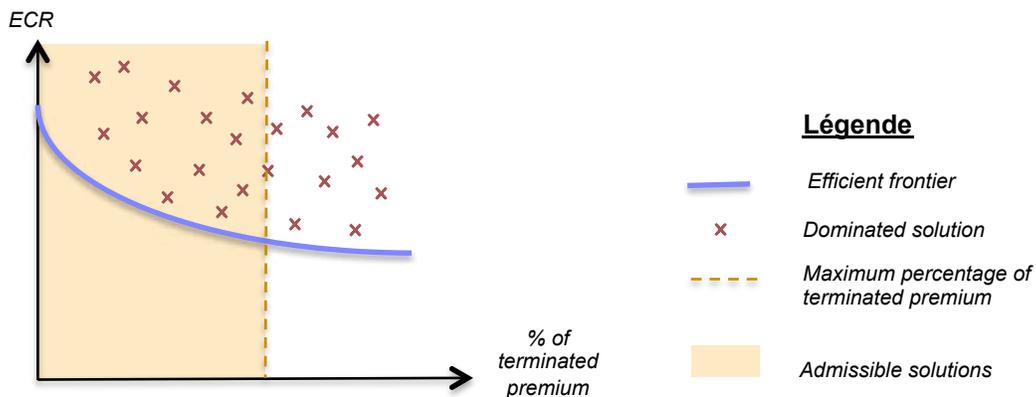
Let n be the total number of policies. Our problem is to determine a set of n -tuples $b = (b_1, \dots, b_n)$, with $b_i \in \{0; 1\}$, that optimizes the following system:

$$\left\{ \begin{array}{l} \min ECR \\ \min \text{pruned premium} \\ \text{with percentage of pruned premium} \leq 4\% \end{array} \right.$$

For each n -tuple b , we are able to:

- Compute the percentage of terminated premium because the hurricane risk is covered using dedicated policies: for each policy, we therefore have the premium information.
- Calculate the ECR of the portfolio after termination using a method developed in this thesis. This method quantifies the non-proportional effects (non-proportional reinsurance treaty, cost of capital).

As shown in the figure below, no constraints are applied on premium. The admissible solutions are determined a posteriori. This method allows for a significant level of flexibility.



Our problem belongs to the NP complexity class: its exact solution requires a computation time exponential to the problem size. Therefore, this type of problem is traditionally solved using heuristic algorithms. A heuristic algorithm is a strategy for moving intelligently through the solution space, in order to obtain the best possible approximate solution in a reasonable computation time. The advantage of this class of algorithms is that there is no need for the problem to verify certain properties (convexity, etc.). It is simply necessary to be able to evaluate the candidate solutions.

2.2. Genetic algorithm

Our optimization is based on a heuristic algorithm and more particularly on a genetic algorithm. Genetic algorithms are based on a population of chromosomes, themselves composed of genes. Each gene corresponds to a policy (terminated or not) and each chromosome to a termination scenario.

At each iteration of the algorithm, a new population of chromosomes is generated: this is a new "generation". Each generation is composed of the same number of chromosomes and undergoes a kind of natural selection: only the chromosomes most adapted to their environment survive. Over the generations, the chromosomes thus tend towards the optimum of the selective function.

The creation of a new population of chromosomes is based on five principles:

- **Selection:** this step consists in selecting "parent" chromosomes, which will be subsequently combined to produce "children" chromosomes.
- **Crossover:** from the selected chromosomes ("parent" chromosomes), "child" chromosomes are generated. A parallel is drawn here with genetic inheritance in biology: a child has genes from both their mother and father. The underlying objective is to reproduce the best chromosomes in order to improve them over generations.
- **Mutation:** Mutation involves changing a gene on a chromosome. This chromosome is usually chosen randomly. The purpose of this phenomenon is to avoid local extremums.
- **Evaluation:** individuals are evaluated according to certain criteria defined by the user. To survive, chromosomes must be adapted or fit these criteria.

- **Elitism:** during this operation, the algorithm selects the chromosomes that best answer the optimization problem. A parallel is made here with the Darwinian theory of evolution: only the best individuals survive.

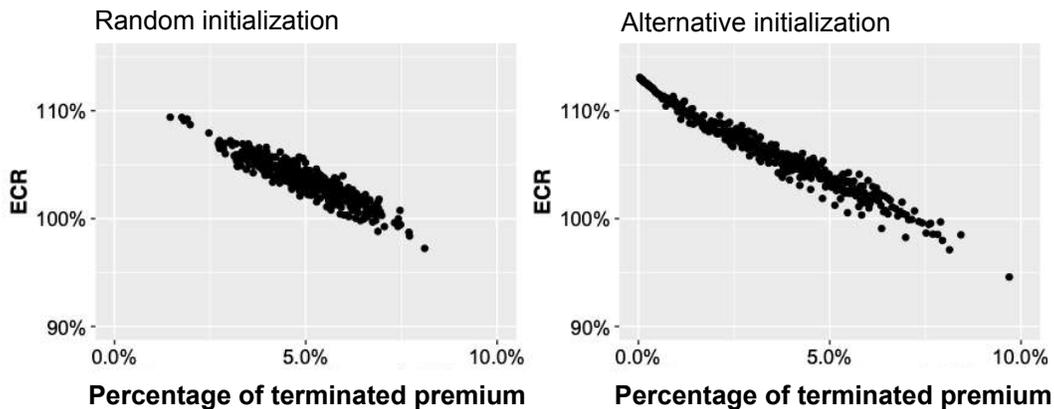
For each step described above, parameters adapted to our problem were chosen. Sensitivity studies were then performed to test the stability of the results obtained.

2.3. Initial population

A method was developed to identify the policies that penalize the most the profitability of the portfolio. The underwriting teams then studied these policies and selected 100 policies that were considered to be terminable. It is assumed that the remaining policies are renewed as expiring.

In this thesis, particular attention was paid to the generation of the initial population. Usually, the initial population is randomly generated: each gene has a 50% probability of being drawn. An "alternative" method was developed in order to generate an initial population with a good level of genetic diversity, allowing to cover uniformly the search space and favoring the convergence speed of the algorithm.

The scatterplots below show the distribution of two populations of 400 chromosomes: the first one was generated randomly and the second one with the so-called alternative method.

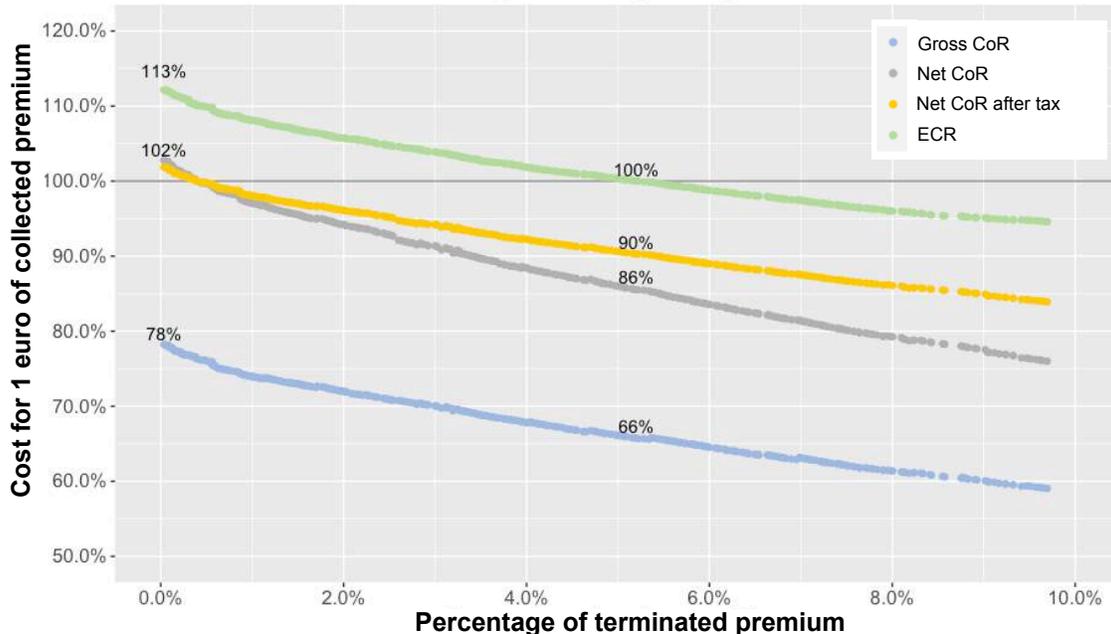


The chromosomes of the random population are concentrated between 2.5% and 7.5% of terminated premium, whereas the alternative population allows to explore the space of premium between 0% and 7.5%. Thus, the alternative initialization allows for better exploration of the space.

3. Results and applications

3.1. Results

The graph below shows the results of the minimization of the ECR and the percentage of terminated premium. The green curve represents the ECR: each point corresponds to an optimal termination scenario. For each scenario, the other curves show the decomposition of the ECR.



To return to equilibrium, i.e. to an ECR of 100%, it would be necessary to terminate at least 5% of premium. With a termination target of maximum 4% of premiums, the entity can hope to reduce its ECR by about 10 points: for one euro of premium collected, the deficit would then be reduced to 3 cents instead of 13.

3.2. Methodology extension

Once the first results have been communicated to the underwriting teams, we assume that the entity wishes to refine its pruning actions with a constraint on premium and then with an additional objective.

Adding a constraint on the percentage of terminated premium

First, the entity wants to add a constraint on the premiums: we want to cancel between 2% and 4% of the gross premiums. To do this, we add a penalty in the evaluation function: the ECR of the termination scenarios that do not respect the premium constraint is penalized.

Adding a premium constraint allows chromosomes to be concentrated between 2% and 4% of terminated premiums without negatively impacting the results over the premium interval studied. This provides more flexibility for underwriting teams to choose from a larger number of termination scenarios that meet the constraint.

Adding a third objective: minimizing the number of terminated policies

In a second step, the entity wants to minimize the number of terminated policies to limit management time. Therefore, we want to add a third objective to our problem: minimizing the number of terminated policies.

The constraints on the premiums have been relaxed and we seek to optimize the following system:

$$\begin{cases} \min ECR \\ \min \text{percentage of pruned premium} \\ \min \text{number of pruned premium} \end{cases}$$

The genetic algorithm used in this dissertation is multi-objective and easily accommodates the addition of an additional objective: simply add the third objective into the evaluation function.

Conclusion

The first objective of this thesis was to develop a method to quantify the impact of any termination scenario on the portfolio economic combined ratio. Methods for estimating the reinsurance and capital costs have been developed to take into account the non-proportional effects. It is now possible to quantify the impact of terminations on the profitability of the portfolio in less than a second whereas similar results would have been obtained in several hours with a CAT model. This method has allowed us to strengthen our knowledge of the portfolio: it has made it possible to identify the policies that weigh most heavily on reinsurance and capital costs. This method also makes it possible to pilot the pruning actions and to anticipate their impacts on reinsurance and capital costs.

The second objective of our study was to propose an optimization of the policies to be terminated in order to minimize the economic combined ratio while minimizing the percentage of premiums terminated. To solve this multi-objective discrete optimization problem, our study relied on a genetic algorithm. The results are presented in the form of an efficiency frontier composed of a set of termination scenarios that satisfy the problem. The algorithm is flexible and stable. It also allows to take into account the different business constraints.

However, this thesis has some limitations. First of all, only policies covering hurricane risk have been considered: earthquake and fire policies have not been taken into account because we did not have their modeling. The reader wishing to apply the methods developed here will have to consider all the policies in the property portfolio.

Moreover, we assume in our study that the policies that have not been terminated are renewed as expiring and that there are no new subscriptions. In practice, not all policies will be renewed and new policies will be written. To address this problem, it is possible to develop a method to quantify the impact of new policy on the portfolio profitability by restructuring the stochastic event catalog and using a database management system (El Alj, 2019).

Finally, this dissertation has the benefit of presenting methods that can be used in other contexts. For example, the methods could be applied to optimize a retrocession portfolio or to anticipate the impact of line sizing actions within a portfolio.

Remerciements

En premier lieu, je souhaite remercier mon tuteur en entreprise Jérôme Crétien pour son encadrement, sa relecture et plus généralement pour m'avoir formée tout au long de mes années passées chez AXA.

Je tiens ensuite à remercier Hamza El Hassani pour ses conseils, ses nombreuses idées et son soutien.

Merci également à David Billod de Filiquier pour les nombreuses discussions qui ont permis de faire avancer mon étude.

Merci enfin à tous mes collègues qui ont contribué de près ou de loin à ce mémoire.

Ce mémoire marque la fin de mon parcours au CNAM. Je souhaite donc remercier mon tuteur académique Oliver Desmettre pour son encadrement, sa relecture et ses remarques. Je souhaite aussi remercier toute l'équipe pédagogique du CNAM.

Je tiens également à remercier mes relecteurs : Alexandre Costa et Anne Jovet.

Merci enfin à mes proches, et notamment mes parents et ma sœur, pour leur soutien sans faille et leurs encouragements tout au long de mes études.

Sommaire

Résumé	i
Note de synthèse	iii
Remerciements	xv
Sommaire	xvi
Liste des abréviations	xix
Introduction	1
1. Assurance et risque ouragan au Mexique	3
1.1 Le Mexique : un pays exposé au risque ouragan	4
1.1.1 Le risque catastrophe.....	4
1.1.2 Le risque ouragan	5
1.1.3 La situation du Mexique	6
1.2 Modélisation des ouragans et métriques d'intérêt	9
1.2.1 Architecture d'un modèle CAT.....	9
1.2.2 Résultats des modèles et métriques d'intérêt	11
1.3.2.1 Sorties des modèles	12
1.3.2.2 Métriques d'intérêt.....	13
1.3 Le ratio combiné économique : un indicateur de performance de la souscription	16
1.3.1 Ratio combiné	17
1.3.2 Coût du capital	19
1.3.3 Les impôts	21
1.3.4 Le ratio combiné économique.....	22
1.3.5 La création de valeur économique	22
2. Analyse du portefeuille et impact d'une politique de résiliation sur le ratio combiné économique	
2.1 Présentation du portefeuille étudié	24
2.1.1 Données à disposition	24
2.1.1.1 Données d'exposition	24
2.1.1.2 Résultats d'un modèle CAT ouragan mexicain.....	26
2.1.1.3 Données de réassurance : traité	27
2.1.1.4 Données de STEC.....	27
2.1.2 Analyse descriptive du portefeuille.....	28

2.1.2.1	Revue globale du portefeuille.....	28
2.1.2.2	Conclusion.....	29
2.1.3	Ratio combiné économique du portefeuille	31
2.2	Quantification de l'impact d'un scénario de résiliation sur l'ECR.....	32
2.2.1	Impact sur la sinistralité brute	33
2.2.2	Impact sur les coûts de réassurance	36
2.2.2.1	Etape 1 : adaptation de la limite du traité	36
2.2.2.2	Etape 2 : application du traité de réassurance	37
2.2.2.3	Etape 3 : estimation de la prime de réassurance.....	37
2.2.2.4	Etape 4 : calcul du ratio combiné net	42
2.2.3	Impact sur le coût du capital	42
2.2.3.1	Calcul du STEC CAT standalone.....	43
2.2.3.2	Calcul du STEC CAT diversifié.....	44
2.2.3.3	Estimation du taux de diversification du STEC CAT dans le STEC total	46
2.2.3.4	Calcul du coût du capital	48
2.2.4	Synthèse et calcul de l'ECR.....	48
2.2.5	Exemple	50
2.3	Application : identification des polices pénalisant le plus la rentabilité du portefeuille	52
3.	Optimisation d'une politique de résiliation de risques à l'aide d'un	
	algorithme génétique	54
3.1	Cadrage du problème	56
3.1.1	Définition du problème	56
3.1.2	Contraintes opérationnelles.....	57
3.1.3	Démarche de résolution.....	58
3.2	Algorithmes génétiques et méthodologie choisie.....	61
3.2.1	Principes généraux	61
3.2.2	Initialisation de la population.....	63
3.2.2.1	Diversité génétique.....	63
3.2.2.2	Couverture uniforme de l'espace de recherche	66
3.2.2.3	Ajout des meilleures solutions dans la population initiale	70
3.2.2.4	Méthode retenue	71
3.2.3	Evaluation de la population.....	72
3.2.4	Sélection.....	74
3.2.5	Croisement	76
3.2.6	Mutations	78

3.2.7	Génération suivante.....	79
3.2.8	Arrêt de l'algorithme.....	79
3.2.9	Synthèse	80
3.3	Résultats et limites	81
3.3.1	Optimisation de l'ECR et du pourcentage de primes résiliées.....	81
3.3.1.1	Performance des résultats	81
3.3.1.2	Impact sur les primes nettes	84
3.3.1.3	Comportement des différents composants de l'ECR.....	86
3.3.1.4	Complexité algorithmique	89
3.3.1.5	Conclusion	90
3.3.2	Adaptation de l'algorithme à de nouvelles contraintes métier.....	91
3.3.2.1	Ajout d'une contrainte sur le pourcentage de primes résiliées.....	91
3.3.2.2	Ajout d'un troisième objectif : minimisation du nombre de polices résiliées.....	93
3.3.2.3	Conclusion.....	95
3.3.3	Etudes de sensibilité.....	96
3.3.3.1	Nombre de gènes	96
3.3.3.2	Taille de la population et nombre de générations.....	98
3.3.3.3	Génération de la population initiale	101
3.3.4	Limites de l'étude.....	103
	Conclusion	105
	Bibliographie	107
	Table des figures	109
	Annexes	111
	Annexe 1 : la réassurance : outil de pilotage du risque CAT	112
	Annexe 2 : analyse descriptive de la situation géographique du portefeuille	121
	Annexe 3 : analyse descriptive des caractéristiques des sites du portefeuille	124
	Annexe 4 : sensibilité des résultats aux paramètres de reproduction et de mutation	128

Liste des abréviations

AEP : *Aggregate Exceedance Probability*

CoR : *Combined Ratio*, ratio combiné

CoV : *Coefficient of Variation*, coefficient de variation

ECR : *Economic Combined Ratio*, ratio combiné économique

LoB : *Line of Business*, branche d'assurance

LoL : *Loss on Line*, sinistres moyens normalisés par la limite d'un traité

LR : *Loss Ratio*, Ratio de Sinistres sur Primes

NIV : *Net Insured Value*, somme assurée nette

OEP : *Occurrence Exceedance Probability*

RoL : *Rate on Line*, prime de réassurance normalisée par la limite d'un traité

SCR : *Solvency Capital Requirement*

STEC : *Short Term Economic Capital*, SCR en modèle interne AXA

YELT: *Year Event Loss Table*, table de sinistralité par année et par événement

YLT : *Year Loss Table*, table de sinistralité par année

Introduction

En 2005, l'ouragan Wilma frappe le Mexique, causant près de 3 milliards de dollars (*Hofmann, 2005 [1]*) de pertes assurantielles, soit environ 23% des primes d'assurance encaissées au Mexique cette année-là (*OCDE, 2021 [2]*). Cet événement illustre le risque d'accumulation des assureurs au Mexique et l'importance de la qualité de la souscription.

Avec un ratio combiné économique de 113%, le portefeuille de polices ouragan mexicain étudié dans ce mémoire est déficitaire et les coûts de réassurance et du capital représentent 35% des primes souscrites.

Dans ce contexte, l'entité ayant souscrit le portefeuille de risques ouragan souhaite effectuer une politique de *pruning*, c'est-à-dire des résiliations ciblées, pour optimiser la rentabilité globale du portefeuille.

Le portefeuille est modélisé à l'aide d'un modèle CAT ouragan, considéré comme complexe et très consommateur en temps de calcul. Dans les sociétés d'assurance, la modélisation CAT est un outil essentiel de *Risk Management*. À la suite de la mise en place de la directive Solvabilité II², les assureurs doivent disposer d'un montant minimum de fonds propres, calculé sur la base d'une déviation bicentenaire de leur résultat. Pour quantifier cette déviation, ils doivent modéliser chacun de leurs risques. Dans ce contexte, la modélisation du risque catastrophe est un défi car la profondeur historique des sinistres est souvent insuffisante pour permettre l'utilisation des méthodes statistiques classiques. De ce fait, les assureurs ont recours à des modèles CAT qui s'appuient sur la connaissance physique du risque sous-jacent.

L'utilisation actuelle de ces modèles est adaptée aux besoins du *Risk Management* – pour quantifier les besoins en capital – mais pas de la souscription : en raison des temps de calcul, il est impossible de lancer un grand nombre de fois les modèles. Sans outil pour pouvoir estimer l'impact des actions de *pruning* sur les coûts de réassurance et du capital, les souscripteurs ne sont pas en mesure d'identifier les combinaisons de polices dont la résiliation permettrait de maximiser la rentabilité du portefeuille.

Ce mémoire a pour objectif de développer des méthodes actuarielles pour aider l'entité dans ses décisions de résiliation.

Dans un premier temps, le contexte de ce mémoire sera présenté ainsi que toutes les notions nécessaires à sa compréhension : analyse et modélisation du risque ouragan mexicain, intérêt de la réassurance et définition du ratio combiné économique.

Dans un deuxième temps, on développera une méthode permettant de quantifier l'impact d'un ensemble de résiliations sur le ratio combiné économique du portefeuille, et ce, sans refaire tourner le modèle CAT. En effet, avant d'optimiser le portefeuille, il est nécessaire de pouvoir quantifier l'impact des résiliations envisagées : impact sur les sinistres, la réassurance, les frais, les impôts et le coût du capital.

Pour ce faire, il est nécessaire de disposer des résultats de la modélisation CAT par police, par événement et par année simulée. Des méthodes d'estimation de la prime de réassurance devront également être développées.

Le portefeuille sera ensuite analysé dans le but d'identifier les polices dont la résiliation a le plus d'impact sur le ratio combiné économique. Ces polices feront partie du périmètre de l'optimisation.

² Solvabilité II est une directive européenne et ne s'applique donc pas au Mexique. Elle est utilisée dans ce mémoire car c'est un cadre conceptuel intéressant et car AXA est un groupe européen qui reporte sous Solvabilité II.

Dans un troisième temps, une optimisation de la politique de *pruning* sera effectuée à l'aide d'un algorithme génétique. L'objectif est d'identifier les combinaisons de polices dont la résiliation optimise le ratio combiné économique du portefeuille tout en minimisant le pourcentage de primes résiliées. Il s'agit d'un problème d'optimisation combinatoire bi-objectif, reposant sur la détermination d'une frontière efficiente. La méthode de quantification d'impact développée dans la partie précédente permettra d'évaluer les différentes solutions candidates.

L'optimisation s'appuie sur des algorithmes heuristiques et plus particulièrement des algorithmes génétiques. Les algorithmes seront adaptés pour prendre en compte les différentes contraintes opérationnelles et une attention particulière sera accordée à la génération de la population initiale de l'algorithme génétique.

L'objectif sous-jacent de ce mémoire est de changer de paradigme sur l'utilisation des modèles CAT et d'élargir leur utilisation : d'un outil de *Risk Management*, on souhaite exploiter leurs résultats pour qu'ils deviennent des outils d'aide aux décisions de résiliation des polices catastrophe spécifiques. Par l'intermédiaire de méthodes actuarielles, ce mémoire a pour objectif de permettre aux équipes de souscription d'anticiper les évolutions de résultats des modèles CAT.

Les méthodes développées se veulent utilisables opérationnellement : on souhaite que les temps de calcul soient raisonnables et que les méthodes s'adaptent facilement aux différentes contraintes métier.

L'ensemble des données figurant dans ce mémoire a été anonymisé. De ce fait, il est suggéré au lecteur d'apprécier la pertinence de la démarche, plus que les résultats chiffrés.

1. Assurance et risque ouragan au Mexique

1. Assurance et risque ouragan au Mexique

Cette première partie a pour but de présenter le contexte et les notions indispensables à la compréhension de ce mémoire.

Dans un premier temps, la situation géographique risquée du Mexique sera décrite. En effet, le pays est situé entre les océans Atlantique et Pacifique, lieu d'activité intense d'ouragans.

Dans un troisième temps, la modélisation des ouragans sera présentée ainsi que les principales métriques permettant d'analyser les résultats des modèles.

Dans un quatrième et dernier temps, un indicateur permettant de quantifier la rentabilité d'un portefeuille de risques ouragan sera présenté : le ratio combiné économique.

1.1 Le Mexique : un pays exposé au risque ouragan

Le Mexique est un pays fortement exposé au risque catastrophes naturelles. Sa situation géographique, entre deux océans et au niveau de la faille de Cocos, le rend vulnérable aux risques climatiques et sismiques.

Cette partie a pour objectif d'expliquer pourquoi le Mexique est exposé au risque catastrophe. Elle sera structurée en trois temps. **Dans un premier temps**, le risque catastrophe et ses enjeux seront définis. **Dans un deuxième temps**, le risque ouragan sera étudié. **Dans un troisième temps**, le profil de risque du Mexique sera présenté.

Dans ce mémoire, seul le risque ouragan sera étudié et le portefeuille tremblement de terre ne sera pas considéré. Cette distinction entre les deux portefeuilles est possible car sur le marché mexicain de l'assurance, chaque péril est couvert indépendamment avec des polices dédiées. On dispose donc d'information détaillée sur les primes ouragan de chaque police du portefeuille.

1.1.1 Le risque catastrophe

Il y a eu près de 317 catastrophes dans le monde en 2019 (Bevere et al., 2020 [3]), dont 202 sont des catastrophes naturelles. Les 115 catastrophes restantes sont d'origine humaine.

Les catastrophes naturelles ont fait 8000 victimes en 2019, un des chiffres les plus faibles depuis les années 1980.

Le code des assurances définit une catastrophe naturelle comme étant « caractérisée par l'intensité anormale d'un agent naturel lorsque les mesures habituelles à prendre pour prévenir ces dommages n'ont pu empêcher leur survenance ou n'ont pu être prises »³.

Il s'agit de phénomènes d'origine climatique (inondations, tempêtes, ouragans, typhons, gel, sécheresses) ou sismique (tremblements de terre, tsunamis...).

Comme on le voit sur les graphiques ci-dessous, il y a eu près de 52 milliards de dollars de pertes assurantielles en 2019, ce qui représente environ 0,17% du PIB mondial. Historiquement, la majorité des pertes assurantielles sont dues à des catastrophes climatiques.

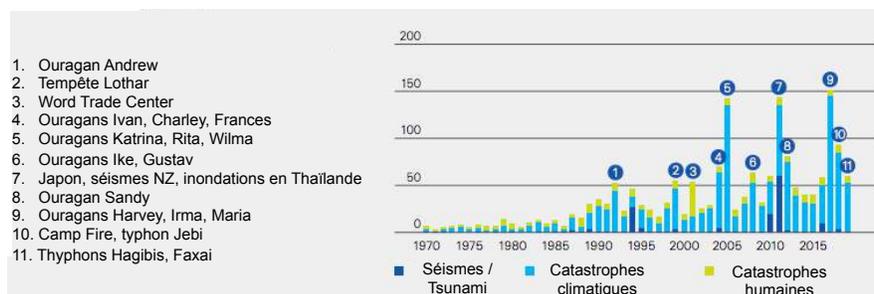


Figure 1 : pertes assurantielles liées aux catastrophes entre 1970 et 2019

Source : Bevere et al., 2020 [3]

Une tendance haussière se dégage depuis les années 1980. Ce phénomène s'explique par la croissance de la population, le développement économique et l'urbanisation, qui ont engendré une forte croissance des expositions.

D'après un rapport de l'ONU, une part de la hausse de la fréquence des catastrophes naturelles peut également être attribuée au changement climatique (Parry, 2020 [4]).

³ Article L. 125-1 alinéa 3 du Code des assurances

1.1.2 Le risque ouragan

Un ouragan est une dépression d'origine tropicale dont la vitesse de vent en rafale a dépassé les 116 km par heure. Elle naît sous l'influence de fortes chaleurs combinées aux eaux très chaudes des latitudes tropicales.

Les ouragans se forment généralement entre juin et novembre dans l'hémisphère nord et entre avril et novembre dans l'hémisphère sud.

Formation d'un ouragan

Comme on le voit sur le schéma ci-dessous, la formation d'un ouragan requiert une température de l'océan supérieure à 26 degrés. L'eau chaude va alors s'évaporer, rendant l'air chaud et humide et le faisant s'élever entre 9000 et 12 500 mètres. Lors de sa montée, une partie de l'air est liquéfiée : des nuages se créent. L'air est progressivement refroidi et redescend. Une fois retourné au niveau de la mer, il se réchauffe et le cycle se répète. De ce fait, chaque cycle entraîne la création de nuages, qui sont de plus en plus nombreux. La rotation de la Terre a pour conséquence de faire tourner les masses d'air et de nuages sur eux-mêmes : c'est la force de Coriolis.

Le système créé n'est pas immédiatement un ouragan. La dénomination dépend de la vitesse du vent.

L'ouragan commence à se dissiper quand il est coupé de sa source de chaleur et d'humidité : lorsqu'il touche la terre ferme ou qu'il atteint des eaux plus froides.

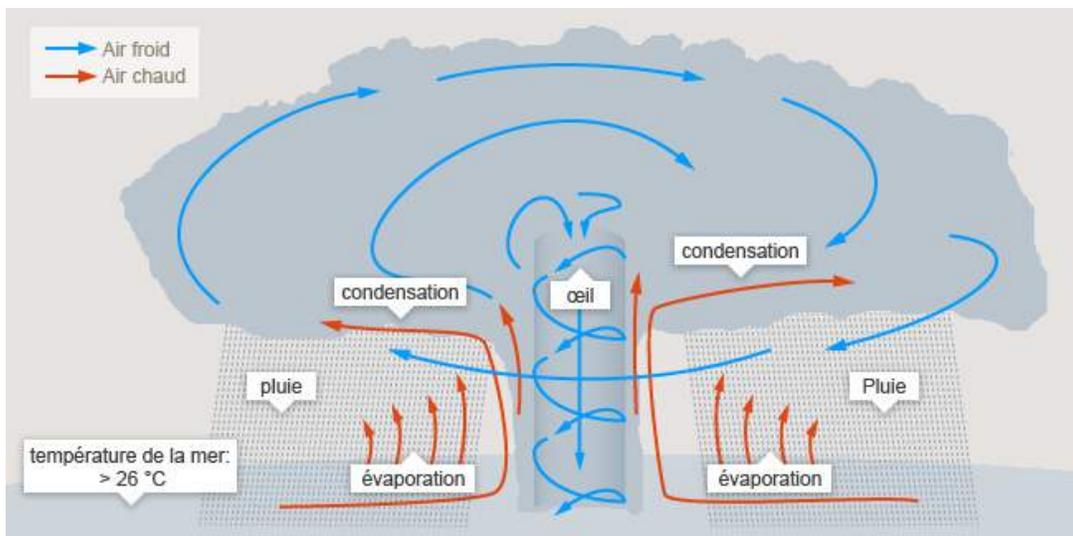


Figure 2 : coupe transversale d'une dépression tropicale

Source : Le Touzé, 2017 [5]

Un ouragan peut avoir plusieurs conséquences physiques :

- **Vents** : un ouragan se caractérise par une vitesse de vent supérieure à 116 km par heure et peut atteindre 270 km par heure avec des rafales à 350 km par heure.
- **Inondations** : du fait de la présence de nuages dans le système, les ouragans peuvent être accompagnés de fortes précipitations. Pendant l'ouragan Harvey, il est par exemple tombé plus de 1,3 mètre de précipitations. Les noyades sont la principale cause de mortalité des ouragans.
- **Submersions marines** : le niveau de la mer peut augmenter de 1 à 7 mètres. Ce phénomène s'explique par la diminution de la pression atmosphérique : le poids de l'air diminue, ce qui entraîne mécaniquement la hausse du niveau de la mer.

Lors de la survenance d'un ouragan, il est parfois difficile de distinguer les dommages engendrés par chacun de ces trois phénomènes.

Evaluation de la capacité destructrice d'un ouragan

Un ouragan est caractérisé par deux paramètres principaux : son intensité et sa durée.

1. Intensité

L'intensité est généralement mesurée à l'aide de l'échelle de Saffir-Simpson. Cette échelle distingue différentes catégories d'ouragan en fonction de la vitesse de vent.

Le tableau ci-dessous présente les différentes catégories ainsi que les dégâts qui leur sont habituellement associés. Pour rappel, pour qu'une dépression soit considérée comme un ouragan, il faut une vitesse de vent supérieure à 116 km par heure.

Intensité de l'ouragan	Vitesse de vent	Augmentation du niveau d'eau	Dégâts potentiels
Dépression tropicale	environ 62km/h		L'ouragan se forme.
Tempête tropicale	entre 62 et 116km/h		La dépression s'est formée et s'intensifie.
Ouragan de catégorie 1	entre 116 et 153km/h	De 1,2 à 1,8 mètres	Les plantes et les maisons mobiles non-ancrées peuvent être endommagées.
Ouragan de catégorie 2	entre 153 et 177km/h	De 1,8 à 2,7 mètres	Des dommages importants sont faits aux plantes et aux maisons mobiles. Quelques toits, fenêtres et portes peuvent être soufflés, mais aucun dommage majeur ne devrait être fait sur les édifices.
Ouragan de catégorie 3	entre 177 et 209km/h	De 2,7 à 4,0 mètres	Les maisons mobiles ainsi que les petites structures peuvent être détruites. Certaines structures plus volumineuses peuvent être endommagées.
Ouragan de catégorie 4	entre 209 et 249km/h	De 4,0 à 5,5 mètres	Des arbres peuvent être déracinés et d'importants dommages peuvent être faits sur les toits, fenêtres et portes.
Ouragan de catégorie 5	supérieur à 249km/h	Plus de 5,5 mètres	Certains édifices peuvent être détruits et d'importants dommages peuvent être faits aux toits, fenêtres et portes. Les toits de plusieurs maisons et édifices peuvent s'effondrer

Figure 3 : échelle de Saffir-Simpson

2. Durée

Un ouragan peut durer d'un jour à un mois, avec une moyenne autour d'une semaine.

La durée d'un ouragan est importante car les contrats d'assurance et de réassurance catastrophe prévoient généralement une clause horaire dans la définition d'un événement, généralement 72 heures (3 jours) ou 168 heures (7 jours).

1.1.3 La situation du Mexique

Le Mexique est situé entre les océans Atlantique et Pacifique, ce qui en fait une zone particulièrement vulnérable au risque ouragan. En effet, le pays est touché à la fois par des dépressions prenant naissance dans l'Atlantique et dans le Pacifique.

La carte ci-après présente les trajectoires des ouragans ayant touché le Mexique entre 1851 et 2019. Les couleurs correspondent aux différentes catégories de l'échelle de Saffir-Simpson.

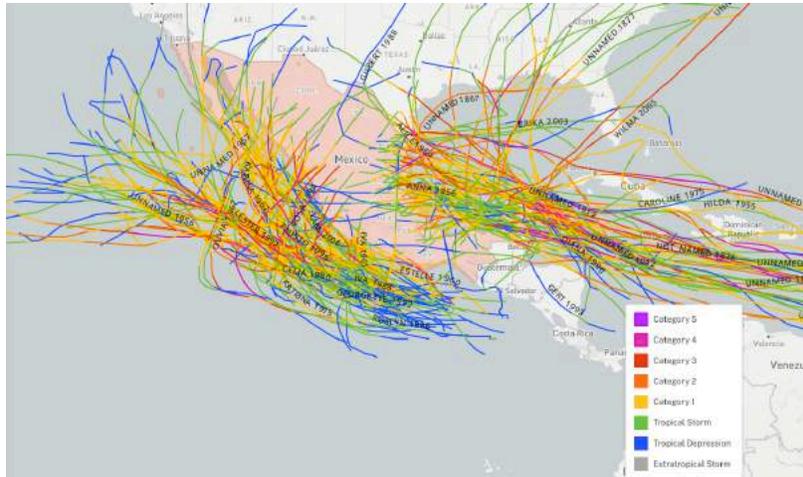


Figure 4 : trajectoires des ouragans ayant touché le Mexique entre 1851 et 2019

Source : NOAA

La carte ci-dessus montre qu'un ouragan peut se déplacer sur plusieurs milliers de kilomètres et changer plusieurs fois de catégorie. Une fois dans les terres, les ouragans se dissipent : les rares trajectoires qui traversent entièrement le Mexique correspondent à des dépressions ou des tempêtes tropicales. Un même ouragan fait donc rarement des dégâts sur les deux côtes du Mexique.

Le Mexique est particulièrement vulnérable au risque ouragan car ses activités et sa population sont concentrées sur les côtes. La carte ci-dessous présente la densité de population du Mexique.

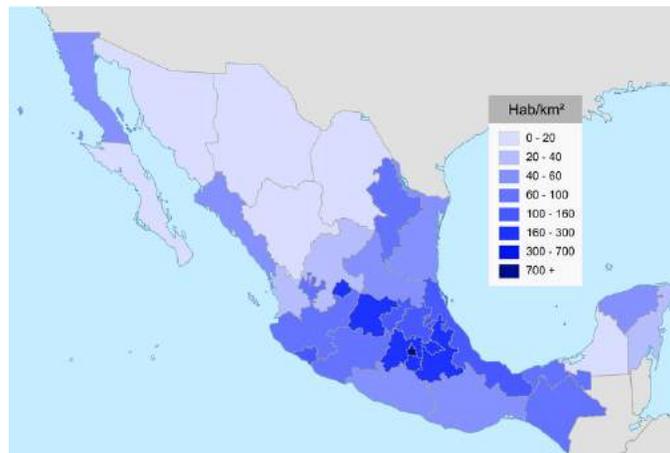


Figure 5 : densité de population au Mexique en 2010

Le Mexique a été touché par des ouragans dévastateurs. Wilma (2005) et Odile (2014) en sont deux exemples. La carte ci-après présente les trajectoires des deux ouragans.

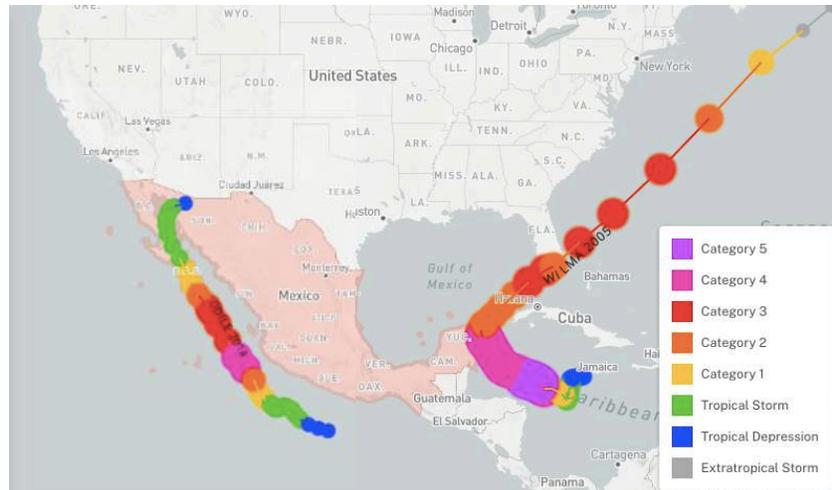


Figure 6 : trajectoires des ouragans Wilma et Odile

Source : NOAA

Wilma a touché la côte atlantique du Mexique, et plus particulièrement l'île de Cozumel et la péninsule du Yucatan. Des dégâts considérables ont affecté les infrastructures du secteur touristique. La marée de tempête, mesurée à 1,8 mètre, a causé de fortes inondations dans les secteurs proches de Cancun.

Odile a touché la côte pacifique du Mexique et plus particulièrement l'Etat de Basse-Californie du Sud. Des nombreux arbres et poteaux électriques ont été abattus, bloquant les routes et privant d'électricité près de 92% de la population de l'Etat.

1.2 Modélisation des ouragans et métriques d'intérêt

Les parties précédentes ont mis en évidence la forte exposition du Mexique au risque ouragan. Dès lors, quelles sont les méthodes disponibles pour modéliser le risque catastrophe et plus particulièrement le risque ouragan ?

Cette partie a pour objectif de présenter les méthodes permettant de modéliser les ouragans. **Dans un premier temps**, les différents modules d'un modèle catastrophe seront présentés. **Dans un second temps**, les métriques obtenues en sortie des modèles seront présentées.

1.2.1 Architecture d'un modèle CAT

Dans le milieu de l'assurance, les risques sont généralement divisés en plusieurs catégories, chacune étant habituellement modélisée indépendamment :

- **Les risques attritionnels** correspondent à des risques individuels fréquents et peu sévères. Ils correspondent à la majeure partie de l'activité d'assurance non-vie.
- **Les risques atypiques** représentent des risques individuels peu fréquents mais ayant une sévérité importante. Un seuil est généralement défini pour distinguer les sinistres atypiques des sinistres attritionnels.
- **Les risques CAT** correspondent à des risques collectifs causés par une catastrophe de forte ampleur. Ces risques sont caractérisés par une faible fréquence et une forte sévérité. Il peut s'agir de catastrophes naturelles ou humaines.

Contrairement aux risques attritionnels et atypiques, les méthodes statistiques habituelles ne sont pas adaptées à la modélisation du risque catastrophe naturelle. Cela s'explique par la profondeur historique des sinistres, trop faible pour être réellement représentative du risque sous-jacent. Dès lors, une connaissance physique du risque est requise pour modéliser ce type de risque : c'est ce qu'on appelle les « modèles CAT ». Des événements fictifs mais probables sont générés et viennent compléter la vision historique.

En raison de la complexité des modèles et de l'investissement qu'ils représentent, beaucoup de compagnies d'assurance et de réassurance utilisent des modèles conçus par des entreprises externes. Le marché est actuellement dominé par trois acteurs principaux : Risk Management Solution (RMS), AIR Worldwide et EQECAT. On note cependant un effort croissant des compagnies pour développer leurs propres modèles CAT. On parle alors de modèle *in-house* (maison).

Comme on le voit sur le schéma ci-dessous, un modèle CAT se compose généralement de quatre modules interdépendants : les modules exposition, aléa, vulnérabilité et financier.

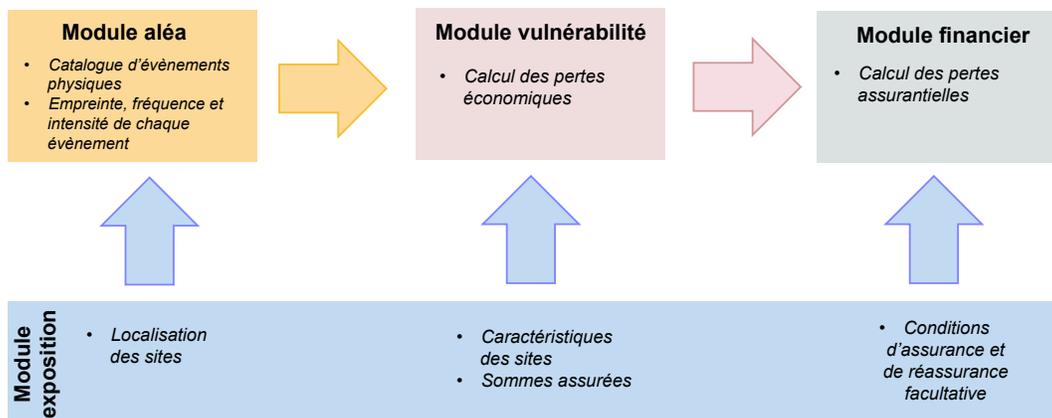


Figure 7 : architecture d'un modèle CAT

Module exposition

Le module exposition apporte des informations sur le portefeuille modélisé. Il sera utilisé dans tous les autres modules du modèle CAT.

Le module exposition contient généralement les informations suivantes :

- **Localisation des sites assurés** : la localisation de chaque site du portefeuille est donnée, ainsi que son niveau de précision (longitude/latitude, rue, ville, région, pays...). Comme on le verra par la suite, ces informations permettent de faire le lien avec le module aléa.
- **Caractéristiques physiques des sites assurés** : des informations sont données sur les matériaux de construction, le type de toit, l'année de construction, le nombre d'étages, etc. Ces données seront utilisées par le module vulnérabilité pour déterminer un taux de destruction.
- **Valeur des sites assurés** : il s'agit de la valeur de remplacement du bâtiment si celui-ci était détruit à 100%. Cette information sera croisée avec le taux de destruction pour obtenir la perte économique.
- **Conditions financières** : des informations sont données sur les conditions d'assurance et de réassurance facultative des sites du portefeuille. Ces données seront utilisées par le module financier pour convertir la perte économique en perte assurantielle.

A noter que toutes ces informations ne sont pas toujours disponibles. De ce fait, les modèles CAT disposent généralement de modules de gestion des valeurs manquantes. La qualité des données est importante : moins elle est bonne et plus il y aura d'incertitudes sur les résultats des modèles.

Module aléa

Ce module a pour objectif de créer un catalogue d'événements pouvant toucher le portefeuille. Le catalogue peut se composer d'événements historiques et d'événements fictifs mais probables.

Chaque événement du catalogue est défini par :

- **Une fréquence** : il s'agit de la fréquence annuelle de survenance de l'événement.
- **Une empreinte géographique** : il s'agit de la représentation géographique des zones qui ont été touchées par l'événement. La figure 14 donne un exemple d'empreinte : il s'agit de l'empreinte de l'ouragan Irma, survenu en 2017.
- **Une intensité** : en chaque point de l'empreinte, on dispose de l'intensité de l'ouragan, exprimée en vitesse de vent.

Si on croise le module aléa avec le module exposition, on sait, pour chaque événement du catalogue, quels sites du portefeuille ont été touchés et avec quelle intensité.

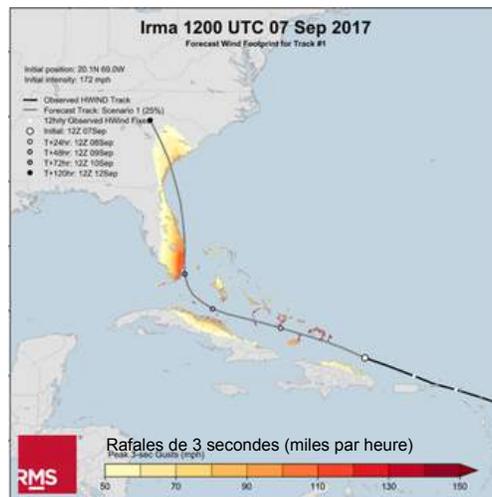


Figure 8 : empreinte de l'ouragan Irma, Source : RMS [6]

Module vulnérabilité

Le module vulnérabilité a pour objectif de traduire un événement en perte économique pour un portefeuille. Pour chacun des sites impactés par un événement, un taux de destruction est calculé et est ensuite multiplié par la somme assurée.

Le taux de destruction d'un site A pour un événement B se définit comme :

$$\frac{\text{Dommages subis par le site A suite à l'événement B}}{\text{Somme assurée totale du site A}}$$

En combinant le module aléa et exposition, on sait quels sites sont touchés par un événement et avec quelle intensité. On connaît également les caractéristiques de chacun des sites (type de toit, matériaux de construction, etc). Le module vulnérabilité traduit l'intensité en taux de destruction, en fonction des caractéristiques de chaque site.

Les courbes de vulnérabilité sont calibrées pour chaque caractéristique : pour un même niveau d'intensité, deux bâtiments peuvent être détruits très différemment. Par exemple, un *mobile home* peut être totalement détruit lors du passage d'un ouragan de catégorie 3 alors qu'un immeuble en béton armé ne sera pas endommagé.

Le module vulnérabilité se compose donc d'un ensemble de courbes qui donnent un taux de destruction en fonction de l'intensité. La figure ci-dessous présente un exemple de courbes de vulnérabilité pour trois matériaux de construction différents.

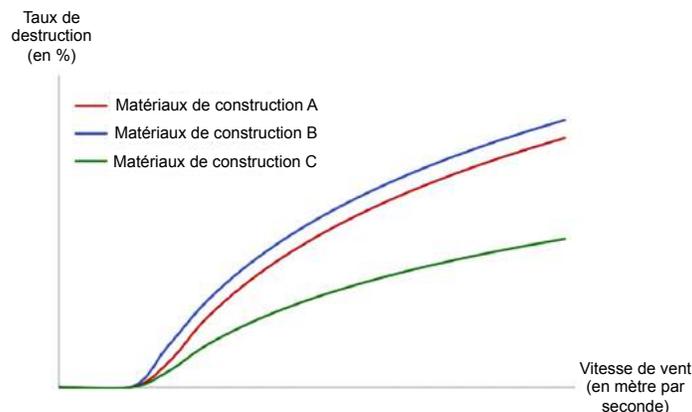


Figure 9 : exemple de courbes de vulnérabilité

Dans le module vulnérabilité, on calcule un taux de destruction puis une perte économique pour chaque événement du catalogue stochastique et pour chaque site du portefeuille touché par ces événements. Les pertes obtenues sont appelées FGU (*From Ground Up*) ou « pertes au premier euro ».

Module financier

Le module financier permet de convertir les pertes économiques en pertes assurantielles puis en pertes nettes. Ce module consiste à appliquer les différentes conditions financières d'assurance (franchises, limites et co-assurance) et de réassurance (facultatives proportionnelles et non proportionnelles).

La principale difficulté de ce module réside dans l'ordre d'application des différentes conditions financières, notamment pour les polices multi-sites : ordre d'application au niveau des garanties, des sites et de la police.

1.2.2 Résultats des modèles et métriques d'intérêt

Une fois que les différents modules du modèle CAT ont été calculés, on obtient un montant de pertes pour chaque événement du catalogue stochastique.

L'objectif de cette partie est de présenter le format habituel des résultats du modèle puis d'étudier les principales métriques d'intérêt.

1.2.2.1 Sorties des modèles

Dans cette section, le format habituel des résultats des modèles CAT sera présenté, ainsi que leur granularité et les différentes visions du risque disponibles.

YELT (Year Event Loss Table) : Table Année Evénement Sinistre

Les résultats de modèle CAT se présentent généralement sous la forme d'une table YELT. Cette table indique, pour chaque année simulée, les événements qui ont été générés ainsi que leurs pertes associées.

Généralement, entre 10 000 et un million d'années indépendantes sont simulées : chaque année simulée est une réalisation possible de la sinistralité à horizon un an.

Le tableau ci-dessous présente un exemple de table YELT.

Année	Événement	Perte
1	A	300
1	B	100
1	C	2
3	D	15
4	A	350
4	F	1
...
10 000	Z	25

Figure 10 : exemple d'une table YELT

Le nombre d'événements varie selon les années simulées : il y a trois événements en année 1 mais aucun en année 2.

Par ailleurs, un même événement peut être simulé plusieurs fois : en année 1 et 4, le même événement a été généré (événement A). En fonction de sa fréquence, un même événement peut ainsi être généré plusieurs fois au cours des années simulées.

A noter qu'un même événement touchant deux fois le portefeuille peut engendrer un montant de pertes différent. C'est par exemple le cas pour l'événement A, qui génère une perte de 300 en année A contre 350 en année 4. Cela s'explique par le fait que la vulnérabilité est stochastique : un même événement, simulé deux fois, ne va pas détruire de la même façon les sites situés dans son empreinte.

Granularité de l'YELT

En sortie des modèles CAT, il est en théorie possible d'obtenir l'YELT à la granularité par site : pour chaque année et chaque événement du portefeuille, on connaît le montant de pertes par site. Ce niveau de granularité demande de manipuler des tables de grande taille : pour chaque site du portefeuille, on dispose de 10 000 voire un million d'années de sinistralité possible. De ce fait, les YELT sont généralement agrégées au niveau du portefeuille ou de la branche d'assurance. En effet, les modèles CAT étant habituellement utilisés à des fins de *Risk Management*, le niveau de granularité correspond généralement à la granularité du traité de réassurance.

Vision du risque

Nous avons vu précédemment que le module financier permettait de convertir une perte économique en perte assurantielle.

Dans la table d'YELT, plusieurs visions du risque sont généralement disponibles. L'objectif est de différencier les conditions d'assurance et de réassurance pour pouvoir quantifier leur impact respectif. Trois visions sont généralement présentes :

- **Les pertes *GUL* (*Ground Up Loss*)** : il s'agit des pertes économiques au premier euro.
- **Les pertes *gross*** : il s'agit des pertes économiques auxquelles les conditions d'assurance ont été appliquées (franchise, limite et coassurance).
- **Les pertes *net*** : il s'agit des pertes *gross* auxquelles les éventuelles conditions de réassurance facultative ont été appliquées (facs proportionnelles et non proportionnelles).

En conclusion, les résultats d'un modèle CAT se présentent généralement sous la forme d'une table d'YELT. Pour limiter la taille de la table, les résultats sont habituellement agrégés au niveau de granularité des traités de réassurance. Différentes visions de la sinistralité sont disponibles (*GUL*, *gross* et *net*) et permettent de quantifier l'impact des conditions d'assurance et de réassurance.

1.2.2.2 Métriques d'intérêt

A partir de la table d'YELT, il est possible de calculer différentes métriques pour analyser les résultats obtenus. Cette partie a pour objectif de présenter les métriques les plus couramment utilisées : l'AAL, l'écart type, les *EP curves* et les *Loss Frequency Curves*.

L'AAL (*Average Annual Loss*)

L'AAL correspond aux pertes annuelles moyennes. Soit n le nombre d'années simulées, n_i le nombre d'événements de l'année i et $S_{i,j}$ le montant du sinistre pour l'événement j de l'année i . L'AAL se calcule comme suit :

$$AAL = \frac{\sum_{i=1}^n \sum_{j=1}^{n_i} S_{i,j}}{n}$$

Cet indicateur correspond à la sinistralité moyenne attendue sur l'année. Il peut être calculé à différents niveaux de granularité : par site, par police, par compte, voire au niveau du portefeuille entier. Il peut être calculé pour les pertes *GUL*, *gross* ou *net*.

A noter que l'AAL quantifie la sinistralité annuelle moyenne et non le montant moyen d'un sinistre.

L'écart type et coefficient de variation

L'écart type représente la variation moyenne des pertes autour de l'AAL, normalisée par le nombre d'années simulées. Plus la valeur est élevée et plus la dispersion des pertes probables autour de l'AAL est grande.

On note : $SD = \sqrt{\frac{\sum_{i=1}^n (\sum_{j=1}^{n_i} S_{i,j} - AAL)^2}{n}}$

Pour comparer la volatilité de deux portefeuilles n'ayant pas le même AAL, l'écart type n'est pas une mesure pertinente. En effet, un portefeuille peut avoir un écart type plus élevé qu'un autre mais être en réalité moins volatile. Fort de ce constat, le coefficient de variation sera préféré car il permet de comparer la volatilité de deux portefeuilles n'ayant pas le même AAL. Il se calcule comme suit : $\frac{SD}{AAL}$.

EP Curves : *Exceedance Probability Curves*

EP curves (*Exceedance Probability Curves*) signifie « courbes de probabilité de dépassement ». Ces courbes mesurent la probabilité qu'une charge de sinistre excède un certain seuil. Elles sont couramment utilisées pour évaluer l'exposition d'un portefeuille d'assurance à un péril sous-jacent. Il existe deux principales *EP curves* :

- **L'AEP (*Aggregate Exceedance Probability*)** : il s'agit de la probabilité que la charge annuelle des sinistres dépasse un certain seuil. On note $AEP(x) = \mathbb{P}(\sum_{j=1}^{n_i} S_j \geq x)$.

- **L'OEP (Occurrence Exceedance Probability)** : il s'agit de la probabilité que le plus grand sinistre annuel dépasse un certain seuil. On note $AEP(x) = \mathbb{P}(\max_{1 \leq j \leq n_i} S_j \geq x)$.

L'OEP et l'AEP sont généralement définis en fonction d'une période de retour plutôt qu'en terme de probabilité. La période de retour correspond à la périodicité entre deux occurrences d'un événement supérieur ou égal à un seuil. Une période de retour peut être déduite d'une probabilité et inversement :

$$\text{Période de retour} = \frac{1}{\text{probabilité de dépassement}}$$

Par exemple, si la probabilité d'avoir un sinistre supérieur ou égal à un million d'euros est de 0,5%, cela signifie qu'un sinistre d'un million d'euros a une période de retour de $\frac{1}{0,5\%} = 200$ ans.

Soient F_{Agg} et F_{Max} les fonctions de répartition des pertes annuelles et maximales d'un portefeuille donné. Pour une période de retour PR donnée, on a :

$$AEP(PR) = F_{Agg}^{-1}\left(1 - \frac{1}{PR}\right)$$

$$OEP(PR) = F_{Max}^{-1}\left(1 - \frac{1}{PR}\right)$$

La période de retour donne une indication sur la probabilité annuelle d'occurrence d'un événement : un ouragan de période de retour 5 ans qui a eu lieu il y a quatre ans ne se produira pas avec certitude l'année prochaine. Il a la même probabilité annuelle d'occurrence, qu'il ait eu lieu il y a quatre ans ou l'année dernière.

Le graphe ci-dessous présente un exemple de courbes d'OEP et d'AEP.

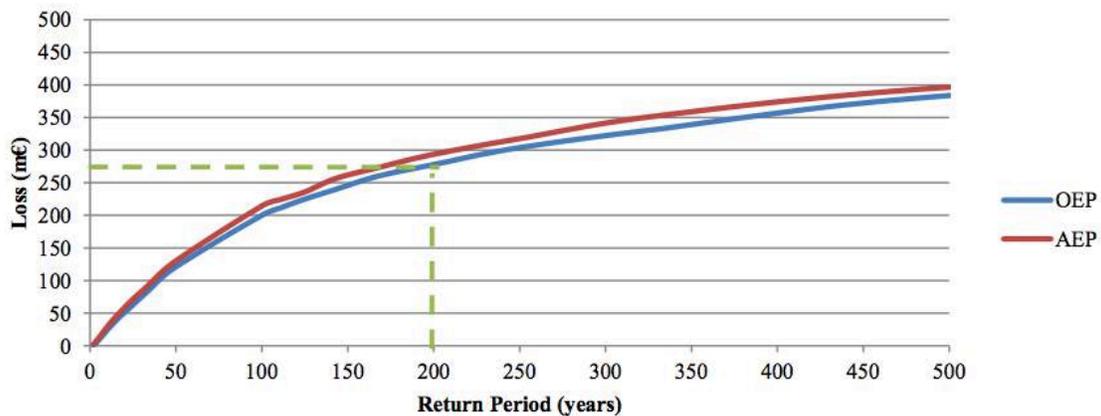


Figure 11 : exemple de courbes d'OEP et d'AEP

Ce graphe s'interprète de la manière suivante : en moyenne tous les 200 ans, le portefeuille subit un sinistre supérieur ou égal à 275 euros et une perte annuelle supérieure ou égale à 300 euros.

L'OEP est toujours inférieure ou égale à l'AEP : cela s'explique par le fait que la perte causée par l'événement le plus sévère de l'année est toujours inférieure ou égale à la sinistralité totale de l'année considérée.

Les *EP curves* peuvent être utilisées pour calibrer la réassurance. Les courbes d'OEP sont généralement utilisées pour calibrer la réassurance par événement (excédents de sinistre par événement) : la limite du traité va être choisie de telle sorte qu'elle couvre les pertes potentielles du portefeuille à une période de

retour donnée. Les courbes d'AEP peuvent aider à structurer les traités annuels (traités en excédents de pertes ou limite annuelle d'une quote-part) : ceux-ci sont calibrés sur la base de la sinistralité annuelle totale.

Par ailleurs, la forme des *EP curves* donne des informations sur le type de périls couverts ainsi que sur la fréquence et la sévérité des événements. C'est par exemple le cas pour les deux courbes d'AEP présentées dans la figure ci-dessous.

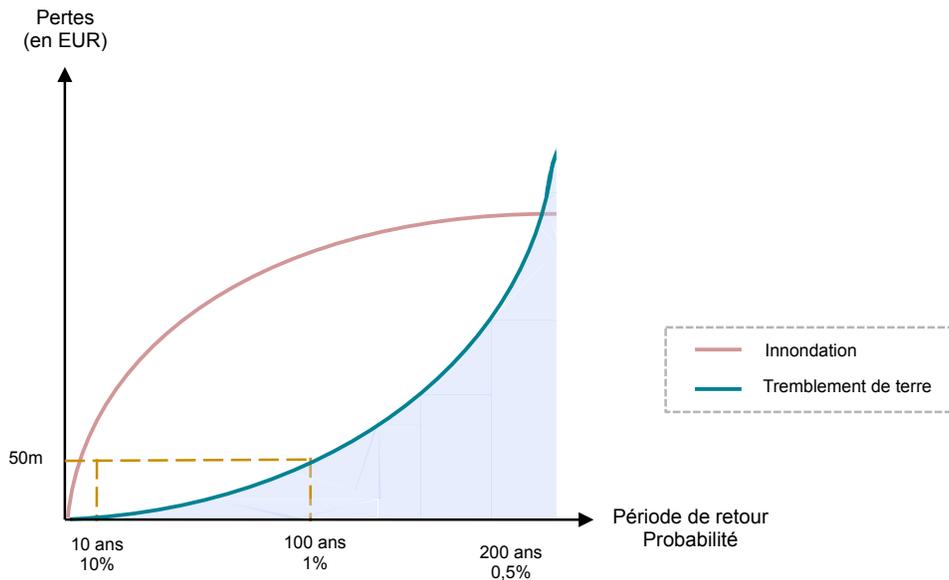


Figure 12 : courbes d'AEP de deux périls

La courbe bleue est convexe : cela signifie que le montant de l'AAL est expliqué principalement par les pertes des hautes périodes de retour. Cette forme d'AEP correspond généralement à des périls peu fréquents, comme par exemple les tremblements de terre.

A contrario, la courbe rose est concave : cela signifie que le montant de l'AAL s'explique principalement par les basses périodes de retour. Ce type de courbe est généralement associé à des périls de fréquence, comme par exemple les inondations ou les tempêtes. Ces périls se caractérisent par une majorité de périls fréquents et d'intensité relativement faible.

1.3 Le ratio combiné économique : un indicateur économique de performance de la souscription

Les parties précédentes ont présenté l'intérêt de la réassurance ainsi que les modèles permettant de quantifier le risque catastrophe. Grâce à ces outils, il est possible de modéliser la sinistralité du portefeuille. Cette partie a pour objectif de présenter une métrique permettant de quantifier la rentabilité d'un portefeuille d'assurance catastrophe : il s'agit du ratio combiné économique, aussi appelé ECR (*Economic Combined Ratio*).

Le ratio combiné économique permet de donner une vision économique de la performance de la souscription du point de vue de l'assureur. Comme on le voit sur le schéma ci-dessous, il prend en compte les sinistres bruts, les coûts de réassurance, les coûts du capital, les impôts et les frais.

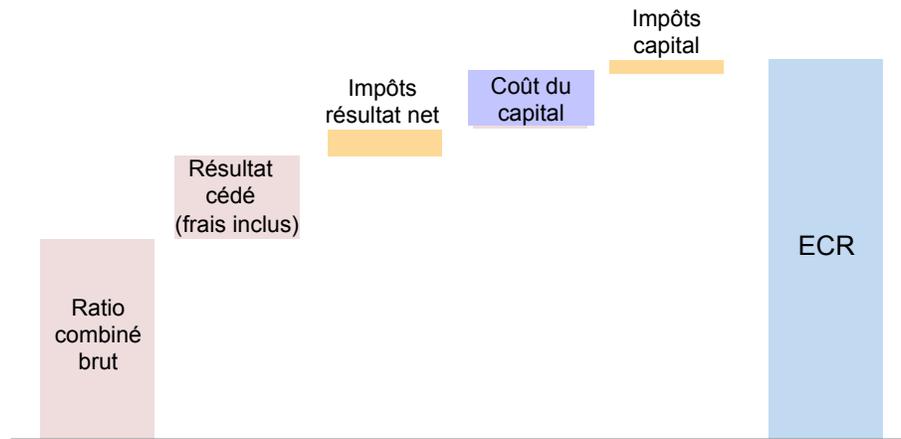


Figure 13 : illustration de l'ECR

Le ratio combiné économique permet de piloter le résultat technique. Cet indicateur quantifie l'impact d'un euro de prime souscrite sur les fonds propres, après rémunération des actionnaires. Il est donc souhaitable d'avoir un ratio inférieur à 100%.

Il existe deux principaux types d'ECR :

- L'ECR prospectif, calculé sur la base du modèle pour les branches CAT et sur le budget pour les autres branches.
- L'ECR rétrospectif, calculé sur la base du *Loss Ratio* réalisé.

Dans ce mémoire, le ratio combiné économique considéré est l'ECR prospectif. Le portefeuille étant composé de polices couvrant le risque ouragan, l'ECR est calculé sur la base des résultats de modélisation CAT.

Cette partie a pour objectif de présenter le ratio combiné économique. Pour ce faire, **dans un premier temps**, le ratio combiné sera présenté. **Dans un deuxième et troisième temps**, le coût du capital et les impôts seront étudiés. Enfin **dans un quatrième et cinquième temps**, le ratio combiné économique et la création de valeur économique seront présentés.

1.3.1 Ratio combiné

Comme on le voit sur le schéma ci-dessous, le ratio combiné correspond au rapport entre les décaissements de l'assureur et ses encaissements.

Contrairement au ratio de sinistres sur primes, le ratio combiné prend en compte la sinistralité, mais également les frais d'exploitation. Les frais pris en compte sont de deux natures : internes et externes.

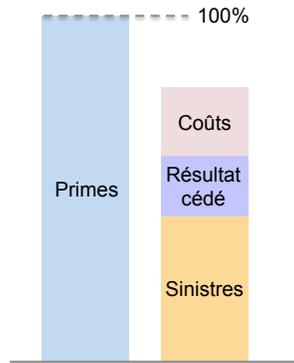


Figure 14 : illustration du ratio combiné

Ratio combiné brut de réassurance

Le ratio combiné brut quantifie la rentabilité du portefeuille avant prise en compte de la réassurance obligatoire. Les éléments suivants sont considérés :

1. Sinistralité brute

Par sinistralité brute, on entend sinistralité nette de facts mais brute des récupérations du traité.

2. Frais internes

Les frais internes correspondent aux frais généraux. Il s'agit principalement des frais de gestion et d'administration. Ils sont alloués au portefeuille et sont fixes.

3. Frais externes

Les frais externes se composent principalement des frais d'acquisition et de gestion. Ces frais correspondent aux frais engagés pour rémunérer les délégataires : commissions versées aux apporteurs d'affaires, rémunérations des courtiers gestionnaires, etc.

Ces frais sont variables et se calculent sur la base des primes souscrites.

Le ratio combiné se calcule comme suit :

$$CoR_{brut} = \frac{S + \pi * \sigma^{brut} + f^{brut}}{\pi}$$

$$avec \left\{ \begin{array}{l} \pi : \text{primes brutes} \\ S : \text{sinistres bruts} \\ \sigma^{brut} : \text{chargement au titre des frais externes} \\ f^{brut} : \text{montant de frais internes} \end{array} \right.$$

Le ratio combiné brut ne donne cependant qu'une image incomplète de la rentabilité de la souscription. En effet, la réassurance obligatoire n'est pas prise en compte. Pour remédier à cet écueil, il est nécessaire de calculer un ratio combiné net de réassurance.

Ratio combiné net de réassurance

Pour passer du ratio combiné brut au ratio combiné net, il est nécessaire de prendre en compte le résultat cédé, mais également les frais. Ceux-ci sont de deux types : externes et internes.

1. Résultat cédé

Le périmètre du résultat cédé est celui de la réassurance obligatoire : la réassurance facultative a déjà été prise en compte dans le ratio combiné brut.

On parle de résultat cédé car la réassurance obligatoire entraîne habituellement une détérioration du résultat moyen : le ratio combiné net moyen est généralement supérieur au ratio combiné brut moyen. Cela s'explique par le fait que la prime de réassurance est supérieure aux récupérations moyennes générées par le traité. En effet, le réassureur demande une rémunération au titre de la volatilité, de ses frais, etc.

Le résultat cédé se calcule comme la différence entre les primes de réassurance et les récupérations moyennes générées par le traité. Les primes de réassurance se composent de la prime *upfront* ainsi que des éventuelles primes de reconstitution.

2. Frais externes

Les frais externes sont les frais de courtage. En effet, sur le marché de la réassurance, une grande partie des traités est placée par l'intermédiaire de courtiers. Ceux-ci font l'intermédiaire entre les cédantes et les réassureurs. Il est souvent nécessaire de passer par un courtier car il peut y avoir jusqu'à plusieurs dizaines de réassureurs sur un même traité. Le courtier centralise les questions des différents réassureurs et négocie les prix et les conditions avec eux.

Le courtier est généralement rémunéré par une commission dont l'assiette est la prime du traité. La prime considérée est généralement la prime *upfront* : la commission ne s'applique pas sur les éventuelles primes de reconstitution.

3. Frais internes

Les frais internes sont les frais généraux et les frais d'administration. Ces frais sont un montant fixe, correspondant aux frais générés par la structuration, le placement, la gestion courante et la gestion des sinistres des traités.

Le ratio combiné net se déduit à partir du ratio combiné brut, auquel s'ajoute le résultat cédé ainsi que les frais internes et externes :

$$CoR_{net} = CoR_{brut} + \frac{P_{upfront} * (1 + \sigma^{réass}) + P_{rec} - R + f^{réass}}{\pi}$$

$$\text{avec } \left\{ \begin{array}{l} P_{upfront} : \text{prime de réassurance upfront} \\ P_{rec} : \text{primes de reconstitution} \\ R : \text{récupérations moyennes} \\ \sigma^{réass} : \text{chargement au titre des frais externes de réassurance} \\ f^{réass} : \text{frais fixes liés à la réassurance} \end{array} \right.$$

Ce ratio quantifie la charge nette de l'assureur pour un euro de prime souscrite. Pour prendre en compte toute la charge de l'assureur, il est cependant nécessaire d'intégrer un autre élément : le coût du capital, qui traduit l'exigence de rémunération des actionnaires.

1.3.2 Coût du capital

Suite à la mise en vigueur de la directive Solvabilité II⁴, les assureurs doivent désormais disposer de fonds propres leur permettant de faire face à la plus importante déviation de sinistralité susceptible d'avoir lieu dans les deux cents prochaines années.

Ces fonds propres sont placés sur des actifs très sûrs, rémunérés généralement au taux sans risque. Les actionnaires ont une exigence de rémunération supérieure à celle du taux sans risque et immobiliser du capital coûte donc de l'argent à l'assureur.

On appelle le *spread* la différence entre le taux de rémunération attendu par les actionnaires et le taux sans risque généré par les fonds propres. Le coût du capital se calcule comme suit :

$$\text{Coût du Capital} = \text{Fonds propres} * \text{Spread}$$

Qu'entend-on exactement par « capital immobilisé » ? Pour comprendre cette notion, il est nécessaire de s'intéresser à la composition du bilan Solvabilité II.

Le bilan Solvabilité II

Dans les sociétés d'assurance, le bilan est évalué tous les trimestres à l'occasion de la clôture.

La figure ci-dessous présente la composition du bilan Solvabilité II. Face aux actifs, valorisés en valeur de marché à la date de clôture, on trouve trois éléments de passif :

- **le BE (*Best Estimate*)** : la meilleure estimation des réserves. Il s'agit de l'espérance des cash-flows des futurs passifs actualisés avec la courbe des taux sans risque
- **la Risk Margin** : la marge pour risque, évaluée à partir du coût d'immobilisation du capital si le portefeuille était en run-off.
- **Les EOF (*Eligible Own Funds*)** : les fonds propres après vérification des critères d'éligibilité. Il s'agit du montant de capital immédiatement disponible afin d'honorer les engagements en cas de contexte défavorable.

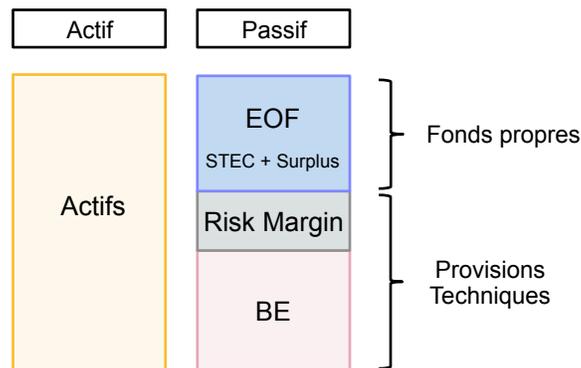


Figure 15 : exemple d'un bilan en vision solvabilité II

Comme on le voit sur la figure ci-dessus, les fonds propres correspondent aux EOF. Ceux-ci se composent du STEC et du surplus. Le STEC (*Short Term Economic Capital*) est le SCR (*Solvency Capital Requirement*) en modèle interne AXA.

⁴ Solvabilité II est une directive européenne et ne s'applique donc pas au Mexique. Elle est utilisée dans ce mémoire car c'est un cadre conceptuel intéressant et car AXA est un groupe européen qui reporte sous Solvabilité II.

Le STEC correspond au montant de fonds propres nécessaires pour ne pas être ruiné à horizon un an avec une probabilité de 99,5%. En d'autres termes, il s'agit du montant de fonds propres permettant de faire face à une déviation bicentenaire du résultat par rapport au budget. Le STEC se calcule comme suit :

$$STEC = Var_{99,5\%}(Résultat) - E(Résultat)$$

Une fois le STEC obtenu, le surplus se déduit à partir des différents éléments du bilan :

$$Surplus = Actifs - Risk Margin - BE - STEC$$

Le surplus permet également de calculer le ratio de solvabilité :

$$Ratio\ de\ solvabilité = \frac{Surplus + STEC}{STEC}$$

Ce ratio permet de quantifier le nombre de chocs bicentennaires pouvant être absorbés compte tenu des fonds propres disponibles.

Poids du portefeuille étudié dans le STEC

Les EOF se composent du STEC total et du surplus. Comme on le voit dans les schémas ci-dessous, dans le modèle interne AXA, le STEC total est l'agrégation de cinq STEC : Marché, Crédit, Vie, IARD et Opérationnel.

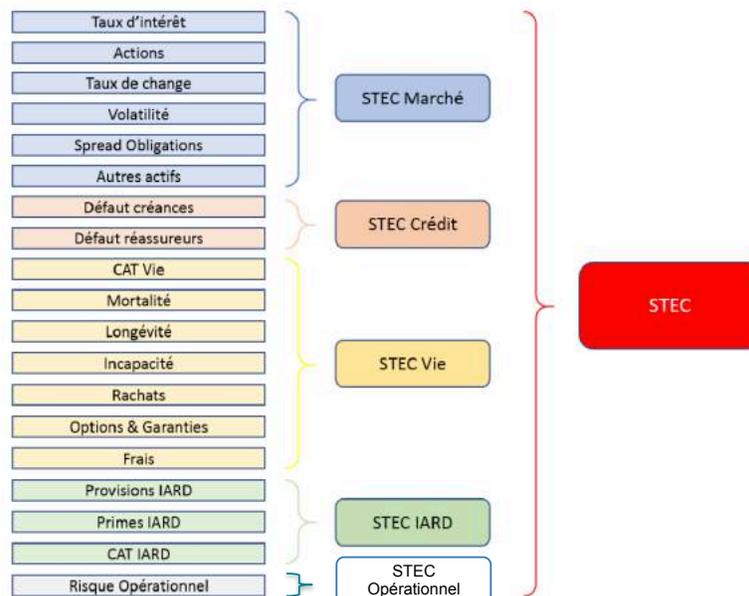


Figure 16 : cartographie des risques dans le modèle interne d'AXA

Source : Benhamou, 2018 [7]

Le portefeuille de notre étude étant composé de risques ouragan mexicain, on considère qu'il n'impacte matériellement que le STEC IARD. Celui-ci se compose de trois risques :

- **STEC CAT** : risque de déviation de la sinistralité CAT à horizon un an. Le portefeuille de notre étude, couvrant le risque d'ouragan, impacte matériellement ce STEC.
- **STEC Réserve** : risque de déviation des réserves à horizon un an. 95% des sinistres du portefeuille de notre étude sont réglés en première année : la contribution du portefeuille au STEC réserve est donc considérée comme non matérielle.

- **STEC Prime** : risque de déviation de la sinistralité non CAT à horizon un an. Le portefeuille de notre étude ne contribue donc pas à ce STEC.

Seul le STEC CAT est donc matériellement impacté par le portefeuille étudié dans ce mémoire.

Lorsque les différents modules de risques sont agrégés, les STEC se diversifient : le STEC total est inférieur à la somme des STEC. L'effet de diversification s'explique par le fait que les modules de risque ont des moteurs de risque différents : tous les scénarios défavorables ne se produisent pas en même temps. On appelle « STEC diversifiés » les STEC auxquels une partie de la diversification a été allouée : la somme des STEC diversifiés correspond au STEC total.

Le poids du portefeuille ouragan dans le STEC total correspond donc au STEC CAT diversifié dans le STEC total.

Conclusion et calcul du coût du capital

Le poids du portefeuille ouragan dans les EOF correspond au STEC CAT diversifié multiplié par le ratio de solvabilité. Le coût du capital se calcule donc comme suit :

$$\text{Coût du Capital} = \text{STEC}_{\text{CAT}}^{\text{Diversifié}} * \text{Ratio de solvabilité} * \text{Spread}$$

Pour rappel, le *spread* est un montant forfaitaire : c'est la différence entre le rendement des fonds propres attendu par les actionnaires et le taux sans risque généré par les fonds propres.

1.3.3 Les impôts

Le calcul de l'ECR prend en compte deux sources d'imposition : les impôts sur le résultat technique et sur les résultats financiers des placements du capital.

Impôts sur le résultat technique du portefeuille

La première imposition est basée sur le résultat net du portefeuille. Soit τ le taux d'imposition, les impôts se calculent comme suit :

$$\text{Impôts} = \tau * \pi * (1 - \text{CoR}_{\text{net}})$$

Ce montant est du même signe que le résultat. En effet, dans le cas où le résultat est négatif, deux cas de figures se présentent :

- **Le résultat global de l'assureur est positif** : le déficit permet de faire diminuer la base imposable globale et entraîne donc un gain d'impôt. C'est le cas dans notre étude.
- **Le résultat global de l'assureur est négatif** : il est possible de reporter le déficit sur les années suivantes, ce qui vient réduire leur base imposable et entraîne donc un gain d'impôt.

Le résultat cédé induit une dégradation du résultat : la prime de réassurance est supérieure aux récupérations moyennes générées par le traité car le réassureur demande une rémunération au titre de la volatilité, de ses frais, etc. De ce fait, la réassurance permet donc de réduire le montant d'impôts payés :

$$\text{Economie}_{\text{impôt}} = \tau * \pi * (1 - \text{CoR}_{\text{net}} - (1 - \text{CoR}_{\text{brut}}))$$

$$\text{Economie}_{\text{impôt}} = \tau * \pi * (-\text{CoR}_{\text{net}} + \text{CoR}_{\text{brut}})$$

$$\text{Economie}_{\text{impôt}} = \tau * \pi * \left(-(\text{CoR}_{\text{brut}} + \frac{P_{\text{upfront}} * (1 + \sigma^{\text{réass}}) + P_{\text{rec}} - R + f^{\text{réass}}}{\pi}) + \text{CoR}_{\text{brut}} \right)$$

$$\text{Economie}_{\text{impôt}} = \tau * (P_{\text{upfront}} * (1 + \sigma^{\text{réass}}) + P_{\text{rec}} - R + f^{\text{réass}})$$

Impôts sur le résultat financier

Si les fonds propres apportent un rendement financier, ces rendements sont alors assujettis à l'impôt sur les plus-values. Soit φ le taux de rendement financier, le montant d'impôts se calcule comme suit :

$$\text{Impôts} = \tau * STEC_{CAT}^{Diversifié} * \text{Ratio de solvabilité} * \varphi$$

A noter que dans le contexte de taux bas, les fonds propres peuvent avoir des rendements financiers proches de zéro voire négatifs.

1.3.4 Le ratio combiné économique

Le ratio combiné économique fournit une vision économique de la performance de souscription. Il prend en compte tous les éléments présentés dans les parties précédentes :

$$ECR = CoR_{net} + \tau * (1 - CoR_{net}) + \frac{\text{Coût du Capital} + \tau * STEC_{CAT}^{Diversifié} * \text{Ratio de solvabilité} * \varphi}{\pi}$$

Le schéma ci-dessous illustre les différentes composantes de l'ECR.

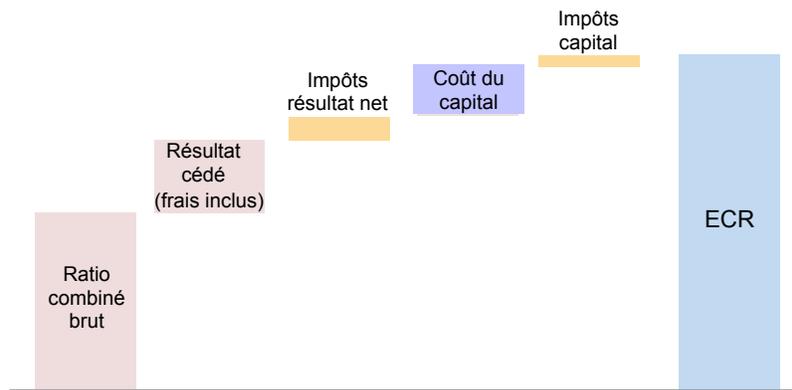


Figure 17 : illustration de l'ECR

A noter que le ratio combiné économique prend habituellement en compte l'effet temps : les différents flux sont actualisés. Dans ce mémoire, l'effet temps n'a pas été pris en compte car les sinistres du portefeuille étudiés sont réglés à 95% en première année. L'effet d'actualisation n'a pas un impact matériel sur les résultats.

1.3.5 La création de valeur économique

Le pendant de l'ECR en montant est appelé « création de valeur économique ». Cet indicateur vaut $ECR * \pi$.

2. Analyse du portefeuille et impact d'une politique de résiliation sur le ratio combiné économique

Avec un ratio combiné économique de 113%, le portefeuille de risques ouragan mexicain étudié dans ce mémoire est déficitaire et les coûts de réassurance et du capital pèsent lourdement sur sa rentabilité. Dans ce contexte, il a été décidé d'effectuer des résiliations ciblées pour optimiser le ratio combiné économique du portefeuille.

Avant de pouvoir optimiser la politique de résiliation, il est nécessaire d'analyser le portefeuille et de pouvoir évaluer l'impact d'un scénario de résiliation sur l'ECR.

L'objectif de cette partie est de développer des méthodes permettant d'analyser le portefeuille et de quantifier les impacts d'une politique de résiliation sur le ratio combiné économique du portefeuille. On souhaite pouvoir identifier les polices qui feront partie du périmètre de l'optimisation et développer une fonction d'évaluation permettant de quantifier l'impact de n'importe quel scénario de résiliation sur l'ECR du portefeuille.

Cette partie est structurée en trois temps.

Dans un premier temps, une présentation des données disponibles puis une analyse descriptive du portefeuille seront effectuées.

Dans un deuxième temps, des méthodes seront développées pour quantifier l'impact d'un scénario de résiliation sur le ratio combiné économique du portefeuille. L'impact sur le résultat brut, additif, est facilement quantifiable. Ce n'est pas le cas du résultat cédé et du coût du capital qui requièrent le développement de méthodes d'estimation.

Dans un troisième temps, la méthode d'estimation du ratio combiné économique sera appliquée à chaque police du portefeuille. Le but est d'identifier celles qui pénalisent le plus la rentabilité du portefeuille.

2.1 Présentation du portefeuille étudié

Cette partie a pour objectif de présenter le portefeuille étudié. **Dans un premier temps**, une analyse critique des données disponibles sera effectuée. **Dans un deuxième temps**, une analyse descriptive du portefeuille sera réalisée. **Dans un troisième temps**, le ratio combiné économique du portefeuille sera étudié.

2.1.1 Données à disposition

Notre étude porte sur un portefeuille d'assurance dommage composé de risques ouragan localisés au Mexique. On suppose que ce portefeuille a été souscrit par une entité du groupe AXA. Cette entité a par ailleurs des activités en assurance IARD et Vie.

Ce portefeuille a été choisi car il correspond à notre problématique : il est fortement exposé au risque catastrophe, son exposition et ses coûts de réassurance augmentent depuis plusieurs années, alors que les primes encaissées diminuent. Forte de ce constat, l'entité souhaite mettre en place une politique de résiliation ciblée pour optimiser la rentabilité de son portefeuille.

Les polices tremblement de terre et incendie sont exclues du périmètre de ce mémoire. Cette exclusion est possible car les polices d'assurance mexicaines ont la spécificité de couvrir chaque péril indépendamment avec des polices dédiées.

Ces polices ont été exclues car ce mémoire a pour objectif de montrer la faisabilité de l'optimisation d'une politique de *pruning*. Une fois la faisabilité prouvée, il sera toutefois nécessaire de prendre en compte les polices tremblement de terre et incendie.

Plusieurs sources de données sont disponibles : données d'exposition du portefeuille, résultats de modélisation CAT, informations sur la réassurance et sur le capital.

2.1.1.1 Données d'exposition

La première source de données est la base de données d'exposition du portefeuille. Elle donne des informations sur les sites et les polices du portefeuille. Par ailleurs, elle a l'avantage d'avoir une structure robuste et de disposer d'un niveau de formalisme important : elle requière donc peu de retraitements.

Cette base de données est remplie chaque année par l'entité et partagée avec le *Group Risk Management*, département qui pilote les expositions CAT et les accumulations du groupe AXA. La base de données étudiée ici se compose de toutes les polices en vigueur en date de juillet 2018.

Description de la base de données

La base de données d'exposition est structurée en sept fichiers. Les quatre premiers fichiers donnent des informations sur les polices du portefeuille :

- **Fichier 1 - périls couverts** : liste des polices couvrant le risque ouragan.
- **Fichier 2 - polices** : informations sur les polices du portefeuille : prime, période de couverture, branche d'assurance.
- **Fichier 3 - conditions financières par police** : conditions financières d'assurance de chacune des polices (part souscrite, franchise et limite).
- **Fichier 4 - conditions de réassurance facultative par police** : conditions financières de réassurance de chacune des polices (facultative proportionnelle et non proportionnelle, part, franchise et limite).

Les trois autres fichiers donnent des informations sur les sites :

- **Fichier 1 - sites** : liste des sites et de la police qui leur est associée.
On y trouve également des informations sur les caractéristiques géographiques des sites (latitude / longitude) ainsi que sur la précision de la géolocalisation (par exemple rue, ville, code postal, zone CRESTA...).
- Cette table fournit aussi des informations sur les caractéristiques physiques des sites assurés : matériau de construction, nombre d'étages du bâtiment, étage du site assuré, type de toit, année de construction...
- On trouve enfin des informations sur les sommes assurées totales par type de garantie (bâtiment, contenu, perte d'activité).
- **Fichier 2 - conditions financières par site** : conditions financières d'assurance de chacun des sites (part souscrite, franchise et limite)
- **Fichier 3 - conditions de réassurance facultative par site** : conditions financières de réassurance de chacun des sites (facultative proportionnelle et non proportionnelle, part, franchise et limite)

Qualité des données

Le tableau ci-dessous présente le taux de remplissage de chacune des caractéristiques physiques des sites :

	Taux de remplissage
Matériau de construction	100%
Activité	100%
Nombre d'étages	5%
Branche d'assurance	100%
Géolocalisation	100%
Etage	2%
Type de toit	3%
Année de construction	1%

Figure 18 : Taux de remplissage par caractéristique

L'exhaustivité de l'information est très hétérogène : 100% de remplissage pour le matériau de construction, l'activité, la branche d'assurance et la géolocalisation mais moins de 5% pour le nombre d'étages du bâtiment, l'étage du bien assuré, le type de toit et l'année de construction.

Cette hétérogénéité s'explique par le fait que les systèmes de souscription ne sont pas harmonisés : l'information collectée lors de la souscription d'une police en direct est différente de celle collectée par un courtier.

Dans le cadre de notre étude, il aurait été souhaitable d'avoir davantage d'informations sur le type de toit. Il s'agit d'une variable importante pour quantifier le risque ouragan : lors d'un événement climatique, c'est généralement le toit qui est endommagé en premier.

A contrario, le peu d'information sur l'étage du bien assuré et l'année de construction a peu de conséquences : ces variables sont peu corrélées avec le taux de destruction d'un bâtiment lors d'un ouragan. En revanche, ces informations sont utiles dans le cadre de la modélisation des inondations (étage du bien assuré) et des tremblements de terre (nombre d'étages et année de construction).

Lors de la soumission de la base de données au Groupe, l'entité donne un avis sur l'exactitude des données entrées dans la base. La qualité des données est estimée bonne concernant la géolocalisation, la branche d'assurance et le matériau de construction. La qualité des données est estimée moyenne pour l'activité. En effet, en cas d'information manquante, des valeurs par défaut ont été utilisées. Ces valeurs ont été déterminées en fonction de la branche d'assurance : par exemple pour la branche industrie, les activités manquantes ont été remplacées par « industrie – général ».

Anonymisation

Pour des raisons de confidentialité, le portefeuille a été anonymisé et seule une partie des polices du portefeuille a été retenue.

2.1.1.2 Résultats d'un modèle CAT ouragan mexicain

Les risques présents dans la base de données d'exposition ont été modélisés à l'aide d'un modèle CAT ouragan mexicain. Ce modèle est un modèle *inhouse* développé au sein d'AXA.

Description des données disponibles

Les résultats du modèle ont la forme d'une table contenant les pertes par site, par année et par événement : il s'agit d'une table YELT (*Year Event Loss Table*). 10 000 années indépendantes ont été simulées.

Le tableau ci-dessous présente un exemple de table YELT par site :

Site	Année	Événement	Perte nette
1	1	A	0
1	1	B	100
1	3	C	50
...
...
1	10 000	Z	0
...
...
117 909	1	A	1000
117 909	1	B	0
117 909	3	C	20
...
...
117 909	10 000	Z	2

Figure 19 : exemple d'une table YELT à la granularité par site

Les résultats ont été produits au niveau de granularité le plus fin possible : par site et par événement. Pour chaque site du portefeuille, la distribution complète de la sinistralité est donc disponible. Les sinistres modélisés sont nets de conditions d'assurance et de réassurance facultative.

L'YELT est de très grande taille : le portefeuille se compose de près de 120 000 sites et le catalogue d'événements stochastiques d'environ 40 000 événements. La table se compose donc de près de 4,8 milliards de lignes.

Pertinence du modèle utilisé

Trois principaux acteurs dominent le marché des modèles CAT : RMS, AIR et RQE. Pour un même péril, il existe une grande variabilité entre les modèles : à une période de retour donnée, les pertes modélisées d'un même portefeuille peuvent varier très significativement.

Pour limiter le risque de volatilité entre modèles, le risque CAT du groupe AXA est modélisé avec un *blending* de plusieurs modèles externes et *inhouse*. Dans le cadre de notre étude, seule la modélisation *inhouse* est disponible. En effet, pour obtenir les résultats avec une granularité par site, il est nécessaire de lancer un *run* entier du modèle CAT, ce qui est très coûteux en temps de calcul.

Cette étude ayant pour objectif de montrer la faisabilité de l'optimisation d'une politique de *pruning*, il a été jugé acceptable de se baser uniquement sur le modèle *inhouse*. Une fois la faisabilité prouvée, il sera nécessaire d'obtenir les résultats des autres modèles pour que la vision du risque corresponde à celle du modèle interne d'AXA.

Cohérence des résultats

Pour vérifier l'exactitude des résultats, le taux de destruction par site et par événement a été calculé. Pour ce faire, il a été nécessaire d'appliquer les conditions financières sur le portefeuille et de calculer la NIV de chaque site assuré. Les taux de destruction obtenus sont toujours inférieurs ou égaux à 100%, ce qui est cohérent.

Comme on l'a vu précédemment, certaines caractéristiques de site sont absentes dans la base d'exposition (type de toit, etc). Le modèle CAT *inhouse* prévoit ce cas de figure et dispose d'un module de gestion des valeurs manquantes : une valeur prudente est affectée aux valeurs manquantes. Les résultats obtenus avec le modèle sont donc sensibles à l'évolution de la qualité des données d'une année sur l'autre.

Par ailleurs, le modèle CAT utilisé s'appuie sur des événements physiques fictifs mais probables. On note des écarts importants avec la vision historique. Cela peut s'expliquer par le fait que la vision historique a une profondeur historique limitée (environ 30 ans) alors que la vision physique a pour objectif de modéliser la sinistralité sur des périodes de temps très longues.

2.1.1.3 Données de réassurance : traité

Le portefeuille est couvert par un traité de réassurance composé de quatre tranches en excédent de sinistres par événement. Pour chaque tranche, nous disposons des informations suivantes :

- Conditions de réassurance : priorité, limite, nombre de reconstitutions
- Prime de réassurance et taux de reconstitution
- Modélisation des sinistres à charge des traités

Ces informations sont issues du renouvellement 2019.

2.1.1.4 Données de STEC

Pour rappel, le STEC (*Short Term Economic Capital*) est le SCR (*Solvency Capital Requirement*) en modèle interne AXA.

Dans le cadre de notre étude, nous disposons du STEC total ainsi que les valeurs de chacun des sous-STECS. Les STECS sont vus au 31/12/2018.

Les STECS utilisés dans notre étude sont le STEC Total, les STECS Marché, Crédit, Vie, IARD, Opérationnel ainsi que les STECS du modèle de risques IARD (Provisions, Primes, CAT).

2.1.2 Analyse descriptive du portefeuille

Un risque se définit par un croisement entre un aléa et une vulnérabilité. L'aléa est associé à la localisation et la vulnérabilité aux caractéristiques physiques d'un site.

Après avoir revu le portefeuille dans son ensemble, il sera donc analysé selon deux axes : la situation géographique puis les caractéristiques physiques des sites. Seule la revue globale du portefeuille sera présentée ici. Les résultats de l'analyse de la situation géographique et des caractéristiques des sites du portefeuille sont présentés en annexe 2 et 3.

L'objectif est de faire une analyse descriptive du portefeuille et d'identifier si les polices non rentables ont des caractéristiques communes. Les analyses sont basées sur la rentabilité brute du portefeuille. A ce stade, la réassurance et le coût du capital n'ont pas été considérés.

2.1.2.1 Revue globale du portefeuille

Le portefeuille étudié se compose de risques ouragan mexicain. Le portefeuille se compose de près de 34 000 polices pour une somme assurée totale de 27 milliards d'euros.

Comme on le voit dans le graphique ci-dessous, plus de la moitié des sommes assurées couvrent du risque dommage aux bâtiments.

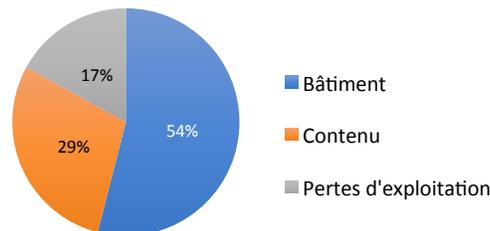


Figure 20 : poids de chaque garantie dans la somme assurée totale

Certaines polices disposent de facs proportionnelles et non proportionnelles. En moyenne, 17% des primes brutes sont cédées par l'intermédiaire de facs.

Profil du portefeuille

Quatre catégories de rentabilité ont été définies pour analyser le profil du portefeuille :

- **Bonne rentabilité** : *Loss Ratio* (ratio de sinistres sur primes) inférieur à 70%
- **Rentabilité moyenne** : *Loss Ratio* entre 70% et 100%
- **Non rentables** : *Loss Ratio* entre 100% et 400%
- **Polices les moins rentables du portefeuille** : *Loss Ratio* supérieur à 400%

Le graphe ci-dessous met en évidence le poids de chaque catégorie dans le volume de polices, la somme assurée (NIV), les sinistres moyens (AAL) et les primes (nettes de facs).

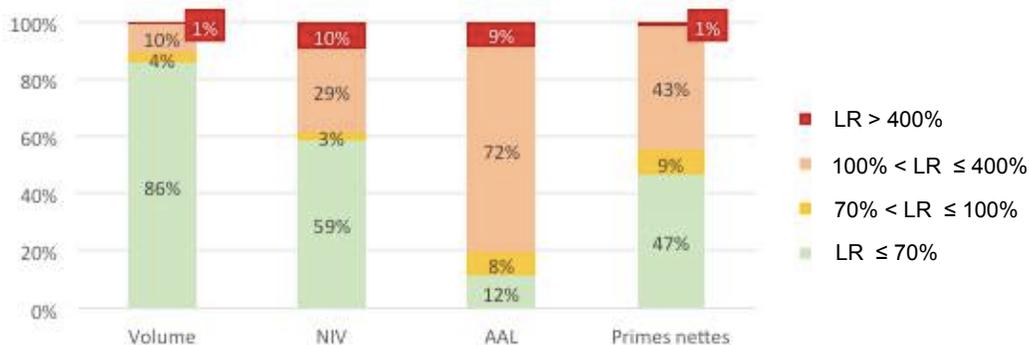


Figure 21 : poids de chaque catégorie de Loss Ratio

On remarque qu'un petit nombre de polices concentrent la majorité des sinistres du portefeuille : 11% des polices représentent 81% des sinistres. Ces polices ne représentent que 44% des primes totales et ont un *Loss Ratio* supérieur à 100%.

Cet effet s'accroît pour les polices avec un *Loss Ratio* supérieur à 400% : 1% des polices représente 10% des sommes assurées, 9% des sinistres mais seulement 1% des primes. Ces polices pénalisent la rentabilité du portefeuille.

Le reste des polices est très rentable : 86% des polices ne représentent que 12% des sinistres mais 47% des primes.

Analyse des polices ayant les plus grosses primes

L'analyse porte ensuite sur les dix polices concentrant les plus grosses primes du portefeuille. Ces polices représentent 23% de la prime totale.

Deux métriques sont introduites pour analyser ces polices : le taux de prime et le taux de destruction.

Le taux de prime permet de quantifier le prix d'un euro de somme assurée. Il se calcule comme suit :

$$TP \text{ (en } \text{‰})} = \frac{\text{Prime}}{\text{NIV}} * 1000$$

Le taux de destruction permet de quantifier le niveau de risque d'une police. Il se calcule comme suit :

$$TD \text{ (en } \text{‰})} = \frac{\text{AAL}}{\text{NIV}} * 1000$$

ID Police	<i>Loss Ratio</i>	Taux de prime (en ‰)	Taux de destruction (en ‰)
1	110%	0,61	0,67
2	25%	5,02	1,25
3	57%	21,45	12,19
4	127%	26,52	33,76
5	118%	0,45	0,53
6	53%	5,20	2,75
7	267%	10,74	28,70
8	138%	10,93	15,03
9	25%	0,69	0,17
10	114%	19,53	22,25

Figure 22 : polices ayant les dix plus grosses primes du portefeuille

On remarque une grande hétérogénéité sur le niveau de risque des polices : les taux de destruction varient entre 0,17‰ et 33,76‰.

Cette hétérogénéité se retrouve sur les taux de primes : les polices 7 et 8 ont le même taux de prime alors que la police 7 est presque deux fois plus risquée que la police 8.

Sur les dix polices ayant les plus grosses primes, 6 polices ne sont pas rentables. Ces contrats représentent une perte de 1,7 million d'euros. Leur profil est varié : avec un taux de destruction de 0,67‰, la police 1 est peu risquée, alors que la police 10 est très risquée (taux de destruction de 22,25‰).

Si ces 6 polices étaient résiliées, cela représenterait une amélioration du *Loss Ratio* de près de 9 points.

2.1.2.2 Conclusion

Le portefeuille étudié est très hétérogène. Une petite partie du portefeuille est non rentable et pénalise la rentabilité globale du portefeuille.

La quasi-totalité du risque du portefeuille est localisée sur les côtes. Les polices non rentables sont majoritairement situées dans les zones CRESTA de Quintana Roo et de la Basse-Californie du Sud.

Les polices les moins rentables relèvent des lignes de business résidentiel et commercial et plus particulièrement des activités hôtel, commerce de détail et maison individuelle. L'activité permet de

mieux identifier les polices les moins rentables que la branche d'assurance car elle donne des informations supplémentaires sur le risque : localisation, type de bâtiment, etc.

Enfin, les matériaux de construction des polices les moins rentables sont principalement le béton, le béton armé et la maçonnerie armée. Le béton est le matériel de construction principal des hôtels et le béton armé celui des maisons individuelles.

Ces conclusions sont basées sur les résultats de l'analyse descriptive du portefeuille. Pour rappel, ces analyses sont basées sur l'étude de la rentabilité brute du portefeuille. A ce stade, la réassurance et le coût du capital n'ont pas été considérés.

2.1.3 Ratio combiné économique du portefeuille

Dans le chapitre 1, un indicateur de rentabilité a été présenté : le ratio combiné économique, aussi appelé ECR (*Economic Combined Ratio*). Pour rappel, cet indicateur quantifie la charge de l'assureur pour chaque euro de prime souscrite et a l'avantage de considérer l'exigence de rémunération des actionnaires ainsi que les impôts.

Ratio combiné économique du portefeuille

Le graphe ci-dessous présente la décomposition de l'ECR du portefeuille de notre étude.

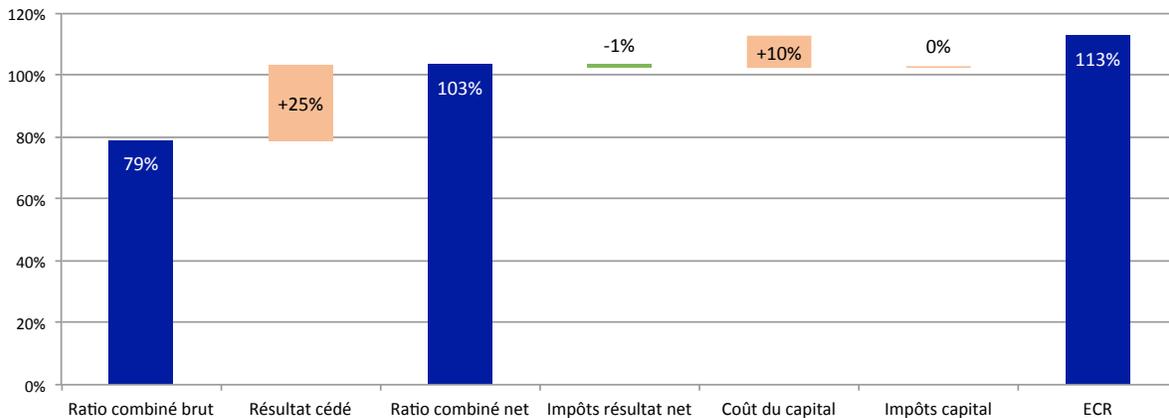


Figure 23 : décomposition de l'ECR

Le portefeuille n'est pas rentable : un ECR de 113% signifie qu'après rémunération des actionnaires, chaque euro de prime encaissé entraîne une détérioration de fonds propres de 13 centimes.

Sur un euro de prime encaissé, la charge brute représente 79 centimes. La réassurance coûte 25 centimes, soit un ratio combiné net de 103%. Après prise en compte de la réassurance, le portefeuille n'est donc pas rentable. De ce fait, il bénéficie d'un crédit d'impôt de 1%.

Le coût du capital représente 10 centimes pour chaque euro de prime souscrite, soit un ECR de 113%. On remarque qu'il n'y a pas d'impôts sur les rendements financiers du capital : le taux sans risque est proche de zéro.

Conclusion

Avec un ECR de 113%, le portefeuille de risques ouragan mexicain n'est pas rentable. Le portefeuille est notamment pénalisé par des coûts de réassurance et du capital importants (près de 35 points d'ECR).

Fort de ce constat, l'entité souhaiterait mettre en place une politique de *pruning* et résilier les polices qui pénalisent la rentabilité du portefeuille.

Dès lors, comment identifier les polices les moins rentables ? Comment calculer leur impact marginal sur l'ECR du portefeuille ? Le résultat brut étant additif, il est facilement possible de quantifier l'impact marginal sur le ratio combiné brut. Ce n'est pas le cas pour le résultat cédé et pour le coût du capital, qui requièrent de développer des méthodes d'estimation.

2.2 Quantification de l'impact d'un scénario de résiliation sur l'ECR

L'entité fictive souhaite mettre en place une politique de résiliation de risques dans le but d'optimiser son ratio combiné économique. Pour ce faire, elle doit être capable d'évaluer l'impact de n'importe quel scénario de résiliation sur l'ECR.

Résilier une police peut avoir un impact significatif sur le résultat brut du portefeuille, mais aussi sur ses courbes d'OEP, AEP, son résultat cédé et ses coûts du capital.

L'entité est capable de calculer l'impact sur le résultat brut : il est additif. Cependant, elle n'a pas les moyens de quantifier l'impact sur les autres métriques sans refaire tourner intégralement le modèle CAT, ce qui impliquerait plusieurs heures de calcul par scénario de résiliation.

L'objectif de cette partie est de développer une méthode permettant de quantifier l'impact de n'importe quel scénario de résiliation sur l'ECR.

Cette méthode rendra par exemple possible l'identification des polices qui pèsent le plus sur les coûts de réassurance et du capital. L'entité pourra alors choisir de façon éclairée les polices qu'elle souhaite résilier.

Cette méthode permettra également de piloter la politique de résiliation de risques et d'anticiper ses impacts sur les coûts de réassurance et sur le STEC.

Le but est de proposer une méthode utilisable opérationnellement : les temps de calcul doivent être raisonnables et le modèle CAT ne sera pas relancé.

Dans cette partie, les notions suivantes sont utilisées :

- **Portefeuille de référence** : portefeuille initial, avant résiliation.
- **Polices résiliées** : polices résiliées dans le scénario de résiliation.
- **Nouveau portefeuille** : portefeuille après prise en compte des résiliations.
- **Impact marginal** : impact des polices résiliées sur le portefeuille de référence.

On distingue bien les polices résiliées de l'impact marginal. Les polices résiliées sont vues indépendamment du portefeuille de référence, alors que la vision marginale quantifie l'impact des polices résiliées sur le portefeuille de référence.

Cette partie est structurée en cinq parties : les trois premières parties traitent de l'impact sur la sinistralité brute, les coûts de réassurance puis le coût du capital. La quatrième partie explicite les calculs nécessaires à l'obtention de l'impact final sur l'ECR. Dans une cinquième partie, un exemple d'utilisation de la méthode sera fourni.

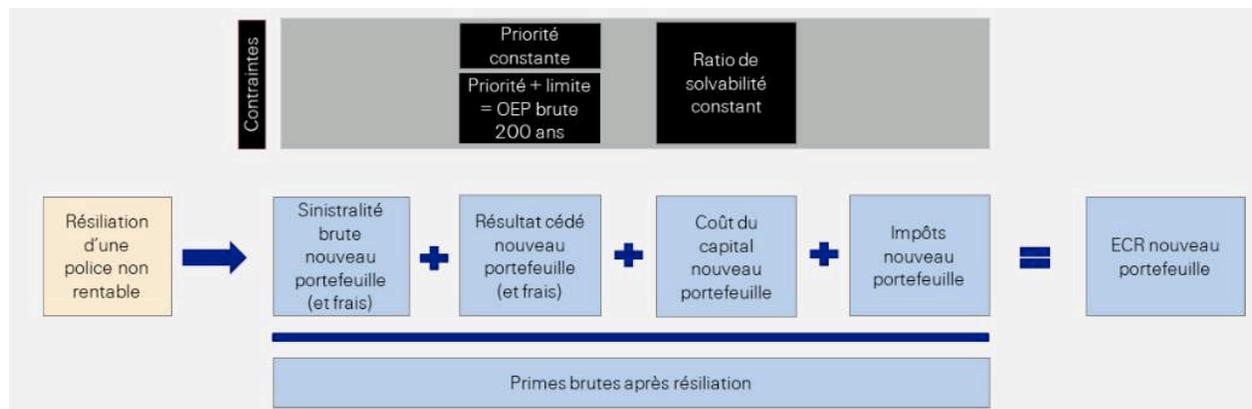


Figure 24 : processus de quantification de l'impact d'un scénario de résiliation

Pour que les impacts de deux scénarios de résiliation différents soient comparables, des indicateurs d'appétence au risque et de solvabilité constants sont utilisés :

- **Appétence au risque** : le traité de réassurance doit permettre de couvrir le sinistre bicentenaire. La priorité du traité est constante mais sa limite est ajustée pour chaque scénario.
- **Ratio de solvabilité** : pour que le ratio de solvabilité soit constant, on suppose que du capital est libéré en cas de baisse du STEC. A l'inverse, du capital est injecté en cas d'augmentation du STEC.

2.2.1 Impact sur la sinistralité brute

Cette partie a pour objectif de quantifier les impacts sur les indicateurs relatifs à la sinistralité brute.

Dans un premiers temps, l'impact sur le ratio combiné brut sera calculé. **Dans un second temps**, une méthode sera développée pour calculer l'impact sur les métriques brutes non additives : OEP et AEP.

Ratio combiné brut

Comme l'AAL et les primes brutes sont additives, il est possible de quantifier l'impact d'un scénario de résiliation sur le ratio combiné sans utiliser la table d'YELT.

Soit un scénario de résiliation $B = (b_i)_{i \in \llbracket 1, p \rrbracket}$. La prime π et les sinistres moyens S du nouveau portefeuille se calculent comme suit :

$$\pi^{Nouveau\ ptf} = \sum_{i=1}^p \pi_i * b_i$$

$$S^{Nouveau\ ptf} = \sum_{i=1}^p S_i * b_i$$

$$avec \begin{cases} b_i \in \{0; 1\}, & \begin{cases} b_i = 0 \text{ si la police } i \text{ est résiliée} \\ b_i = 1 \text{ si la police } i \text{ n'est pas résiliée} \end{cases} \\ & \pi_i \text{ la prime de la police } i \\ & S_i \text{ la sinistralité moyenne de la police } i \\ & p \text{ nombre de polices du portefeuille} \end{cases}$$

Soit σ^{brut} le taux de chargement au titre des frais externes et f^{brut} le montant des frais internes alloués au portefeuille.

Le montant des frais internes est fixe : le montant sera le même dans tous les scénarios de résiliation.

Les frais externes correspondent au frais de gestion et d'acquisition. Ils dépendent du montant de la prime. Si une police est résiliée, l'assureur ne paiera plus ses frais d'acquisition, mais il devra cependant s'acquitter de frais de gestion au titre de la résiliation. De ce fait, on peut décomposer β en deux taux :

- σ_{acq}^{brut} : taux de chargement au titre des frais d'acquisition. Ce taux s'applique uniquement sur les polices du nouveau portefeuille, i.e. les polices non résiliées.
- $\sigma_{gestion}^{brut}$: taux de chargement au titre des frais de gestion. Ce taux s'applique sur toutes les polices du portefeuille de référence, résiliées ou non.

Le ratio combiné brut se calcule comme suit :

$$COR_{brut}^{Nouveau\ ptf} = \frac{S^{Nouveau\ ptf} + f^{brut} + \sigma_{acq}^{brut} * \pi^{Nouveau\ ptf} + \sigma_{gestion}^{brut} * \pi^{ptf\ réf}}{\pi^{Nouveau\ ptf}}$$

Non additivité de l'OEP et de l'AEP

L'OEP et l'AEP sont des quantiles : ces mesures ne sont donc pas additives.

Dans l'exemple ci-dessous, on suppose que le portefeuille de référence et les polices résiliées ne sont touchés que par 5 événements, répartis sur 4 années.

Le plus gros sinistre du portefeuille de référence a lieu lors de l'événement 3 alors qu'il a lieu lors de l'événement 1 pour les polices résiliées. De même, l'année la plus sinistrée est l'année 2 pour le portefeuille de référence et l'année 1 pour les polices résiliées.

Année	Evénement	Portefeuille de référence	Polices résiliées	Nouveau portefeuille
1	1	300	200	100
1	2	50	50	0
2	3	500	0	500
3	4	15	0	15
4	5	20	5	15

Figure 25 : exemple d'un scénario de résiliation

Dans cet exemple, on voit que les polices résiliées n'ont pas d'impact sur le plus gros sinistre et sur l'année la plus sinistrée : ces deux indicateurs sont inchangés dans le portefeuille de référence et dans le nouveau portefeuille.

L'OEP et l'AEP des polices résiliées ne donnent donc pas d'information sur l'OEP et l'AEP du nouveau portefeuille : les plus gros sinistres du portefeuille de référence et des polices résiliées peuvent avoir lieu sur des événements différents.

Cet exemple prend son sens dans le cas du risque ouragan au Mexique. En effet, le Mexique est exposé à deux types d'ouragans : les ouragans prenant naissance dans l'Atlantique et dans le Pacifique.

Si on prend le cas de deux polices, l'une située à Quitana Roo et l'autre en Basse-Californie du Sud, la plupart des événements impacteront soit une police, soit l'autre, mais extrêmement rarement les deux polices en même temps. C'est ce phénomène qui peut expliquer pourquoi dans notre exemple, les polices résiliées n'ont pas d'impact sur le plus grand événement du portefeuille de référence. Le plus gros événement du portefeuille de référence pourrait correspondre à un sinistre à Quintana Roo alors que toutes les polices résiliées seraient localisées en Basse-Californie du Sud.

Calcul d'une nouvelle table d'YELT

L'OEP et l'AEP n'étant pas des métriques additives, il est donc nécessaire de calculer l'YELT du nouveau portefeuille pour calculer l'OEP et l'AEP du nouveau portefeuille. Cela Pour ce faire, il faut disposer de l'YELT du portefeuille de référence et de celle des polices résiliées.

Comme on le voit dans la figure ci-dessous, on obtient l'YELT du nouveau portefeuille en soustrayant, pour chaque événement du catalogue, les pertes des polices résiliées aux pertes du portefeuille de référence.

Portefeuille de référence			Polices résiliées			Nouveau portefeuille		
Année	Evénement	Perte	Année	Evénement	Perte	Année	Evénement	Perte
1	A	50	1	A	0	1	A	50
1	B	100	1	B	10	1	B	90
2	C	2	2	C	0	2	C	2
3	D	15	3	D	2	3	D	13
4	E	350	4	E	17	4	E	333
4	F	1	4	F	0	4	F	1
...
10 000	Z	25	10 000	Z	1	10 000	Z	24

Figure 26 : mécanisme de calcul de l'YELT du nouveau portefeuille

Soit $Brut^{Nouveau\ ptf} = (Brut_{i,j}^{Nouveau\ ptf})_{i \in \llbracket 1, n \rrbracket, j \in \llbracket 0, n_i \rrbracket}$ la distribution des pertes brutes du nouveau portefeuille, n le nombre d'années simulées et n_i le nombre d'événements de l'année i . On note :

$$Brut_{i,j}^{Ptf\ nouveau} = Brut_{i,j}^{Ptf\ ref} - Brut_{i,j}^{Polices\ résiliées}$$

A partir des trois YELT, il est possible de calculer l'OEP *standalone* des polices résiliées et l'impact marginal sur le portefeuille de référence.

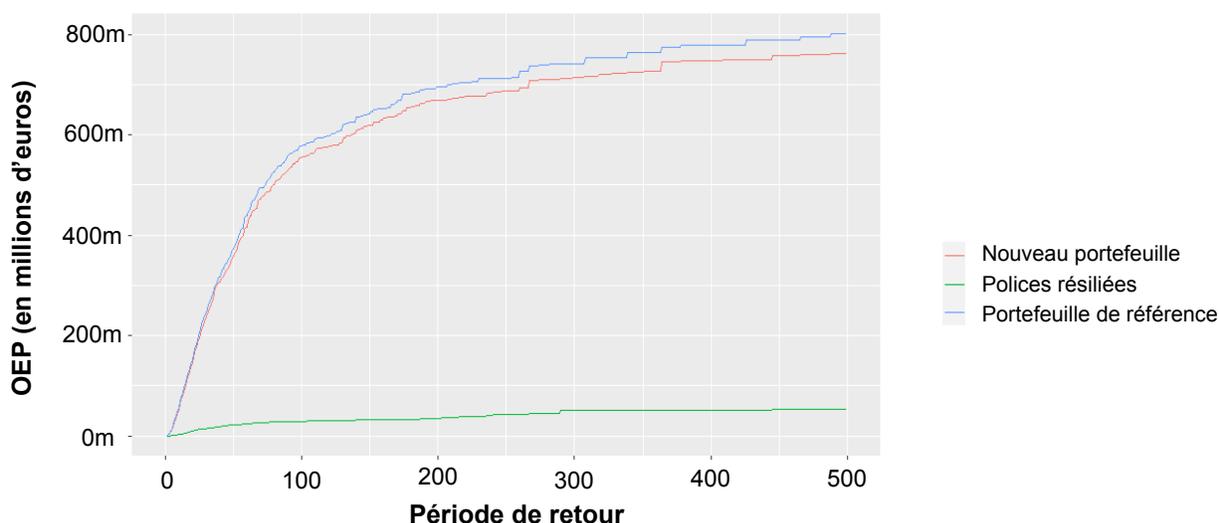


Figure 27 : impact d'un scénario de résiliation sur l'OEP

Pour un scénario de résiliation donné, le graphe ci-dessus montre l'OEP du portefeuille de référence, des polices résiliées et du nouveau portefeuille. A partir de ce graphe, il est possible de calculer l'impact marginal des polices résiliées sur l'OEP pour chaque période de retour PR .

$$OEP\ marginale_{PR} = OEP_{PR}(Brut^{Ptf\ référence}) - OEP_{PR}(Brut^{Nouveau\ ptf})$$

Le tableau ci-dessous présente les impacts d'un scénario de résiliation.

	AAL	OEP 50y	OEP 200y	AEP 50y	AEP 200y
Portefeuille de référence	30,0m	373,1m	695,8m	401,7m	715,7m
Nouveau portefeuille	28,4m	357,4m	669,1m	381,9m	686,0m
Impact marginal des polices résiliées	1,6m	15,7m	26,7m	19,8m	29,6m
Polices résiliées standalone	1,6m	21,7m	35,4m	22,5m	39,4m
Impact marginal - polices résiliées standalone	0,0m	-6,0m	-8,6m	-2,7m	-9,8m

Figure 28 : exemple de l'impact marginal d'un scénario de résiliation

On remarque que l'AAL est additif : l'impact marginal des polices résiliées est égal à l'AAL des polices résiliées vues indépendamment du portefeuille de référence.

Ce n'est pas le cas pour les OEP et les AEP : l'impact marginal est inférieur à l'OEP et l'AEP des polices résiliées. Le portefeuille a un effet diversifiant sur les polices résiliées.

2.2.2 Impact sur les coûts de réassurance

La quantification de l'impact sur le résultat cédé est réalisée en plusieurs étapes :

- Tout d'abord, la limite du traité de réassurance doit être ajustée : pour des raisons d'appétence au risque, l'entité souhaite se couvrir à la hauteur de son événement bicentenaire. La priorité reste fixe dans tous les scénarios de résiliation.
- Les conditions du traité sont ensuite appliquées sur l'YELT du nouveau portefeuille.
- Une estimation de la prime de réassurance est réalisée.
- Enfin, les frais sont pris en compte et le ratio combiné net est calculé.

Les sections suivantes détaillent chacune des étapes décrites précédemment.

2.2.2.1 Etape 1 : adaptation de la limite du traité

Pour des raisons d'appétence au risque, l'entité souhaite acheter de la réassurance à la hauteur de son événement bicentenaire. De ce fait, il est nécessaire d'adapter la limite du traité de réassurance pour chaque scénario de résiliation.

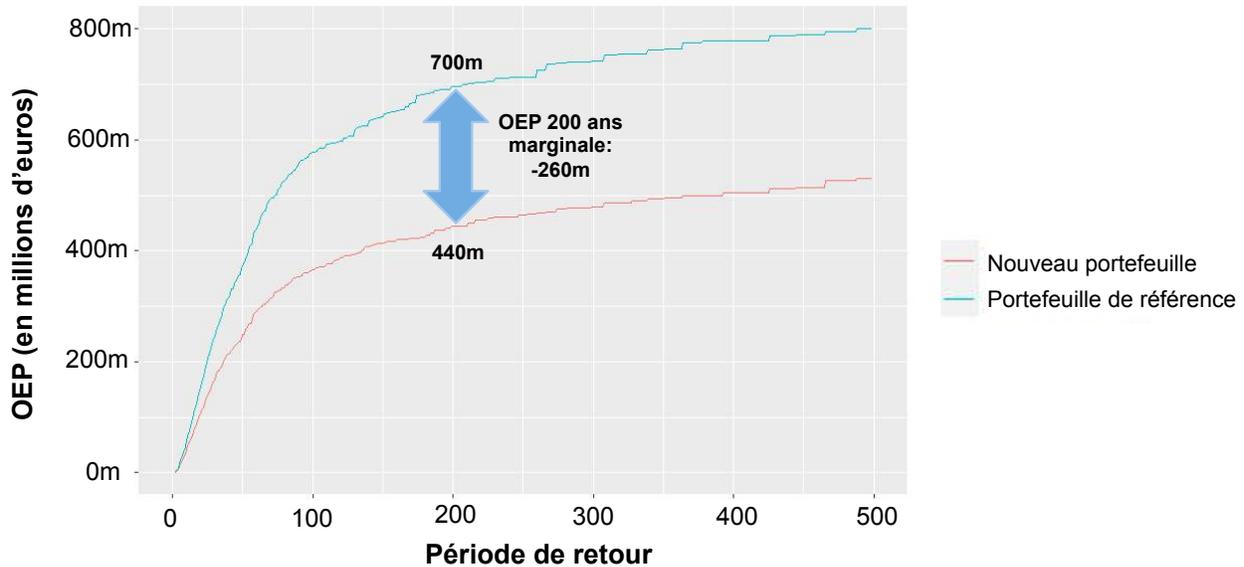


Figure 29 : adaptation de la limite du traité pour un scénario de résiliation

Pour un scénario de résiliation donné, le graphe ci-dessus représente la distribution de l'OEP du portefeuille de référence et du nouveau portefeuille. Dans cet exemple, le scénario de résiliation a un impact marginal de -260 millions sur l'OEP bicentenaire. La limite du traité sera donc réduite de 260 millions d'euros.

$$\text{Limite} = OEP_{200y} - \text{priorité}$$

A noter que la limite du traité est calibrée sur la base de l'OEP car le traité de réassurance est un traité par événement.

2.2.2.2 Etape 2 : application du traité de réassurance

Après avoir ajusté la limite du traité, les conditions du traité sont appliquées sur l'YELT du nouveau portefeuille. Le traité étant par événement, les sinistres vont être agrégés par événement et par année : il n'y a donc plus de granularité par site.

Application du traité et calcul de l'YELT nette

Les récupérations annuelles $Recup = (Recup_i)_{i \in \llbracket 1, n \rrbracket}$ à charge du traité sont calculées pour chaque année simulée i :

$$Recup_i^{Nouveau\ ptf} = \min((nb\ rec + 1) * Limite^{Nouveau\ ptf}, \sum_{j=1}^{n_i} \min(Limite^{Nouveau\ ptf}, \max(Brut_{i,j}^{Nouveau\ ptf} - Priorité, 0)))$$

$$\text{avec } \begin{cases} n : \text{le nombre d'années simulées} \\ n_i : \text{nombre d'évènements de l'année } i \\ nb\ rec : \text{nombre de reconstitutions} \end{cases}$$

On en déduit les sinistres nets annuels $Net^{Nouveau\ ptf} = (Net_i^{Nouveau\ ptf})_{i \in \llbracket 1, n \rrbracket}$. A ce stade, les événements sont agrégés par année et il n'y a plus d'événement : la table est donc une YLT (*Year Loss Table*).

$$Net_i^{Ptf\ nouveau} = Brut_i^{Ptf\ nouveau} - Recup_i^{Ptf\ nouveau}$$

Calcul du Loss on Line

Le *Loss on Line* mesure la concentration de récupérations dans la limite. En d'autres termes, il s'agit de la part de la limite qui sera en moyenne consommée. Le LoL se calcule comme suit :

$$LoL = \frac{E(Recup)}{Limite} \quad \text{avec } E(Recup) = \frac{\sum_{i=1}^n Recup_i}{n}$$

Coefficient de variation

Le coefficient de variation $CoV(Recup)$ mesure la volatilité des récupérations autour de leur moyenne. Il se calcule comme le rapport entre l'écart type et la moyenne :

$$CoV(Recup) = \frac{SD(Recup)}{E(Recup)} \quad \text{avec } SD(Recup) = \sqrt{\frac{\sum_{i=1}^n (Recup_i - E(Recup))^2}{n}}$$

2.2.2.3 Etape 3 : estimation de la prime de réassurance

L'objectif de cette partie est de calibrer des courbes de prix marché. Ces courbes seront ensuite utilisées pour prédire le RoL du traité du nouveau portefeuille. Il a été choisi de prédire le RoL et non la prime car cette métrique permet de comparer la concentration de la prime de deux traités n'ayant pas la même limite.

Pour calibrer les courbes, des informations sont disponibles sur quatre traités couvrant le portefeuille. Les indicateurs suivants sont disponibles : RoL, LoL et coefficient de variation. On connaît également la priorité et la limite des quatre traités. Ces informations proviennent du renouvellement 2019 et ont été anonymisées.

Dans cette partie, trois méthodes d'estimation de la prime de réassurance seront présentées.

Méthode 1 : prédiction du RoL en fonction du Tiers-Point

Le *tiers-point* d'un traité correspond à sa priorité à laquelle on ajoute un tiers de la limite.

Le graphe ci-dessous illustre l'évolution du RoL en fonction du *tiers-point*. Le *tiers-point* permet de capter à la fois l'effet de la priorité et de la limite. C'est le *tiers-point* et non le *mid-point* (priorité à laquelle on ajoute 50% de la limite) qui a été choisi car les récupérations sont plus concentrées en bas de la tranche.

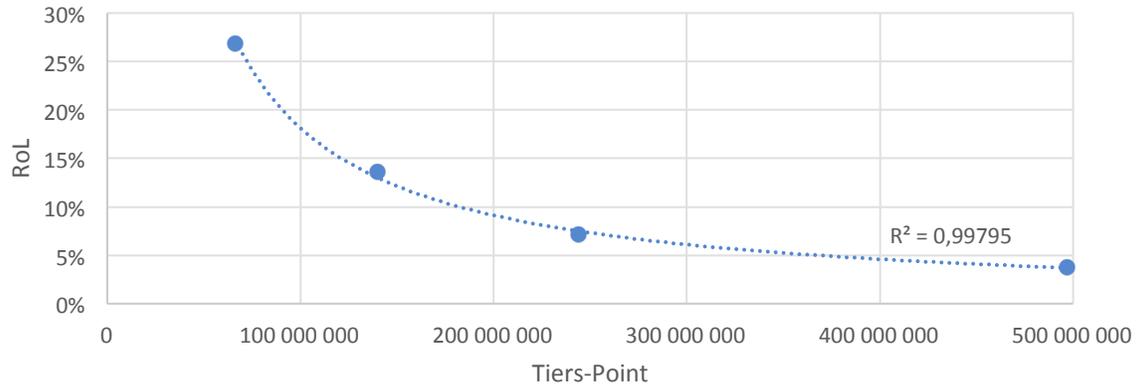


Figure 30 : évolution du RoL en fonction du tiers-point

On remarque que plus le *tiers-point* est grand et plus le RoL est petit. En effet, plus la tranche est haute et moins elle génère de récupérations relativement à sa limite : le *Loss on Line* est décroissant. Cela s'explique simplement par le fait qu'un sinistre touchant une tranche haute touche aussi toutes les tranches situées en-dessous.

Sur la base des quatre traités disponibles, on effectue une régression pour déterminer la fonction f telle que :

$$f\left(\text{Priorité} + \frac{\text{Limite}}{3}\right) = \text{RoL}$$

Le R carré obtenu est de 99,8%.

Cette méthode a l'avantage d'être facile à mettre en œuvre : il suffit de disposer du prix, de la limite et de la priorité des traités. Il n'est pas nécessaire de disposer de la modélisation de la sinistralité.

La méthode par *tiers-point* présente cependant certaines limites : elle ne prend pas en compte l'évolution des récupérations moyennes à charge du traité. De ce fait, le RoL peut être surestimé dans le cas où un scénario de résiliation entraînerait une diminution des récupérations du traité mais n'impacterait pas l'événement bicentenaire : la limite et a fortiori le *tiers-point* seraient donc inchangés. Dans un tel cas de figure, le traité du portefeuille de référence et du nouveau portefeuille auraient le même RoL, alors que le nouveau portefeuille devrait en réalité avoir un RoL plus petit.

Méthode 2 : prédiction du RoL en fonction du LoL

Les courbes *tiers-points* ont le désavantage de ne pas prendre en compte l'évolution du LoL. Pour remédier à cette limite, il est possible de prédire la prime de réassurance à l'aide du *Loss on Line*.

La première étape consiste à étudier l'évolution du RoL en fonction du LoL. Le LoL est une métrique intéressante car elle prend en compte à la fois l'évolution des récupérations moyennes et de la limite.



Figure 31 : évolution du RoL en fonction du LoL

Les tranches hautes sont situées à droite : ce sont celles qui ont les plus petits LoL. Plus le LoL est petit et plus le RoL est petit.

La droite orange est la première bissectrice. Les points situés sur cette droite correspondent à des traités pour lesquels un euro de récupération coûterait un euro de prime de réassurance. Ici, tous les points sont situés au-dessus de la droite : la prime payée est supérieure aux récupérations moyennes du traité. Cet écart entre le RoL et le LoL correspond à la rémunération du réassureur : le réassureur facture un chargement en plus de la moyenne de la sinistralité attendue.

Le multiple permet de mieux comprendre l'évolution de ce chargement. Cet indicateur quantifie le prix d'un euro de récupérations moyennes et se calcule comme suit :

$$\text{Multiple} = \frac{\text{RoL}}{\text{LoL}} = \frac{\text{Prime}}{\text{Récupérations moyennes}}$$

Le nuage de points ci-dessous illustre l'évolution du multiple en fonction du LoL. Cela permet de visualiser l'évolution de la rémunération demandée par le réassureur.

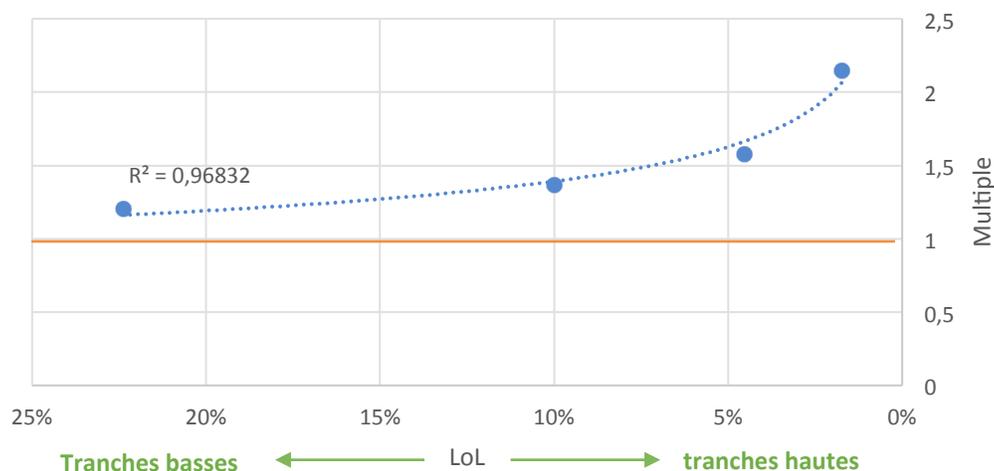


Figure 32 : évolution du prix d'un euro de récupération moyenne

Les points sont situés au-dessus de la droite d'équation $x=1$: un euro de récupération moyenne coûte plus cher qu'un euro de prime.

On remarque que plus le LoL est petit et plus le multiple est grand : un euro de récupération moyenne coûte plus cher pour les tranches hautes que pour les tranches basses. Ce constat peut s'expliquer par la volatilité. En effet, comme on le voit sur le nuage de points ci-dessous, plus les tranches sont hautes et plus leur coefficient de variation est élevé. Les tranches hautes sont moins souvent touchées qu'une tranche basse : de ce fait, il y a plus de volatilité autour de leur moyenne.

Un euro de récupération moyenne coûte plus cher pour les tranches volatiles car le réassureur fait payer à la cédante l'incertitude autour de la moyenne des récupérations. Par ailleurs, comme les tranches hautes sont moins travaillantes que les tranches basses, l'historique sinistre n'est pas représentatif de la sinistralité future. Le réassureur a donc plus d'incertitudes sur la sinistralité attendue de ces tranches.

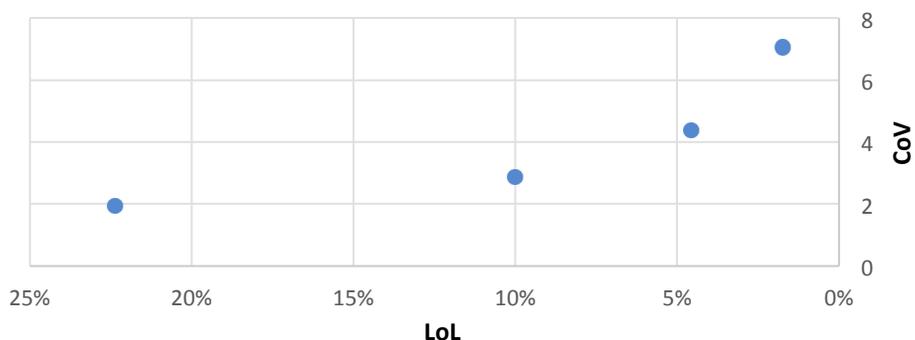


Figure 33 : évolution du coefficient de variation en fonction du LoL

Sur la base des quatre traités disponibles, on effectue une régression pour déterminer la fonction g telle que :

$$g(LoL) * LoL = RoL$$

Le R carré obtenu est de 96,83%.

Cette méthode a l'avantage d'être basée sur le LoL. Cette variable a l'intérêt de prendre en compte à la fois les récupérations moyennes et la limite.

Méthode 3 : prédiction du RoL en fonction du LoL et du coefficient de variation

Cette méthode a pour but de prédire le RoL en fonction du LoL et du coefficient de variation.

Dans cette section, on suppose que le RoL est calculé par le réassureur sur la base des éléments suivants :

- La prime de risque : il s'agit du coût moyen des sinistres à charge du traité : c'est le LoL.
- Un chargement au titre de l'incertitude : ce chargement prend en compte le risque d'avoir sous-estimé la prime de risque. De ce fait, il s'applique sur la base du LoL. On appelle ce chargement α .
- Un chargement au titre de la volatilité : ce chargement est facturé au titre de la volatilité des récupérations autour de leur moyenne. Plus l'incertitude autour de la moyenne est grande et plus le chargement sera grand. De ce fait, ce chargement s'applique sur le coefficient de variation. On l'appelle β .

L'objectif est donc de déterminer des chargements α et β tels que :

$$RoL = (\alpha + 1) * LoL + \beta * CoV$$

On détermine α et β avec pour contraintes $\alpha > 0$ et $\beta > 0$. La prédiction est de bonne qualité : le R carré obtenu est de 99,45%.

Pour chaque tranche, le graphe ci-dessous présente la décomposition du RoL obtenu.

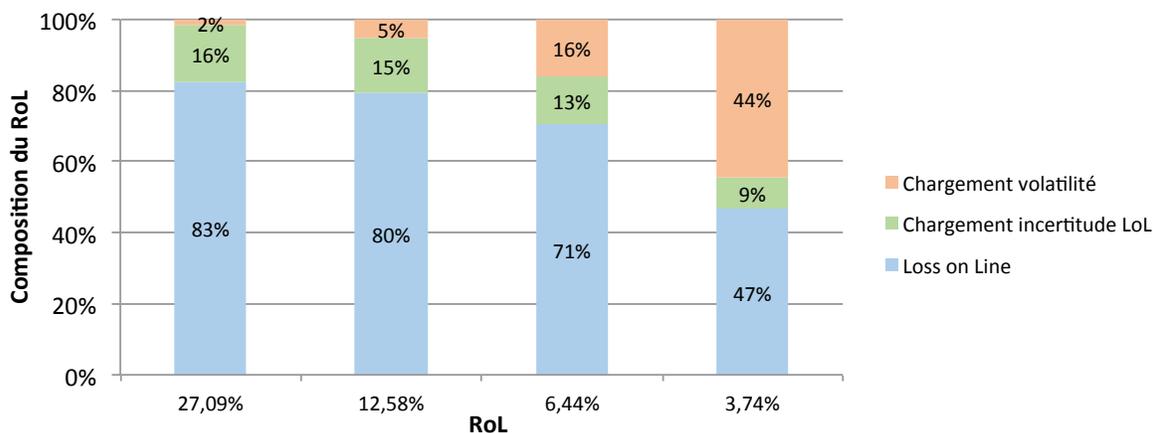


Figure 34 : décomposition du RoL

Plus le RoL est petit et plus la rémunération du réassureur augmente : pour un RoL de 27,09%, la rémunération représente 18% du RoL contre 53% pour un RoL de 3,74%.

Ce constat s'explique par la volatilité : plus la tranche est haute, plus la volatilité et l'incertitude sont importantes. Par conséquent, le réassureur demande une rémunération élevée. Cette observation est cohérente avec celles faites pour la méthode 2.

Conclusions et limites

Comme on le voit dans le tableau ci-dessous, les méthodes 1 et 3 ont des R carré proches de 100%. Leur qualité de prédiction est meilleure que celle de la méthode 2.

	R carré
Méthode 1 : RoL ~ Tiers-Point	99,87%
Méthode 2 : RoL ~ LoL	96,83%
Méthode 3 : RoL ~ LoL + CoV	99,45%

Figure 35 : comparaison entre les différentes méthodes d'estimation du RoL

La méthode 1 a le désavantage de ne pas prendre en compte l'évolution des récupérations moyennes. Les méthodes 2 et 3 sont plus adaptées à notre problématique : elles sont basées sur le LoL, métrique qui a l'avantage de prendre en compte l'évolution des récupérations moyennes et de la limite. La méthode 3 est la méthode la plus adaptée : elle prend en compte à la fois l'évolution du LoL mais aussi de la volatilité. Par ailleurs, elle a un meilleur R carré que la méthode 2. C'est la méthode qui sera retenue pour notre étude.

Pour calibrer les paramètres des méthodes 2 et 3, il est nécessaire de disposer de la modélisation des récupérations à charge des tranches. Cette information n'est malheureusement pas toujours disponible. La méthode par *tiers-point* est une méthode adaptée aux cas où peu d'information est disponible : il suffit de disposer des caractéristiques de la tranche et du montant de la prime pour pouvoir calibrer cette courbe.

Les trois méthodes présentées ont un écueil principal : les courbes ont été calibrées sur la base de seulement quatre traités de réassurance.

Par ailleurs, ces traités sont issus du renouvellement de l'année 2019. Or d'une année sur l'autre, les prix de réassurance peuvent varier significativement. Cela peut s'expliquer par la survenance d'un sinistre, la dégradation des conditions de marché, des négociations entre la cédante et le réassureur, etc. De ce fait, les prix obtenus avec les trois méthodes ne sont que des estimations de la prime de réassurance.

2.2.2.4 Etape 4 : calcul du ratio combiné net

La méthode 3 a été retenue. La prime de réassurance se calcule donc comme suit :

$$p^{\text{Nouveau ptf}} = ((\alpha + 1) * LoL + \beta * CoV) * Limite^{\text{Nouveau ptf}}$$

On en déduit le ratio combiné net :

$$CoR_{net}^{\text{Nouveau ptf}} = CoR_{brut}^{\text{Nouveau ptf}} + \frac{p^{\text{Nouveau ptf}} * (1 + \sigma^{\text{réass}}) - E(Recup^{\text{Nouveau ptf}}) + f^{\text{réass}}}{\pi^{\text{Nouveau ptf}}}$$

2.2.3 Impact sur le coût du capital

Le coût du capital traduit l'exigence de rémunération des actionnaires. Il est calculé sur la base du STEC CAT diversifié et d'un ratio de solvabilité constant.

Dans un premier temps, le STEC CAT *standalone* sera calculé. **Dans un deuxième temps**, ce STEC sera diversifié. **Dans un troisième temps**, une méthode d'estimation du taux de diversification sera développée, dans le but de réduire les temps de calcul. **Dans un quatrième temps**, le coût du capital sera calculé.

2.2.3.1 Calcul du STEC CAT standalone

Formule de calcul

Le STEC CAT *standalone* est la déviation bicentenaire du résultat par rapport à la moyenne :

$$Res^{Nouveau\ ptf} = \pi^{Nouveau\ ptf} - p^{Nouveau\ ptf} - Net^{Nouveau\ ptf}$$

$$STEC_{CAT\ standalone}^{Nouveau\ ptf} = Var_{99,5\%}(Res^{Nouveau\ ptf}) - E(Res^{Nouveau\ ptf})$$

On suppose que les polices sont renouvelées à l'identique : les primes brutes sont donc constantes. Par ailleurs, les primes de reconstitution sont gratuites : les primes nettes du scénario bicentenaire et du scénario moyen sont égales. On en déduit :

$$STEC_{CAT\ standalone}^{Nouveau\ ptf} = Var_{99,5\%}(Net^{Nouveau\ ptf}) - E(Net^{Nouveau\ ptf})$$

Application à un scénario de résiliation

Pour un scénario de résiliation donné, le graphe ci-dessous donne un exemple d'AEP brutes et nettes pour le portefeuille de référence et le nouveau portefeuille.

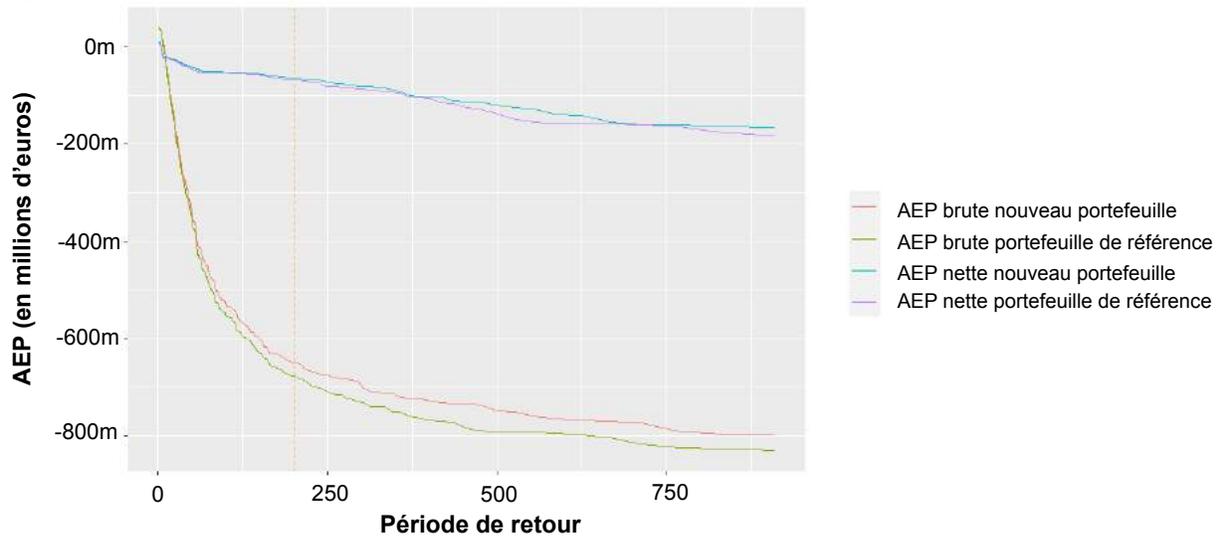


Figure 36 : exemple de courbe d'AEP du résultat brut et net avant et après résiliation

Le tableau ci-dessous présente les différentes métriques obtenues en vision brute et nette pour le portefeuille de référence et le nouveau portefeuille.

		Moyenne du résultat	AEP 200y résultat	STEC CAT standalone
Brut	Portefeuille de référence	7,3m	-677,5m	684,8m
	Nouveau portefeuille	8,9m	-647,9m	656,8m
	Impact marginal des polices résiliées	+ 1,6m	+ 29,6m	-28,0m
Net	Portefeuille de référence	-1,5m	-69,0m	67,5m
	Nouveau portefeuille	0,0m	-63,6m	63,7m
	Impact marginal des polices résiliées	+ 1,5m	+ 5,3m	-3,8m

Figure 37 : impact marginal du STEC CAT standalone

Le scénario de résiliation a permis d'améliorer le résultat net de 1,5 million et de faire diminuer le STEC CAT de 3,8 millions.

Dans cet exemple, le STEC du nouveau portefeuille diminue. Ce n'est pas le cas dans tous les scénarios de résiliation. Les conditions suivantes sont nécessaires pour que le STEC du portefeuille diminue :

- Il faut que l'impact marginal de l'AEP bicentenaire soit positif. Ce n'est pas toujours le cas : l'éventuelle réduction de la limite du traité de réassurance peut avoir comme conséquence une détérioration de l'AEP.
- Le STEC étant la différence entre l'AEP bicentenaire et la moyenne, il faut que l'impact marginal sur l'AEP bicentenaire soit supérieur à celui sur le résultat moyen.

Par ailleurs, cet exemple permet également de mettre en évidence une des fonctions de la réassurance : la réduction du capital immobilisé. En effet, grâce à la réassurance, le STEC a été divisé par près de 10, soit près de 617 millions d'euros de STEC économisé (67,5 millions de STEC net contre 684,8 millions de STEC brut).

2.2.3.2 Calcul du STEC Cat diversifié

Une fois le STEC CAT *standalone* obtenu, on souhaite obtenir le STEC CAT diversifié. Pour ce faire, il est nécessaire d'agréger le STEC CAT avec les autres STEC pour obtenir le STEC total. Cette agrégation se fait en deux étapes : entre les risques du module IARD puis entre les modules de risque.

Le schéma ci-dessous met en évidence en bleu les STEC qui vont être impactés par les scénarios de résiliation : il s'agit des STEC CAT, IARD et total.

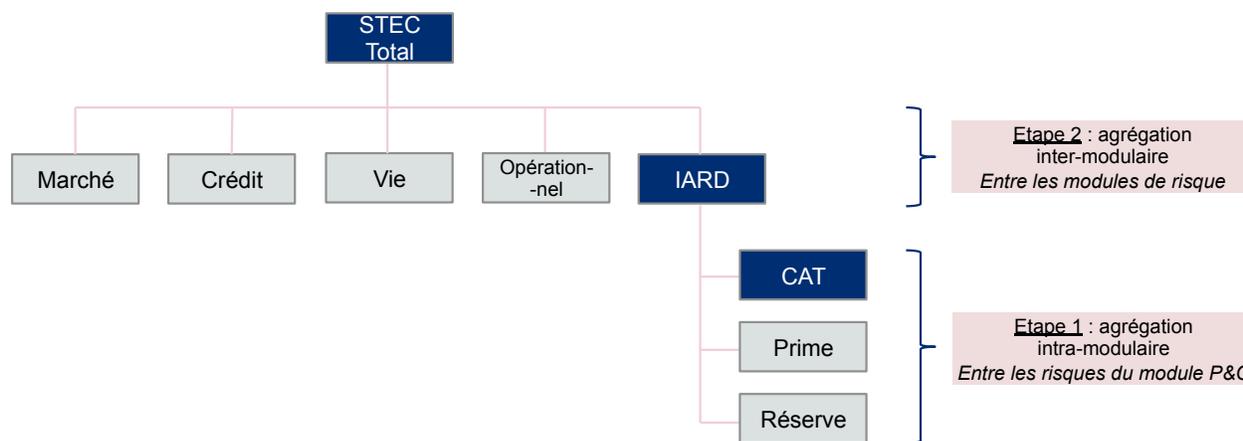


Figure 38 : méthodologie d'agrégation des risques en modèle interne AXA

Lors de l'agrégation, les STEC se diversifient : le STEC total est inférieur à la somme des STEC de chaque module et le STEC de chaque module est inférieur à la somme des STEC de ses différents risques. Cela s'explique par le fait que les modules de risque ont des moteurs de risque différents : tous les scénarios défavorables ne se produisent pas en même temps.

Comme on le voit dans le schéma ci-dessous, la diversification peut ensuite être allouée par type de risque. On appelle « STEC diversifiés » les STEC auxquels une part de la diversification a été allouée : la somme des STEC diversifiés est égale au STEC total.

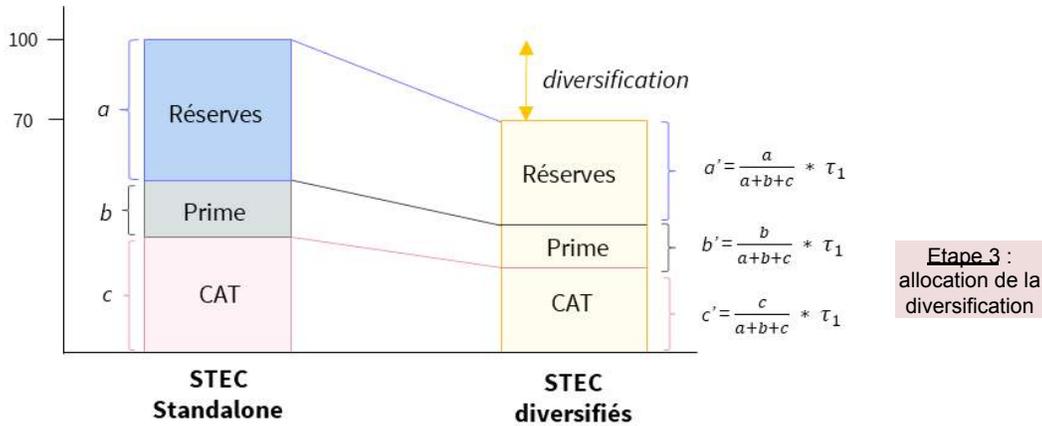


Figure 39 : illustration de la diversification du STEC CAT

Dans cette partie, on va donc raisonner en trois étapes. Tout d'abord, l'agrégation intra-modulaire sera effectuée. On réalisera ensuite l'agrégation inter-modulaire. Une fois le STEC total obtenu, une partie de la diversification sera allouée au STEC CAT.

Etape 1 : agrégation intra-modulaire

La première étape consiste à agréger les risques du module IARD entre eux : risque CAT, réserve et prime. Le STEC IARD se calcule comme suit :

$$STEC_{IARD} = \sqrt{\sum_{(i,j) \in R_{IARD}^2} \rho_{i,j}^{R_{IARD}} * C_i * C_j}$$

avec $\left\{ \begin{array}{l} R_{IARD} : \text{risques du module IARD} \\ C_i : \text{capital au titre du risque } i \\ \rho_{i,j}^{R_{IARD}} : \text{coefficient de corrélation permettant d'agréger les capitaux au titre} \\ \text{des risques } i \text{ et } j \text{ appartenant au module IARD} \end{array} \right.$

Etape 2 : agrégation inter-modulaire

Une fois le STEC IARD calculé, il est agrégé avec les autres modules de risque (marché, crédit, vie, opérationnel) pour former le STEC total :

$$STEC_{Total} = \sqrt{\sum_{(i,j) \in M^2} \rho_{i,j}^M * SCR_i * SCR_j}$$

avec $\left\{ \begin{array}{l} SCR_i : \text{capital économique du module } i \\ M : \text{ensemble des modules} \\ \rho_{i,j}^M : \text{coefficient de corrélation permettant d'agréger les capitaux des modules } i \text{ et } j \end{array} \right.$

Etape 3 : allocation de la diversification

La dernière étape consiste à allouer la diversification. Deux niveaux de diversification seront alloués : la diversification du STEC CAT dans le STEC IARD puis la diversification du STEC IARD dans le STEC total.

$$STEC\ CAT_{diversifié} = STEC\ CAT_{standalone} * \tau_1 * \tau_2$$

$$avec \begin{cases} \tau_1 : \text{taux de diversification du risque CAT dans le module IARD} \\ \tau_2 : \text{taux de diversification du module IARD dans le STEC total} \end{cases}$$

La diversification est allouée proportionnellement au poids des STEC *standalone* :

$$\tau_1 = \frac{STEC_{IARD\ standalone}}{STEC_{CAT\ standalone} + STEC_{Prime\ standalone} + STEC_{Réserve\ standalone}}$$

$$\tau_2 = \frac{STEC_{Total}}{STEC_{IARD\ stand} + STEC_{Marché\ stand} + STEC_{Crédit\ stand} + STEC_{Opérationnel\ stand} + STEC_{Vie\ stand}}$$

2.2.3.3 Estimation du taux de diversification du STEC CAT dans le STEC total

Dans cette partie, une méthode d’estimation du taux de diversification est développée. L’objectif de cette méthode est de réduire les temps de calcul et de comprendre le comportement du taux de diversification.

Diversification du STEC CAT dans le STEC IARD et du STEC IARD dans le STEC total

Le nuage de points ci-dessous montre l’évolution des taux de diversification en fonction de l’évolution du STEC CAT.

Pour rappel, le STEC CAT peut évoluer à la hausse ou à la baisse en fonction des scénarios de résilience. On souhaite que le taux de diversification soit le plus petit possible : plus il est petit et plus la diversification est importante. Un taux de diversification de 40% signifie que le STEC diversifié est égal à 40% du STEC *standalone*.

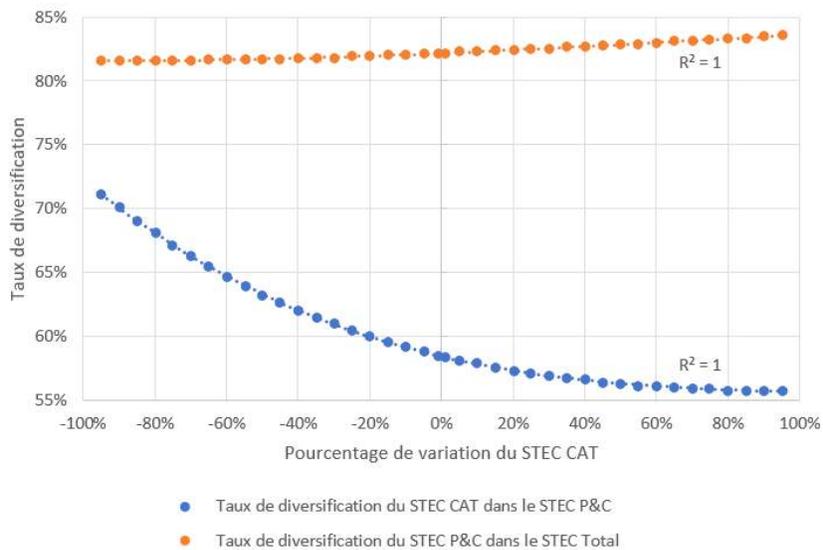


Figure 40 : impact de l’évolution du STEC CAT sur le STEC P&C et sur le STEC total

Une fois diversifié dans le STEC IARD, le STEC CAT vaut entre 55% et 70% du STEC CAT *standalone*. On remarque que plus le STEC CAT augmente et plus le taux de diversification dans le STEC P&C diminue. Le STEC CAT se diversifie de plus en plus dans le STEC IARD car il n'est pas corrélé avec les STEC prime et réserve.

Une fois diversifié dans le STEC P&C, le STEC IARD vaut entre 82% et 84% du STEC IARD *standalone*. Ce montant est plus élevé que le taux de diversification du STEC CAT dans le STEC IARD. En effet, le STEC IARD est corrélé avec tous les STEC à l'exception du STEC Vie. De ce fait, il génère moins de diversification.

Le taux de diversification est stable : les variations du STEC CAT impactent peu le STEC total. Cela s'explique par le fait que les variations du STEC IARD sont peu matérielles en comparaison du volume des autres STEC.

Diversification du STEC CAT dans le STEC total

Le graphe ci-dessous montre le taux de diversification du STEC CAT dans le STEC total en fonction de l'évolution du STEC CAT. Comme vu précédemment, la majorité de la diversification s'explique par la diversification du STEC CAT dans le STEC IARD.

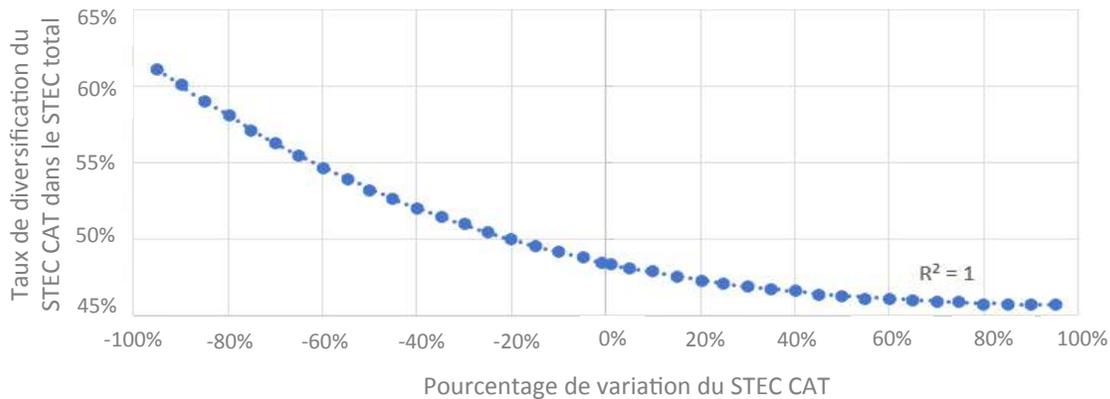


Figure 41 : impact de la réduction du STEC CAT sur le STEC Total

Une régression est effectuée pour prédire le taux de diversification du STEC CAT dans le STEC total en fonction du taux de variation du STEC CAT. La fonction trouvée a une très bonne qualité de prédiction : son R carré est de 1.

On note cette fonction f_{div} et on calcule le STEC CAT diversifié comme suit :

$$STEC_{CAT\ diversifié}^{Nouveau\ ptf} = STEC_{CAT\ standalone}^{Nouveau\ ptf} * f_{div} \left(\frac{STEC_{CAT\ standalone}^{Nouveau\ ptf}}{STEC_{CAT\ standalone}^{ptf\ ref}} - 1 \right)$$

La méthode d'estimation a l'intérêt de faciliter la compréhension par un public non technique, souvent non familier avec les calculs matriciels.

Par ailleurs, cette méthode permet une réduction par 12 des temps de calcul⁵ :

- 0,085 secondes sont nécessaires pour diversifier le STEC CAT avec la méthode exacte matricielle
- 0,007 secondes sont nécessaires pour diversifier le STEC CAT avec la méthode d'estimation.

A titre d'exemple, si 5 000 scénarios de résiliation sont évalués, la méthode d'estimation permet de faire gagner près de 6,5 minutes de calcul.

⁵ Système d'exploitation MacOS Sierra - processeur Intel Core i5 @1,3 GHz - 4Go RAM

2.2.3.4 Calcul du coût du capital

Pour que les différents scénarios de résiliation soient comparables, le ratio de solvabilité est constant. Le coût du capital se calcule comme suit :

$$C^{Nouveau\ ptf} = STEC_{CAT\ standalone}^{Nouveau\ ptf} * f_{div} \left(\frac{STEC_{CAT\ standalone}^{Nouveau\ ptf}}{STEC_{CAT\ standalone}^{ptf\ ref}} - 1 \right) * Ratio\ de\ Solvabilité * Spread$$

2.2.4 Synthèse et calcul de l'ECR

Dans les trois parties précédentes, des méthodes ont été développées pour calculer l'impact d'un scénario de résiliation sur le ratio combiné brut, le résultat cédé et le coût du capital.

Cette partie a pour objectif de présenter un récapitulatif des différentes étapes de calcul présentées et d'explicitier comment obtenir l'impact final sur l'ECR.

Les notations ont déjà été explicitées dans les parties précédentes et ne seront pas rappelées ici.

1. Première étape : choix d'un scénario de résiliation

$$B = (b_1, b_2, \dots, b_p)$$

2. Deuxième étape : calcul du ratio combiné brut

Etape 2.1 : calcul des primes brutes : $\pi^{Nouveau\ ptf} = \sum_{i=1}^p \pi_i * b_i$

Etape 2.2 : calcul des sinistres bruts : $S^{Nouveau\ ptf} = \sum_{i=1}^p S_i * b_i$

Etape 2.2 : calcul du ratio combiné brut

$$CoR_{brut}^{Nouveau\ ptf} = \frac{S^{Nouveau\ ptf} + f^{brut} + \sigma_{acq}^{brut} * \pi^{Nouveau\ ptf} + \sigma_{gestion}^{brut} * \pi^{ptf\ ref}}{\pi^{Nouveau\ ptf}}$$

3. Troisième étape : calcul du ratio combiné net

Etape 3.1 : mise à jour de l'YELT brut : $Brut_{i,j}^{Nouveau\ ptf} = Brut_{i,j}^{ptf\ ref} - Brut_{i,j}^{Polices\ résiliées}$

Etape 3.2 : calcul de la limite du traité : $Limite^{Nouveau\ ptf} = OEP_{200y}(Brut_{i,j}^{Nouveau\ ptf}) - Priorité$

Etape 3.3 : calcul du Loss On Line et du coefficient de variation

$$Recup_i^{ptf\ nouveau} =$$

$$\min((nb\ rec + 1) * Limite^{Nouveau\ ptf}, \sum_{j=1}^{n_i} \min(Limite^{Nouveau\ ptf}, \max(Brut_{i,j}^{Nouveau\ ptf} - Priorité, 0)))$$

$$R = E(Recup^{Nouveau\ ptf}) = \frac{\sum_{i=1}^n Recup_i^{Nouveau\ ptf}}{n}$$

$$LoL^{ptf\ nouveau} = \frac{E(Recup^{Nouveau\ ptf})}{Limite^{Nouveau\ ptf}}$$

$$CoV^{ptf\ nouveau} = \sqrt{\frac{\sum_{i=1}^n (Recup_i - E(Recup^{Nouveau\ ptf}))^2}{n}} * \frac{1}{E(Recup^{Nouveau\ ptf})}$$

Etape 3.4 : estimation du prix du traité

$$p^{\text{Nouveau ptf}} = ((\alpha + 1) * LoL^{\text{Nouveau ptf}} + \beta * CoV^{\text{Nouveau ptf}}) * Limite^{\text{Nouveau ptf}}$$

Etape 3.5 : estimation du ratio combiné net

$$COR_{net}^{\text{Nouveau ptf}} = COR_{brut}^{\text{Nouveau ptf}} + \frac{p^{\text{Nouveau ptf}} * (1 + \sigma^{réass}) - E(Recup^{\text{Nouveau ptf}}) + f^{réass}}{\pi^{\text{Nouveau ptf}}}$$

4. Quatrième étape : calcul du coût du capital

Etape 4.1 : calcul de l'YLT net de réassurance

$$Net_i^{\text{Nouveau ptf}} = Brut_i^{\text{Nouveau ptf}} - Recup_i^{\text{Nouveau ptf}}$$

Etape 4.2 : calcul du STEC CAT standalone

$$STEC_{CAT\ standalone}^{\text{Nouveau ptf}} = Var_{99.5\%}(Net^{\text{Nouveau ptf}}) - E(Net^{\text{Nouveau ptf}})$$

Etape 4.3 : estimation du STEC CAT diversifié

$$STEC_{CAT\ diversifié}^{\text{Nouveau ptf}} = STEC_{CAT\ standalone}^{\text{Nouveau ptf}} * f_{div} \left(\frac{STEC_{CAT\ standalone}^{\text{Nouveau ptf}}}{STEC_{CAT\ standalone}^{ptf\ ref}} - 1 \right)$$

Etape 4.4 : calcul du coût du capital

$$C^{\text{Nouveau ptf}} = STEC_{CAT\ diversifié}^{\text{Nouveau ptf}} * Ratio\ de\ Solvabilité * Spread$$

5. Cinquième étape : calcul des impôts

Etape 5.1 : calcul de l'impôt sur le résultat

$$Impôts_{résultat}^{\text{Nouveau ptf}} = \tau * \pi^{\text{Nouveau ptf}} * (1 - COR_{net}^{\text{Nouveau ptf}})$$

Etape 5.2 : calcul de l'impôt sur les rendements financiers

$$Impôts_{financier}^{\text{Nouveau ptf}} = \tau * STEC_{CAT\ diversifié}^{\text{Nouveau ptf}} * Ratio\ de\ Solvabilité * \varphi$$

6. Sixième étape : calcul de l'ECR

$$ECR^{\text{Nouveau ptf}} = COR_{net}^{\text{Nouveau ptf}} + \frac{C^{\text{Nouveau ptf}} + Impôts_{résultat}^{\text{Nouveau ptf}} + Impôts_{financier}^{\text{Nouveau ptf}}}{\pi^{\text{Nouveau ptf}}}$$

2.2.5 Exemple

Pour illustrer l'intérêt de la méthode développée, prenons l'exemple d'un souscripteur qui aurait la possibilité de résilier l'une des deux polices suivantes :

- Police 1 : localisée à Quintana Roo
- Police 2 : localisée en Basse-Californie du Sud



Figure 42 : localisation des deux sites

Le souscripteur souhaite résilier la police qui lui permettra de maximiser la rentabilité de son portefeuille. Grâce aux résultats du modèle CAT, il dispose des informations suivantes :

	Police 1	Police 2
Sinistres bruts moyens	407 520	315 639
Prime brute	100 000	100 000
Résultat brut moyen	-307 520	-215 639
OEP 200 ans standalone	18 413 851	11 766 899

Figure 43 : métriques obtenues pour les deux polices

Intuitivement, le souscripteur pourrait penser que la meilleure stratégie de souscription serait de résilier la police 1. En effet, les deux polices ont la même prime mais le résultat de la police 1 est 32% plus mauvais que celui de la police 2. Par ailleurs, la police 1 est la plus risquée : elle a une OEP bicentenaire 56% plus élevée que celle de la police 2.

Grâce à la méthode d'estimation de l'impact sur l'ECR, le souscripteur dispose d'une vision complète sur l'impact de la résiliation de chaque police. Il obtient les indicateurs suivants :

		Police 1	Police 2
Impact marginal par rapport au portefeuille de référence	OEP 200 ans	0	- 1 169 385
	COR brut	-0,6%	-0,4%
	COR net	-0,8%	-0,9%
	ECR	-0,5%	-0,8%
	Création de valeur	+ 199 511	+ 308 978

Figure 44 : indicateurs marginaux par rapport au portefeuille de référence

Remarque 1 : Alors que la police 1 a une OEP 56% plus élevée que la police 2, c'est pourtant la police 2 qui permet de faire diminuer l'OEP du portefeuille de référence. En effet, la résiliation de la police 2 entraîne une diminution de l'OEP du portefeuille de référence de près de 1,2 million alors que la police 1 n'a pas d'impact sur l'OEP.

Ce cas de figure pourrait avoir lieu si l'événement bicentenaire du portefeuille de référence était par exemple un ouragan touchant la côte pacifique. La résiliation de la police 1, située sur la côte atlantique, ne vient pas réduire l'événement bicentenaire. Par ailleurs, la résiliation de la police 1 n'a pas assez d'impact sur la distribution des pertes pour qu'il y ait un changement d'événement bicentenaire.

Cet exemple illustre bien la différence entre la vision *standalone* et la vision portefeuille. Vue indépendamment, une police peut être beaucoup plus risquée qu'une autre. Mais si on compare l'impact de chacune de ces polices par rapport au profil du portefeuille, la situation peut s'inverser. L'OEP *standalone* ne permet pas d'anticiper l'impact de la résiliation d'une police sur l'OEP du portefeuille.

Remarque 2 : La police 1 est celle qui a le plus d'impact sur le ratio combiné brut. La résiliation de cette police entraîne une baisse de 0,6% du ratio combiné du portefeuille contre 0,4% pour la police 2.

En revanche, la tendance s'inverse si on prend en compte le résultat cédé et le coût du capital : si la police 2 était résiliée, l'ECR du portefeuille décroîtrait de 0,8% contre 0,5% pour la police 1. Cela représente une différence de création de valeur d'environ 109 000 euros entre les deux scénarios de résiliation.

Cela s'explique par le fait que la police 2 permet de faire diminuer la limite du traité de réassurance et a fortiori les coûts de réassurance. Par ailleurs, la police 2 a plus d'impact sur le STEC.

Cet exemple illustre bien le fait que le ratio combiné brut n'est pas suffisant pour piloter les décisions de souscription. Pour prendre les décisions les plus efficaces d'un point de vue économique, il est nécessaire de pouvoir estimer l'impact sur le ratio combiné économique.

En conclusion, la police qu'il serait la plus judicieux de résilier est la police 2. Pour prendre la meilleure décision de souscription d'un point de vue économique, le souscripteur a donc besoin d'outils pour pouvoir estimer les différents impacts d'un scénario de résiliation.

La méthode d'estimation d'impact a été utilisée ici pour départager deux polices. Dans la partie suivante, elle sera appliquée à chacune des polices du portefeuille dans le but d'identifier les polices qui pénalisent le plus la rentabilité du portefeuille.

2.3 Application : identification des polices pénalisant le plus la rentabilité du portefeuille

Dans la partie précédente, on a développé une méthode d'estimation permettant d'anticiper l'impact d'un scénario de résiliation sur le ratio combiné économique du portefeuille. Dans cette partie, la méthode d'estimation d'impact est appliquée à chacune des polices du portefeuille. L'objectif est d'identifier les polices les moins rentables : ces polices feront partie du périmètre de l'optimisation.

Sélection des polices dont la résiliation est la plus efficace

Pour sélectionner les polices les moins rentables, on définit un indicateur de performance : l'efficacité marginale. Cet indicateur est le rapport entre l'ECR marginal et le pourcentage de primes résiliées. Pour chaque scénario de résiliation, il se calcule comme suit :

$$\text{Efficacité marginale} = \frac{ECR_{\text{portefeuille de référence}} - ECR_{\text{nouveau portefeuille}}}{\text{Pourcentage de primes brutes résiliées}}$$

L'efficacité marginale permet de quantifier l'impact d'un point de pourcentage de primes résiliées sur l'ECR : si l'efficacité est de 2, cela signifie qu'un euro de primes résiliées permet de faire baisser la charge de l'assureur de 2 euros.

L'efficacité marginale permet donc de comparer deux solutions n'ayant pas le même montant de primes résiliées et d'évaluer la performance des résultats obtenus.

Le nuage de points ci-dessous présente l'efficacité marginale des polices dont la résiliation a permis de réduire l'ECR en fonction de la création de valeur.

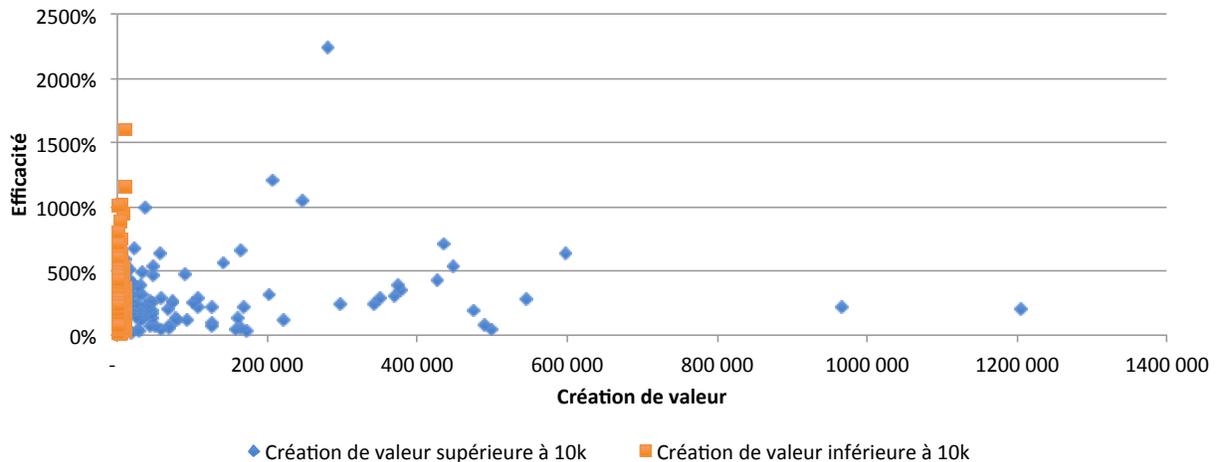


Figure 45 : efficacité des polices permettant de réduire l'ECR

Parmi ces polices, ne sont conservées que celles ayant une efficacité supérieure à 100% et dont l'impact marginal sur la création de valeur est supérieur ou égal à 10 000 euros. En effet, l'entité estime que les coûts de gestion de résiliation ne sont pas rentables pour un impact de moins de 10 000 euros. Ces polices correspondent aux points bleus sur la figure ci-dessus.

Par ailleurs, seules les polices qui ont un impact marginal positif sur le résultat brut sont conservées : les scénarios de résiliation ne doivent pas détériorer la rentabilité brute.

Après avoir appliqué ces deux critères, il reste 127 polices. Elles représentent 8,3 millions de primes, soit près de 22% des primes brutes.

Ces montants sont dans les ordres de grandeur attendus dans notre étude. En effet, l'entité souhaite résilier au maximum 5% des primes brutes, soit environ 2 millions d'euros. Par ailleurs, il est souhaitable d'utiliser un nombre limité de polices dans le périmètre de l'optimisation : chaque police doit être étudiée par la souscription pour déterminer s'il sera effectivement possible de la résilier. De ce fait, il n'est pas possible de choisir un grand nombre de polices.

Caractéristiques des polices choisies

Comme on le voit dans les diagrammes ci-dessous, les 127 polices sont principalement situées dans la région de Quintana Roo et de la Basse-Californie du Sud. Les risques sont majoritairement des maisons individuelles et leur structure en maçonnerie et béton armés.

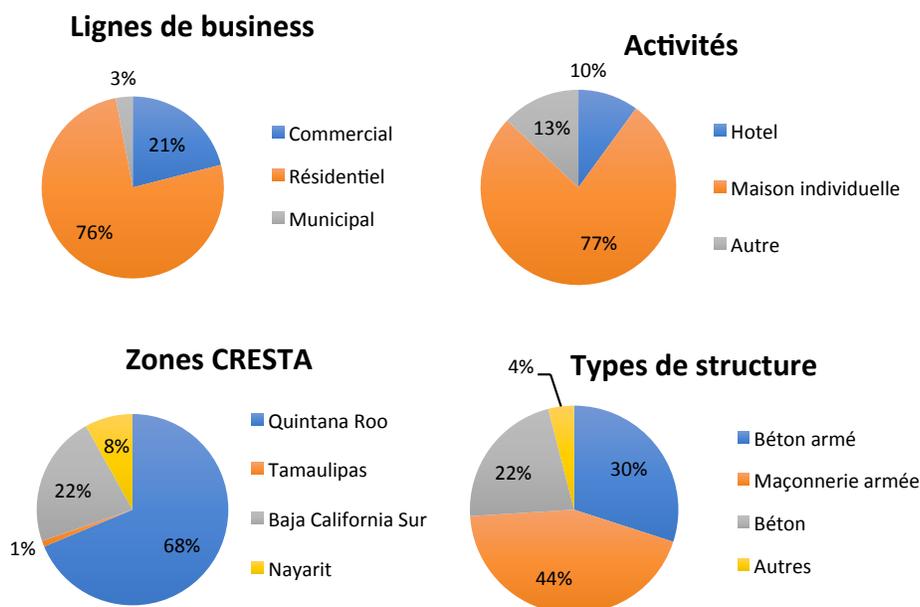


Figure 46 : profil des polices les moins rentables

Les zones CRESTA et les lignes de business correspondent à celles identifiées lors de l'analyse descriptive du portefeuille. Ce n'est pas le cas des types de structure : l'analyse descriptive identifiait la maçonnerie armée comme un type de structure rentable. Par ailleurs, l'activité hôtel représentait une part beaucoup plus importante des sites non rentables.

L'analyse descriptive n'était basée que sur le résultat brut. La prise en compte du résultat cédé et du coût du capital a permis d'identifier des polices non rentables qui n'avaient pas été identifiées précédemment.

Partage des résultats avec les équipes de souscription

Il est nécessaire de partager ces résultats avec les équipes de souscription. Cela leur permet d'identifier les polices qui pèsent sur le résultat cédé et sur le coût du capital.

Les 127 polices seront ensuite étudiées pour identifier celles qui pourraient opérationnellement être résiliées. Ces polices feront partie du périmètre de l'optimisation.

3. Optimisation d'une politique de résiliation de risques à l'aide d'un algorithme génétique

Dans le chapitre précédent, le portefeuille a été analysé. Grâce aux méthodes développées, nous sommes désormais en mesure de quantifier l'impact de n'importe quel scénario de résiliation sur le ratio combiné économique du portefeuille. Par ailleurs, nous avons identifié les polices les moins rentables et les résultats ont été partagés avec les équipes de souscription.

Ce chapitre traite les deuxième et troisième étapes du schéma ci-dessous :

- L'actuarial identifie les scénarios de résiliation optimaux à l'aide d'algorithmes adaptés au problème.
- La faisabilité des scénarios proposés est ensuite étudiée par les équipes de souscription, *Risk Management* et réassurance. Lors de l'étude de faisabilité, de nouvelles contraintes métier peuvent apparaître : contrainte sur le montant de primes à résilier, sur le volume de polices, etc. L'actuarial doit alors être en mesure d'intégrer ces contraintes dans les algorithmes et proposer rapidement de nouveaux scénarios de résiliation optimaux. Les algorithmes utilisés doivent donc être flexibles et permettre de prendre en compte les nouvelles contraintes.

Les étapes 2 et 3 se répètent jusqu'à ce qu'une solution satisfaisante soit obtenue. Elle est alors mise en œuvre : c'est l'étape 4. Cette dernière étape n'est pas traitée ici car ce mémoire est une démonstration de faisabilité et n'a pas été implémentée.

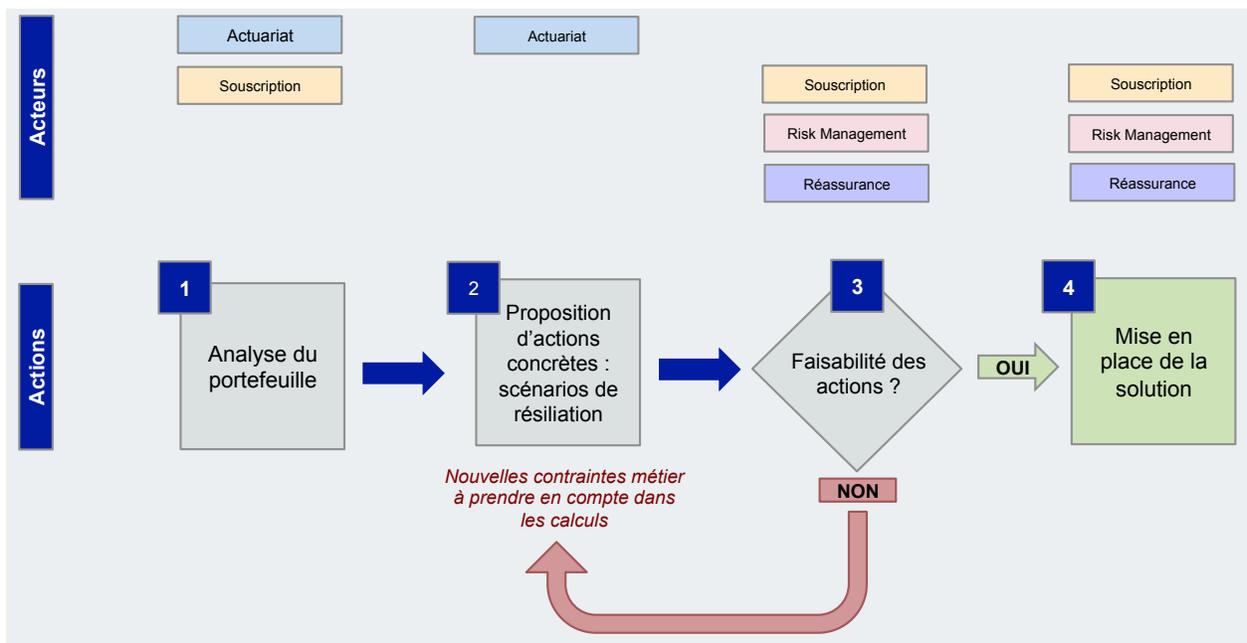


Figure 47 : processus de résolution

Ce chapitre se structure en trois parties.

Dans un premier temps, le problème d'optimisation sera cadré : les objectifs de l'optimisation seront définis ainsi que les différentes contraintes à considérer. Par la suite, une classe d'algorithmes d'optimisation adaptée à notre problème sera choisie : les algorithmes heuristiques et plus particulièrement les algorithmes génétiques.

Dans un deuxième temps, les algorithmes génétiques seront présentés et leurs différents paramètres seront étudiés. Une attention particulière sera accordée à la génération de la population initiale.

Dans un troisième et dernier temps, les résultats de l'optimisation seront analysés. Les algorithmes seront ensuite adaptés pour prendre en compte de nouvelles contraintes métiers : ajout d'une contrainte puis ajout d'un nouvel objectif. Enfin, des études de sensibilité seront effectuées pour étudier la stabilité des résultats.

3.1 Cadrage du problème

Cette partie a pour but de cadrer le problème d'optimisation. **Premièrement**, les objectifs de l'optimisation seront définis. **Deuxièmement**, les différentes contraintes seront identifiées. **Troisièmement**, des algorithmes et méthodologies adaptés au problème seront choisis.

3.1.1 Définition du problème

Une optimisation bi-objectif

Avec un ratio combiné économique de 113%, le portefeuille étudié n'est pas rentable : après rémunération des actionnaires, chaque euro de prime encaissée entraîne une détérioration des fonds propres de 13 centimes. Forte de ce constat, l'entité souhaite résilier des polices pour améliorer la rentabilité du portefeuille.

Plus l'ECR est petit et meilleure est la rentabilité. L'entité souhaite donc minimiser l'ECR tout en minimisant le pourcentage de primes résiliées. Le problème consiste donc à optimiser le système suivant :

$$\begin{cases} \text{Min ECR} \\ \text{Min pourcentage de primes résiliées} \end{cases}$$

L'optimisation est donc bi-objectif. Par ailleurs, les deux objectifs sont contradictoires : plus on va résilier de polices non rentables et plus le ratio combiné économique va s'améliorer. Par conséquent, il est nécessaire de faire un compromis entre les différents objectifs.

La figure ci-dessous présente les deux principales méthodes permettant de résoudre un problème bi-objectif : les méthodes dites « a priori » et « a posteriori ». Contrairement aux méthodes a priori, les méthodes a posteriori ne requièrent pas de connaître le compromis entre les différents objectifs dès le départ.

Chaque point rouge correspond à une solution : il s'agit d'un scénario de résiliation, composé d'une combinaison de polices résiliées. L'axe des abscisses indique le pourcentage de primes résiliées. L'axe des ordonnées indique l'ECR du portefeuille après prise en compte des résiliations.

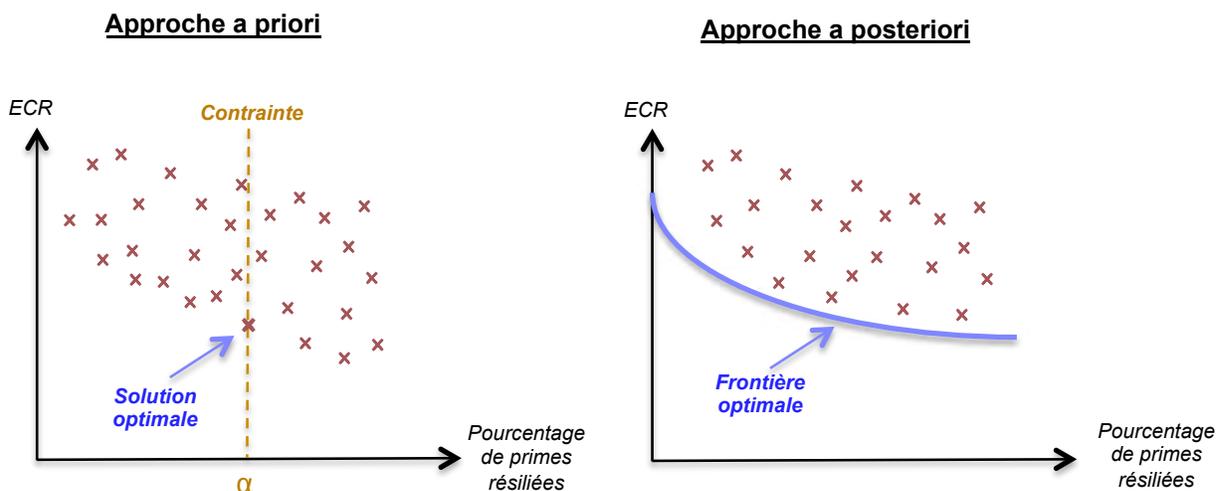


Figure 48 : comparaison des approches a priori et a posteriori

Approche a priori

La méthode a priori a pour but de déterminer une solution optimale parmi l'ensemble des scénarios de résiliation. Cette solution correspond au plus petit ECR obtenu pour $\alpha\%$ de primes résiliées. Cette méthode est adaptée aux problèmes pour lesquels il existe une contrainte sur un des objectifs dès le départ. Ce n'est pas le cas de notre problème : au départ, il n'y a pas de contrainte fixe sur le pourcentage de primes à résilier.

L'approche a priori n'est donc pas adaptée à notre problème et ne sera pas retenue.

Approche a posteriori

Les méthodes a posteriori ne requièrent pas de connaître de compromis entre les différents objectifs dès le départ. Ces méthodes reposent sur la détermination d'une frontière optimale : il s'agit de l'ensemble des solutions pour lesquelles il n'est pas possible d'améliorer un critère sans en dégrader un autre.

La définition mathématique de cette frontière repose sur la notion de domination. Notons C_1, \dots, C_M les critères de l'optimisation et supposons que l'objectif est de les minimiser. Soit x et y deux allocations candidates, alors : x domine $y \Leftrightarrow \forall m \in \llbracket 1, M \rrbracket, C_m(x) \leq C_m(y)$ et l'inégalité est stricte pour au moins un m . La frontière optimale est l'ensemble des allocations non dominées.

Sur la figure ci-dessus, on remarque que tous les points rouges sont situés au-dessus de la frontière efficiente : ce sont des solutions dominées. Pour chaque solution dominée, il existe une ou plusieurs solutions ayant un ECR plus faible, pour un pourcentage de primes résiliées inférieur ou égal.

Les solutions situées sur la frontière efficiente correspondent aux combinaisons de polices dont la résiliation a le plus d'impact sur l'ECR, pour un pourcentage de primes résiliées donné.

Avec les méthodes a posteriori, il est possible de visualiser les différents compromis possibles entre l'ECR et le pourcentage de primes résiliées avant de faire un choix. Ces méthodes permettent d'aider à la décision lorsqu'il n'y a pas de compromis entre les différents objectifs dès le départ. Par conséquent, les méthodes a posteriori sont donc adaptées à notre problème.

3.1.2 Contraintes opérationnelles

Le problème d'optimisation a pour objectif de minimiser le ratio combiné économique et le pourcentage de primes résiliées. Cette optimisation doit cependant respecter certaines contraintes opérationnelles.

Contraintes métier

Tout d'abord, il n'est pas possible de résilier partiellement une police : la police est soit totalement résiliée soit totalement conservée. En effet, résilier une part d'une police obligerait d'avoir recours à la réassurance facultative, jugée trop coûteuse en temps de gestion par l'entité.

Par conséquent, si une police est résiliée, elle sera totalement résiliée. Dans le cas où la police dispose d'une fac, celle-ci sera également résiliée.

Ensuite, la résiliation s'effectue au niveau des polices : si seul un site est non rentable au sein d'une police multi-sites, il ne sera pas possible de le résilier sans résilier la police entière. En effet, les polices multi-sites couvrent généralement différents bâtiments appartenant à un même client : par exemple les différentes usines d'un constructeur automobile. Il n'est donc pas possible de résilier un seul site sans résilier la police entière.

Il est également nécessaire de prêter attention à l'apporteur d'affaires. Dans le cas où les polices ont été souscrites par l'intermédiaire d'un apporteur d'affaires, par exemple un courtier, il faut considérer les polices dans leur ensemble et non pas indépendamment. Il ne sera pas possible de résilier une seule police

parmi toutes les polices d'un apporteur d'affaires : soit toutes les polices seront résiliées, soit aucune ne sera résiliée.

Enfin, toutes les polices ne vont pas être considérées dans le cadre de notre étude : l'entité peut souhaiter ne pas résilier certaines polices, bien que celles-ci ne soient pas rentables. Ainsi, certains clients sont couverts sur plusieurs lignes de business (Cat, Responsabilité Civile, etc). La résiliation de la police cat pourrait être mal perçue par le client et l'amener à changer d'assureur pour tous ses contrats. Une fois les polices non rentables identifiées, il est donc nécessaire d'avoir l'avis de la souscription sur les polices à considérer dans l'étude.

Volume de primes et de polices limité

L'entité souhaite résilier un volume de polices et de primes limité. Le but est de ne pas envoyer de mauvais signal au marché. L'entité souhaite résilier au maximum 4% de ses primes brutes.

Evolution des contraintes

Les contraintes sont susceptibles d'évoluer au fil de l'étude. En effet, le pourcentage de primes à résilier est amené à être affiné au fil de l'avancée de l'étude et de nouvelles contraintes peuvent apparaître.

Temps de calcul

Notre étude a le risque de demander des temps de calcul importants. Même en restreignant l'étude aux cent polices les moins rentables du portefeuille, cela représente 2^{100} scénarios de résiliation possibles. Chaque scénario étant évalué en une $0,143^6$ seconde, plus de $5,7 \times 10^{21}$ années de calcul seraient nécessaires pour évaluer toutes les combinaisons possibles.

Même en restreignant le nombre de polices prises en compte dans l'optimisation, il n'est donc pas possible d'évaluer tous les cas possibles. Dès lors, il est nécessaire d'utiliser des algorithmes adaptés pour limiter les temps de calcul. En effet, pour que les méthodes développées dans ce mémoire puissent être utilisées opérationnellement, il est impératif que les temps de calcul soient raisonnables.

3.1.3 Démarche de résolution

Cette partie a pour objectif de proposer une démarche de résolution à notre problème. Après avoir choisi un format d'encodage et sélectionné les polices qui feront partie du périmètre de l'optimisation, des classes d'algorithmes adaptés à notre problème seront choisies.

Encodage du problème

Il n'est pas possible de résilier une partie d'une police : la police est soit totalement conservée, soit totalement résiliée. De ce fait, un encodage binaire est adapté à notre problème :

- Si une police est 100% résiliée, elle est encodée 1.
- Si une police est conservée (0% résiliée), elle est encodée 0.

Les solutions recherchées dans notre problème sont donc un ensemble de p -uplet (b_1, \dots, b_p)

Chaque gène b_i peut prendre deux valeurs :

- $b_i = 1$: la police i est résiliée. Si la police dispose d'un fac, celle-ci est également résiliée.
- $b_i = 0$: la police i est conservée. Si la police dispose d'une fac, celle-ci est également conservée.

Limitation du périmètre des polices étudiées

⁶ Système d'exploitation MacOS Sierra, processeur Intel Core i5 @1,3 GHz, Mémoire (RAM) 4Go

Dans la partie 2.3, les polices pénalisant le plus la rentabilité du portefeuille ont été identifiées. Ces polices ont un impact marginal d'au moins 10 000 euros sur la création de valeur et permettent d'améliorer la rentabilité brute. Ces polices sont au nombre de 127.

Parmi les 127 polices, toutes ne sont pas résiliables, pour les raisons évoquées précédemment. L'entité a donc étudié la liste et a retenu 100 polices qui pourraient opérationnellement être résiliées.

Les 100 polices choisies représentent 10% des primes totales du portefeuille, ce qui est estimé suffisant puisque l'entité souhaite résilier jusqu'à 4% de primes.

Formalisation

On cherche un ensemble de 100-uplets $b = (b_1, \dots, b_{100})$, avec $b_i \in \{0; 1\}$ qui minimise le système suivant :

$$\begin{cases} \text{Min ECR} \\ \text{Min pourcentage de primes résiliées} \end{cases}$$

L'ECR de chaque nouveau portefeuille est calculé à l'aide de la méthode d'estimation d'impact décrite dans le chapitre 2 partie 2.

Le pourcentage de primes résiliées se calcule comme suit :
$$\frac{\sum_{i=1}^{100} b_i * \pi_i}{\text{Primes du portefeuille de référence}}$$

Avec π_i la prime de la police i . Il s'agit de la prime brute payée par l'assuré de laquelle sont déduites les éventuelles primes de réassurance facultative payées par l'entité.

Comme on le voit sur la figure ci-dessous, aucune contrainte n'est appliquée sur les primes. Les solutions admissibles sont déterminées a posteriori. Cette méthode permet de disposer d'un niveau de flexibilité important. Cela permet à l'entité de pouvoir adapter ses décisions de résiliation à la hausse ou à la baisse en cas de besoin.

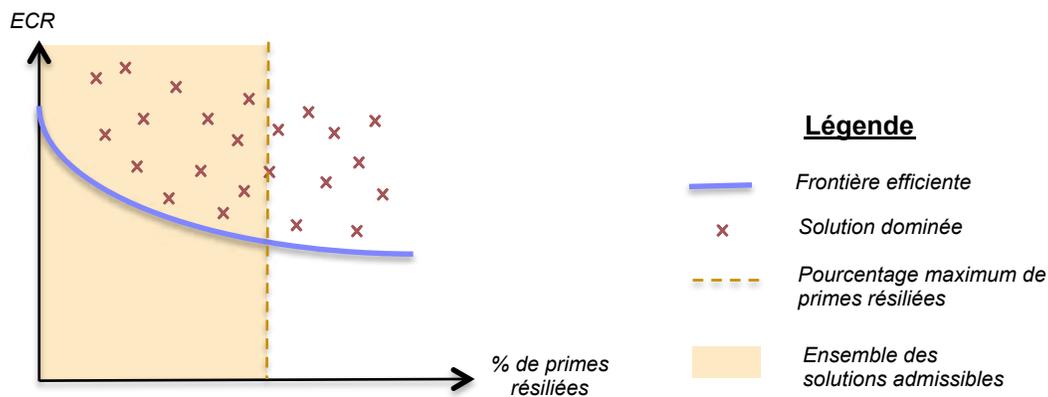


Figure 49 : démarche de résolution

Classe d'algorithmes adaptée au problème

Notre problème est un problème d'optimisation combinatoire multi-objectif : on souhaite identifier les combinaisons de polices dont la résiliation permettrait de minimiser l'ECR du portefeuille tout en minimisant le pourcentage de primes résiliées.

Ce problème est analogue à celui du sac à dos multidimensionnel, un problème très classique d'allocation de ressources qui est étudié depuis plus d'un siècle. Schématiquement, le problème consiste à remplir un

sac à dos avec des objets. Chaque objet a un poids et une valeur et on souhaite maximiser la valeur du sac à dos tout en respectant une contrainte de poids.

Le problème du sac à dos trouve des applications dans divers secteurs, comme par exemple l'industrie (découpe de métaux) ou la finance (optimisation de portefeuille).

Dans notre étude, le poids du sac à dos est l'ECR du portefeuille. Les objets sont les polices : elles ont chacune un poids (leur impact marginal sur l'ECR) et un prix (leur prime). On souhaite vider le sac à dos (résilier des polices du portefeuille) pour minimiser son poids (l'ECR du portefeuille) tout en maximisant sa valeur (résilier le moins de primes possible).

Le problème du sac à dos appartient à la classe de complexité NP. En effet, on sait vérifier rapidement toute solution proposée mais on ne sait pas en trouver efficacement. Les algorithmes connus pour résoudre des problèmes NP ont un temps d'exécution exponentiel en la taille des données d'entrée et sont donc inexploitable en pratique, même pour des cas de taille modérée. C'est le cas pour notre problème : il existe 2^{100} scénarios de résiliation possibles.

Traditionnellement, ces problèmes sont résolus à l'aide de méthodes approchées comme par exemple des algorithmes heuristiques. Une heuristique est une stratégie pour se déplacer intelligemment dans l'espace des solutions, afin d'obtenir une solution approchée la meilleure possible, dans un temps de calcul raisonnable. L'avantage de cette classe d'algorithmes est qu'il n'y a pas besoin que le problème vérifie certaines propriétés (convexité, continuité, etc). Il est simplement nécessaire de pouvoir évaluer les solutions candidates, ce qui est le cas dans notre problème : nous sommes capables de quantifier l'impact d'un scénario de résiliation sur le ratio combiné économique du portefeuille. Les algorithmes heuristiques sont particulièrement adaptés là où les algorithmes mathématiques usuels échouent, sont longs à converger ou aboutissent à un optimum local.

On appelle méta-heuristiques des heuristiques dont le but est de trouver un optimum global. L'idée sous-jacente est de parcourir l'espace de recherche et d'explorer les zones qui paraissent prometteuses, sans être piégé par un optimum local. Ces algorithmes sont généralement inspirés de processus naturels, comme par exemple la théorie de l'évolution ou le refroidissement du métal.

Dans la littérature, l'algorithme méta-heuristique le plus fréquemment utilisé pour résoudre le problème du sac à dos multidimensionnel est l'algorithme génétique. Les algorithmes génétiques appartiennent au groupe des algorithmes évolutionnaires. Comme leur nom le suggère, ils sont inspirés de la théorie darwinienne de l'évolution.

Dans notre étude, l'optimisation de la politique de résiliation sera réalisée à l'aide d'un algorithme génétique.

3.2 Algorithmes génétiques et méthodologie choisie

3.2.1 Principes généraux

Comme leur nom le suggère, les algorithmes génétiques sont inspirés de la théorie darwinienne de l'évolution. Des phénomènes inspirés de la nature tels que la sélection naturelle ou la mutation génétique sont appliqués à une population dans le but de la perfectionner pour répondre à un problème donné.

Les algorithmes génétiques appartiennent au groupe des algorithmes évolutionnaires. Ces algorithmes sont définis comme des méthodes d'optimisation « heuristiques ». L'heuristique désigne toutes les stratégies qui consistent à résoudre un problème dans le cas où il n'est pas possible de rechercher une solution dans tout l'espace des possibilités. Ces solutions tendent à trouver une solution qui s'approchent de l'optimum.

John Holland marque un tournant décisif dans le développement des algorithmes génétiques. Dans son ouvrage *Adaptation in Natural and Artificial Systems (Holland, 1992 [8])* il étudie les phénomènes évolutionnaires dans la nature dans le but de les transcrire dans le domaine informatique. Cette publication est considérée comme fondatrice de la théorie des algorithmes génétiques.

Ce type d'algorithme vise à trouver des optimums en manipulant une population de taille constante. La population est composée de points candidats appelés « chromosomes ». Chaque chromosome est composé de « gènes », qui constituent une solution potentielle au problème d'optimisation.

La taille constante de la population est à l'origine d'un phénomène de compétition entre chromosomes. Ainsi, à chaque itération de l'algorithme, une nouvelle population de chromosomes est générée : il s'agit d'une nouvelle « génération ». Chaque génération est composée du même nombre de chromosomes et subit une sorte de sélection naturelle : seuls les chromosomes les plus adaptés à leur environnement survivent. Au fil des générations les chromosomes tendent ainsi vers l'optimum de la fonction sélective.

Comme on le voit sur le schéma ci-dessous, la création d'une nouvelle population de chromosomes est basée sur cinq principes :

- **La sélection** : cette étape consiste à sélectionner des chromosomes « parents » qui seront par la suite combinés pour produire des chromosomes « enfants ».
- **La reproduction (aussi appelée croisement ou enjambement)** : à partir des chromosomes sélectionnés (chromosomes « parents »), des chromosomes « enfants » sont générés. Un parallèle est fait ici avec l'héritage génétique en biologie : un enfant possède à la fois des gènes de sa mère et de son père. L'objectif sous-jacent est de faire se reproduire les meilleurs chromosomes pour les perfectionner au fil des générations.
- **La mutation** : la mutation consiste à modifier un gène d'un chromosome. Ce chromosome est habituellement choisi aléatoirement. Ce phénomène a pour but de sortir des minimums locaux.
- **L'évaluation** : les individus sont évalués au regard de certains critères définis par l'utilisateur. Pour survivre, les chromosomes doivent être adaptés ou s'adapter à ces critères.
- **L'élitisme** : au cours de cette opération, l'algorithme sélectionne les chromosomes qui répondent le mieux au problème d'optimisation. Un parallèle est fait ici avec la théorie darwinienne de l'évolution : seuls les meilleurs individus survivent.

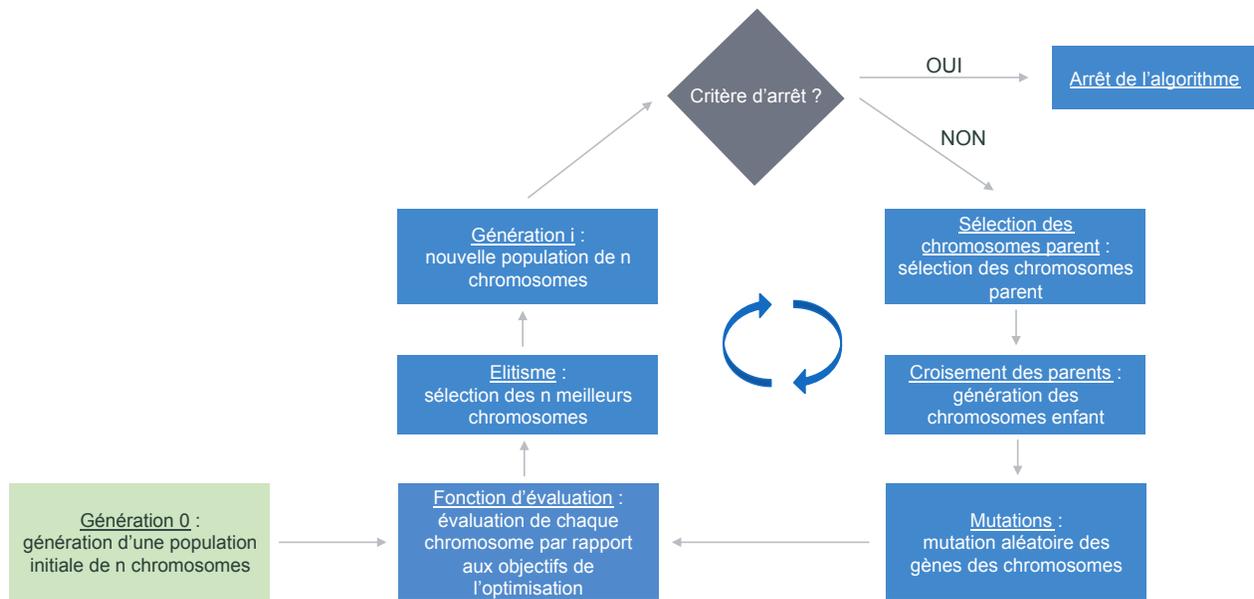


Figure 50 : principes généraux de l'algorithme génétique

A noter que l'algorithme est initialisé à l'aide d'une première population entrée par l'utilisateur ou générée par l'algorithme.

Le mécanisme décrit ci-dessus est ensuite reproduit sur plusieurs générations jusqu'à ce que le critère d'arrêt soit atteint.

Les parties suivantes présentent les différentes étapes de l'algorithme génétique introduites ci-dessus et explicitent les différents paramètres qui ont été choisis pour répondre au problème étudié dans ce mémoire.

3.2.2 Initialisation de la population

La population initiale est composée de 400 chromosomes, chacun composé de 100 gènes b_i . Le schéma ci-dessous présente la structure d'un chromosome.



Figure 51 : structure d'un chromosome

Chaque gène b_i peut prendre deux valeurs :

- $b_i = 1$: la police i est résiliée
- $b_i = 0$: la police i est conservée

Chaque chromosome correspond donc à un scénario de résiliation : $\sum_{i=1}^k b_i$ polices sont résiliées.

On a choisi une population de 400 chromosomes : il est important que la taille de la population soit suffisamment grande pour pouvoir explorer une large partie de l'espace. Par ailleurs, plus la population initiale est grande et plus la probabilité qu'elle contienne des solutions optimales augmente.

On constate que plus la population est grande et meilleurs sont les résultats de l'algorithme génétique (Hassanat et al., 2019, [9]). Cependant, plus la taille de la population augmente et plus l'algorithme est lent à converger, entraînant des coûts de calcul importants.

Dans notre étude, on décide donc d'utiliser une population de 400 chromosomes : suffisamment grande pour explorer largement l'espace mais suffisamment petite pour permettre des temps de calcul raisonnables.

Les algorithmes heuristiques sont fréquemment utilisés dans les problèmes d'optimisation. Cependant, encore peu de recherches ont été faites sur la génération des populations initiales et notamment dans le cas de populations encodées en binaire. Généralement, les populations initiales sont générées de manière aléatoire.

Dans une étude publiée en 2006, Maaranen et al. ([10]) identifient les propriétés recherchées d'une population initiale. Tout d'abord, la population doit disposer d'une bonne diversité génétique : on souhaite que tous les gènes soient représentés. Ensuite, la population doit permettre une couverture uniforme de l'espace de recherche : on souhaite que les chromosomes soient dispersés et qu'il n'y ait pas de zones de l'espace non explorées.

L'objectif de cette partie est de présenter les méthodes qui ont été choisies pour générer une population initiale ayant les propriétés recherchées.

Dans un premier temps, la diversité génétique de la population initiale sera étudiée. **Dans un deuxième temps**, on développera des méthodes pour couvrir de manière uniforme l'espace de recherche. **Dans un troisième temps**, de « bons » chromosomes seront ajoutés dans la population initiale pour accélérer la convergence. **Dans un quatrième temps**, la méthode retenue sera décrite.

3.2.2.1 Diversité génétique

La première propriété recherchée lors de la génération d'une population initiale est la diversité génétique. On souhaite que tous les gènes soient présents dans la population initiale afin que l'algorithme génétique soit capable d'explorer de larges portions de l'espace.

Pour analyser la diversité génétique, nous allons générer aléatoirement une population initiale. La distance entre les chromosomes ainsi que la répartition de chaque gène sera ensuite étudiée.

Génération pseudo-aléatoire d'une population de 400 chromosomes

Dans notre étude, chaque chromosome se compose de 100 gènes et on souhaite générer 400 chromosomes. On va donc générer une séquence binaire pseudo-aléatoire de 40 000 bits et répartir les bits dans les 400 chromosomes.

Une séquence pseudo-aléatoire est déterministe mais elle est aléatoire dans le sens où la valeur d'un bit est indépendante des valeurs de n'importe quel autre bit, comme dans une vraie séquence aléatoire. Ces séquences sont généralement construites à partir d'un aléa, comme par exemple l'heure du système.

Une séquence binaire pseudo-aléatoire est cependant différente d'une véritable séquence aléatoire : la séquence pseudo-aléatoire est cyclique, alors qu'une véritable séquence aléatoire est infinie.

Dans cette étude, on génère un petit nombre de variables, ce type de génération aléatoire est donc considéré suffisant.

On génère une séquence binaire pseudo-aléatoire $B = (b_1, \dots, b_{40\,000})$ à l'aide de la fonction *runif* de R. Cette fonction réplique l'algorithme de Mersenne-Twister. Les valeurs générées sont ensuite attribuées à chaque chromosome.

```
Pour  $i = 1$  à 40 000
  Tirage d'une variable aléatoire uniforme  $u_i$ 
  Si  $u_i \leq 0,5$ , alors  $b_i = 0$ 
  Sinon  $b_i = 1$ 
Fin
Pour  $j = 1$  à 400
  Soit  $C_j$  le chromosome numéro  $j$ 
   $C_j = (b_{(j-1)*100+1}, \dots, b_{j*100})$ 
Fin
```

Distance entre les chromosomes

Pour quantifier la diversité génétique de la population générée, on souhaite calculer la distance génétique entre deux chromosomes.

Pour ce faire, la distance de Hamming est utilisée. Cette métrique permet de calculer la distance entre deux séquences binaires.

Soient deux séquences binaires $A = (a_1, \dots, a_n)$ et $B = (b_1, \dots, b_n)$.

En dimension n , la distance de Hamming $H_{A,B}$ entre les séquences A et B se calcule comme suit :

$$H_{A,B} = \sum_{i=1}^n \text{abs}(a_i - b_i)$$

La distance de Hamming peut être utilisée pour calculer la distance entre deux chromosomes. En effet, comme on le voit sur la figure ci-dessous, calculer la distance de Hamming revient à compter le nombre de gènes différents entre deux chromosomes de la population.

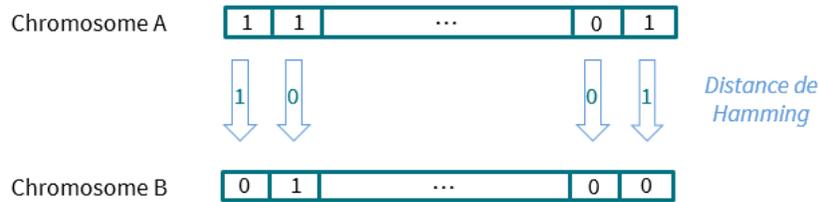


Figure 52 : exemple de calcul de la distance de Hamming entre deux chromosomes

Dans notre étude, on normalise la distance de Hamming par la taille d'un chromosome pour obtenir le pourcentage de gènes différents entre deux chromosomes de la population :

$$H_{A,B}^{normalisé} = \frac{H_{A,B}}{n}$$

Le tableau ci-dessous présente la distance de Hamming normalisée moyenne entre deux chromosomes, ainsi que le coefficient de variation et la distance minimale et maximale.

	Moyenne	CoV	Minimum	Maximum
Distance entre les chromosomes	50%	10%	30%	70%

Figure 53 : distance entre les chromosomes de la population générée aléatoirement

En moyenne, la distance de Hamming entre deux chromosomes est de 50%. Un chromosome étant composé de 100 gènes, deux chromosomes ont en moyenne la moitié de leurs gènes en commun. Cela s'explique par le fait que chaque gène peut prendre deux valeurs (0 et 1), chacune ayant une probabilité de 50% d'être choisie.

Il y a peu de volatilité : le coefficient de variation est de 10%.

La distance de Hamming minimale est de 30, soit un tiers de chromosomes différents entre deux chromosomes.

Le niveau de diversité génétique entre les chromosomes est jugé d'un niveau satisfaisant.

Répartition de chacun des gènes

Après avoir analysé la distance entre les chromosomes, on souhaite étudier la répartition des gènes. L'objectif est de s'assurer que pour chaque gène, chaque variation (0 ou 1) est sélectionnée dans des proportions comparables.

Le tableau ci-dessous présente le pourcentage de sélection des variations de chaque gène.

	Moyenne	CoV	Minimum	Maximum
Pourcentage de sélection d'un gène	50%	7%	40%	60%

Figure 54 : répartition des gènes

En moyenne, chaque gène est sélectionné avec une probabilité de 50%. Cela s'explique par le fait que chaque gène ne peut prendre que deux valeurs : 0 ou 1.

Parmi tous les gènes, la variation la moins sélectionnée a été tirée pour 40% des chromosomes. Chaque gène n'ayant que deux variations possibles, si le pourcentage minimum est de 40%, alors le pourcentage maximum sera son complémentaire : 60%.

En conclusion, la population initiale générée aléatoirement présente un bon niveau de diversité génétique. A ce stade, aucun retraitement de la population initiale n'est nécessaire pour améliorer la diversité génétique.

3.2.2.2 Couverture uniforme de l'espace de recherche

Lors de la génération d'une population initiale, il est souhaitable d'avoir une couverture uniforme de l'espace de recherche. En effet, on ne souhaite pas avoir de *clusters* ou de zones de l'espace non explorées. Cette propriété a pour objectif d'éviter la convergence prématurée de l'algorithme vers un optimum local.

Pour s'assurer que la population initiale dispose d'une couverture uniforme, on étudiera la distance entre les pourcentages de primes résiliées des chromosomes. On définira ensuite des règles pour contraindre une distance minimale entre les primes.

Distance entre les pourcentages de primes résiliées

Le nuage de points ci-dessous présente la répartition des chromosomes de la population initiale générée aléatoirement. En ordonnées, on observe l'évolution de l'ECR et en abscisse le pourcentage de primes résiliées.

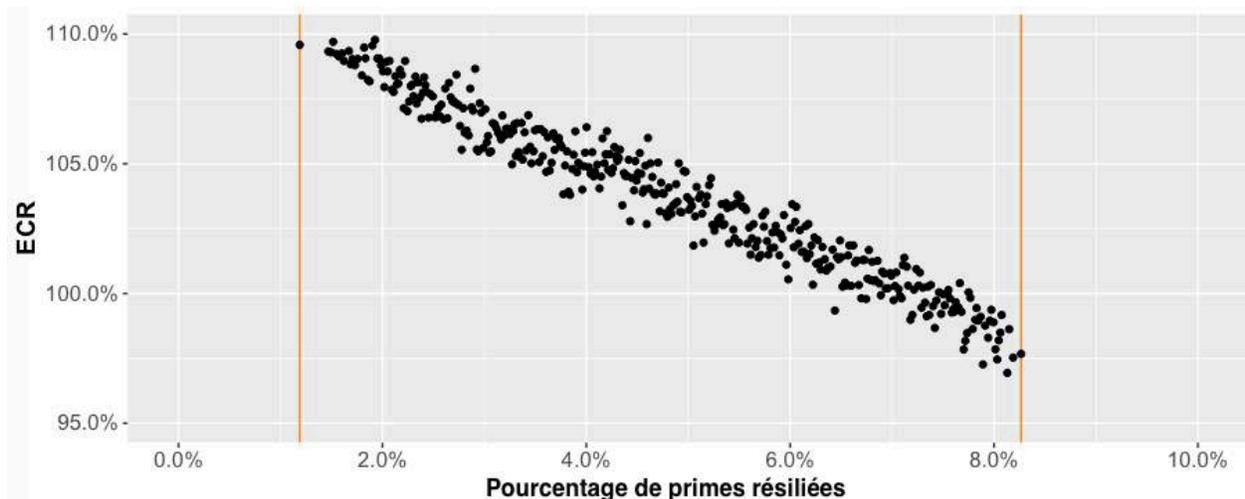


Figure 55 : répartition de la population initiale dans l'espace de recherche

Les axes de couleur orange représentent le pourcentage de primes résiliées minimum et maximum observé. La population initiale est répartie entre 1,8% et 7,8% de primes résiliées.

Dans notre problème, on souhaite résilier jusqu'à 4% des primes. Cette partie de l'espace est sous-explorée : seuls 18% des chromosomes s'y situent et aucun chromosome n'a un pourcentage de primes résiliées inférieur à 1,8%.

La majorité des chromosomes sont concentrés entre 3 et 7% de primes résiliées. Ce constat s'explique par le fait que beaucoup de polices comprises dans le périmètre de l'optimisation ont un montant de prime important. En moyenne, 50% des polices vont être sélectionnées dans un chromosome, ce qui représente en moyenne 5% des primes totales.

Idéalement, on souhaiterait qu'il y ait le plus d'écart de prime possible entre les chromosomes. La population générée se compose de 400 chromosomes et le pourcentage de primes résiliées peut varier entre 0% (chromosome n'ayant que des gènes nuls) et 10% (chromosome ayant tous ses gènes valant 1). Si les 400 chromosomes étaient uniformément répartis, chaque chromosome aurait un écart de 0,025% avec le chromosome juste avant et juste après lui.

Le tableau ci-dessous montre le pourcentage d'écart de primes résiliées entre deux chromosomes. Pour ce faire, les chromosomes générés ont été classés par ordre croissant en fonction du pourcentage de primes résiliées. Le tableau présente l'écart moyen, maximum et minimal entre un chromosome et le chromosome situé juste avant et juste après lui.

	Ecart maximum	Ecart minimum
Pourcentage d'écart de prime	0,28572%	0,00003%

Figure 56 : pourcentage d'écart de prime entre deux chromosomes

L'écart minimum est de 0,00003% : certains chromosomes sont concentrés à un même niveau de primes résiliées. A l'inverse, l'écart maximum est de 0,286%. Cet écart est plus de 11 fois plus grand que l'écart optimal (0,025%) : l'espace de recherche n'est pas uniformément couvert.

En conclusion, la population initiale générée aléatoirement ne couvre pas uniformément l'espace de recherche. Avec cette population, l'algorithme aura besoin des opérations de mutation et de croisement pour espérer explorer les espaces en-dessous de 1,8% de primes résiliées. Il y a un risque que l'algorithme n'explore pas ou pas suffisamment cet espace, alors que c'est l'espace de recherche central de notre étude.

Ajout d'une contrainte sur l'écart minimal de prime entre deux chromosomes

Pour améliorer la couverture uniforme de la population initiale, on propose d'ajouter une contrainte lors de la génération des chromosomes. Seuls les chromosomes assez éloignés des chromosomes déjà créés sont retenus.

On initialise un chromosome C_1 .

Soit $seuil_{prime}$ le seuil de prime minimum à atteindre.

Soit $seuil_{Hamming}$ la distance de Hamming minimum à atteindre.

$i = 2$

Tant que la taille de la population est inférieure à 400 :

 On génère aléatoirement un nouveau chromosome C_i .

 On calcule les distances de Hamming $H = (h_{i, 1}, \dots, h_{i, i-1})$ entre le nouveau chromosome et les $(i - 1)$ chromosomes déjà présents dans la population.

 On calcule les écarts de prime $P = (p_{i, 1}, \dots, p_{i, i-1})$ entre le nouveau chromosome et les $(i - 1)$ chromosomes déjà présents dans la population.

 Si $\min(P) > seuil_{prime}$ et $\min(H) > seuil_{Hamming}$

 On conserve le chromosome

$i = i + 1$

 Sinon, le chromosome C_i est rejeté

Fin

Trois nouvelles populations initiales sont générées en ajoutant une contrainte d'écart sur les primes. Les résultats obtenus se trouvent dans le tableau ci-dessous.

	Ecart de prime entre les chromosomes		Exploration de l'espace	
	Ecart maximum	Ecart minimum	Prime minimum	Prime maximum
Sans contrainte	0,28572%	0,00003%	1,811%	7,821%
Minimum 0,005%	0,25771%	0,00500%	1,723%	7,981%
Minimum 0,01%	0,11920%	0,01001%	1,681%	8,382%
Minimum 0,012%	0,07859%	0,01201%	1,188%	8,266%

Figure 57 : application de contraintes sur l'écart de prime entre les chromosomes

On fait varier la contrainte entre 0,005% et 0,012%. Au-delà de ce montant, il n'est plus possible de générer 400 chromosomes. Plus on augmente la contrainte et plus l'écart maximum entre les chromosomes diminue : les chromosomes sont plus uniformément répartis.

Par ailleurs, augmenter la contrainte sur l'écart de prime permet une meilleure exploration de l'espace : pour une contrainte de 0,012%, les chromosomes ont des pourcentages de primes résiliées entre 1,2% et 8,3%, soit respectivement 0,6% et 0,4% de plus que la population générée sans contrainte.

Le nuage de points ci-dessous présente la répartition dans l'espace de la population générée avec contrainte de 0,012%. Les axes bleus représentent les pourcentages minimaux et maximaux de primes résiliées. Les axes orange représentent les pourcentages obtenus pour la population générée sans contrainte.

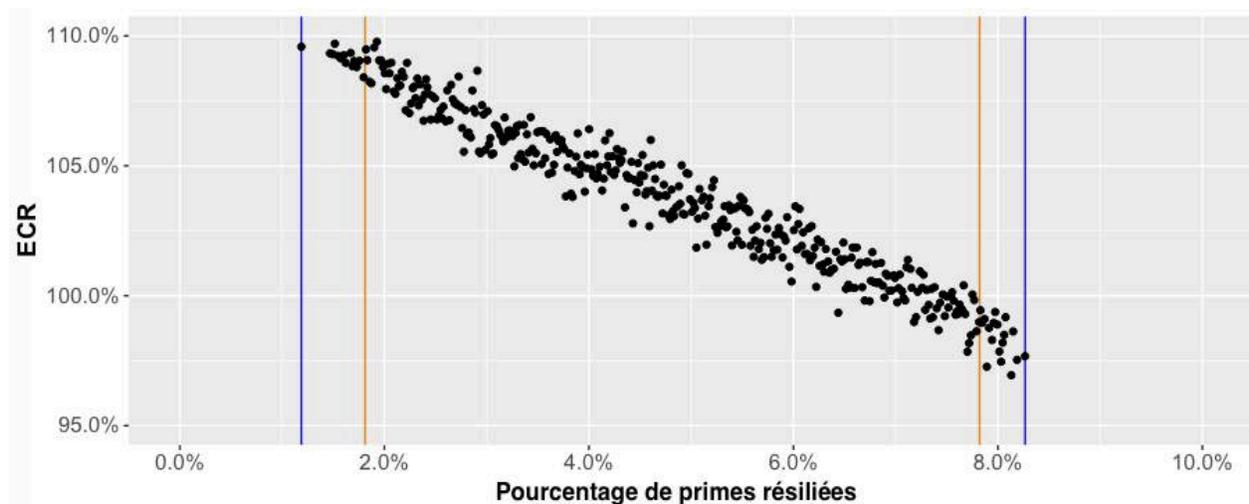


Figure 58 : répartition de la population initiale générée avec contrainte de 0,012%

En ajoutant une contrainte de 0,012%, la nouvelle population initiale explore environ 10% de l'espace en plus.

Par ailleurs, 41% des chromosomes sont inférieurs à 4% contre 18% lorsque la population était générée sans contrainte. On constate que les chromosomes sont plus uniformément répartis.

Cependant, il n'y a pas de chromosomes inférieurs à 1,19% de primes résiliées et un seul est inférieur à 1,4%. Dans notre problème, on souhaite résilier jusqu'à 4% de prime. Il serait donc souhaitable de générer davantage de chromosomes avec un niveau de prime inférieur à 1,4%.

Pénalisation des polices ayant un pourcentage de prime important

Parmi les 100 polices comprises dans le périmètre de l'optimisation, certaines polices ont des montants de primes importants : jusqu'à 1,56% de primes résiliées pour une seule police. Chaque gène ayant une probabilité de 50% d'être sélectionné, il est par conséquent difficile de générer des chromosomes avec des primes inférieures à 1,4%.

Pour remédier à cette situation, on propose de pénaliser les chromosomes ayant un pourcentage de primes résiliées important. Pour ce faire, la procédure suivante est utilisée pour générer un chromosome :

Soient (b_1, \dots, b_{100}) les 100 polices comprises dans le périmètre de l'optimisation.
 On ordonne les polices par pourcentage de primes résiliées croissant. Soit $r_{i,j}$ le rang de la police i .
 On calcule la probabilité p_i que la police i soit sélectionnée : $p_i = 50\% - r_{i,j} * 0,25\%$.
 On génère un chromosome dont chaque gène b_i a une probabilité p_i d'être sélectionné.

La probabilité qu'un gène soit sélectionné varie entre 25% et 50%. Les polices avec un faible pourcentage de primes résiliées sont favorisées.

On génère une nouvelle population initiale. On utilise la procédure de contrainte sur l'écart de prime entre les individus et les chromosomes sont générés en utilisant la méthode de pénalisation décrite ci-dessus.

Les résultats obtenus sont présentés dans le tableau ci-dessous.

Contrainte d'écart entre les primes	Pénalisation des polices ayant un pourcentage de prime important	Ecart de prime entre les chromosomes		Exploration de l'espace	
		Ecart maximum	Ecart minimum	Prime minimum	Prime maximum
Sans contrainte	Non	0,28572%	0,00003%	1,811%	7,821%
Minimum 0,005%	Oui	0,30305%	0,00512%	0,781%	6,432%
Minimum 0,01%	Oui	0,25898%	0,01004%	0,559%	6,690%

Figure 59 : pénalisation des polices en fonction de leur pourcentage de primes résiliées

La pénalisation a permis d'explorer davantage l'espace des petits pourcentages de prime : avec une contrainte d'écart sur les primes de 0,01%, la prime minimale est de 0,6%. Ce montant est raisonnable : l'entité souhaite résilier jusqu'à 4% des primes.

Il n'est pas possible de générer en un temps raisonnable 400 chromosomes avec une contrainte supérieure à 0,01%.

Le nuage de point ci-dessous présente la répartition dans l'espace de la population générée avec contrainte de 0,01% et avec pénalisation. Les axes bleus et orange représentent respectivement les pourcentages minimaux et maximaux de primes résiliées pour la population générée sans contrainte et la population générée avec contrainte de 0,012%.

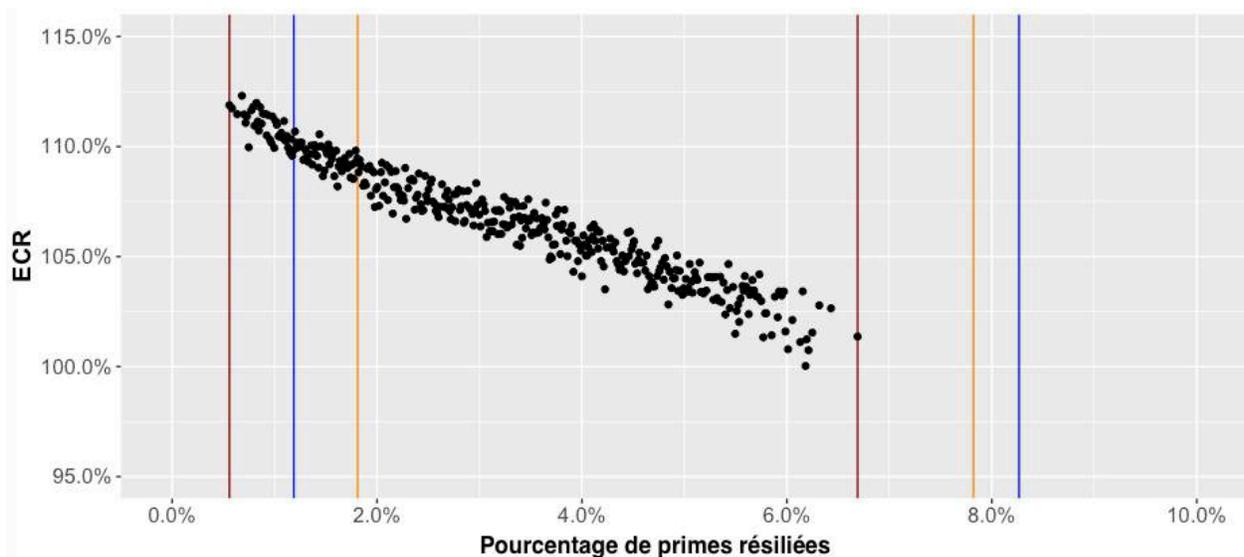


Figure 60 : répartition de la population initiale générée avec pénalisation

On constate que les chromosomes sont concentrés entre 0,6% et 6%. 71% des individus ont une prime inférieure à 4%, soit presque deux fois plus que dans la population générée sans pénalisation. 12% des individus ont une prime inférieure ou égale à 1,4%.

La procédure avec pénalisation permet donc de bien explorer l'espace des primes entre 0,6% et 4%.

On constate que pour un même pourcentage de primes résiliées, l'ECR peut varier d'environ 4% : certains chromosomes sont bien meilleurs que d'autres. Comment optimiser la population initiale en ajoutant manuellement de « bons » chromosomes ?

3.2.2.3 Ajout des meilleures solutions dans la population initiale

Il est possible d'inclure dans la population initiale de « bons » chromosomes de façon à restreindre l'espace de recherche et accélérer la convergence.

Pour ce faire, il a été choisi de générer une population initiale à l'aide d'un algorithme glouton. Il s'agit d'un algorithme généralement utilisé dans la résolution du problème du sac à dos et qui a donc été jugé adapté à notre problème.

L'algorithme glouton est une procédure algorithmique qui construit une solution de manière incrémentale. A chaque étape, l'algorithme prend la décision la plus avantageuse en suivant des règles établies. Ce choix, localement optimal, ne conduit qu'à une solution sous-optimale. Elle est cependant utile lorsqu'on ne connaît pas d'algorithme de résolution exacte comme c'est le cas dans notre problème.

La procédure suivante décrit les étapes permettant de générer 100 chromosomes à l'aide de l'algorithme glouton. A chaque tour, l'algorithme résilie une police : le premier chromosome généré se compose d'un gène (une seule police résiliée) et le dernier chromosome se compose de 100 gènes (toutes les polices sont résiliées). Chaque chromosome est généré sur la base du chromosome précédent, auquel on ajoute la résiliation d'une police supplémentaire. La police résiliée est celle qui a la plus grande efficacité marginale sur l'ECR.

Soit B l'ensemble des gènes b_1, \dots, b_{100} .

Soit $ECR(0)$ l'ECR du portefeuille de référence.

On souhaite créer 100 chromosomes (R_1, \dots, R_{100}) .

Soit $R = \emptyset$

$i = 1$

Tant que $i < 100$

 Soit $B' = B \setminus R$

 Pour tout $b_j \in B'$

$$S_j = R \cup b_j$$

$$\text{Calculer l'efficacité marginale du scénario de résiliation } S_j : E(S_j) = \frac{ECR(0) - ECR(S_j)}{\% \text{ primes résiliées}(S_j)}$$

 On obtient $E = (E(S_1), \dots, E(S_j), \dots, E(S_{100-i+1}))$

 Soit k tel que $E(S_k) = \max(E)$

$$R_i = R \cup b_k$$

$$R = R \cup b_k$$

$$i = i + 1$$

Sortie : (R_1, \dots, R_{100})

Fin

Le nuage de point ci-dessous présente la population générée avec l’algorithme glouton.

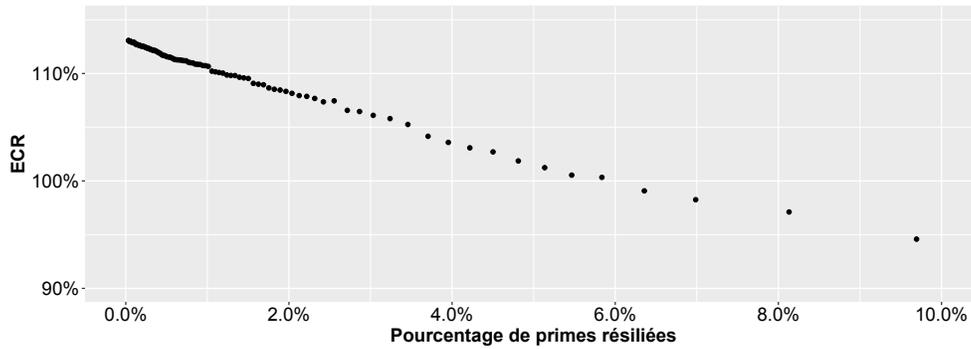


Figure 61 : population générée avec l’algorithme glouton

L’algorithme glouton a permis de générer 100 scénarios de résiliation. Chaque chromosome généré a une police résiliée de plus que le chromosome précédent et une police résiliée de moins que le chromosome suivant. Cela s’explique par le fait que l’algorithme glouton construit ses solutions par voisinage. Chaque scénario est construit à partir du précédent, en ajoutant la police dont la résiliation a la plus grande efficacité marginale sur l’ECR.

Les chromosomes sont concentrés entre 0% et 3% de primes résiliées : les polices ayant des petites primes sont les plus efficaces.

3.2.2.4 Méthode retenue

Dans notre étude, il a été décidé de générer la population initiale à partir des trois méthodes décrites précédemment. La population initiale de 400 chromosomes se compose de :

- 100 chromosomes générés avec l’algorithme glouton. Ces chromosomes permettent d’avoir de « bons » chromosomes dans la population initiale et d’accélérer la convergence.
- 200 chromosomes générés avec pénalisation et avec une contrainte de 0,02%. Ces chromosomes permettent de concentrer la population initiale entre 0% et 6% de primes résiliées.
- 100 chromosomes générés sans pénalisation et avec une contrainte de 0,04%. Ces chromosomes permettent d’explorer l’espace des primes entre 6% et 8%.

Le nuage de points ci-dessous donne un exemple de population initiale générée avec la méthode décrite ci-dessus.

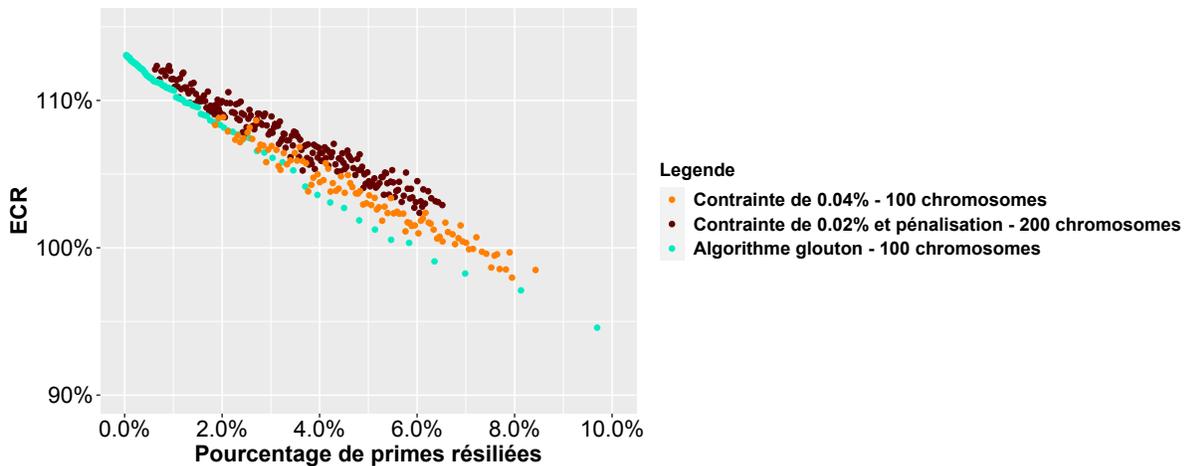


Figure 62 : exemple de population initiale

Les chromosomes générés avec l'algorithme glouton sont les plus performants, notamment pour les pourcentages de primes élevées. Quelques chromosomes surperforment l'algorithme glouton : plus la population est grande et plus la probabilité qu'un « bon » chromosome soit présent dans la population initiale augmente.

La population initiale est répartie de la manière suivante :

- 25% des chromosomes ont un pourcentage de primes résiliées inférieur à 1,5%
- 25% des chromosomes ont un pourcentage de primes résiliées entre 1,5% et 3,2%
- 25% des chromosomes ont un pourcentage de primes résiliées entre 3,2% et 5%
- 25% des chromosomes ont un pourcentage de primes résiliées entre 5% et 10%

Les chromosomes sont concentrés entre 0% et 5%, ce qui a été considéré comme adapté à notre problème.

3.2.3 Evaluation de la population

L'algorithme NSGA-II (*Non-dominated Sorting Genetic Algorithm II*) décrit par Deb et al. en 2002 ([11]) est l'un des algorithmes génétiques les plus populaires pour déterminer des frontières efficaces. Cet algorithme évalue la qualité d'un individu avec une procédure de « *non-dominated sorting* », basée sur la notion de domination.

Notion de domination et de front

Soit K est l'ensemble des objectifs à optimiser. L'individu x est dominé par y si :

$$\forall k \in \llbracket 1, K \rrbracket : O_k(x) \leq O_k(y) \text{ et s'il existe un } k \text{ tel que l'inégalité soit stricte.}$$

On note $x < y$.

Si un individu est non dominé, il est optimal au sens de Pareto et il appartient donc à la frontière efficace.

L'algorithme NSGA-II départage les individus en fonction de leur performance relative. Pour ce faire, les individus sont classifiés en « fronts » :

- Les individus du front F_1 ne sont dominés par aucun individu
- Les individus du front F_2 sont dominés par au moins un individu du front F_1
- Les individus du front F_i sont dominés par moins un individu du front F_{i-1}

De cette manière, chaque individu de la population dispose d'un « rang de domination ». On peut alors comparer les individus : un individu de front i est jugé de meilleure qualité qu'un individu de front $i + 1$.

Deb et al. proposent l'algorithme suivant pour calculer le front de chaque individu :

Soit P l'ensemble des individus.

Pour tout $p \in P$,

Soit S_p l'ensemble des solutions dominées par p . $S_p = \emptyset$

Soit n_p le compteur de solutions dominant p . $n_p = 0$

Pour tout $q \in P$

Si $p < q$, alors $S_p = S_p \cup \{q\}$

Si $q < p$, alors $n_p = n_p + 1$

Si $n_p = 0$

Alors $p_{rank} = 1$ et $\mathcal{F}_1 = \mathcal{F}_1 \cup \{p\}$

$i = 1$

Tant que $\mathcal{F}_i \neq \emptyset$

Soit Q l'ensemble des individus du front $i + 1$. $Q = \emptyset$

Pour tout $p \in \mathcal{F}_i$
 Pour tout $q \in S_p$
 $n_q = n_q - 1$
 Si $n_q = 0$, alors $q_{rank} = i + 1$ et $Q = Q \cup \{q\}$
 $i = i + 1$
 $\mathcal{F}_i = Q$

La figure ci-dessous donne un exemple de population à cinq fronts.

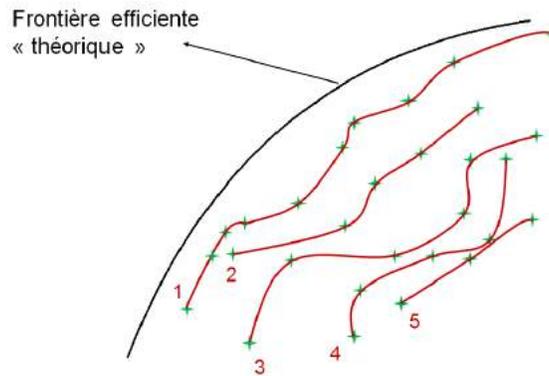


Figure 63 : illustration d'une population à cinq fronts

Source : Pierrat, 2017 [12]

On constate que les individus de front 1 sont ceux qui sont les plus proches de la frontière efficiente théorique. Ces individus sont jugés de meilleure qualité que les individus de front supérieur.

Notion de crowding distance

La notion de front permet de départager des individus de front différent. Mais au sein d'un même front, comment juger de la qualité d'un individu par rapport à un autre ?

La *crowding distance* mesure la taille du vide au voisinage d'un individu. L'objectif de cette mesure est une répartition uniforme des solutions le long de la frontière efficiente.

La crowding distance se calcule de la manière suivante :

Soit P l'ensemble des individus.

Pour chaque critère $k \in \llbracket 1, K \rrbracket$,

Trier les individus par ordre croissant de critère k . Soit i le i ème meilleur individu au regard du critère k .

Calculer f_{min} et f_{max} les valeurs minimales et maximales du critère k .

Attribuer une crowding distance infinie aux 2 individus ayant pour valeur f_{min} et f_{max}

Pour les autres individus $i \in P$, calculer la *crowding distance* : $i_{CD, k} = \frac{O_k(i+1) - O_k(i-1)}{f_{max} - f_{min}}$

La *crowding distance* permet de départager deux individus d'un même front : celui qui a la plus grande *crowding distance* sera préféré. Cette méthode permet de favoriser une bonne répartition des individus sur la frontière efficiente.

La figure ci-dessous présente une illustration de calcul de *crowding distance* avec deux critères. Les points noirs appartiennent au front 1 et les points blancs au front 2.

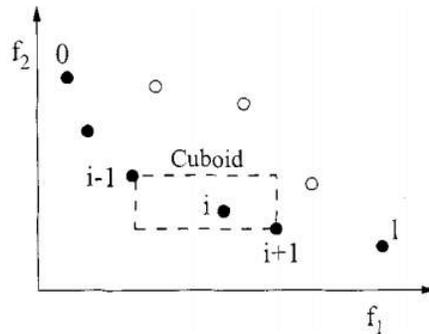


Figure 64 : illustration de la crowding distance avec deux critères
Source : Deb et al, 2002 [12]

Comme on le voit sur la figure, l'individu i est celui qui a la plus grande *crowding distance* du front 1. C'est donc l'individu du front 1 qui est jugé le plus performant.

La *crowding distance* permet donc de départager deux individus d'un même front. C'est l'individu qui a la plus grande *crowding distance* qui sera privilégié.

Classement des individus : procédure de non-dominated sorting

En utilisant la notion de front et de *crowding distance*, on peut ordonner tous les individus de la population. On note $p <_n q$ si p est préféré à q .

Soit p_{front} le front de p
 Soit p_{CD} la *crowding distance* de p
 Si $p_{front} < q_{front}$ ou si $((p_{front} = q_{front}) \text{ et } (p_{CD} = q_{CD}))$
 Alors $p <_n q$

De cette manière, il est possible d'ordonner tous les individus de la population.

3.2.4 Sélection

L'opérateur de sélection est une étape importante car elle permet de déterminer à partir de quels individus les nouveaux individus seront créés.

L'enjeu consiste à trouver le juste équilibre entre sélection trop forte et trop faible. Le premier cas peut entraîner un risque de perte de diversité génétique. Le second cas peut avoir pour conséquence une convergence trop lente.

Dans la littérature, on trouve trois principales formes de sélection : la sélection par roulette, sur le rang et par tournoi.

Sélection par roulette

La méthode par roulette consiste à sélectionner les chromosomes parents avec une probabilité proportionnelle à leur qualité : plus un chromosome est « bon » au titre d'un objectif et plus il a une forte probabilité d'être choisi.

Soit n le nombre de chromosomes, $O_k(i)$ la valeur du chromosome i pour l'objectif k et $O_k(0)$ la valeur de l'objectif k pour le portefeuille de référence.

Soit l le nombre d'objectifs. Pour chaque objectif k , on sélectionne $\frac{n}{l}$ chromosomes parents. Chaque chromosome i a une probabilité $p_{k, i}$ d'être sélectionné pour l'objectif k :

$$p_{k, i} = \frac{O_k(0) - O_k(i)}{\sum_{j=1}^n (O_k(0) - O_k(j))}$$

Une fois les chromosomes sélectionnés, les parents sont ensuite répartis aléatoirement par couple.

Cette technique permet aux « bons » chromosomes au titre d'un objectif de se reproduire avec les « bons » chromosomes au titre d'un autre objectif et de potentiellement générer des chromosomes enfants qui seront un compromis entre les deux objectifs.

Dans le cas où les deux « bons » chromosomes choisis correspondent au même objectif, le croisement permettra potentiellement de générer de meilleurs individus au regard de l'objectif, en combinant certaines caractéristiques des deux chromosomes parents.

Cette méthode fait face à trois critiques principales :

- Elle est coûteuse en temps de calcul, notamment dans le cas d'une grande population.
- Elle peut entraîner une convergence prématurée de l'algorithme dans le cas où un chromosome domine très largement la population : la pression sélective sera très forte et ce chromosome sera presque toujours sélectionné, entraînant le risque que la population n'évolue plus.
- Elle peut entraîner une stagnation dans le cas où tous les chromosomes sont de qualité comparable. En effet, les probabilités de sélection seront alors peu dispersées et la pression sélective sera faible, ce qui se rapproche d'une sélection aléatoire des parents.

Sélection sur le rang

La sélection par le rang consiste à ordonner les chromosomes grâce à la procédure de *non-dominated sorting*. Chaque chromosome i a ensuite une probabilité p_i d'être sélectionné proportionnelle à son rang r_i :

$$p_i = \frac{1}{r_i + 1}$$

Cette méthode permet de contrôler la pression sélective : tous les chromosomes ont une chance d'être sélectionnés. La sélection par le rang permet d'éviter les problèmes de stagnation ou de convergence prématurée.

Cependant, cette méthode a le désavantage d'entraîner une convergence plus lente car les meilleurs individus sont sélectionnés moins souvent qu'avec des méthodes comme celle de la roulette.

Sélection par tournoi

La sélection par tournoi consiste à choisir aléatoirement T chromosomes parmi la population. Ces chromosomes sont ensuite ordonnés selon la procédure de *non-dominated sorting* et le meilleur chromosome est retenu. L'opération est répétée jusqu'à la sélection de tous les couples de chromosomes parents.

Cette méthode a plusieurs avantages :

- Elle permet de contrôler la pression sélective en choisissant la taille de T . Plus la taille de T est grande et plus la pression sélective sera forte.
- Elle favorise la diversité génétique : les chromosomes moins bons ont une probabilité plus élevée d'être tirés par rapport aux sélections par roulette et sur le rang.
- Elle est peu coûteuse en temps de calcul.

Dans la littérature (*Lavinas, 2018 [13]*), la valeur de T la plus utilisée est 2. C'est également la valeur utilisée dans l'algorithme NSGA-II.

Opérateur de sélection retenu

La sélection par roulette n'est pas retenue en raison des risques de convergence prématurée et de stagnation.

La sélection par tournoi est préférée à la sélection sur le rang car elle est moins coûteuse en temps de calcul et permet un meilleur contrôle de la pression sélective.

La valeur de T choisie est 2 : cela permet de favoriser de la diversité génétique et c'est la valeur la plus couramment utilisée dans la littérature.

3.2.5 Croisement

A l'issue des opérations de sélection, on obtient un couple de chromosomes. A partir de ce couple, on va réaliser un « croisement » ou « enjambement » : on crée deux nouveaux chromosomes « enfants » à partir des deux chromosomes « parents ».

Le croisement est un mécanisme central de l'algorithme génétique : ce mécanisme part du principe que le croisement des meilleurs individus permet de générer des individus encore meilleurs. En effet, les enfants vont combiner des caractéristiques de chacun des parents et peuvent donner de meilleures solutions.

Il existe un grand nombre de méthodes pouvant être utilisées pour réaliser un enjambement. Dans la littérature, les méthodes les plus utilisées pour les chromosomes encodés en binaire sont le *single-point crossover* (croisement en un point), le *N-point crossover* (croisement en N points) et l'*uniform crossover* (croisement uniforme).

Croisement en un point

Pour réaliser un croisement en un point, on choisit aléatoirement un entier i entre 1 et le nombre de gènes k du chromosome. On a $i \in \llbracket 1, k - 1 \rrbracket$ et i est appelé le « point de croisement ».

Avec cette méthode, le chromosome de l'enfant 1 sera composé des i premiers gènes du parent 1 et des $k - i$ derniers gènes du parent 2.

Inversement, le chromosome de l'enfant 2 sera composé des i premiers gènes du parent 2 et des $k - i$ derniers gènes du parent 1.

La figure ci-dessous donne un exemple de mutation en un point pour des chromosomes de 13 gènes et ayant un point de croisement de 5. Les gènes des enfants colorés en rouge correspondent aux gènes issus du parent 1 et ceux colorés en bleu au parent 2.

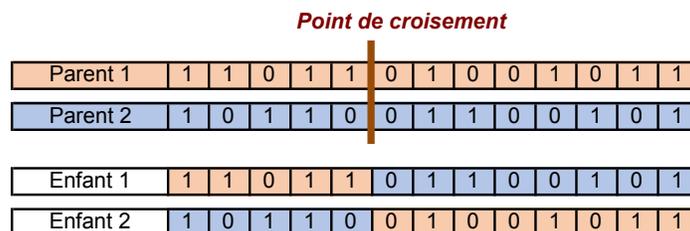


Figure 65 : exemple d'un croisement en un point

On remarque que l'enfant 1 se compose des 5 premiers gènes du parent 1 et des 6 derniers gènes du parent 2. Inversement, l'enfant 2 se compose des 5 premiers gènes du parent 2 et des 6 derniers gènes du parent 1.

Croisement en N points

Le croisement en N points ressemble au croisement en un point. La différence repose sur le fait que N points de croisements sont utilisés au lieu d'un seul.

Il faut d'abord définir N , avec $N \in \llbracket 2, k \rrbracket$. On choisit ensuite aléatoirement les N points de croisement tels que $N_1 \in \{1, k - N\}$, $N_2 \in \{N_1 + 1, k - N + 1\}$, ... , $N_i \in \{N_{i-1} + 1, k - N + i - 1\}$

Les gènes compris entre chaque point de croisement du parent 1 seront répartis à tour de rôle entre l'enfant 1 et l'enfant 2. Le mécanisme inverse sera effectué avec les gènes du parent 2.

La figure ci-dessous donne un exemple de croisement en 3 points pour des chromosomes de 13 gènes. Les points de croisement sont 3, 8 et 12.

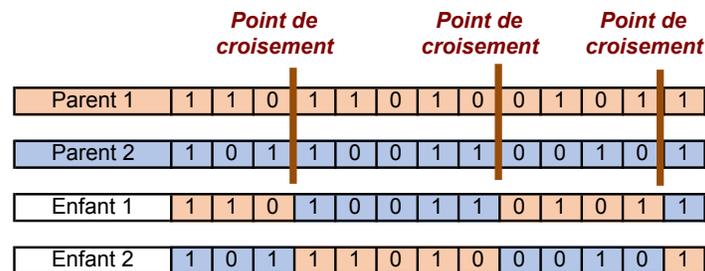


Figure 66 : exemple d'un croisement en N points

On constate que l'enfant 1 est composé des 3 premiers gènes du parent 1, des gènes 4 à 8 du parent 2, des gènes 9 à 12 du parent 1 et du dernier gène du parent 2.

A l'inverse, l'enfant 2 se compose des 3 premiers gènes du parent 2, des gènes 4 à 8 du parent 1, des gènes 9 à 12 du parent 2 et du dernier gène du parent 1.

Pour un montant de N relativement petit, par exemple 2, ce type de croisement a l'avantage d'éviter de trop changer la structure des chromosomes. Cette méthode a cependant le désavantage de limiter l'exploration de l'espace une fois que la population devient homogène.

Croisement uniforme

Le croisement uniforme consiste à sélectionner aléatoirement les gènes d'un parent et de prendre les gènes restant de l'autre parent.

La première étape de cette méthode consiste à réaliser N tirages uniformes U . On choisit ensuite une probabilité $p \in [0, 1]$, correspondant à la probabilité que chaque gène de l'enfant 1 soit un gène du parent 1 :

- Si $U_i \geq p$ alors le gène i de l'enfant 1 est le gène i du parent 1. Inversement, le gène i de l'enfant 2 est le gène i du parent 2.
- Si $U_i < p$ alors le gène i de l'enfant 1 est le gène i du parent 2. Inversement, le gène i de l'enfant 2 est le gène i du parent 1.

La figure ci-dessous donne un exemple de croisement uniforme dont la valeur de p est de 70%.

Parent 1	1	1	0	1	1	0	1	0	0	1	0	1	1
Parent 2	1	0	1	1	0	0	1	1	0	0	1	0	1
Tirage uniforme	0,43	0,57	0,86	0,78	0,59	0,15	0,19	0,42	0,70	0,12	0,10	0,24	0,22
Probabilité (p=0,7)	<0,7	<0,7	>=0,7	>=0,7	<0,7	<0,7	<0,7	<0,7	>=0,7	<0,7	<0,7	<0,7	<0,7
Enfant 1	1	0	0	1	0	0	1	1	0	0	1	0	1
Enfant 2	1	1	1	1	1	0	1	0	0	1	0	1	1

Figure 67 : exemple de croisement uniforme

Chaque gène de l'enfant 1 a une probabilité de 70% d'être un gène du parent 1 et de 30% d'être un gène du parent 2. Sur les 13 tirages uniformes, 3 sont supérieurs ou égaux à 70% : l'enfant 1 aura 3 gènes du parent 1. Les gènes tirés sont les gènes 3, 4 et 9. Les autres gènes de l'enfant 1 seront les gènes du parent 2.

Ce type de croisement a l'avantage de permettre une large exploration de l'espace mais présente le risque de fortement changer la structure des chromosomes, au risque de les rendre moins performants.

Méthodologie choisie

Parmi les algorithmes génétiques encodés en binaire, les moins performants sont ceux qui n'effectuent pas d'opération de croisement, suivis de ceux qui effectuent des croisements en un point (*Picek et Golub, 2010 [14]*). Les méthodes de croisement les plus performantes sont les croisements uniformes et en deux points : elles permettent d'obtenir de meilleurs résultats et une convergence plus rapide de l'algorithme.

Au regard de cette étude, la méthode de croisement uniforme a été choisie. Cette méthode a été préférée au croisement en 2 points car elle permet une plus large exploration de l'espace.

La valeur de p choisie est de 50%. C'est la valeur la plus couramment utilisée dans la littérature.

3.2.6 Mutations

Les opérations de mutation permettent de garantir une diversité génétique au sein des générations de solutions candidates. Cette étape consiste à choisir aléatoirement un gène dans un chromosome et à remplacer sa valeur en suivant un principe de mutation. A l'image des phénomènes biologiques, le chromosome peut bénéficier d'un avantage ou d'un handicap, la mutation étant un phénomène impartial.

Durant le processus évolutif, les mutations se produisent à un niveau qui peut être prédéfini. Habituellement, un faible niveau de mutation est choisi : il ne faut pas transformer l'algorithme en une simple recherche aléatoire. Un niveau de mutation modéré permet d'accélérer l'évolution tout en évitant que les chromosomes ne deviennent trop similaires les uns aux autres.

Inversion de bit

Les mutations réalisées sur les chromosomes codés en binaire sont généralement des inversions de bit.

Pour chaque gène, on réalise un tirage uniforme $r_i \in [0, 1]$. On définit également une probabilité $p \in [0, 1]$ qu'un gène subisse une mutation.

- Si $r_i \leq p$, alors le gène i est inversé : si le gène vaut 1, il est muté en 0 et inversement.
- Si $r_i > p$, alors le gène i est inchangé.

La figure ci-dessous présente un exemple de mutation par inversion de bit avec une probabilité de 1%.

Chromosome i avant mutation	1	1	0	1	1	0	1	1	0	0	1	0	1
Tirage uniforme	0,474	0,360	0,888	0,005	0,872	0,055	0,602	0,125	0,878	0,271	0,916	0,818	0,072
Probabilité (p=0,01)	non muté	non muté	non muté	muté	non muté	non muté	non muté	non muté	non muté	non muté	non muté	non muté	non muté
Chromosome i après mutation	1	1	0	0	1	0	1	1	0	0	1	0	1

Figure 68 : exemple de mutation par inversion de bit

La valeur r_4 est de 0,005, ce qui est inférieur à la valeur de p . Le gène 4 est donc muté : sa valeur 1 est changée en 0. Les autres gènes sont inchangés car les tirages uniformes ont donné des valeurs supérieures à 1%.

Méthodologie choisie

Dans la littérature, il n'y a pas de consensus sur le taux de mutation à utiliser. Il est recommandé d'effectuer des études de sensibilité pour déterminer le taux de mutation adapté à chaque problème.

Deb et al. suggèrent d'utiliser un taux de mutation de $\frac{1}{n}$ avec n le nombre de gènes d'un chromosome. Dans notre étude, on choisit donc d'effectuer des mutations par inversion de bits avec un pourcentage de 1% car les chromosomes sont composés de 100 gènes. En moyenne, chaque chromosome aura donc une mutation.

Une étude de sensibilité sera réalisée par la suite pour déterminer si ce pourcentage est adapté à notre problème.

3.2.7 Génération suivante

Une fois les croisements et les mutations effectués, on obtient une population G_i'' composée de :

- La génération actuelle G_i
- Les enfants E_i issus des opérations de croisement
- Les chromosomes mutant M_i obtenus en effectuant des mutations sur G_i et E_i

La population G_i'' est ensuite évaluée. Le rang et la *crowding distance* de chaque chromosome sont calculés. Les chromosomes sont ordonnés du plus performant au moins performant en fonction de la procédure de *non-dominated sorting*.

Une opération d'élitisme est alors effectuée : les N chromosomes les plus performants sont sélectionnés pour faire partie de la génération suivante G_{i+1} .

Cette procédure permet de conserver les meilleurs individus de la génération G_i'' dans la génération G_{i+1} .

3.2.8 Arrêt de l'algorithme

A chaque fois qu'une nouvelle génération est obtenue, l'algorithme a le choix entre lancer une nouvelle itération ou s'arrêter.

Dans notre problème, l'algorithme s'arrête lorsque tous les chromosomes sont sur le même front. En effet, on estime que de nouvelles itérations ne permettront pas d'améliorer de manière significative la qualité des résultats.

Dans les faits, un très grand nombre de générations sont nécessaires pour que tous les chromosomes soient sur le même front, ce qui implique des temps de calcul très longs. Il est également possible que les chromosomes ne soient jamais tous sur le même front. Pour pallier cette situation, l'algorithme prévoit un nombre maximum d'itérations à partir duquel l'algorithme s'arrête. Ce nombre est un paramètre entré par l'utilisateur et doit correspondre à une estimation prudente du nombre de générations nécessaires pour obtenir des résultats stables. Dans notre étude, on choisit un critère d'arrêt de 50 générations.

3.2.9 Synthèse

Cette partie a pour objectif de réaliser une synthèse des différentes données utilisées par l'algorithme et d'en récapituler les différentes étapes.

Données utilisées par l'algorithme

Avant le lancement de l'algorithme, il est nécessaire que l'utilisateur indique les informations suivantes :

- Taille d'un chromosome C : nombre de gènes présents dans un chromosome.
- Taille de la population N : nombre de chromosomes présents dans une génération.
- Nombre de générations G : l'algorithme s'arrête au maximum au bout de G générations. Il peut s'arrêter avant si tous les chromosomes se situent sur le front 1.
- Population initiale P₀ : population initiale de N chromosomes composés chacun de C gènes. Si l'utilisateur ne fournit pas cette information, la population initiale est générée aléatoirement.

Pour rappel, les informations suivantes sont nécessaires pour pouvoir évaluer l'impact d'un scénario de résilience sur le ratio combiné économique :

- Polices résiliables : identifiant et prime des polices étant dans le périmètre de l'optimisation
- Modélisation du risque : YELT par année, par événement et par police.
Pour optimiser les temps de calcul, la modélisation des polices hors du périmètre de la politique de résilience ont été agrégées par année et par événement.
- Traité de réassurance : priorité du traité de réassurance, nombre de reconstitutions, taux de reconstitution, fonction de calcul de la prime.
- Capital : taux de diversification du STEC CAT dans le STEC total, ratio de solvabilité, *spread*.
- Chargements : taux d'imposition, frais internes et externes.

Synthèse des différentes étapes de l'algorithme génétique

Les différentes étapes de l'algorithme génétique utilisé dans notre étude sont les suivantes :

1. Si P_0 n'est pas fournie par l'utilisateur, génération d'une population initiale P_0 de N chromosomes composés chacun de C gènes.
2. Evaluation de P_0 :
 - 2.1 Calcul du front de chaque chromosome
 - 2.2 Calcul de la *crowding distance* de chaque chromosome
 - 2.3 Ordonnancement des chromosomes par *non-dominated sorting*.
3. $i = 1$. Soit G_i la i ème génération.
4. Tant que ($G_i < G$) ou que (toutes les solutions ne sont pas sur le front 1)
 - 3.1 Sélection des parents par tournoi
 - 3.2 Reproduction des parents par croisement uniforme. On appelle E_i les chromosomes « enfant » obtenus et G'_i une population telle que $G'_i = G_i + E_i$
 - 3.3. Mutation des chromosomes de la population G'_i . On appelle M_i les chromosomes mutants et G''_i une population telle que $G''_i = G'_i + M_i$
 - 3.4 Evaluation de la population G''_i
 - 3.4.1 Calcul du front de chaque chromosome
 - 3.4.2 Calcul de la *crowding distance* de chaque chromosome
 - 3.4.3 Ordonnancement des chromosomes de G''_i par *non-dominated sorting*.
 - 3.5 Création de la génération G_{i+1} : sélection des N meilleurs chromosomes parmi G''_i .
 $i = i + 1$
5. Arrêt de l'algorithme. La sortie est une frontière efficiente.

3.3 Résultats et limites

L'objectif de ce chapitre est de proposer une optimisation de la politique de *pruning* du portefeuille. On souhaite minimiser le ratio combiné économique du portefeuille tout en minimisant le pourcentage de primes résiliées.

Dans la partie précédente, l'algorithme génétique et ses différents paramètres ont été présentés. Cette partie a pour objectif de présenter les résultats obtenus et d'en faire une analyse critique.

L'analyse sera guidée par les questions suivantes : les résultats sont-ils satisfaisants ? Les méthodes utilisées sont-elles assez flexibles pour prendre en compte de nouvelles contraintes métier ? Sont-elles robustes ?

Une partie sera dédiée à chacune de ces questions.

Les résultats présentés dans cette partie ont été calculés avec R. Aucune parallélisation n'a été réalisée et la machine suivante a été utilisée :

- Système d'exploitation MacOS Sierra
- Processeur Intel Core i5 @1,3 GHz
- Mémoire (RAM) 4Go

Pour rappel, les paramètres utilisés pour l'algorithme génétique sont les suivants :

- Nombre de gènes : 100
- Taille de la population : 400
- Nombre de générations : 50
- Sélection : par tournoi de taille 2
- Croisement : uniforme, avec une probabilité de 50%
- Mutation : inversion de bit, avec une probabilité de 1%
- Population initiale : générée selon la méthode alternative

3.3.1 Optimisation de l'ECR et du pourcentage de primes résiliées

Cette première partie présente une analyse critique des résultats de l'optimisation. L'objectif est dual : on souhaite comprendre les résultats obtenus et évaluer leur qualité.

Dans un **premier temps**, les résultats seront présentés et leur performance sera évaluée.

Dans un **deuxième temps**, l'impact sur les primes nettes de réassurance obligatoire sera analysé.

Dans un **troisième temps**, le comportement de chaque composant de l'ECR sera étudié.

Enfin, dans un **quatrième temps**, la complexité algorithmique et les temps de calcul seront évalués.

3.3.1.1 Performance des résultats

L'objectif de cette partie est d'analyser les résultats de l'optimisation. Pour ce faire, on étudiera l'évolution de l'ECR en fonction du pourcentage de primes résiliées. On analysera ensuite l'efficacité des solutions obtenues.

Pour évaluer la performance des résultats obtenus, les résultats seront comparés avec ceux du front 1 de la population initiale. En effet, on souhaite comparer les résultats obtenus avec les meilleurs chromosomes de la population initiale pour évaluer les progrès réalisés par l'algorithme génétique entre la première et la dernière génération.

Présentation des résultats

Pour rappel, les résultats se présentent sous la forme d'une frontière efficiente : les solutions situées sur la frontière minimisent l'ECR pour un montant de primes résiliées donné. Les solutions sont des scénarios de résiliation construits à partir de 100 polices non rentables : chaque scénario est une combinaison parmi les 2^{100} combinaisons possibles. La frontière présente des solutions pour un montant de primes résiliées allant de 0 (aucune police résiliée) à 3,5 millions d'euros (les cent polices sont résiliées). Cela correspond à un intervalle entre 0 et 10% des primes.

Le graphe ci-dessous présente les résultats de la minimisation de l'ECR et du pourcentage de primes résiliées. Chaque point représente un scénario de résiliation. La frontière efficiente obtenue avec l'algorithme génétique est présentée en rouge et le front 1 de la population initiale en bleu.

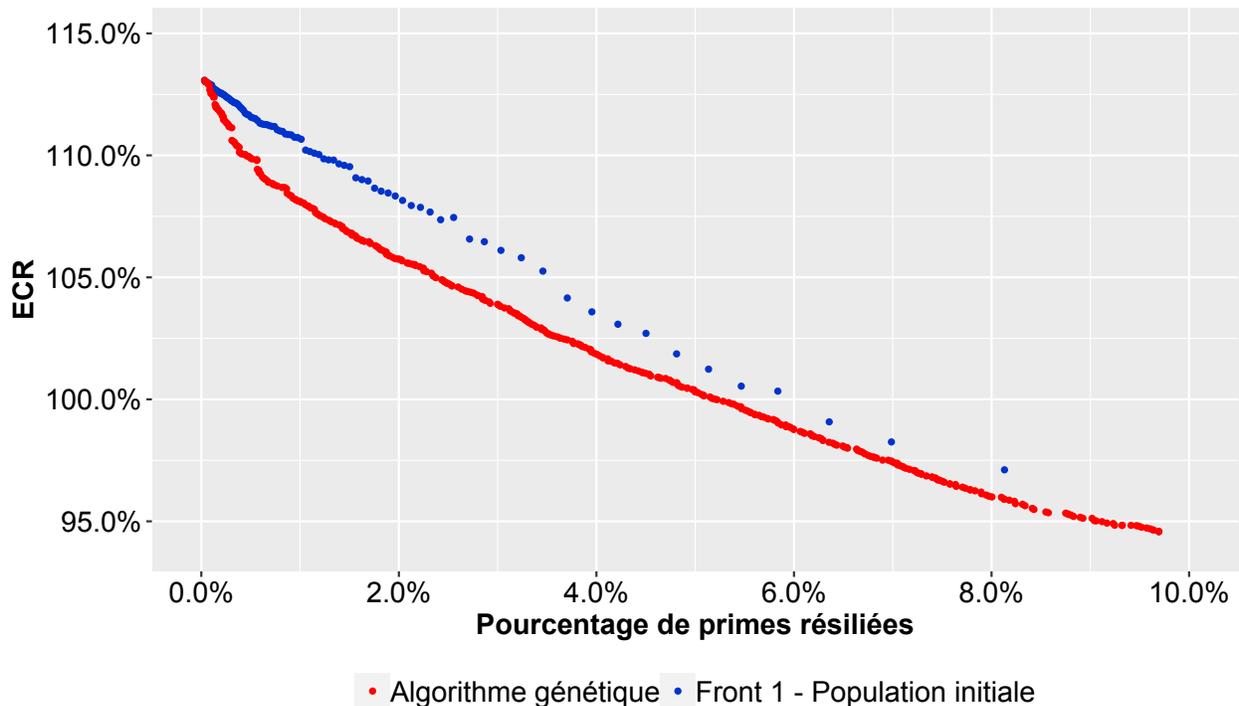


Figure 69 : évolution de l'ECR en fonction du pourcentage de primes résiliées

Remarque 1 : tout d'abord, on remarque qu'avec un ECR de 113%, le portefeuille de référence (0% de primes résiliées) n'est pas rentable : après rémunération des actionnaires, chaque euro de prime encaissé entraîne une détérioration des fonds propres de 13 centimes. Pour revenir à l'équilibre, c'est à dire à un ECR de 100%, il serait nécessaire de résilier au minimum 5% des primes.

Avec un objectif de résiliation de maximum 4% des primes, l'entité peut espérer réduire son ECR d'environ 10 points : pour un euro de prime encaissé, la détérioration de fonds propres serait alors réduite à 3 centimes.

Remarque 2 : l'algorithme génétique surperforme la population initiale d'environ 2,5 points d'ECR entre 1% et 4% de primes résiliées. Cela représente un écart d'environ 900 000 euros sur la création de valeur. L'algorithme génétique permet donc d'améliorer significativement la population initiale.

L'écart entre les deux populations se réduit à partir de 5% de primes résiliées : il est d'environ 1 point d'ECR, soit un impact d'environ 350 000 euros sur la création de valeur.

Efficacité marginale des résultats

La frontière efficiente présente des résultats compris entre 0% et 10% de primes brutes résiliées. Dès lors, comment arbitrer entre deux scénarios n'ayant pas le même montant de primes résiliées ? Quel indicateur définir pour comparer la performance des deux scénarios ?

Pour répondre à cette problématique, on définit un indicateur de performance qu'on appelle l' « efficacité marginale ». Cet indicateur a été défini en partie 2.3. Pour rappel, il s'agit du rapport entre l'ECR marginal et le pourcentage de primes résiliées.

$$\text{Efficacité marginale ECR} = \frac{ECR_{\text{portefeuille de référence}} - ECR_{\text{nouveau portefeuille}}}{\text{Pourcentage de primes résiliées}}$$

L'efficacité marginale permet de quantifier l'impact d'un point de pourcentage de primes résiliées sur l'ECR : si l'efficacité est de 2, cela signifie qu'un point de pourcentage de primes résiliées permet de faire baisser l'ECR de deux points. En d'autres termes, un euro de primes résiliées permet de faire baisser la charge de l'assureur de 2 euros.

L'efficacité marginale permet donc de comparer deux solutions n'ayant pas le même montant de primes résiliées et d'évaluer la performance des résultats obtenus. On souhaite que l'efficacité marginale soit la plus grande possible : plus elle est grande et plus l'impact d'un point de pourcentage de primes résiliées sur l'ECR est important.

Le graphe ci-dessous présente l'efficacité marginale des résultats obtenus avec l'algorithme génétique et la population initiale.

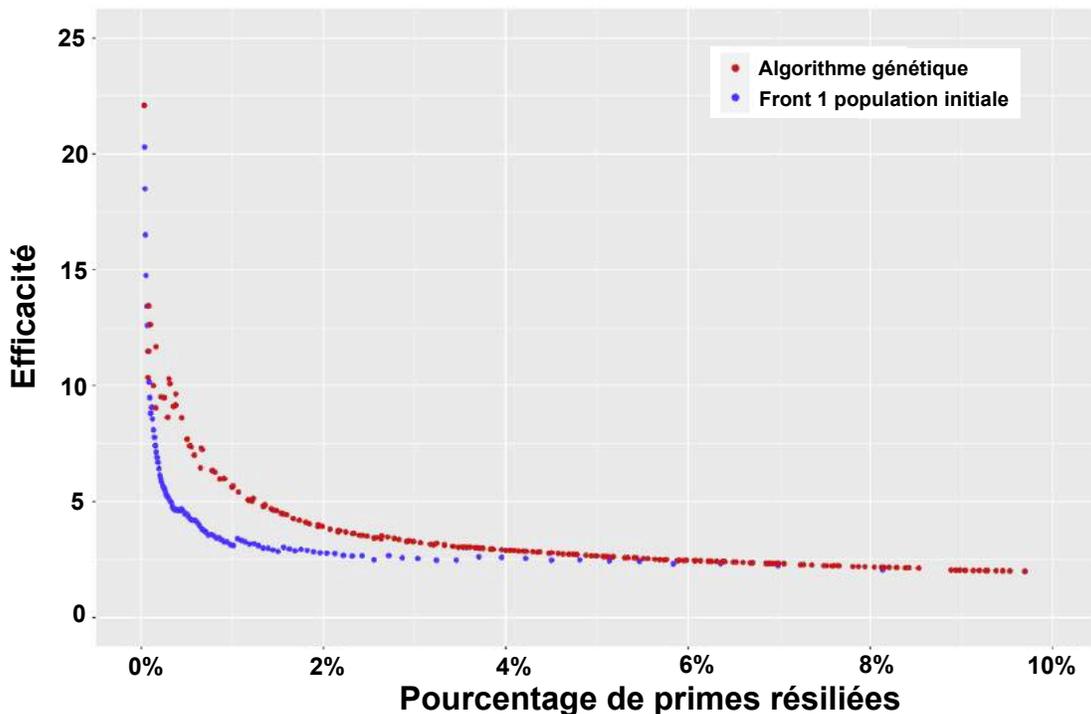


Figure 70 : évolution de l'efficacité marginale en fonction du pourcentage de primes résiliées

Remarque 1 : l'efficacité est décroissante : plus le montant de primes résiliées est important et plus l'impact relatif sur l'ECR diminue. En effet, les polices les moins rentables relativement à leur prime sont sélectionnées en premier par l'algorithme génétique.

L'efficacité est positive : tous les scénarios de résiliation présentés permettent de faire baisser l'ECR.

Le graphe présenté ici ne montre que des scénarios basés sur les 100 polices les moins rentables du portefeuille. Si on réalisait un graphe similaire avec toutes les polices du portefeuille, on remarquerait que l'efficacité est décroissante puis devient négative : à partir d'un certain montant de primes résiliées, le montant de primes encaissées ne permet plus de faire face aux frais fixes et résilier une police supplémentaire entraîne une augmentation de l'ECR.

Remarque 2 : on remarque que l'efficacité se stabilise autour de 2,5 à partir de 2% de primes résiliées pour la population initiale et à partir de 5% pour l'algorithme génétique. Les résultats sont donc de bonne qualité : si l'on résilie 4% de prime, l'impact marginal sur l'ECR sera de 4 % x 2,6, soit une réduction de l'ECR d'environ 10,4 points.

Remarque 3 : l'algorithme génétique a une meilleure efficacité que le front 1 de la population initiale jusqu'à 6% de primes résiliées. Au-delà de ce montant, l'efficacité des deux populations est proche. En effet, à partir de 6%, les scénarios de résiliation ont des profils similaires : ils contiennent tous les deux les polices avec les plus grosses primes. Il est nécessaire d'inclure ces polices dans les scénarios de résiliation pour réussir à dépasser les 6% de primes résiliées.

3.3.1.2 Impact sur les primes nettes

Dans le cadre de cette étude, on cherche à minimiser l'ECR et le pourcentage de primes brutes résiliées. Cependant, il est aussi intéressant de s'intéresser aux primes nettes. Par primes nettes, on entend les primes brutes réduites du montant des primes du traité de réassurance :

$$Primes_{nettes} = Primes_{brutes} - Primes_{traité}$$

Pour rappel, les primes brutes sont nettes des primes de réassurance facultative. Les primes nettes sont donc les primes encaissées par l'assureur après avoir financé tous les coûts de réassurance (facultative et obligatoire).

Le graphe ci-dessous montre l'évolution en montant des primes nettes par rapport aux primes brutes résiliées. Les axes secondaires permettent d'analyser les évolutions en pourcentage.

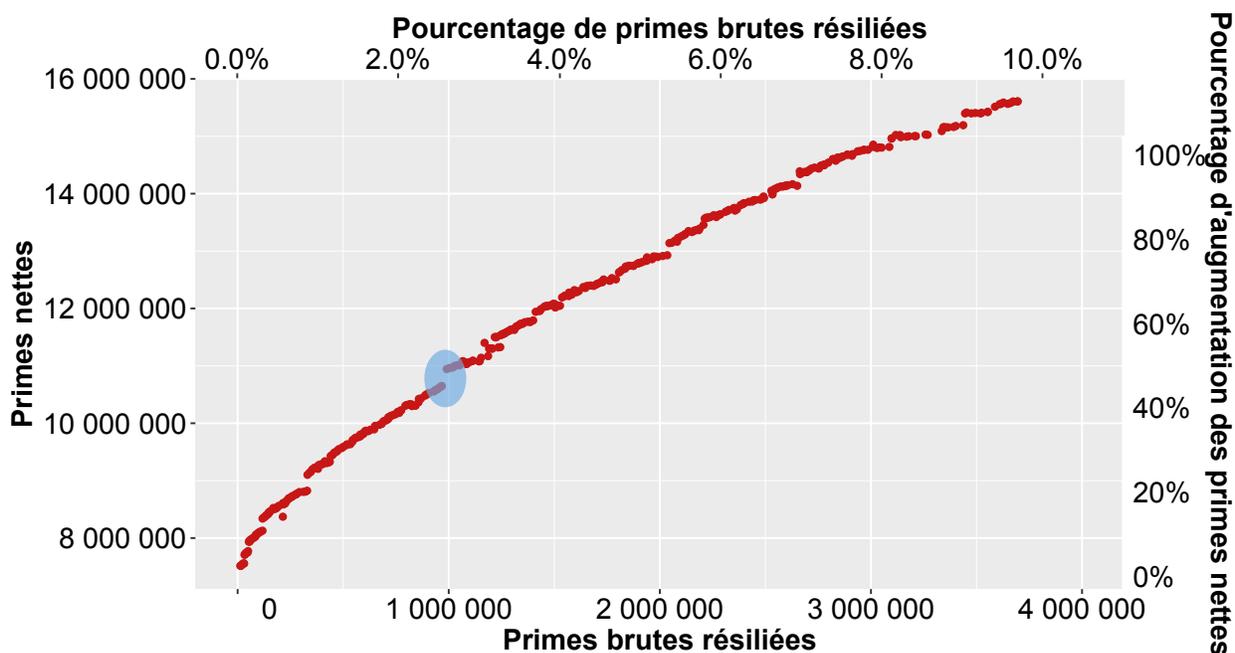


Figure 71 : évolution des primes nettes en fonction des primes brutes résiliées

Remarque 1 : Les scénarios de résiliation proposés ont un impact matériel sur les primes nettes. En effet, une réduction de 2% des primes brutes entraîne une hausse des primes nettes d'environ 40%, soit une efficacité marginale de 20. Une réduction de prime de 8% entraîne une hausse des primes nettes de 105%, soit une efficacité marginale de 13,1.

Dans tous les scénarios de résiliation, on constate que les primes nettes augmentent par rapport au portefeuille de référence : les primes de réassurance baissent davantage que les primes brutes résiliées.

En effet, dans tous les scénarios de résiliation, on constate une baisse de prime de réassurance par rapport au portefeuille de référence. Cela s'explique par deux effets :

- Pour chaque scénario de résiliation, la limite du traité est ajustée de telle manière à ce que le traité couvre l'événement bicentenaire. Les nouveaux portefeuilles ont un événement bicentenaire inférieur ou égal à celui du portefeuille de référence : leur traité a donc une limite inférieure ou égale à celle du portefeuille de référence.
- Les récupérations des nouveaux portefeuilles sont inférieures ou égales à celles du portefeuille de référence car des polices ont été résiliées et la limite du traité est inférieure ou égale à celle du traité du portefeuille de référence.

Remarque 2 : Les courbes présentent des sauts comme on le voit par exemple dans le cercle bleu. Au niveau d'un million de primes résiliées, on trouve par exemple un saut d'environ trois cent mille euros de primes nettes. Le traité des scénarios de résiliation situé juste avant et après le saut ont respectivement une limite de 596 millions et de 577 millions, soit une baisse de 19 millions.

Cet écart important s'explique par le fait que la limite est calculée à partir de l'événement bicentenaire. La résiliation d'une police peut avoir plusieurs impacts sur l'événement bicentenaire :

- L'événement bicentenaire peut être inchangé.
- L'événement bicentenaire peut être inchangé mais être réduit si les polices résiliées sont touchées par cet événement.
- Un changement d'événement bicentenaire peut avoir lieu.

Ici, c'est le troisième cas qui a lieu : le scénario de résiliation a changé l'ordre d'importance des événements et c'est un nouvel événement qui est l'événement bicentenaire.

On remarque que les récupérations à charge du traité ont baissé de 113 000 euros. 93% de cette baisse s'explique par la baisse de la limite du traité. En effet, sur les 10 000 années simulées, le traité est épuisé 55 fois : une baisse de la limite de 19 millions a donc un impact de $19^6 \times 55 / 10\ 000 = 105\ 000$ euros sur les récupérations moyennes.

Remarque 3 : si on cherche à optimiser les primes nettes, l'entité aurait intérêt à résilier un maximum de primes brutes. En effet, en résiliant 10% de ses primes brutes, l'entité maximise ses primes nettes (près de 16 millions de primes nettes, soit plus du double par rapport au portefeuille de référence) et minimise son ECR (ECR de 95% soit 18% de moins que le portefeuille de référence).

Cependant, les primes nettes sont un indicateur qui est généralement piloté par le *Risk Management* et la réassurance, alors que les primes brutes vont être pilotées par la souscription.

De ce fait, les différents services peuvent avoir des objectifs différents : la souscription peut souhaiter maximiser les primes brutes alors que la réassurance et le *Risk Management* peuvent avoir comme objectif de réduire les coûts de réassurance. De ce fait, il est nécessaire de piloter la politique de résiliation à l'aide de ces deux indicateurs.

Remarque 4 : il est important d’avoir un regard critique sur les primes nettes obtenues ici. En effet, les primes du traité ont été obtenues sur la base d’une estimation, basée sur les primes de réassurance du précédent renouvellement. Or, rien n’assure que le traité pourra être renouvelé aux mêmes conditions que celles de l’année précédente. En effet, sur le marché de la réassurance, les prix peuvent varier sensiblement d’une année sur l’autre comme par exemple à la suite d’un sinistre ou une hausse des prix sur le marché de la rétrocession.

3.3.1.3 Comportement des différents composants de l’ECR

Cette partie a pour objectif d’approfondir la compréhension des résultats en analysant l’évolution des différents composants de l’ECR : ratio combiné brut, coûts de réassurance, impôts et coût du capital.

Dans un premier temps, l’ECR sera décomposé fonction de ses différents composants. **Dans un second temps**, l’efficacité de chaque composant sera étudié.

Décomposition de l’ECR

Le graphe ci-dessous montre l’évolution de chacun des composants de l’ECR en fonction du pourcentage de primes résiliées. Le ratio combiné brut montre l’impact du résultat brut, le ratio combiné net l’impact des coûts de réassurance, le ratio combiné net après impôts l’impact des impôts et l’ECR l’impact du coût du capital.

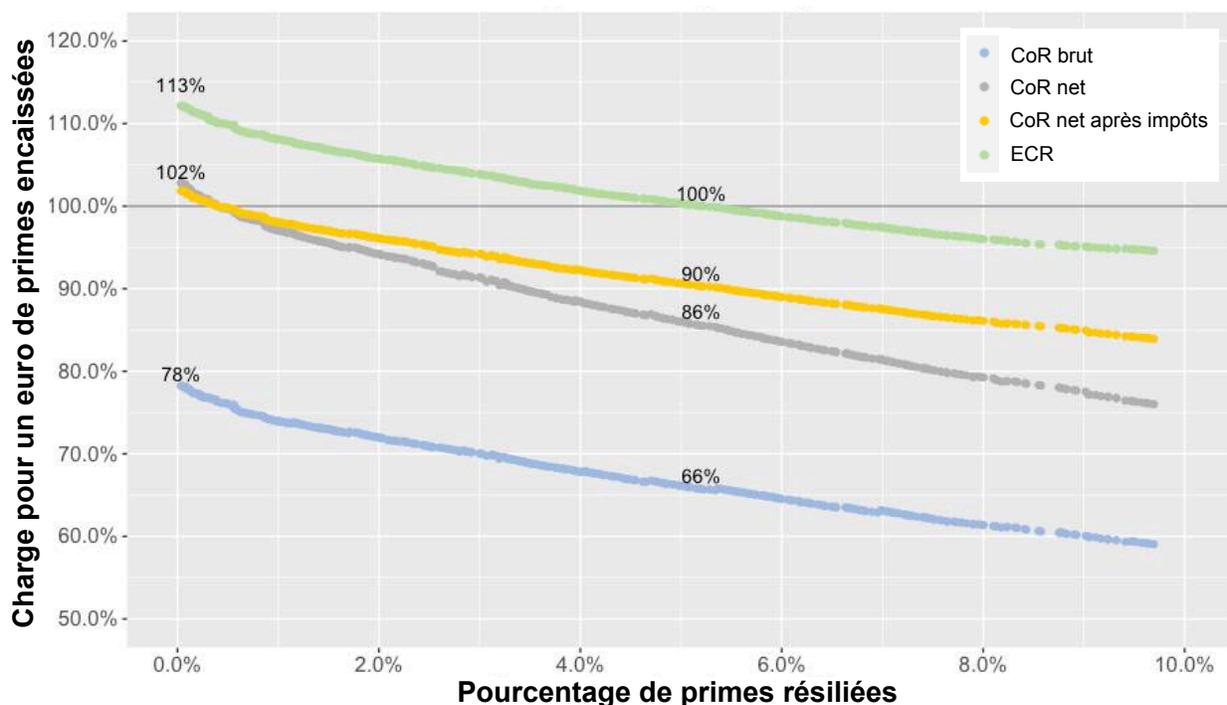


Figure 72 : décomposition de l’ECR

Remarque 1 : l’ECR du portefeuille de référence est de 113% : pour un euro de prime encaissée, l’assureur a une charge de 1,13 euro. Sur cet 1,13 euro, les pertes brutes représentent 78 centimes, le résultat cédé 24 centimes, et le coût du capital 11 centimes.

Le tableau ci-dessous présente le poids constaté de chaque composant de l’ECR parmi les différents scénarios de résiliation obtenus. Pour rendre les montants comparables, le poids de chaque composant a été normalisé par l’ECR :

$$Poids\ normalisé_{Composant\ A}^{Scenario\ i} = \frac{Poids_{composant\ A}}{ECR^{Scenario\ i}}$$

	CoR brut	Résultat cédé	CoC	Impôts
Minimum	62,4%	17,9%	9,0%	-0,8%
Maximum	69,8%	21,9%	11,3%	8,4%
Moyenne	66,6%	20,1%	9,6%	3,6%

Figure 73 : poids de chaque composant de l'ECR

Le ratio combiné brut représente entre 62% et 70% de l'ECR, le résultat cédé entre 18% et 22% et le coût du capital entre 9 et 11%. Le ratio combiné brut est donc le plus gros contributeur de l'ECR, suivi du résultat cédé et du coût du capital.

Remarque 2 : il y a un crédit d'impôts pour le portefeuille de référence : le résultat net est négatif. Des impôts commencent à être payés à partir du moment où le ratio combiné net est inférieur à 100% : à partir d'environ 0,5% de primes résiliées. Les impôts augmentent ensuite à mesure que le résultat net s'améliore jusqu'à atteindre 8,4% pour 9,7% de primes résiliées.

Efficacité des différents composants de l'ECR

Dans la section précédente, le poids de chaque composant de l'ECR a été analysé. Qu'en est-il de leur efficacité ? Quels composants sont les moteurs de la baisse de l'ECR ?

Dans cette section, on analysera l'efficacité marginale de chaque composant. Pour rappel, l'efficacité marginale se calcule comme suit :

$$efficacité\ marginale = \frac{Indicateur_{portefeuille\ de\ référence} - Indicateur_{nouveau\ portefeuille}}{Pourcentage\ de\ primes\ résiliées}$$

Comme vu précédemment, tous les composants de l'ECR n'ont pas le même poids : le ratio combiné brut est largement supérieur aux autres composants. Par conséquent, il est intéressant de quantifier l'impact de chaque composant relativement à son montant. On appelle cet indicateur l'efficacité. Il se calcule comme suit :

$$efficacité = \frac{Indicateur_{portefeuille\ de\ référence} / Indicateur_{nouveau\ portefeuille} - 1}{Pourcentage\ de\ primes\ résiliées}$$

Le graphe ci-dessous permet de comparer la performance de chaque composant de l'ECR. Il présente l'efficacité marginale et l'efficacité du ratio combiné brut, des coûts de réassurance et du capital. L'efficacité des impôts n'a pas été représentée.

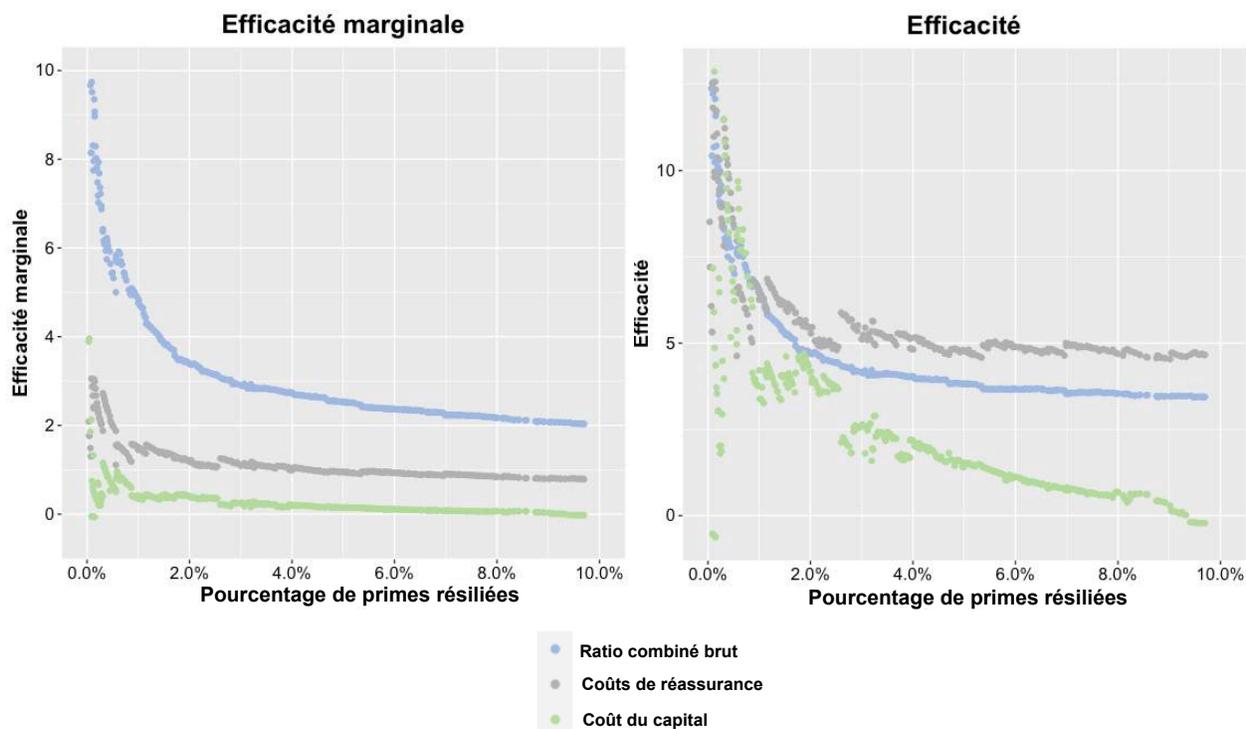


Figure 74 : efficacité marginale et efficacité des composants de l'ECR

Remarque 1 : le composant qui a de loin la plus grande efficacité marginale est le ratio combiné brut. C'est le principal moteur de la baisse de l'ECR. Viennent ensuite les coûts de réassurance puis le coût du capital. Ce résultat s'explique par le fait que le ratio combiné brut est le composant de l'ECR qui a le plus de poids, suivi par les coûts de réassurance.

Si on analyse la variation des composants relativement à leur montant, c'est les coûts de réassurance qui sont les plus efficaces, suivi du ratio combiné brut. Ainsi, l'efficacité du résultat cédé est d'environ 5 pour un pourcentage de primes résiliées entre 2% et 10%. Cela signifie qu'une baisse des primes de $\alpha\%$ entraîne une réduction de $5\alpha\%$ du résultat cédé.

Remarque 2 : le coût du capital a une efficacité marginale proche de 0, voire parfois négative. Certains scénarios de résiliation entraînent donc une hausse du coût du capital. Cela peut s'expliquer par deux phénomènes :

- La limite du traité de réassurance est fixée au niveau de l'événement bicentenaire. Dans certains scénarios, la baisse de la limite du traité peut entraîner une hausse de l'AEP nette : c'est par exemple le cas si l'année bicentenaire est inchangée mais que l'événement bicentenaire et a fortiori la limite du traité ont baissé.
- Un scénario de résiliation peut avoir un plus gros impact sur la moyenne nette que sur l'AEP nette. Le STEC étant une déviation par rapport à la moyenne, un tel cas entraînerait une hausse du coût du capital.

Remarque 3 : on note la présence de nombreux sauts dans l'efficacité et l'efficacité marginale du résultat cédé et du coût du capital. Cela s'explique par le fait que ces deux métriques sont non additives : calculs basés sur des quantiles et application des conditions de réassurance non proportionnelles. De ce fait, ces métriques sont beaucoup plus instables que le résultat brut, qui est une moyenne.

Conclusion

En conclusion, c'est le ratio combiné brut qui est le moteur de l'évolution de l'ECR : c'est le composant qui a le plus de poids dans l'ECR. Cependant, c'est le résultat cédé qui a le plus d'impact relativement à son montant.

3.3.1.4 Complexité algorithmique

La complexité algorithmique est l'étude de la quantité de ressources nécessaire à l'exécution d'un algorithme. Cette partie porte sur l'étude comparative de la complexité de notre problème. La complexité de l'algorithme génétique, par force brute et glouton sera étudiée.

Algorithme génétique

L'algorithme génétique tourne en 48 minutes pour une population de 400 chromosomes et 50 générations. En effet, l'algorithme évalue 400 scénarios de résiliation sur 50 générations, chaque scénario étant évalué en 0,143 seconde.

La complexité se calcule donc comme suit :

$$\text{complexité (en secondes)} = \text{Nombre}_{\text{chromosomes}} * \text{Nombre}_{\text{générations}} * 0,143$$

Etude par force brute

En comparaison, l'étude par force brute de tous les scénarios de résiliation possibles demanderait $2^{100} * 0,143$, soit $5,7 \times 10^{21}$ années de calcul.

Soit nb le nombre de polices dans le périmètre de l'optimisation. La complexité est exponentielle et se calcule comme suit :

$$\text{complexité (en secondes)} = 2^{nb} * 0,143$$

Algorithme glouton

L'algorithme glouton a été utilisé pour générer les meilleurs chromosomes de la population initiale de l'algorithme génétique. Il requiert 13 minutes pour générer une population de solutions. En effet, l'algorithme évalue $\sum_{i=1}^{100} i$ scénarios de résiliation différents, soit $\frac{100 * (100+1)}{2} = 5\,050$ scénarios et chaque scénario est évalué en 0,143 seconde.

Soit nb le nombre de polices, la complexité se calcule comme suit :

$$\text{complexité (en secondes)} = \frac{nb * (nb + 1)}{2} * 0,143$$

Conclusion

Pour 100 polices, une population de 400 chromosomes et 50 générations, l'algorithme glouton est 3,7 fois plus rapide que l'algorithme génétique.

Cependant, la complexité des algorithmes glouton et génétique est négligeable en comparaison de celle de l'étude par force brute. Ces algorithmes sont donc des solutions efficaces pour répondre à notre problème tout en limitant les temps de calcul.

3.3.1.5 Conclusion

Les solutions de l'algorithme génétique ont des caractéristiques adaptées à notre problème : il y a 400 scénarios de résiliation possibles uniformément répartis entre 0% et 10% de primes résiliées. Cela offre de la souplesse aux équipes de souscription dans le choix des polices à résilier.

L'algorithme génétique permet d'améliorer sensiblement les résultats par rapport à la population initiale. Les résultats obtenus ont une efficacité au moins égale à 2,5 : chaque euro de prime résiliée a un impact d'au moins 2,5 euros sur la charge totale de l'assureur. Ces performances s'expliquent principalement par une forte sensibilité du résultat brut et du résultat cédé aux résiliations.

Cependant, il serait nécessaire de résilier au moins 5% des primes brutes pour obtenir un ECR de 100% et que le portefeuille soit à l'équilibre. C'est un montant conséquent à l'échelle d'un portefeuille mais qui entraînerait une hausse de 13 millions des primes nettes.

Une fois ces résultats obtenus, il est nécessaire que les équipes de souscription, *Risk Management* et réassurance étudient la faisabilité des scénarios de résiliation proposés. A mesure que l'étude avance, l'intervalle de primes à considérer sera réduit et de nouvelles contraintes pourront apparaître. De ce fait, il est nécessaire de pouvoir adapter les algorithmes pour répondre efficacement aux nouvelles contraintes métier.

3.3.2 Adaptation de l'algorithme à de nouvelles contraintes métier

Une fois les premiers résultats communiqués aux équipes de souscription, on suppose que l'entité souhaite affiner sa politique de résiliation avec une nouvelle contrainte puis un objectif supplémentaire.

Dans un **premier temps**, l'entité souhaite ajouter une contrainte sur les primes : on souhaite résilier entre 2% et 4% des primes brutes.

Dans un **second temps**, l'entité souhaite minimiser le nombre de polices résiliées pour limiter les temps de gestion des résiliations.

3.3.2.1 Ajout d'une contrainte sur le pourcentage de primes résiliées

Initialement, il n'y avait pas de contrainte sur les primes. Les scénarios de résiliation allaient entre 0% de prime résiliées (aucune police résiliée) à 10% (100 polices sont résiliées).

Les équipes de souscription commençant à affiner le plan de résiliation, il est demandé d'ajouter une contrainte sur les primes. L'objectif est de résilier entre 2% et 4% des primes brutes.

Pénalisation des chromosomes ne respectant pas la contrainte

Pour prendre en compte la contrainte de prime, on ajoute une pénalisation dans la fonction d'évaluation : on pénalise l'ECR des scénarios de résiliation qui ne respectent pas la contrainte de prime.

Pour ce faire, on ajoute la procédure suivante à la fonction d'évaluation.

Soit C_{min} le pourcentage minimum de primes résiliées et C_{max} le pourcentage maximum.

On calcule π_i le pourcentage de primes résiliées du chromosome i .

On calcule ECR_i , l'ECR du chromosome i .

Si $\pi_i < C_{min}$ ou $\pi_i > C_{max}$

On calcule la pénalité $P_i = (\max(0, C_{min} - \pi_i) + \max(0, \pi_i - C_{max}) + 1)^2$

$ECR_i = ECR_i + P_i$

On renvoie (P_i, ECR_i)

Fin

Cette méthode permet de prendre en compte les contraintes de prime sans modifier les paramètres de l'algorithme génétique.

Résultats obtenus

Le graphe ci-dessous présente les résultats obtenus avec et sans contrainte sur le pourcentage de primes résiliées.

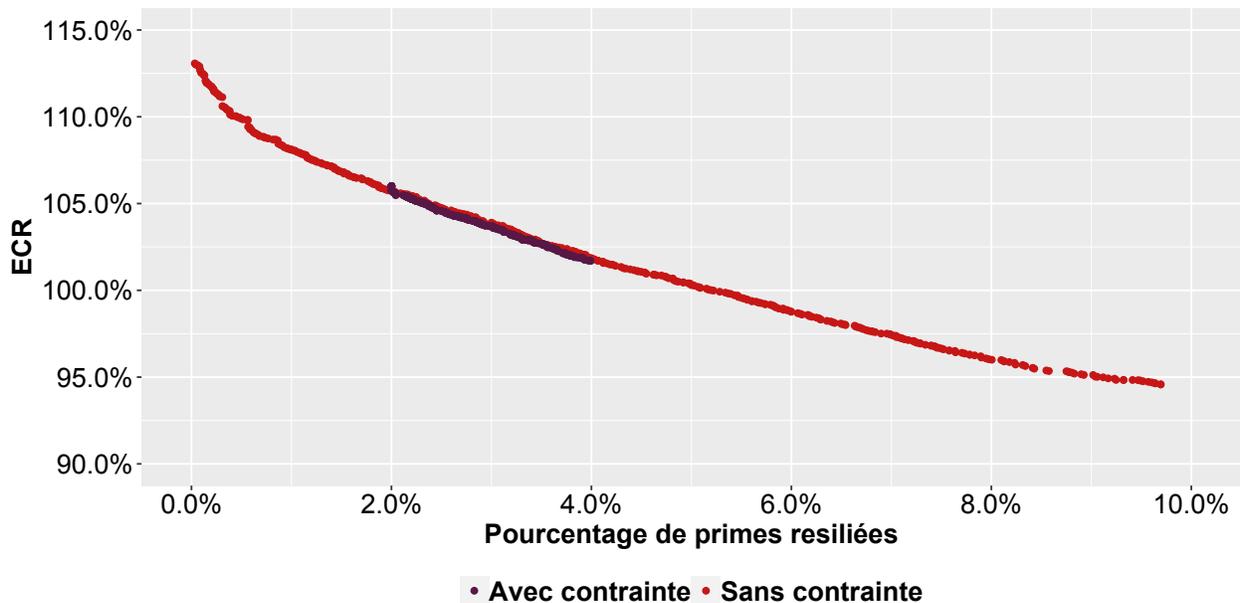


Figure 75 : résultats obtenus avec et sans contrainte sur le pourcentage de primes résiliées

Remarque 1 : pour les pourcentages de primes entre 2% et 4%, les résultats obtenus avec contrainte sont légèrement meilleurs que ceux obtenus sans contrainte. La différence n'est cependant pas matérielle et les résultats sont très proches. L'ajout d'une contrainte n'a donc pas permis d'améliorer sensiblement l'ECR.

Remarque 2 : l'option avec contrainte n'a des résultats qu'entre 2% et 4% de primes. En effet, les scénarios qui ne respectent pas la contrainte ont été pénalisés et leur ECR est trop élevé pour être visible sur le graphe.

Avec cette option, 347 chromosomes sur 400 se situent entre 2% et 4% de primes résiliées, contre 125 pour l'option sans contrainte. C'est le principal avantage de l'option avec contrainte : l'entité dispose d'un plus grand nombre de scénarios de résiliation respectant la contrainte fixée.

Conclusion

En conclusion, l'algorithme génétique permet facilement d'ajouter une contrainte sur le pourcentage de primes résiliées.

On constate que l'ajout d'une contrainte sur les primes permet de concentrer les chromosomes entre 2% et 4% de primes résiliées sans impacter négativement les résultats sur l'intervalle de primes étudié. Cela offre davantage de souplesse aux équipes de souscription qui peuvent faire leur choix parmi un plus grand nombre de scénarios de résiliation respectant la contrainte.

3.3.2.2 Ajout d'un troisième objectif : minimisation du nombre de polices résiliées

Initialement, l'optimisation était bi-objectif : elle portait sur la minimisation de l'ECR et du pourcentage de primes résiliées.

Les équipes de souscription souhaitent ajouter un troisième objectif : minimiser le nombre de polices résiliées. L'objectif est de résilier le nombre minimum de polices possibles pour limiter les temps de gestion liés aux résiliations.

Les contraintes sur les primes ont été relâchées et on cherche à optimiser le système suivant :

$$\begin{cases} \min ECR \\ \min \text{pourcentage de primes résiliées} \\ \min \text{nombre de polices résiliées} \end{cases}$$

L'algorithme génétique NSGA-II est multi-objectif et il s'adapte facilement à l'ajout d'un objectif supplémentaire : il suffit d'ajouter le troisième objectif dans la fonction d'évaluation.

Les graphes ci-dessous présentent les résultats obtenus. Pour une meilleure visualisation des résultats, ceux-ci seront présentés en trois graphes à deux dimensions.

ECR et pourcentage de primes résiliées

Le nuage de points ci-dessous présente l'évolution de l'ECR en fonction du pourcentage de primes résiliées.

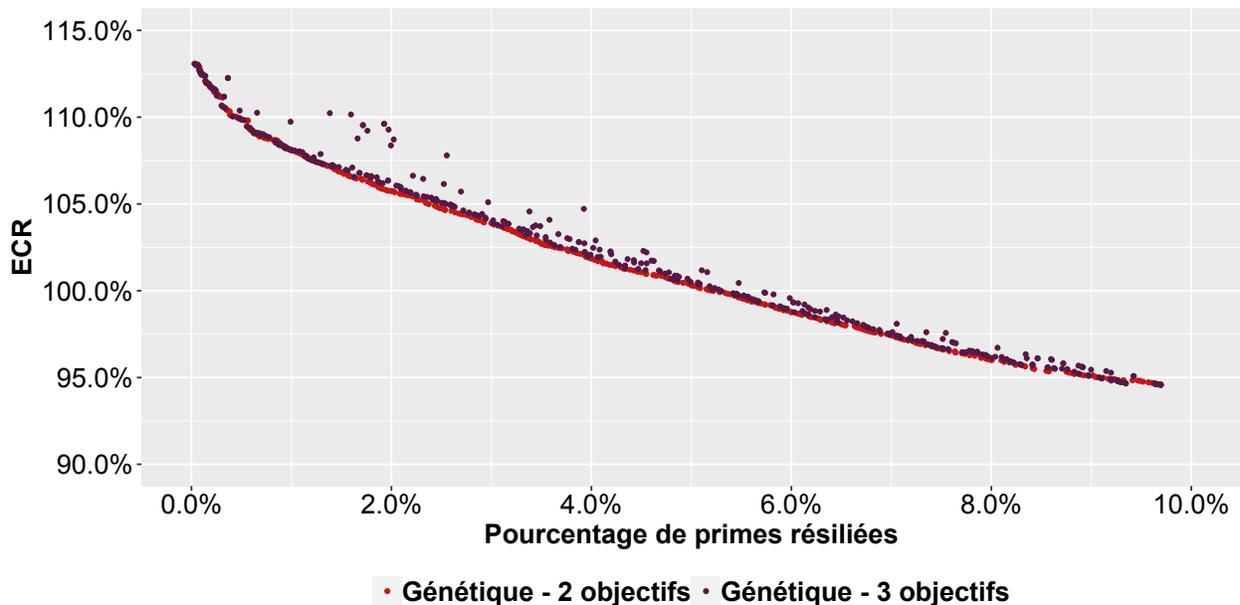


Figure 76 : évolution de l'ECR en fonction du pourcentage de primes résiliées

L'algorithme génétique avec deux objectifs surperforme l'algorithme génétique avec trois objectifs. En effet, ajouter un troisième objectif entraîne une dégradation des deux premiers critères.

Les résultats obtenus pour l'optimisation basée sur trois objectifs sont cependant de bonne qualité : la majorité d'entre eux sont très proches de la frontière efficiente de l'optimisation bi-objectif.

Nombre de polices résiliées et pourcentage de primes résiliées

Le nuage de points ci-dessous présente l'évolution du pourcentage de primes résiliées en fonction du nombre de polices résiliées.

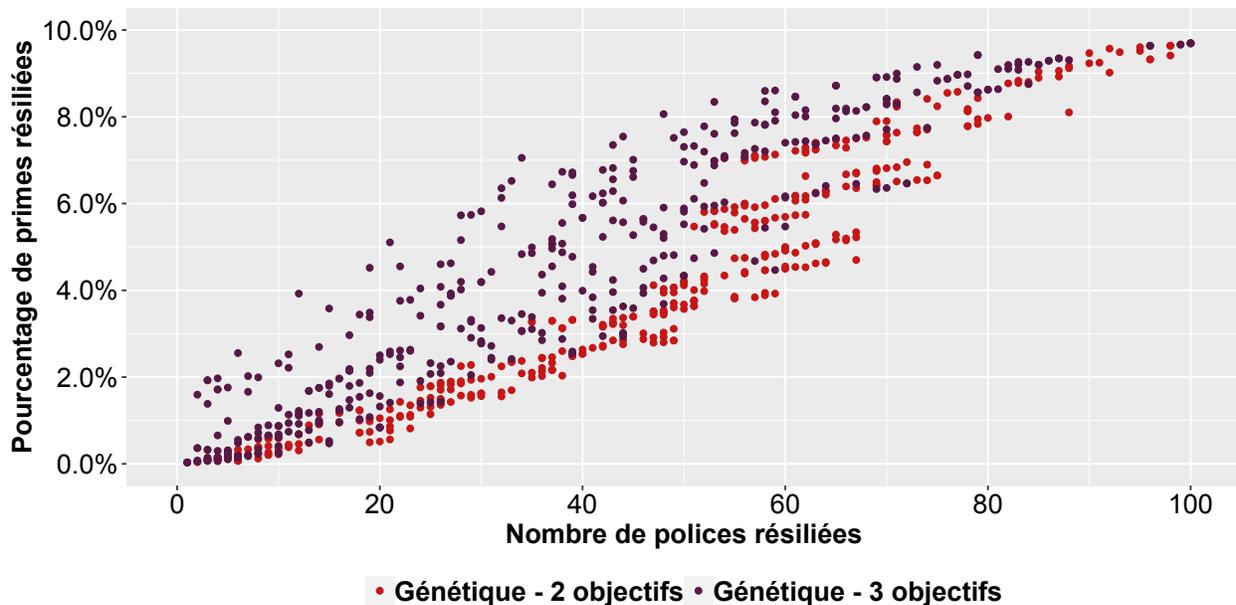


Figure 77 : évolution du pourcentage de primes résiliées en fonction du nombre de polices

Remarque 1 : la taille des polices est très hétérogène : pour un même nombre de polices résiliées, le pourcentage de primes résiliées peut varier de plus de 5 points. Cet écart se réduit à partir de 80 polices résiliées : la majorité des polices sont résiliées et il y a donc moins d'écart entre les différents scénarios de résiliation possibles.

Remarque 2 : les scénarios de résiliation générés avec trois objectifs sont composés de plus grosses polices que ceux générés avec deux objectifs. En effet, on constate que pour un même nombre de polices résiliées, l'optimisation tri-objectif résilie des pourcentages de primes beaucoup plus important que l'optimisation bi-objectif. L'optimisation tri-objectif minimise le nombre de polices résiliées : pour un même pourcentage de primes résiliées, on souhaite avoir le moins de polices possible.

Remarque 3 : pour un nombre de polices résiliées entre 0 et 18, on constate que certains pourcentages de primes résiliées sont très proches entre les deux optimisations. Cela s'explique par le fait que certaines petites polices sont très efficaces : elles sont choisies par l'optimisation tri-objectif car elles ont un bon impact sur l'ECR malgré la hausse du nombre de polices qu'elles entraînent. L'algorithme tri-objectif fait des compromis entre les différents critères.

ECR et nombre de polices résiliées

Le nuage de points ci-dessous présente l'évolution de l'ECR en fonction du nombre de polices résiliées.

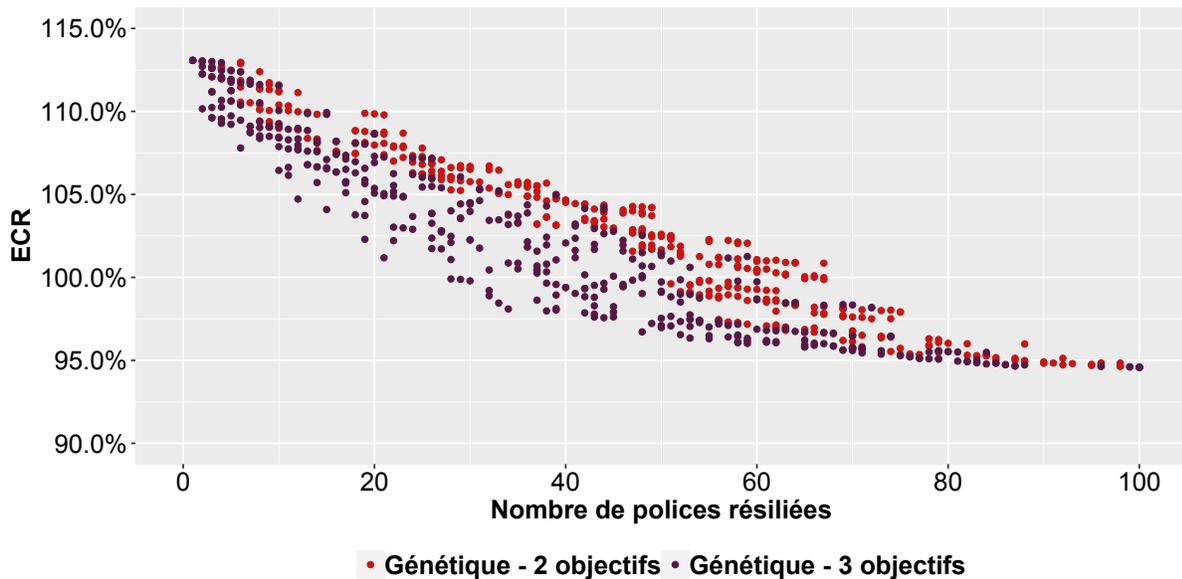


Figure 78 : évolution de l'ECR en fonction du nombre de polices résiliées

Remarque 1 : pour un même nombre de polices résiliées, l'ECR peut varier de près de 10 points. Cet écart se réduit à partir de 80 polices résiliées. Comme vu précédemment, cela s'explique par le fait que les polices ont des montants de primes très hétérogènes jusqu'à 80 polices résiliées.

Remarque 2 : l'optimisation tri-objectif permet d'obtenir des montants d'ECR plus petits, à nombre égal de polices résiliées. La minimisation du nombre de polices résiliées a donc bien été prise en compte par l'algorithme génétique.

On constate cependant que certains chromosomes de l'optimisation tri-objectif ont un ECR plus grand que les chromosomes bi-objectif, pour un même nombre de polices résiliées. Ces cas correspondent à des scénarios pour lesquels un compromis entre les trois objectifs a été fait par l'algorithme. Ces scénarios contiennent des polices avec de petites primes mais étant très efficaces sur l'ECR.

Conclusion

L'algorithme génétique permet d'ajouter facilement un nouvel objectif. Les résultats obtenus permettent d'obtenir des scénarios de résiliation qui minimisent le nombre de polices tout en conservant un niveau d'ECR comparable.

3.3.2.3 Conclusion

En conclusion, l'algorithme génétique est flexible et s'adapte facilement à de nouvelles contraintes métier.

Ainsi, une pénalisation peut être ajoutée à la fonction d'évaluation pour prendre en compte une contrainte sur le pourcentage de primes résiliées. Cette pénalisation ne permet pas d'améliorer matériellement les résultats mais a l'avantage de concentrer les chromosomes dans l'intervalle respectant la contrainte. De cette manière, les équipes de souscription peuvent faire leur choix parmi davantage de scénarios de résiliation.

Il est également possible d'ajouter un objectif supplémentaire, comme par exemple la minimisation du nombre de polices résiliées. Les résultats obtenus permettent d'obtenir des scénarios de résiliation qui minimisent le nombre de polices tout en conservant un niveau d'ECR comparable.

3.3.3 Etudes de sensibilité

L'objectif de cette partie est de présenter les résultats d'études de sensibilité aux différents paramètres de l'algorithme génétique. On souhaite évaluer la stabilité des résultats présentés dans la partie précédente.

Pour rappel, les paramètres suivants ont été utilisés :

- **Nombre de gènes** : 100
- **Taille de la population** : 400
- **Nombre de générations** : 50
- **Population initiale** : générée selon la méthode alternative
- **Reproduction** :
 - Sélection des parents : par tournoi
 - Croisement : uniforme avec un taux de 50%
- **Mutation** :
 - Type de mutation : inversion de bit
 - Probabilité de mutation : 1%

Dans un premier temps, la sensibilité au nombre de gènes choisis sera analysée. **Dans un deuxième temps**, on étudiera la sensibilité des résultats à la taille de la population et au nombre de générations. **Dans un troisième temps**, on s'intéressera à la sensibilité des résultats à la génération de la population initiale.

La sensibilité des résultats aux paramètres de reproduction et de mutation ne seront pas développés ici. Ils sont présentés en Annexe 4.

3.3.3.1 Nombre de gènes

Dans notre étude, nous avons utilisé des chromosomes composés de 100 gènes. Chaque gène correspond à une police non rentable et a été estimé résiliable par les équipes de souscription. Chaque chromosome correspond donc à un scénario de résiliation composé de 0 à 100 polices correspondant à un pourcentage entre 0% et 10% de primes résiliées.

Cette partie a pour objectif de comprendre si les résultats auraient été plus performants si on avait utilisé un nombre de gènes différent. Pour ce faire, nous avons comparé les résultats obtenus avec des chromosomes de 50, 100 et 250 gènes.

Analyse des résultats obtenus avec des chromosomes composés de 50, 100 et 250 gènes

Les nuages de points ci-dessous présentent les résultats obtenus pour des chromosomes composés de 50 (option 1), 100 (option 2) et 250 gènes (option 3).

L'option 2 correspond aux résultats présentés dans la partie précédente. Les gènes des options 1 et 3 ont été choisis en fonction de leur efficacité. Chaque option a été initialisée aléatoirement et 50 générations de 400 chromosomes ont été simulées.

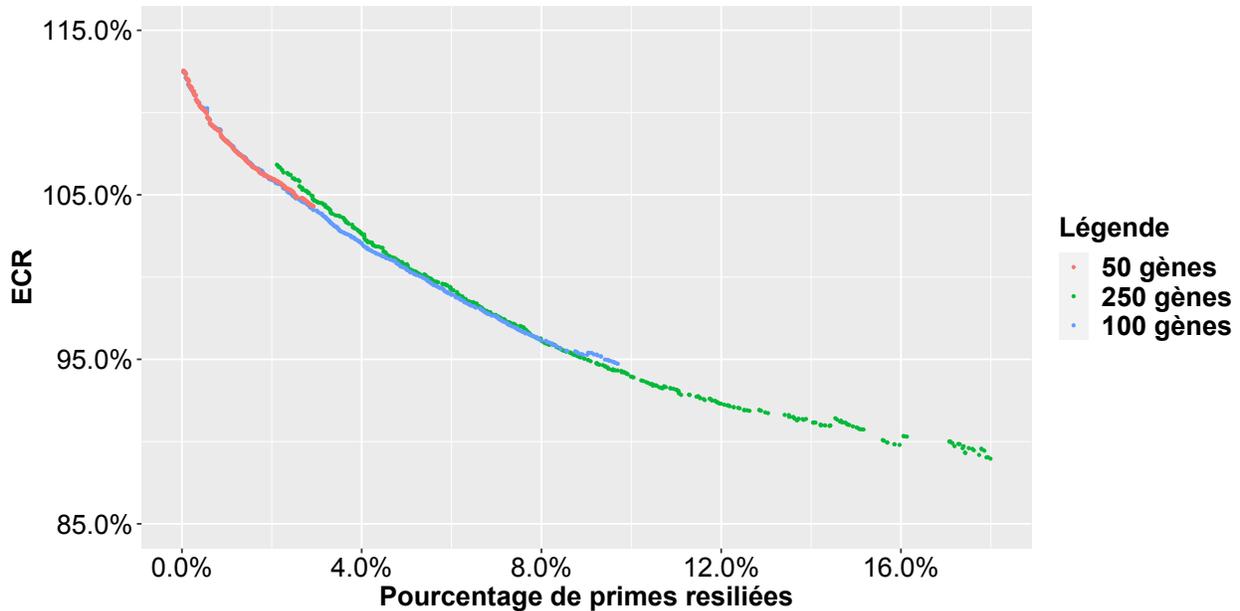


Figure 79 : sensibilité des résultats au nombre de gènes présents dans un chromosome

On remarque que le nombre de gènes influence l'espace des primes résiliées exploré par l'algorithme :

- Avec 50 gènes, les scénarios de résiliation obtenus sont compris entre 0% et 3% de primes résiliées. L'espace exploré est trop restreint aux vues des contraintes de l'entité (résilier jusqu'à 4% des primes).
- Avec 100 gènes, les scénarios de résiliation obtenus sont compris entre 0% et 10% de primes résiliées. L'espace exploré est satisfaisant : il inclut les contraintes de l'entité et donne de la flexibilité dans l'hypothèse où l'entité souhaiterait résilier un montant supérieur à 4% des primes.
- Avec 250 gènes, les scénarios de résiliation obtenus sont compris entre 2% et 18% de primes résiliées. L'espace exploré est vaste mais n'est pas concentré autour des contraintes de l'entité. En effet, aucun scénario de résiliation n'est inférieur à 2% de primes résiliées. Les scénarios supérieurs à 10% ne sont pas utiles car trop éloignés des contraintes de l'entité.

Les résultats obtenus avec 50 et 100 gènes sont très proches. Les résultats obtenus avec 250 gènes sont moins performants que ceux obtenus avec 100 gènes entre 2% et 8% de primes résiliées. A partir de 8%, la tendance s'inverse : entre 8% et 9%, les meilleurs résultats sont obtenus avec 250 gènes.

Ces constats s'expliquent principalement par deux phénomènes :

- Dans les chromosomes de 100 et 250 gènes, les 50 gènes de l'option 1 sont présents. De même, les chromosomes de 250 gènes contiennent les 100 gènes de l'option 2. En théorie, pour un même montant de primes résiliées, les chromosomes de 250 gènes peuvent donc donner des résultats de qualité similaire à ceux des chromosomes de 50 et 100 gènes. Dans les faits, ce n'est pas le cas. Cela s'explique par le fait que la population de 250 chromosomes, du fait de sa taille, met plus de temps à converger. Dès lors, il aurait été nécessaire de générer davantage de générations pour obtenir des résultats de qualité équivalente.
- Au-delà de 8% de primes résiliées, les chromosomes de 250 gènes donnent de meilleurs résultats que les chromosomes de 100 gènes. Cela s'explique par le fait qu'avec 250 gènes, il existe un très grand nombre de combinaisons de polices permettant de résilier 8% de primes. Ce n'est pas le cas pour les chromosomes de 100 gènes : pour atteindre 8% de primes résiliées, il est nécessaire de résilier presque intégralement les 100 polices.

Conclusion

Des chromosomes de 100 gènes sont adaptés à notre problème. Avec des chromosomes de 50 et 250 gènes, l'exploration de l'espace n'est pas concentrée entre 0 et 4% de primes.

3.3.3.2 Taille de la population et nombre de générations

Dans notre étude, la population utilisée se compose de 400 chromosomes. Le critère d'arrêt de l'algorithme est de 50 générations.

Cette section a pour but d'effectuer des études de sensibilité sur la taille de la population et sur le nombre de générations. Ces deux paramètres sont étudiés ensemble car leur montant a une influence similaire sur les temps de calcul : une population de 400 chromosomes sur 50 générations a un temps de calcul équivalent à celui d'une population de 50 chromosomes sur 400 générations.

L'objectif est donc de déterminer s'il est plus avantageux de favoriser le nombre de chromosomes ou le nombre de générations.

Dans un premier temps, la sensibilité des résultats à la taille de la population sera étudiée. **Dans un second temps**, l'impact du nombre de générations sera analysé.

Taille de la population

La taille de la population est le nombre de chromosomes présents dans chaque génération. Pour rappel, la taille de la population est constante. La taille de la population utilisée dans notre étude est de 400 chromosomes.

Les nuages de points ci-dessous présentent en bleu les résultats obtenus pour des populations de 50, 100 et 200 chromosomes. Les points rouges correspondent aux résultats obtenus dans notre étude, soit une population de 400 chromosomes. Pour que les résultats soient comparables, nous avons adapté le nombre de générations pour que les temps de calcul de chaque option soient identiques :

- 50 chromosomes : 400 générations ont été simulées contre 50 pour la population de 400 chromosomes.
- 100 chromosomes : 200 générations ont été simulées contre 50 pour la population de 400 chromosomes.
- 200 chromosomes : 100 générations ont été simulées contre 50 pour la population de 400 chromosomes.

Chaque option a été initialisée aléatoirement.

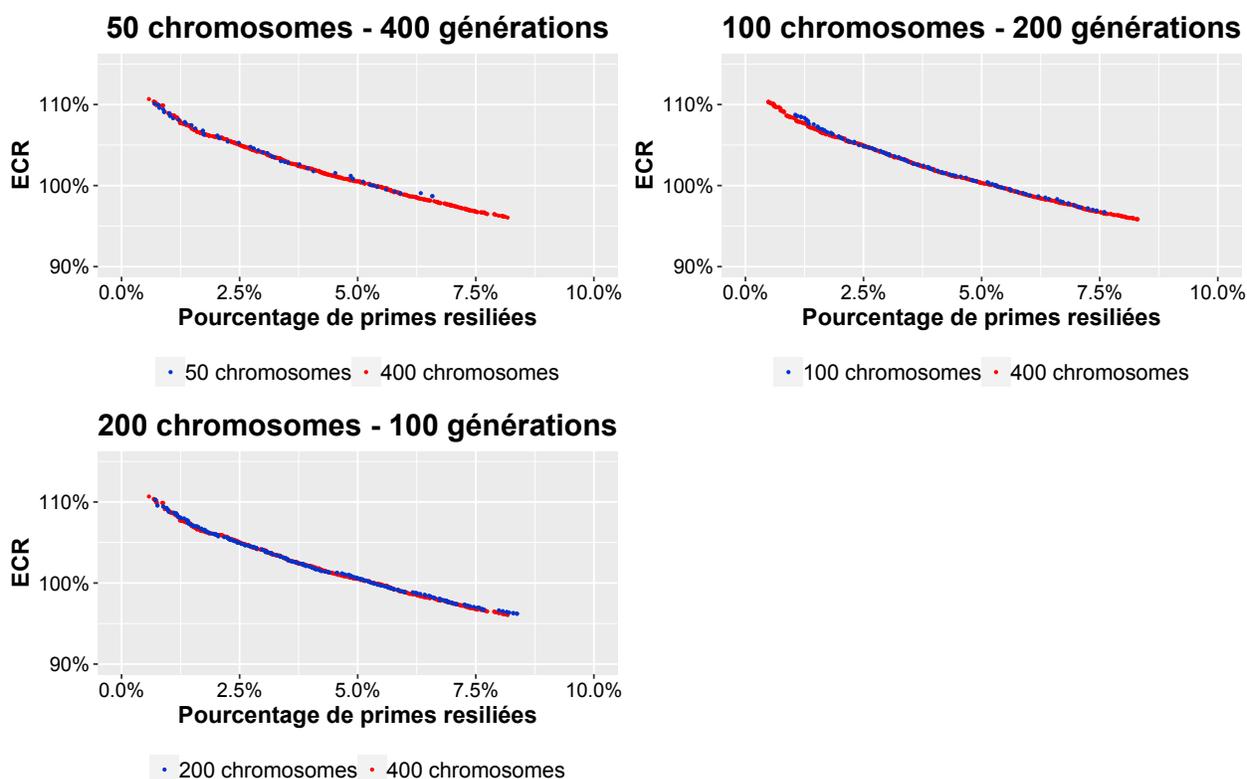


Figure 80 : sensibilité des résultats à la taille de la population

Les options avec 200 et 400 chromosomes par génération permettent une meilleure exploration de l'espace des primes que les options avec 50 et 100 chromosomes.

L'option avec 400 chromosomes explore l'espace de manière équivalente à l'option avec 200 chromosomes.

Le nombre de chromosomes par génération étant constant, les options avec 50, 100 et 200 chromosomes donnent respectivement 50, 100 et 200 scénarios de résiliation possibles. Parmi ces scénarios de résiliation, le nombre de chromosomes suivants respecte les contraintes de l'entité, c'est-à-dire résilier jusqu'à 4% de primes :

- 50 chromosomes : 33 chromosomes entre 0% et 4% de primes résiliées
- 100 chromosomes : 61 chromosomes entre 0% et 4% de primes résiliées
- 200 chromosomes : 118 chromosomes entre 0% et 4% de primes résiliées
- 400 chromosomes : 125 chromosomes entre 0% et 4% de primes résiliées

Les options avec 50 et 100 chromosomes donnent peu de scénarios de résiliation situés entre 0% et 4% de primes résiliées. Les options avec 200 et 400 chromosomes donnent un nombre de scénarios de résiliation jugé satisfaisant. Par rapport à l'option avec 200 chromosomes, l'option avec 400 ne permet pas d'améliorer significativement le nombre de chromosomes respectant la contrainte de prime.

On en conclut qu'un nombre de chromosomes entre 200 et 400 permet d'explorer l'espace et de donner un nombre satisfaisant de solutions respectant la contrainte de l'entité. L'option avec 200 chromosomes a des performances proches de celle avec 400 chromosomes et aurait pu être utilisée dans notre étude.

Nombre de générations

Le nombre de générations correspond au critère d'arrêt de l'algorithme : si l'algorithme n'a pas convergé au bout d'un nombre prédéfini d'itérations, il est automatiquement arrêté. Dans notre étude, on utilise un nombre de générations maximum de 50. Pour rappel, la taille de la population est de 400 chromosomes.

Dans cette section, la sensibilité des résultats au nombre de générations est étudiée. Les nuages de points ci-dessous comparent les résultats obtenus pour un nombre de générations variant entre 5 et 200. Les points bleus correspondent aux options avec 5, 10, 25, 50, 75 et 100 générations. Les points rouges sont les mêmes sur toutes les figures : il s'agit de l'option avec 200 générations, qui permet de comparer la convergence des différentes options.

Toutes les options ont une population de 400 chromosomes et ont été initialisées avec la même population aléatoire.

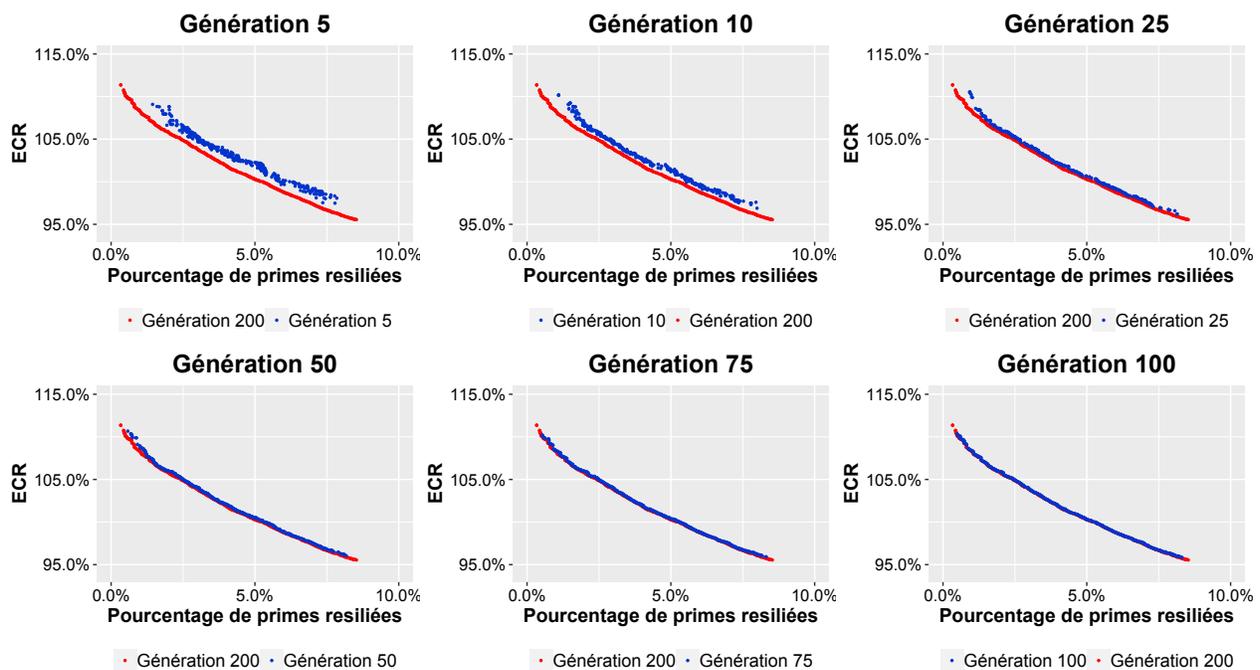


Figure 81 : sensibilité des résultats au nombre de générations

On remarque qu'entre les générations 50 et 200, les résultats sont stables et la population n'évolue que très peu.

Conclusion

En conclusion, plus la population est grande et moins on aura besoin d'un grand nombre de générations. A temps de calcul égal, la taille de la population a été privilégiée au nombre de générations. En effet, plus la population initiale est grande et plus la probabilité qu'elle contienne de bons chromosomes augmente. Par ailleurs, une grande population permet d'obtenir davantage de scénarios de résiliation sur la frontière efficiente et de donner de la flexibilité à l'entité dans ses choix de résiliation.

3.3.3.3 Génération de la population initiale

Cette partie a pour objectif d'effectuer des études de sensibilité sur la génération de la population initiale. Les résultats obtenus dans ce mémoire ont été obtenus avec une génération « alternative » de la population initiale. La méthode alternative a pour objectif de favoriser la diversité génétique, la couverture uniforme de l'espace de recherche et d'accélérer la vitesse de convergence. Cette méthode s'oppose à la génération aléatoire de la population initiale, qui est considérée comme la méthode de génération la plus courante. Dans cette partie, on comparera l'initialisation alternative avec l'initialisation aléatoire. **Dans un premier temps**, les résultats obtenus seront comparés. **Dans un second temps**, la vitesse de convergence sera étudiée.

Comparaison des résultats obtenus

Le nuage de points ci-dessous présente les résultats obtenus pour une population initiale générée aléatoirement et pour une population initiale générée avec la méthode dite « alternative ».

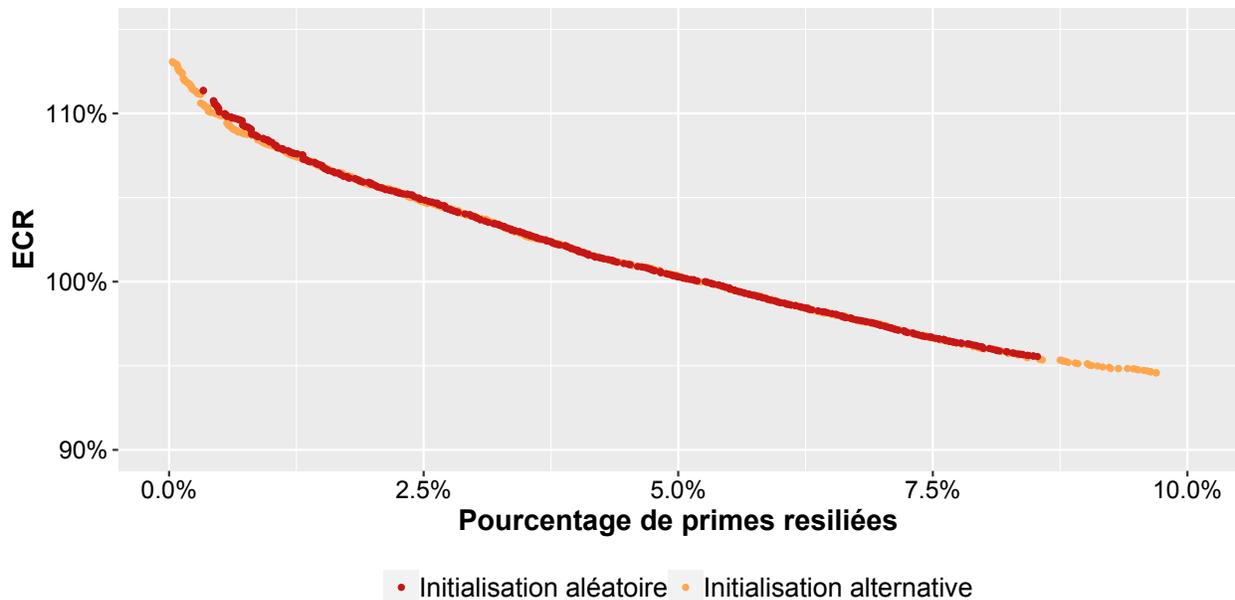


Figure 82 : comparaison des résultats avec initialisation aléatoire et alternative

Tout d'abord, on remarque que les résultats obtenus avec la méthode alternative ont permis d'explorer davantage l'espace de recherche : les chromosomes ont des montants de primes résiliées entre 0% et 9,7% contre entre 0,3% et 8,5% pour les chromosomes obtenus à partir de l'initialisation aléatoire.

Par ailleurs, les résultats obtenus avec la méthode alternative sont de meilleure qualité pour les pourcentages de primes résiliées entre 0,3% et 1%. Les résultats sont de qualité équivalente à partir de 1% de primes résiliées.

En conclusion, l'entité souhaitant résilier entre 1% et 4% de prime, l'initialisation alternative ne permet pas d'obtenir de meilleurs résultats que l'initialisation aléatoire.

Comparaison de la vitesse de convergence

De « bons » chromosomes ont été intégrés dans la population initiale alternative dans le but d'accélérer la convergence de l'algorithme génétique : il s'agit des chromosomes obtenus avec l'algorithme glouton. Dans cette section, nous allons comparer la vitesse de convergence des méthodes aléatoire et alternative.

Les nuages de points ci-dessous présentent différentes générations de l'algorithme génétique pour l'initialisation aléatoire et alternative.

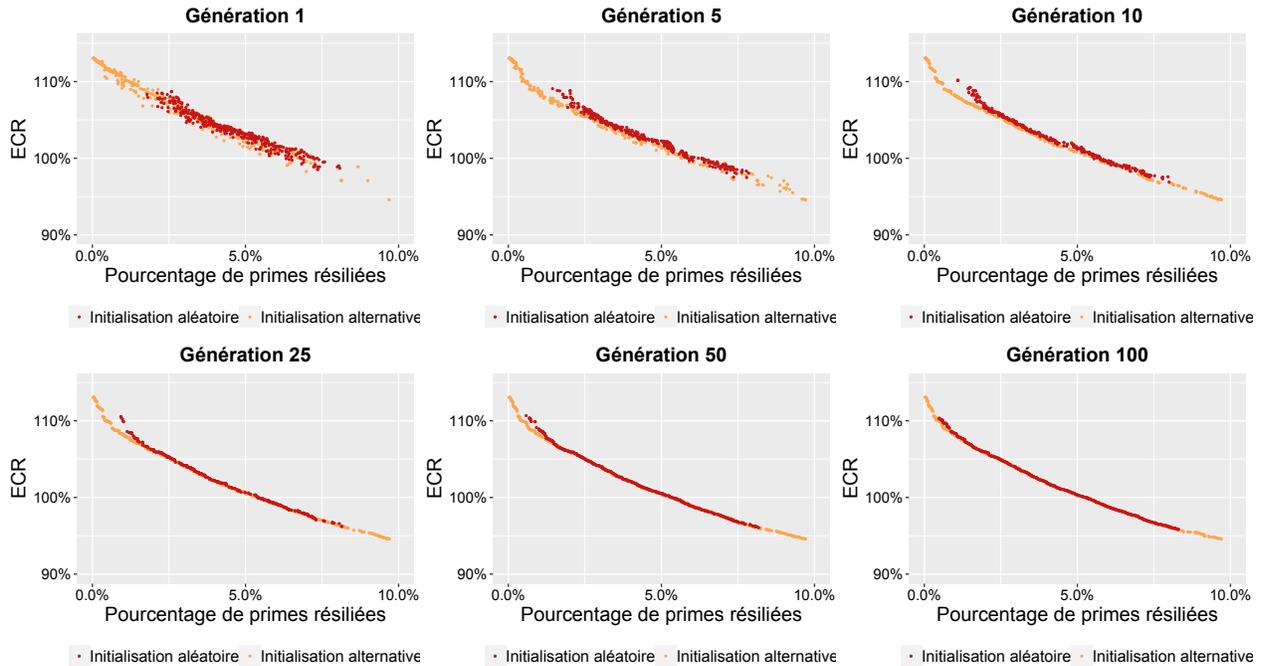


Figure 83 : comparaison de la vitesse de convergence des résultats obtenus avec initialisation aléatoire et alternative

On constate que les résultats obtenus avec l'initialisation alternative convergent plus vite que les résultats obtenus avec initialisation aléatoire. En effet, les résultats de l'initialisation alternative commencent à être stables à partir de la génération 10, contre la génération 50 pour l'initialisation aléatoire. L'initialisation alternative permet donc de simuler un plus petit nombre de générations et de réduire les temps de calcul.

Conclusion

L'initialisation alternative ne permet pas d'améliorer les résultats entre 0% et 4% de primes résiliées mais elle permet d'accélérer la vitesse de convergence et de limiter les temps de calcul.

3.3.4 Limites de l'étude

L'objectif de cette partie est de présenter les limites de notre étude et de proposer des pistes de réflexion.

La première limite de notre étude repose sur les données utilisées. En effet, seules les polices couvrant le risque ouragan ont été considérées : les polices tremblement de terre et incendie n'ont pas été prises en compte car nous ne disposons pas de leur modélisation. Ce choix se justifie par le fait que notre étude a pour vocation de prouver la faisabilité de l'optimisation de la politique de *pruning* et se restreindre aux ouragans fait donc sens dans un premier temps. Par ailleurs, ce choix était opérationnellement possible car au Mexique, chaque péril est couvert indépendamment par l'intermédiaire d'une police dédiée. Les informations de primes étaient donc disponibles par péril.

Le lecteur souhaitant mettre en pratique les méthodes développées ici devra considérer toutes les polices du portefeuille dommage.

Une deuxième limite réside dans le modèle CAT utilisé. En effet, seuls les résultats de modélisation *inhouse* étaient à disposition. Or au sein du groupe AXA, le risque ouragan est modélisé par un *blending* de plusieurs modèles CAT (*inhouse* et externes). Les modèles CAT ayant un niveau d'incertitude important, il peut y avoir des différences matérielles entre les résultats de deux modèles. Dès lors, l'intérêt du *blending* est de limiter le risque de volatilité entre les modèles. Chaque modèle se voit attribuer une note, et on obtient la distribution du modèle final en faisant, à chaque période de retour, une moyenne pondérée des résultats des différents modèles. Le lecteur souhaitant mettre en pratique les méthodes développées ici devra obtenir les résultats des différents modèles utilisés dans le *blending* et calculer l'YELT *blendée* par site.

Par ailleurs, ce mémoire part de l'hypothèse que les résultats de modélisation CAT par site sont fiables. Or les modèles CAT sont généralement calibrés pour donner des résultats pertinents à l'échelle d'un portefeuille et on peut se poser la question de la pertinence des résultats à un niveau de granularité par site.

Une troisième limite de notre étude concerne la sensibilité des résultats obtenus.

Tout d'abord, l'impact de la politique de *pruning* sur le ratio combiné économique est très sensible à un éventuel changement de version des modèles utilisés dans le modèle interne. En effet, une nouvelle version peut avoir des impacts matériels sur les résultats des modèles (changement du module vulnérabilité, du catalogue d'événements stochastiques, etc.) et avoir des impacts différents sur les coûts de réassurance et du capital.

Les résultats sont également très sensibles à la qualité des données d'exposition. Les modèles CAT disposent de modules de gestion des valeurs manquantes et leur assignent des valeurs par défaut. Une forte hausse ou forte baisse de la qualité des données peut avoir un impact matériel sur les résultats de modélisation et donc biaiser les impacts attendus par la politique de résiliation.

Une quatrième limite porte sur les polices non résiliées. Dans notre étude, on suppose que les polices non résiliées sont renouvelées à l'identique et qu'il n'y a pas de nouvelles souscriptions. Or en pratique, toutes les polices ne seront pas renouvelées, certaines polices renouvelées subiront des changements tarifaires et de nouvelles polices seront souscrites.

Avec les méthodes décrites dans ce mémoire, il est possible de prendre en compte les polices résiliées mais il n'est pas possible de quantifier l'impact des nouvelles souscriptions sur la rentabilité du portefeuille car on ne dispose pas de leur modélisation. Des méthodes pourraient être développées pour prendre en compte l'impact des nouvelles souscriptions sur le ratio combiné économique. Le lecteur intéressé pourra se référer au mémoire « Sophistication de la tarification du risque catastrophe naturelle » (*El Alj, 2019, [15]*) qui développe une méthode de quantification d'impact d'une nouvelle souscription sur

la table YELT d'un portefeuille. La méthode repose sur une restructuration du catalogue d'événements stochastiques du modèle CAT et sur un système de gestion de bases de données.

Une cinquième limite porte sur l'absence d'analyse ex post des méthodes utilisées. Il aurait été intéressant d'analyser si les polices résiliées ont vraiment permis d'améliorer l'ECR pour comprendre si l'algorithme génétique a permis de faire une proposition pertinente.

Conclusion

Avec un ratio combiné économique de 113%, le portefeuille de risques ouragan étudié dans ce mémoire était déficitaire. Les coûts de réassurance et du capital pesaient lourdement sur la rentabilité et représentaient près de 35% des primes brutes.

Pour améliorer sa rentabilité, l'entité souhaitait mettre en place une politique de résiliation ciblée de certaines polices de son portefeuille. Elle n'était cependant pas en mesure d'identifier les polices qui pesaient le plus sur la rentabilité et d'anticiper l'impact de leur résiliation sur le ratio combiné économique. En effet, cela aurait nécessité de lancer le modèle CAT pour évaluer l'impact de chaque option sur les coûts de réassurance et du capital, ce qui est opérationnellement impossible.

Le **premier objectif** de ce mémoire était de développer une méthode permettant de quantifier l'impact de n'importe quel scénario de résiliation sur le ratio combiné économique du portefeuille. Des méthodes d'estimation du résultat cédé et du coût du capital ont été développées pour prendre en compte les effets non-proportionnels. Il est désormais possible de quantifier en moins d'une seconde l'impact des résiliations sur la rentabilité du portefeuille. Des résultats similaires auraient été obtenus en plusieurs heures avec un modèle CAT. Cette méthode a permis de renforcer la connaissance du portefeuille : elle a permis d'identifier les polices qui pèsent le plus sur les coûts de réassurance et du capital. Cette méthode permet également de piloter la politique de résiliation de risques et d'anticiper ses impacts sur les coûts de réassurance et sur le STEC.

Le **second objectif** de notre étude consistait à proposer une optimisation des combinaisons de polices à résilier pour minimiser le ratio combiné économique tout en minimisant le pourcentage de primes résiliées. Pour résoudre ce problème d'optimisation combinatoire multi-objectif, notre étude s'est appuyée sur un algorithme génétique. Les résultats se présentent sous la forme d'une frontière d'efficacité composée d'un ensemble de scénarios de résiliation répondant au problème. L'algorithme est flexible et stable. Il permet aussi de prendre en compte les différentes contraintes métier.

Ce mémoire permet d'aider l'entité à identifier les combinaisons de polices à résilier qui permettraient d'optimiser le ratio combiné économique tout en minimisant le pourcentage de primes résiliées.

Ce faisant, ce mémoire permet d'élargir l'utilisation des modèles CAT : d'un simple outil de *Risk Management*, il devient possible d'exploiter leurs résultats pour piloter et optimiser les résiliations des polices catastrophes spécifiques. Les méthodes développées permettent aux équipes de souscription de s'approprier davantage les modèles CAT : elles sont maintenant en mesure de mieux anticiper les évolutions de leurs résultats.

Notre étude a l'intérêt de présenter des méthodes adaptées à un environnement professionnel complexe et à des contraintes opérationnelles. Par exemple : toutes les polices ne peuvent pas être résiliées car certaines sont souscrites en bouquet ; on souhaite résilier un montant limité de prime pour ne pas envoyer un mauvais signal au marché ; il n'est pas possible de ne résilier qu'une partie d'une police car cela impliquerait la mise en place de facs, coûteuses en temps de gestion, etc.

Pour prendre en compte ces contraintes, les méthodes sont développées en partenariat avec les équipes de souscription, pour s'assurer qu'elles sont utilisables opérationnellement.

Les algorithmes d'optimisation utilisés dans ce mémoire sont des algorithmes génétiques, généralement utilisés pour des problèmes d'application industrielle, financière ou d'allocation de ressources. Ce mémoire propose d'élargir l'utilisation de ces algorithmes pour optimiser un portefeuille d'assurance.

Les algorithmes ont dû être adaptés pour prendre en compte les différentes contraintes opérationnelles. L'algorithme génétique a été choisi car il est flexible : il est possible de le modifier pour prendre en compte de nouvelles contraintes ou un objectif supplémentaire.

Notre étude a également l'intérêt de présenter des méthodes innovantes pour générer la population initiale d'un l'algorithme génétique encodé en binaire.

Un des défis de ce mémoire concerne les temps de calcul. En effet, même en restreignant l'étude aux cent polices les moins rentables du portefeuille, cela représente 2^{100} scénarios de résiliation possibles. Chaque scénario étant évalué en 0,143 seconde⁷, $5,7 \times 10^{21}$ années de calcul seraient nécessaires pour évaluer toutes les combinaisons possibles.

Les algorithmes choisis ont l'avantage de présenter un temps de calcul raisonnable : environ 48 minutes sont nécessaires pour obtenir une frontière efficiente.

Enfin, ce mémoire a l'intérêt de présenter des méthodes qui peuvent être utilisées dans d'autres contextes.

Les méthodes pourraient par exemple être appliquées pour optimiser un portefeuille de rétrocession : l'algorithme génétique permettrait d'identifier les taux de cessions optimaux des couvertures à rétrocéder pour optimiser un indicateur de risque (OEP, AEP, etc.) tout en minimisant le montant de primes rétrocédées.

La méthode de calcul d'impact sur le ratio combiné économique pourrait également permettre d'anticiper la mise en place d'une politique de *line sizing*, i.e. un changement dans les limites et franchises des polices souscrites. Cette application demanderait le développement d'un module d'application des conditions financières et le développement de courbes de rabais pour estimer les évolutions de prix.

⁷ Machine utilisée : MacOS Sierra, processeur Intel Core i5 @1,3 GHz, Mémoire (RAM) 4Go

Bibliographie

- [1] <https://www.businessinsurance.com/article/20051030/ISSUE01/100017815/wilma-losses-rise-as-industry-reels> (estimation des pertes assurantielles de l'ouragan Wilma), HOFMANN M., *Business Insurance*, site consulté le 20 mars 2021.
- [2] <https://data.oecd.org/insurance/gross-insurance-premiums.htm#indicator-chart> (primes d'assurance brutes encaissées au Mexique entre 2000 et 2019), OCDE, site consulté le 20 mars 2021.
- [3] BEVERE L. et al. (2020) « *Natural catastrophes in times of economic accumulation and climate change* ». *Swiss Re Sigma Research*.
- [4] <https://news.un.org/fr/story/2020/10/1079642> (le changement climatique, moteur du doublement des catastrophes naturelles au cours des 20 dernières années (ONU), PARRY D., 2020), site consulté le 15 octobre 2020.
- [5] <https://www.dw.com/fr/des-ouragans-de-plus-en-plus-forts-dans-les-prochaines-d%C3%A9cennies/a-40420062> (coupe transversale d'une dépression tropicale, 2017, LE TOUZE A.), site consulté le 20 octobre 2020.
- [6] <https://www.rms.com/blog/2019/08/23/rms-hwind-forecasting-untangling-the-spaghetti> (prédiction des ouragans, exemple d'empreintes), site consulté le 15 octobre 2020.
- [7] BENHAMOU J. (2018) *Les indicateurs de souscription et la tarification économique en réassurance. Mémoire d'actuariat*, ENSAE ParisTech.
- [8] HOLLAND J. (1992) *Adaptation in Natural and Artificial Systems*. MIT Press.
- [9] HASSANAT A. et al. (2019) « *Choosing Mutation and Crossover Ratios for Generic Algorithms – A review with a New Dynamic Approach* », *Information*.
- [10] MAARANEN H. et al. (2007) « *On initial populations of a genetic algorithm for continuous optimization problems* ». *Journal of Global Optimization*, vol. 37, 405-436.
- [11] DEB K. et al. (2002) « *A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II* ». *IEEE Transactions on Evolutionary Computation*, vol. 6.
- [12] PIERRAT M. (2017) *Allocation stratégique optimale d'actifs : une approche multi-objectifs et interactive. Mémoire d'actuariat*, Université Paris-Dauphine.
- [13] LAVINAS Y. et al. (2018) « *Experimental Analysis of the Tournament Size on Genetic Algorithms* ». *IEEE International Conference on Systems, Man, and Cybernetics*.
- [14] PICEK S. et GOLUB M. (2010) « *Comparison of a crossover operator in binary-coded genetic algorithms* ». *WSEAS Transactions on Computers*, vol. 9.
- [15] EL ALJ M. (2019) *Sophistication de la tarification du risque catastrophe naturelle. Mémoire d'actuariat*, ENSAE ParisTech.
- [16] S&P Global Ratings (2020) « *Global Reinsurance Highlights 2020* », p. 19.

- [17] https://www.institutdesactuaires.com/global/gene/link.php?doc_id=9981&fg=1 (sondage RGA sur les motivations du recours à la réassurance, slide 5), site consulté le 2 février 2021.
- [18] AIR Worldwide (2014) « The AIR Tropical Cyclone Model for Mexico ».
- [19] SCHLIERKAMP-VOOSEN D. (1993) « Optimal interaction of mutation and crossover in the breeder genetic algorithm ». *International Conference on Genetic Algorithms*, Morgan Kaufmann Publishers inc.

Table des figures

<u>Figure 1</u>	: pertes assurantielles liées aux catastrophes entre 1970 et 2019.....	4
<u>Figure 2</u>	: coupe transversale d'une dépression tropicale	5
<u>Figure 3</u>	: échelle de Saffir-Simpson	6
<u>Figure 4</u>	: trajectoires des ouragans ayant touché le Mexique entre 1851 et 2019.....	7
<u>Figure 5</u>	: densité de population au Mexique en 2010.....	7
<u>Figure 6</u>	: trajectoires des ouragans Wilma et Odile	8
<u>Figure 7</u>	: architecture d'un modèle CAT.....	9
<u>Figure 8</u>	: empreinte de l'ouragan Irma.....	10
<u>Figure 9</u>	: exemple de courbes de vulnérabilité.....	11
<u>Figure 10</u>	: exemple d'une table YELT.....	12
<u>Figure 11</u>	: exemple de courbes d'OEP et d'AEP	14
<u>Figure 12</u>	: courbes d'AEP de deux périls.....	15
<u>Figure 13</u>	: illustration de l'ECR.....	16
<u>Figure 14</u>	: illustration du ratio combiné	17
<u>Figure 15</u>	: exemple d'un bilan en vision solvabilité II.....	19
<u>Figure 16</u>	: cartographie des risques dans le modèle interne d'AXA.....	20
<u>Figure 17</u>	: illustration de l'ECR.....	22
<u>Figure 18</u>	: Taux de remplissage par caractéristique.....	25
<u>Figure 19</u>	: exemple d'une table YELT à la granularité par site	26
<u>Figure 20</u>	: poids de chaque garantie dans la somme assurée totale	28
<u>Figure 21</u>	: poids de chaque catégorie de Loss Ratio	28
<u>Figure 22</u>	: polices ayant les dix plus grosses primes du portefeuille	29
<u>Figure 23</u>	: décomposition de l'ECR.....	31
<u>Figure 24</u>	: processus de quantification de l'impact d'un scénario de résilience	32
<u>Figure 25</u>	: exemple d'un scénario de résilience.....	34
<u>Figure 26</u>	: mécanisme de calcul de l'YELT du nouveau portefeuille.....	34
<u>Figure 27</u>	: impact d'un scénario de résilience sur l'OEP	35
<u>Figure 28</u>	: exemple de l'impact marginal d'un scénario de résilience	35
<u>Figure 29</u>	: adaptation de la limite du traité pour un scénario de résilience	36
<u>Figure 30</u>	: évolution du RoL en fonction du tiers-point.....	38
<u>Figure 31</u>	: évolution du RoL en fonction du LoL	39
<u>Figure 32</u>	: évolution du prix d'un euro de récupération moyenne	40
<u>Figure 33</u>	: évolution du coefficient de variation en fonction du LoL	40
<u>Figure 34</u>	: décomposition du RoL.....	41
<u>Figure 35</u>	: comparaison entre les différentes méthodes d'estimation du RoL	42
<u>Figure 36</u>	: exemple de courbe d'AEP du résultat brut et net avant et après résilience	43
<u>Figure 37</u>	: impact marginal du STEC CAT standalone.....	43
<u>Figure 38</u>	: méthodologie d'agrégation des risques en modèle interne AXA	44
<u>Figure 39</u>	: illustration de la diversification du STEC CAT.....	45
<u>Figure 40</u>	: impact de l'évolution du STEC CAT sur le STEC P&C et sur le STEC total	46
<u>Figure 41</u>	: impact de la réduction du STEC CAT sur le STEC Total	47
<u>Figure 42</u>	: localisation des deux sites.....	50

<u>Figure 43</u> : métriques obtenues pour les deux polices.....	50
<u>Figure 44</u> : indicateurs marginaux par rapport au portefeuille de référence	50
<u>Figure 45</u> : efficacité des polices permettant de réduire l'ECR.....	52
<u>Figure 46</u> : profil des polices les moins rentables	53
<u>Figure 47</u> : processus de résolution	54
<u>Figure 48</u> : comparaison des approches a priori et a posteriori.....	56
<u>Figure 49</u> : démarche de résolution	59
<u>Figure 50</u> : principes généraux de l'algorithme génétique	62
<u>Figure 51</u> : structure d'un chromosome.....	63
<u>Figure 52</u> : exemple de calcul de la distance de Hamming entre deux chromosomes	65
<u>Figure 53</u> : distance entre les chromosomes de la population générée aléatoirement.....	65
<u>Figure 54</u> : répartition des gènes	65
<u>Figure 55</u> : répartition de la population initiale dans l'espace de recherche	66
<u>Figure 56</u> : pourcentage d'écart de prime entre deux chromosomes.....	67
<u>Figure 57</u> : application de contraintes sur l'écart de prime entre les chromosomes.....	68
<u>Figure 58</u> : répartition de la population initiale générée avec contrainte de 0,012%.....	68
<u>Figure 59</u> : pénalisation des polices en fonction de leur pourcentage de primes résiliées	69
<u>Figure 60</u> : répartition de la population initiale générée avec pénalisation.....	69
<u>Figure 61</u> : population générée avec l'algorithme glouton.....	71
<u>Figure 62</u> : exemple de population initiale	71
<u>Figure 63</u> : illustration d'une population à cinq fronts	73
<u>Figure 64</u> : illustration de la crowding distance avec deux critères	74
<u>Figure 65</u> : exemple d'un croisement en un point	76
<u>Figure 66</u> : exemple d'un croisement en N points.....	77
<u>Figure 67</u> : exemple de croisement uniforme	78
<u>Figure 68</u> : exemple de mutation par inversion de bit.....	79
<u>Figure 69</u> : évolution de l'ECR en fonction du pourcentage de primes résiliées	82
<u>Figure 70</u> : évolution de l'efficacité marginale en fonction du pourcentage de primes résiliées	83
<u>Figure 71</u> : évolution des primes nettes en fonction des primes brutes résiliées.....	84
<u>Figure 72</u> : décomposition de l'ECR	86
<u>Figure 73</u> : poids de chaque composant de l'ECR	87
<u>Figure 74</u> : efficacité marginale et efficacité des composants de l'ECR.....	88
<u>Figure 75</u> : résultats obtenus avec et sans contrainte sur le pourcentage de primes résiliées.....	92
<u>Figure 76</u> : évolution de l'ECR en fonction du pourcentage de primes résiliées	93
<u>Figure 77</u> : évolution du pourcentage de primes résiliées en fonction du nombre de polices.....	94
<u>Figure 78</u> : évolution de l'ECR en fonction du nombre de polices résiliées.....	95
<u>Figure 79</u> : sensibilité des résultats au nombre de gènes présents dans un chromosome.....	97
<u>Figure 80</u> : sensibilité des résultats à la taille de la population	99
<u>Figure 81</u> : sensibilité des résultats au nombre de générations	100
<u>Figure 82</u> : comparaison des résultats avec initialisation aléatoire et alternative.....	101
<u>Figure 83</u> : comparaison de la vitesse de convergence des résultats obtenus avec initialisation aléatoire et alternative	102
<u>Figure 84</u> : classification des formes et types de réassurance	112
<u>Figure 85</u> : fonctionnement d'une fac proportionnelle.....	114
<u>Figure 86</u> : fonctionnement d'une fac non proportionnelle.....	115
<u>Figure 87</u> : réduction de la somme assurée d'un risque cédé à un traité	116
<u>Figure 88</u> : différence entre un quota share et un surplus.....	117

<u>Figure 89</u>	: différents types de traités non-proportionnels.....	118
<u>Figure 90</u>	: principales motivations du recours à la réassurance	120
<u>Figure 91</u>	: poids de chaque zone CRESTA	121
<u>Figure 92</u>	: taux de destruction et taux de prime moyens par zone CRESTA.....	122
<u>Figure 93</u>	: répartition du résultat négatif par zone CRESTA	123
<u>Figure 94</u>	: poids de chaque branche d'assurance et activité.....	124
<u>Figure 95</u>	: résultat par branche d'assurance et activité	125
<u>Figure 96</u>	: poids de chaque matériau de construction	126
<u>Figure 97</u>	: répartition du résultat par matériau de construction	127
<u>Figure 98</u>	: sensibilité des résultats au type de sélection	129
<u>Figure 99</u>	: sensibilité des résultats au type de croisement.....	129
<u>Figure 100</u>	: sensibilité des résultats au taux de croisement uniforme.....	130
<u>Figure 101</u>	: sensibilité des résultats au taux de mutation	131

Annexe 1 :

La réassurance : outil de pilotage du risque CAT

Les réassureurs ont payé près de 25% des sinistres d'assurance catastrophe naturelle survenus en 2019 (*S&P Global Ratings, 2020, [16]*). La réassurance joue donc un rôle important sur le marché du risque catastrophe. L'objectif de cette partie est d'expliquer pourquoi les assureurs ont recours à la réassurance et comment celle-ci permet de piloter leurs expositions au risque catastrophe.

Pour ce faire, **dans un premier temps**, la réassurance sera définie et ses différents types et formes seront présentés. **Dans un deuxième et troisième temps**, les deux principales formes de réassurance seront présentées : réassurance facultative puis obligatoire.

a) Formes et types de réassurance traditionnelle

La réassurance moderne est apparue en Allemagne au XIXème siècle. En effet, suite à la révolution industrielle, les assureurs avaient de plus en plus de mal à couvrir les complexes industriels contre le risque d'incendie. Pour limiter leur exposition, les assureurs avaient alors recours à la coassurance. Ces opérations avaient généralement lieu avec des assureurs étrangers pour éviter de divulguer des informations aux concurrents.

Suite à l'incendie de Hambourg en 1842, le marché prend conscience que la coassurance ne suffit plus et le besoin de réassurance devient évident. La première société de réassurance est créée en 1846 : c'est la Kölnische Rück.

La réassurance est une opération de transfert de risques entre une cédante et un réassureur : la cédante cède du risque et le réassureur accepte ce risque en échange du paiement d'une prime.

Comme on le voit sur le schéma ci-dessous, il existe deux formes de réassurance (facultative et obligatoire) et deux types de réassurance (proportionnelle et non proportionnelle).

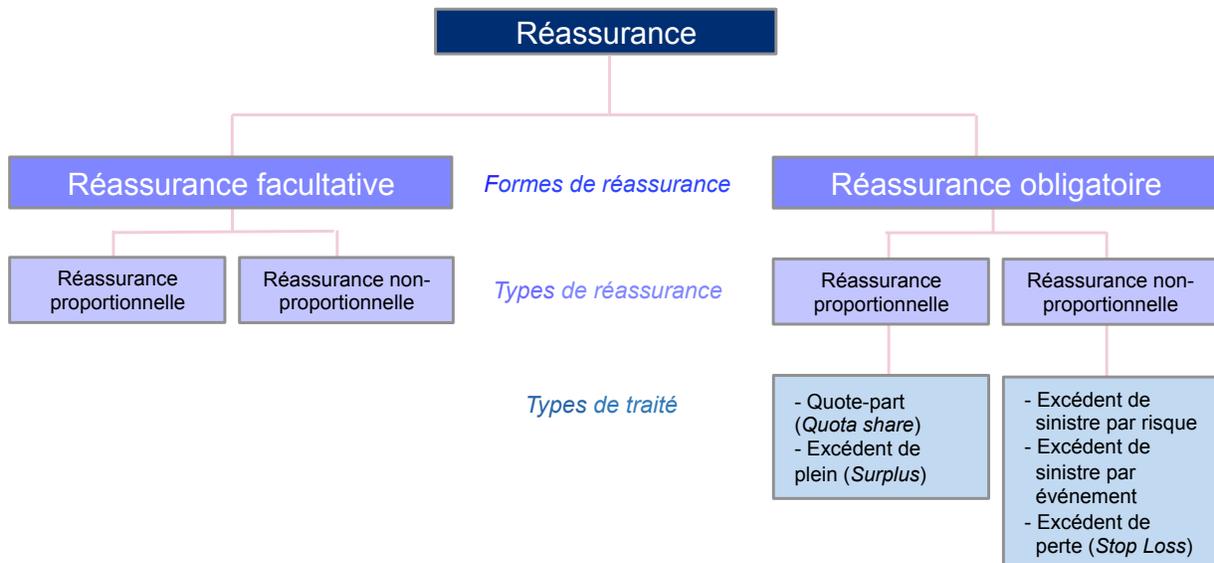


Figure 84 : classification des formes et types de réassurance

Formes de réassurance

La principale différence entre les deux formes de réassurance est leur champ d'application : la réassurance facultative s'applique au niveau d'un site ou d'une police, alors que la réassurance obligatoire s'applique généralement au niveau d'un portefeuille.

On parle de réassurance « facultative » car chaque risque est présenté individuellement au réassureur. Celui-ci a le droit d'accepter ou de refuser le risque.

On parle de réassurance « obligatoire » car une fois le traité conclu, la cédante doit céder tous les risques de son portefeuille et le réassureur a l'obligation de les accepter. A noter que les risques cédés au traité doivent respecter les conditions prévues au contrat, comme par exemple ne pas dépasser un certain montant de somme assurée, être localisé dans une zone géographique donnée, couvrir certaines garanties, etc.

Types de réassurance

Pour chaque forme de réassurance, il existe deux types : la réassurance proportionnelle et non proportionnelle. La principale différence entre ces deux types réside dans le partage du sort ou non entre la cédante et le réassureur.

La réassurance proportionnelle prévoit un partage du sort entre la cédante et le réassureur. Ainsi, le réassureur reçoit un certain pourcentage des primes d'un risque et paie en contrepartie le même pourcentage des sinistres. L'assureur cède une partie de son résultat et est rémunéré par le réassureur par l'intermédiaire d'une commission.

Il existe deux types de réassurance obligatoire proportionnelle : la quote-part (*quota share*) et l'excédent de plein (*surplus*). Ces deux types de réassurance se distinguent par la méthode choisie pour calculer le taux de cession des primes et des sinistres :

- **Quote-part (*quota share*)** : le taux de cession est le même pour tous les risques du portefeuille
- **Excédent de plein (*surplus*)** : le taux de cession est variable. Il est calculé pour chaque risque en fonction de sa somme assurée.

Le problème du calcul du taux de cession ne se pose pas pour la réassurance facultative car cette forme de réassurance ne couvre qu'un seul risque. Il n'y a donc qu'un seul type de réassurance facultative proportionnelle.

La réassurance non proportionnelle prévoit le paiement des sinistres si ceux-ci excèdent un certain seuil fixé au contrat : on parle de traité en excédent de sinistre ou en excédent de perte. En échange, la cédante paie une prime au réassureur.

Il existe trois types de réassurance obligatoire non proportionnelle : excédent de sinistre par risque, par événement et *Stop Loss* (excédent de perte). La différence entre ces trois traités repose dans l'agrégation des sinistres présentés au traité :

- **Excédent de sinistre par risque** : chaque sinistre est présenté individuellement au traité.
- **Excédent de sinistre par événement** : les sinistres sont agrégés par événement avant d'être présentés au traité. La notion d'événement est donc clé pour ce type de traité.
- **Excédent de perte (*Stop Loss*)** : les sinistres sont agrégés par unité de temps (généralement par année) avant d'être présentés au traité.

La question de l'agrégation des sinistres ne se pose pas dans le cas de la réassurance facultative car un seul risque est couvert. Il n'y a donc qu'un seul type de réassurance facultative non proportionnelle.

b) Réassurance facultative

Contrairement à la réassurance obligatoire, généralement souscrite au niveau d'un portefeuille, la réassurance facultative est souscrite au niveau d'une police, d'un site, voire même d'une garantie. On appelle généralement un contrat de réassurance facultative une « fac ».

Pour rappel, une fac est un contrat conclu entre une cédante et un réassureur. La cédante est libre de proposer le risque à son réassureur et celui-ci est libre de l'accepter ou de le refuser.

Cette partie se compose de trois sections. Les deux premières décrivent le fonctionnement des facs proportionnelles puis non proportionnelles. La troisième section présente l'intérêt que peut avoir la réassurance facultative pour un assureur.

Facs proportionnelles

Les facs proportionnelles consistent à céder une part verticale d'un site ou d'une police : la cédante cède une certaine proportion de la prime d'un risque et le réassureur paie la même proportion des sinistres éventuels.

Le schéma ci-dessous illustre le fonctionnement d'une fac proportionnelle. La police réassurée dispose d'une franchise et d'une limite d'assurance. Si le risque sous-jacent est totalement détruit, le réassureur paiera un certain pourcentage de la limite. Si le risque n'est pas totalement détruit, le réassureur paiera un certain pourcentage du sinistre, en excès de la franchise.

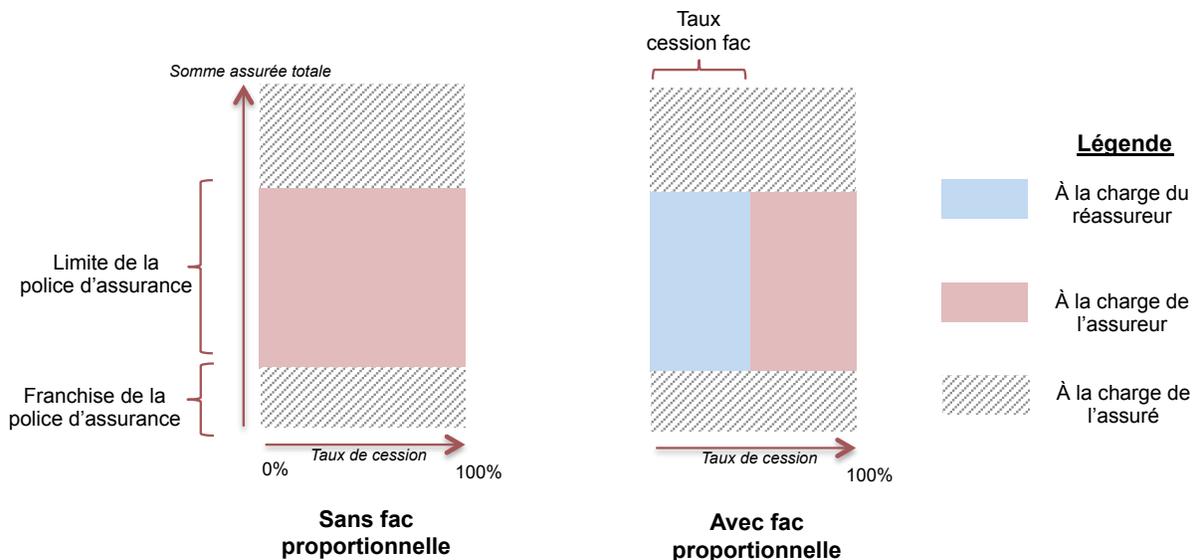


Figure 85 : fonctionnement d'une fac proportionnelle

On appelle récupérations (*recoveries*) la part des sinistres à charge du réassureur. Les récupérations se calculent comme suit :

$$\text{Récupérations} = \text{Taux de cession}_{fac} * \min(\text{Limite}, \max(\text{sinistre} - \text{franchise}, 0))$$

Une fac proportionnelle a donc un fonctionnement similaire à celui de la coassurance. Cependant, la fac diffère de la coassurance par le fait que le réassureur n'est pas engagé envers l'assuré :

- **Une fac est un engagement d'un réassureur envers une cédante** : le réassureur n'est pas engagé envers l'assuré. Si une fac est souscrite et que le réassureur fait faillite, l'assureur sera tenu de payer l'intégralité du sinistre seul.
- **La coassurance engage un assureur envers un assuré** : si un risque est coassuré par deux assureurs et que l'un d'eux fait faillite, l'assureur restant a le droit de ne payer que sa part.

Facs non proportionnelles

Le schéma ci-dessous illustre le fonctionnement d'une fac non proportionnelle.

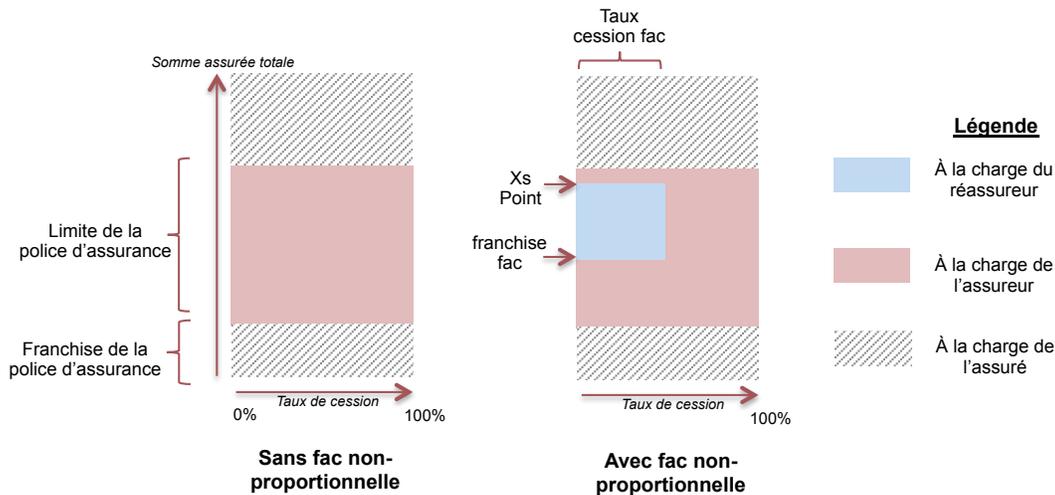


Figure 86 : fonctionnement d'une fac non proportionnelle

Une fac non proportionnelle se caractérise par sa franchise, son *Xs point* et son taux de cession. Le réassureur paie la part des sinistres située entre la franchise de la fac et l'*Xs point*, à la hauteur du taux de cession.

Les récupérations se calculent comme suit :

$$\text{Récupérations} = \text{Taux de cession}_{fac} * \min(\text{Xs point}_{fac}, \max(\text{Brut} - \text{franchise}_{fac}, 0))$$

$$\text{avec Brut} = \min(\text{Limite}_{assurance}, \max(\text{sinistre} - \text{franchise}_{assurance}, 0))$$

Intérêt de la réassurance facultative

Les deux sections précédentes ont présenté le fonctionnement de la réassurance facultative. Cette section a pour objectif d'expliquer comment la réassurance facultative peut permettre aux assureurs de piloter leur risque CAT.

Lors de la souscription de facs, chaque risque est proposé séparément et le réassureur peut accepter ou décliner le risque. Souscrire des facs est donc consommateur en temps à la fois pour l'assureur et le réassureur. De ce fait, les facs sont souvent utilisées quand seuls quelques risques requièrent d'acheter de la réassurance. Cela arrive généralement dans l'une des situations suivantes :

- **Risques dont la somme assurée dépasse la limite de souscription de l'entité⁸** : pour réduire la somme assurée, il est alors nécessaire d'avoir recours à la coassurance ou à une fac.
- **Sites risqués** : les facultatives aident à gérer les accumulations sur les sites les plus risqués du portefeuille. Dans le cas du Mexique, les assureurs vont par exemple souscrire des facs pour les grands complexes hôteliers situés en front de mer, fortement exposés au risque ouragan.
- **Risques exclus du traité de réassurance obligatoire** : les risques de pointe sont souvent exclus des traités de réassurance. En effet, les réassureurs essaient de contrôler ces risques en les excluant des traités pour les gérer avec des facultatives.

⁸ La limite de souscription correspond au montant maximal qu'un assureur est prêt à assurer pour une police donnée. Ce montant est généralement défini dans les standards et toute déviation requiert une dérogation.

Les facultatives proportionnelles ont l'intérêt de pouvoir remplacer la coassurance. Cela peut être souhaitable car la coassurance a le désavantage d'obliger l'assureur à partager des informations sur le risque assuré avec un concurrent. Par ailleurs, la coassurance rend la conclusion du contrat plus complexe pour l'assuré, qui doit négocier les conditions avec plusieurs acteurs. De ce fait, avoir recours à la coassurance peut faire perdre une affaire : un assuré peut préférer s'assurer chez un concurrent qui est capable de souscrire le risque seul.

Les facs proportionnelles permettent donc d'éviter d'avoir recours à la coassurance. Elles permettent aux assureurs d'augmenter leur part et de maintenir leur prestige et réputation sur les marchés.

Les facs proportionnelles sont particulièrement utilisées au Mexique car la coassurance y est interdite.

Les facs proportionnelles sont également utilisées dans le cas où la somme assurée d'un risque dépasse la somme assurée maximale admise par un traité. Pour qu'un tel risque puisse être cédé au traité, les réassureurs imposent généralement la souscription d'une fac proportionnelle.

Prenons l'exemple d'un traité qui prévoit une somme assurée maximale de 30 par risque. Une cédante souhaite céder un risque de 40. Pour ce faire, elle va souscrire une fac proportionnelle de telle sorte à réduire la somme assurée au montant maximal admis par le traité. Comme on le voit sur la figure ci-dessous, la cédante va souscrire une fac de 25% pour réduire la somme assurée à 30.

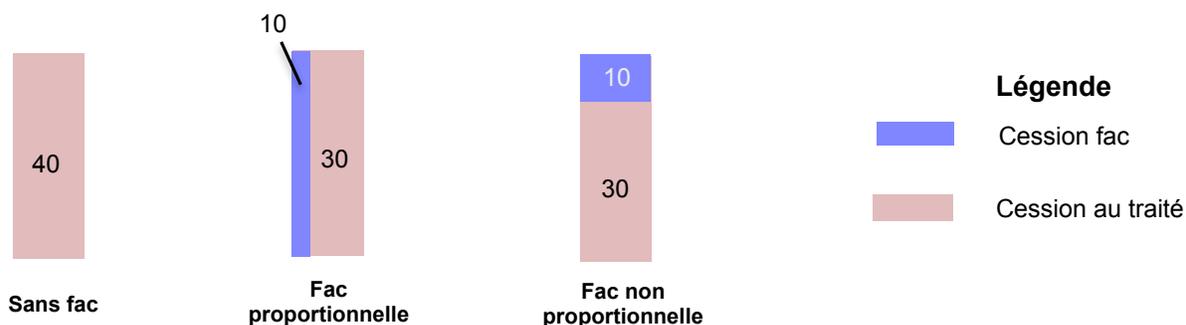


Figure 87 : réduction de la somme assurée d'un risque cédé à un traité

Certaines cédantes préfèrent souscrire une fac non proportionnelle en-deçà de la somme assurée maximale. Dans notre exemple, la cédante souscrit une fac 10 xs 30 : la somme assurée cédée au traité est alors réduite à 30. L'avantage principal réside dans le fait qu'une fac non proportionnelle coûtera moins chère qu'une fac proportionnelle. A noter que cette solution n'est généralement pas appréciée par les réassureurs car davantage de risque sera cédé au traité.

Enfin, les facultatives non proportionnelles sont utiles pour couvrir les risques de pointe du portefeuille. Prenons l'exemple d'un portefeuille de 100 risques dont deux polices dépassent de loin la somme assurée des autres risques. L'assureur peut choisir de calibrer son traité de réassurance obligatoire sur la base des 98 polices et de céder avec une fac non proportionnelle les deux risques de pointe. La fac sera généralement structurée de telle sorte à avoir la même priorité que le traité mais aura une limite plus haute.

c) Réassurance obligatoire

Contrairement à la réassurance facultative, la réassurance obligatoire couvre généralement un portefeuille. Cette forme de réassurance est appelée obligatoire ou automatique : le réassureur couvre tous les risques de la cédante qui rentrent dans les conditions du traité de réassurance.

La période de couverture d'un traité est généralement de 12 mois. Les sinistres des polices sous-jacentes sont rattachées à un traité par l'une des trois méthodes suivantes :

- **Par année de survenance** (*Loss Occuring*) : les sinistres survenant au cours de la période de couverture du traité sont couverts.
- **Par année de souscription** (*Risk Attaching*) : les sinistres des polices souscrites pendant la période de couverture du traité sont couverts.
- **Par année de déclaration** (*Claims Made*) : les sinistres déclarés pendant la période de couverture du traité sont couverts.

Cette partie se compose de trois sections. Les deux premières décrivent le fonctionnement des traités de réassurance proportionnels puis non-proportionnels. La troisième section présente l'intérêt que peut avoir la réassurance obligatoire pour un assureur.

Traités proportionnels

Il existe deux principaux types de réassurance obligatoire proportionnelle : le *quota share* et le *surplus*.

Le schéma ci-dessous illustre la différence entre ces deux types de traités :

- **Le taux de cession du *quota share* est fixe** : c'est le même pour tous les risques du portefeuille.
- **Le taux de cession du *surplus* est variable** : le taux de cession est calculé pour chaque risque. C'est le rapport entre la somme assurée d'un risque et le « plein » de réassurance. Schématiquement, plus la somme assurée est importante et plus le taux de cession sera grand, de façon à ce que la rétention de la cédante ne dépasse pas le plein. Le taux de cession est généralement limité par l'intermédiaire d'un nombre de pleins maximum par risque. Comme on le voit sur le schéma ci-dessous, le surplus permet de céder moins sur les petits risques et davantage sur les gros risques. Certains petits risques peuvent même être retenus à 100% par la cédante comme c'est par exemple le cas pour le sinistre 2.

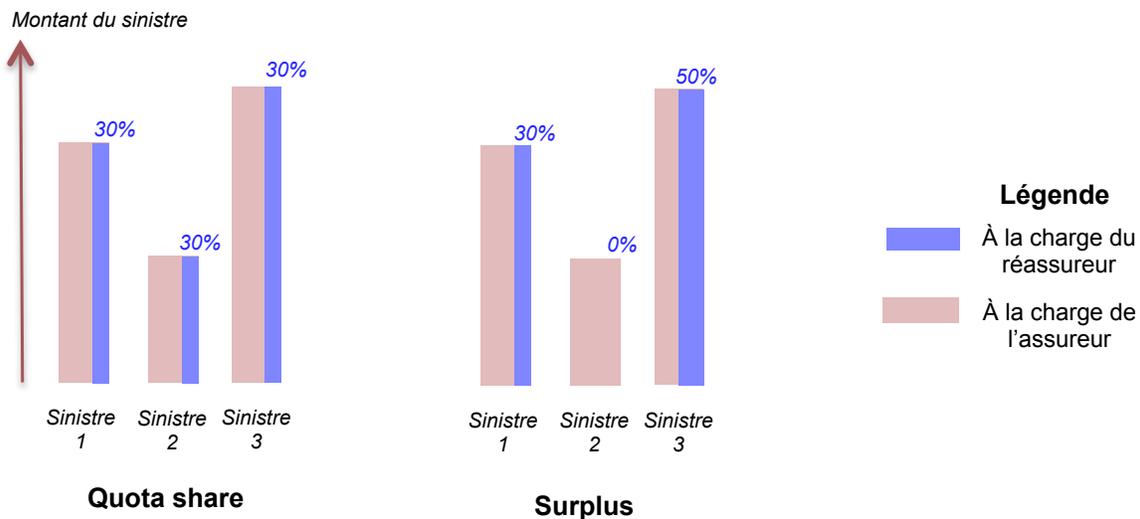


Figure 88 : différence entre un quota share et un surplus

Les récupérations du *quota share* et du *surplus* se calculent comme suit :

$$\text{Récupérations}_{\text{Quota share}} = \text{taux de cession} * \text{Sinistre}$$

$$\text{Récupérations}_{\text{Surplus}} = \frac{\min(\text{Nombre de pleins max} * \text{Plein}, \max(\text{Somme assurée} - \text{Plein}, 0))}{\text{Somme assurée}} * \text{Sinistre}$$

Traités non-proportionnels

Contrairement à la réassurance proportionnelle, la réassurance non proportionnelle ne se caractérise pas par un partage du sort entre la cédante et le réassureur. La cédante transfère une partie de son risque en excès d'un montant donné et dans une certaine limite.

Le schéma ci-dessous présente les différences entre les trois grands types de réassurance non proportionnelle : l'excédent de sinistre par risque, l'excédent de sinistre par événement et l'excédent de perte. Leur principale différence réside au niveau de l'agrégation des sinistres avant présentation au traité :

- Excédent de sinistre par risque : les sinistres sont présentés individuellement au traité.
Le sinistre 1 dépasse la priorité et génère des récupérations. Le sinistre 2 est inférieur à la priorité et ne donne pas lieu à un paiement du réassureur. Le sinistre 3 dépasse la priorité et la limite : le montant des récupérations est égal à la limite.
- Excédent de sinistre par événement : les sinistres sont agrégés par événement avant d'être présentés au traité.
Sur le schéma, on voit que les sinistres 1 et 2, qui relèvent du même événement, ont été agrégés. Le traité par événement génère davantage de récupérations pour les sinistres 1 et 2 que le traité par risque. L'événement 2 se compose d'un seul sinistre : le sinistre 3. Pour ce sinistre, les traités par risque et par événement génèrent donc le même montant de récupérations.
- Excédent de perte : les sinistres sont agrégés par année avant d'être présentés au traité.
Sur le schéma, on voit que les sinistres 1, 2 et 3 ont été agrégés. De ce fait, un traité en excédent de perte sera généralement calibré différemment d'un excédent de sinistre par événement.

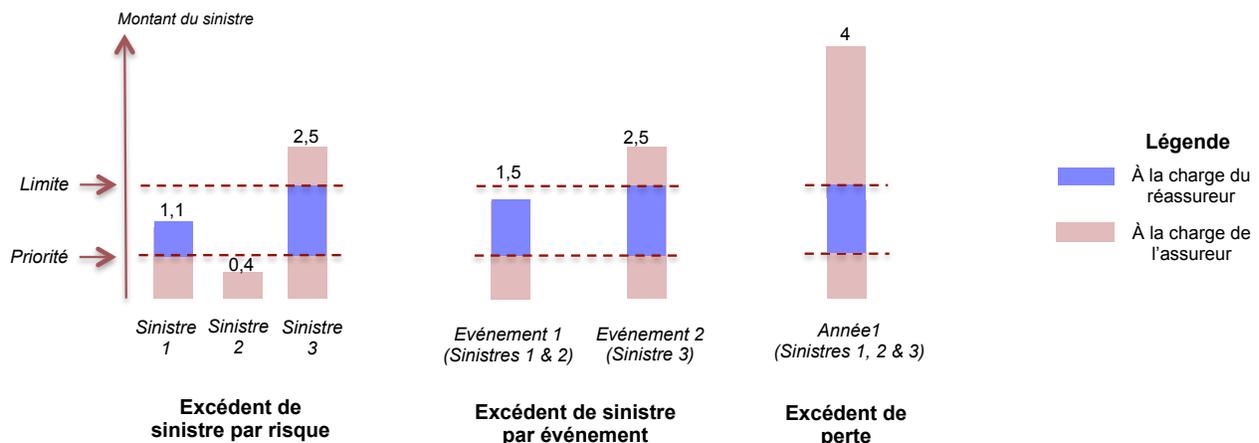


Figure 89 : différents types de traités non-proportionnels

Un traité non-proportionnel est défini par plusieurs paramètres :

- **La priorité** : c'est le montant de sinistre à partir duquel le réassureur commence à payer. Le principe de fonctionnement est celui d'une franchise déduite : si un sinistre dépasse la franchise, le réassureur paie la part des sinistres qui dépasse la priorité.
- **La portée (aussi appelée limite)** : c'est la part du sinistre qui est prise en charge par le réassureur au-delà de la priorité.

- **Le taux de cession** : la part du risque acceptée par le réassureur.
- **Le RoL (Rate on Line)** : c'est la prime de réassurance normalisée par la limite du traité. La prime peut également être exprimée en pourcentage des primes d'assurance : on parle alors de taux de prime.
- **Clauses de reconstitution et Annual Aggregate Limit (AAL)** : ces clauses ont pour but de limiter le montant total des récupérations.
Le nombre de reconstitutions correspond au nombre de portées que la cédante peut consommer sur une année. L'AAL est le montant maximal de récupérations qu'un traité peut générer sur une année.
Il est possible de convertir le nombre de reconstitutions en AAL : $AAL = Limite * (nb\ rec + 1)$.
- **Annual Aggregate Deductible (AAD)** : certains traités prévoient des AAD. Il s'agit du montant de récupérations à la charge de l'assuré avant que le réassureur ne commence à payer.

Soient n le nombre d'événements de l'année, n_i le nombre de sinistres de l'événement i et $S_{i,j}$ le montant du sinistre j de l'événement i . Pour chaque type de traité, le montant payé par le réassureur sur l'année est :

$$\text{Récupérations} = \min(AAL, \max(\text{Brut} - AAD, 0))$$

$$\text{avec} \begin{cases} \text{Brut}_{\text{excédent de sinistre par risque}} = \sum_{i=1}^n \sum_{j=1}^{n_i} \min(Limite, \max(S_{i,j} - Priorité, 0)) \\ \text{Brut}_{\text{excédent de sinistre par événement}} = \sum_{i=1}^n \min(Limite, \max(\sum_{j=1}^{n_i} S_{i,j} - Priorité, 0)) \\ \text{Brut}_{\text{excédent de perte}} = \min(Limite, \max(\sum_{i=1}^n \sum_{j=1}^{n_i} S_{i,j} - Priorité, 0)) \end{cases}$$

Dans le cas des excédents de sinistre, la cédante peut bénéficier de reconstitutions de la portée. Pour ce faire, la cédante paie une prime additionnelle pour bénéficier de reconstitutions : on parle de prime de reconstitution. Dans le cas des primes de reconstitution payantes, la prime est payée au fil de l'eau, au prorata de la capacité consommée par chaque sinistre :

$$\text{Prime reconstitution} = \min(nb\ rec * \text{Prime}_{\text{upfront}} * \text{Taux}_{\text{rec}}, \frac{\text{Récupérations}}{\text{Limite}} * \text{Prime}_{\text{upfront}} * \text{Taux}_{\text{rec}})$$

Si les primes de reconstitution sont gratuites, la prime upfront coûtera plus cher car elle contient l'espérance des primes de reconstitution :

$$\text{Prime}_{\text{upfront}}^{\text{recs gratuites}} = \text{Prime}_{\text{upfront}}^{\text{rec payantes}} * \frac{\text{Récupérations totales}}{\text{Récupérations au titre de la première portée}}$$

On note généralement un traité en excédent de sinistres « limite XS priorité – nombre de reconstitutions @ taux de reconstitution ».

Intérêt de la réassurance obligatoire

Le diagramme ci-dessous présente les principales motivations des clients qui ont recours à la réassurance (Sondage RGA [17]).

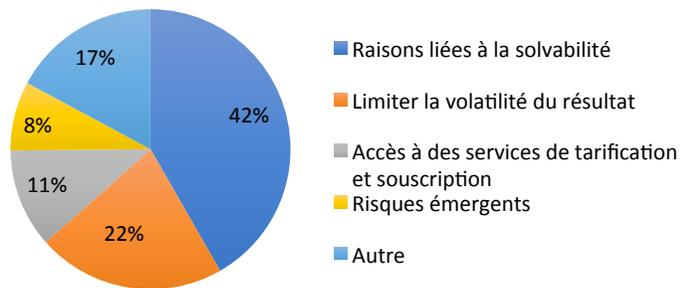


Figure 90 : principales motivations du recours à la réassurance

La principale motivation des cédantes pour avoir recours à la réassurance concerne la solvabilité. En effet, dans le contexte de la directive Solvabilité II, la réassurance joue un rôle clé. Schématiquement, les sociétés d'assurance doivent posséder des fonds propres leur permettant de faire face à leurs engagements en cas de déviation bicentenaire de leur résultat à horizon un an. Or l'immobilisation des fonds propres coûte de l'argent aux assureurs : c'est le coût du capital. La réassurance est une opération de transfert de risque : elle permet donc à la cédante de réduire le capital réglementaire et a fortiori de réduire le coût du capital.

Par ailleurs, la réassurance permet également de limiter la volatilité du résultat et de protéger le bilan. Avec la réassurance, un assureur cédera du résultat les années peu sinistrées mais limitera ses pertes en cas de forte sinistralité. La réassurance permet donc de lisser le résultat et d'en limiter la volatilité.

Les excédents de sinistre par événement sont particulièrement adaptés pour limiter la volatilité d'un portefeuille d'assurance catastrophe. Ce type de traité permet de se couvrir contre le risque d'accumulation.

La réassurance permet également aux cédantes d'avoir accès à des services. Certains réassureurs mettent ainsi à disposition de leurs cédantes des outils de souscription et de tarification, comme par exemple des cartes de risques catastrophes naturelles, des outils pour quantifier leurs accumulations ou des outils de tarification. Ces outils leur permettent de se différencier sur le marché et de négocier des places de *leader* sur des traités. Certains réassureurs proposent même d'aider la cédante à lancer un nouveau produit : ils peuvent aider à choisir les garanties adéquates, conseiller sur le niveau de tarification et sur les clauses du contrat.

Enfin, la réassurance est particulièrement utilisée pour les risques émergents, comme par exemple le risque *cyber*. Pour couvrir ce type de risque, des quote-parts sont généralement utilisées. La réassurance permet aux assureurs de limiter leurs expositions sur les marchés sur lesquels ils n'ont pas encore suffisamment d'expérience.

Annexe 2 :

Analyse descriptive de la situation géographique du portefeuille

L'objectif de cette partie est d'analyser la situation géographique du portefeuille. Pour ce faire, une granularité par zone CRESTA est utilisée.

CRESTA signifie « évaluation du risque catastrophe et standardisation des objectifs d'accumulation » (*Catastrophe Risk Evaluation and Standardizing Target Accumulations*). Créé à l'initiative de Swiss Re et Munich Re, son objectif est d'établir un standard pour le contrôle des accumulations du risque catastrophe naturelle. Les zones CRESTA sont aujourd'hui largement utilisées dans l'industrie de l'assurance et de la réassurance.

Poids de chaque zone CRESTA

Le graphe ci-dessous présente le poids de chaque zone CRESTA dans le nombre de sites, la NIV, les sinistres moyens (AAL) et les primes.

Les primes des polices multi-sites situées dans des zones CRESTA différentes ont été allouées par site au prorata de leurs pertes moyennes modélisées.

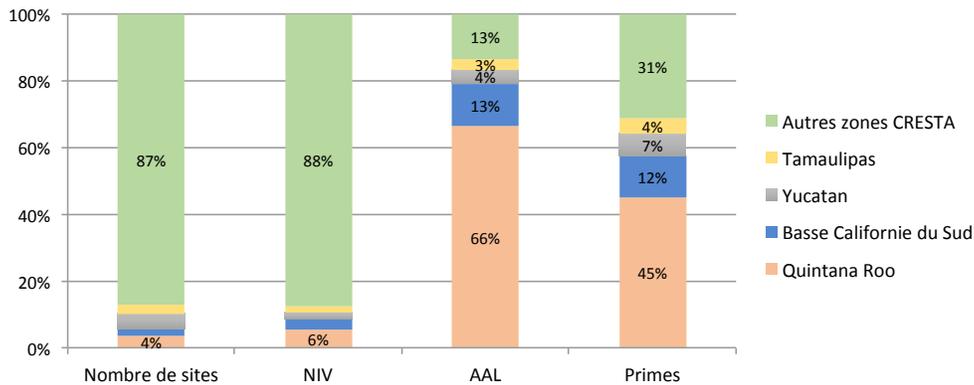


Figure 91 : poids de chaque zone CRESTA

On remarque une grande hétérogénéité entre zones CRESTA. Quintana Roo représente seulement 4% du nombre de sites mais 66% des sinistres. Sa part dans les primes du portefeuille (45%) est beaucoup plus faible que sa part dans les sinistres.

On constate le même phénomène pour les zones CRESTA de la Basse-Californie du Sud.

A l'inverse, les autres zones CRESTA sont très profitables : 87% des sites représentent 13% des sinistres mais 31% des primes.

Il y a donc une forte hétérogénéité entre zones CRESTA. Cette hétérogénéité s'explique en partie par leur situation géographique : côtière ou non côtière. La carte ci-dessous présente les taux de destruction et le taux de prime moyen par zone CRESTA.

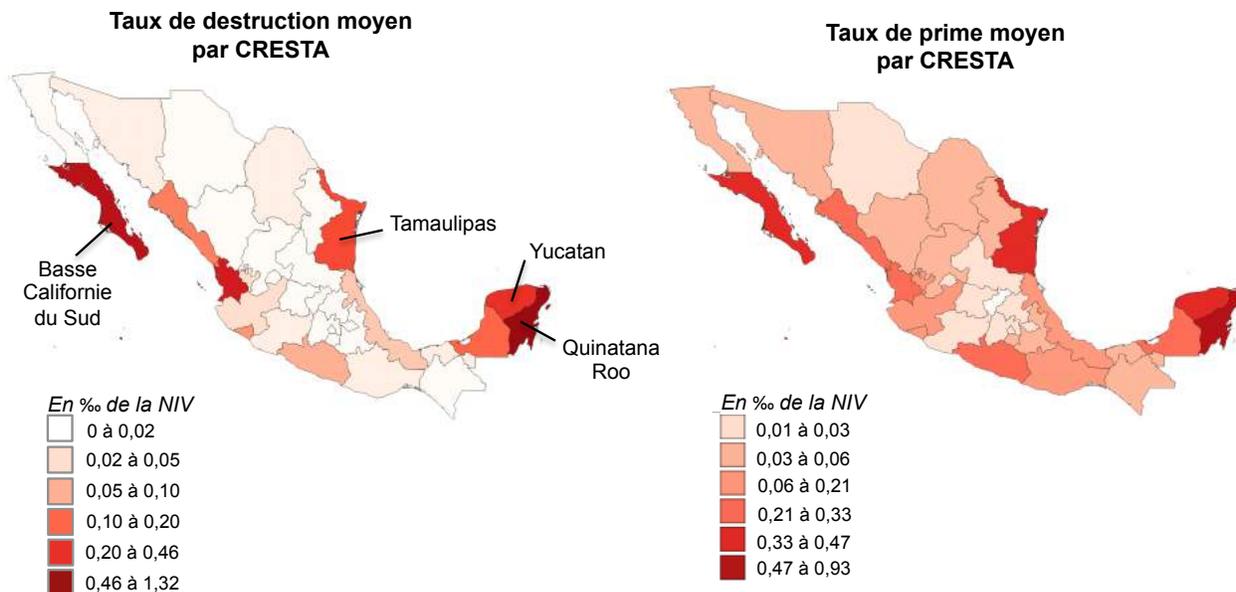


Figure 92 : taux de destruction et taux de prime moyens par zone CRESTA

On remarque une très forte hétérogénéité entre zones côtières et non côtières.

Le cas de Quintana Roo est particulièrement frappant. Avec un taux de destruction moyen de 1,32‰ et un taux de prime de 0,93‰, cette zone est de loin la zone CRESTA la plus risquée et la plus chère. Elle est cependant sous-tarifée par rapport à son niveau de risque et présente le plus mauvais *Loss Ratio* du portefeuille : 142%.

La zone CRESTA de la Basse Californie du Sud est également sous-tarifée par rapport à son niveau de risque : elle a un taux de destruction moyen de 0,47‰ pour un taux de prime de 0,46‰, soit un *Loss Ratio* de 102%.

A l'inverse, les zones non côtières concentrent le résultat positif du portefeuille : alors que ces zones affichent des taux de destruction moyen de moins de 0,01‰, leur taux de prime moyen est de 0,034‰, soit un *Loss Ratio* d'environ 30%.

Zones CRESTA pénalisant le résultat du portefeuille

Le *Loss Ratio* donne une indication de rentabilité relativement à une prime. Or les polices aux petites primes, même avec un *Loss Ratio* très élevé, ont un impact très limité sur la rentabilité globale du portefeuille.

Le résultat, pendant du *Loss Ratio* en montant, est donc un indicateur plus adapté pour identifier les polices qui pénalisent le plus la rentabilité du portefeuille.

La carte ci-dessous présente la répartition géographique des polices ayant un résultat négatif. Seules les polices ayant un résultat négatif ont été conservées : de cette manière, il n'y a pas d'effet de compensation avec les polices ayant un résultat positif. Les résultats ont ensuite été agrégés par zone CRESTA pour identifier les zones qui sont à l'origine des plus fortes pertes.

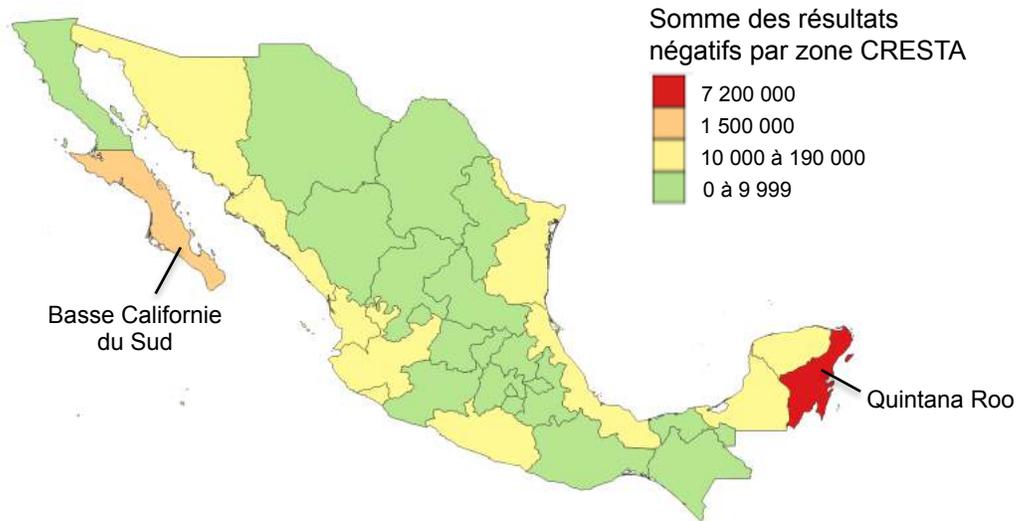


Figure 93 : répartition du résultat négatif par zone CRESTA

On remarque que les pertes sont concentrées sur les côtes et particulièrement dans les CRESTA de Quintana Roo et de la Basse-Californie du Sud : 83% du résultat négatif est réalisé dans ces deux régions. A l'inverse, les zones intérieures représentent une part négligeable du résultat négatif global.

Annexe 3 :

Analyse descriptive des caractéristiques des sites du portefeuille

Un risque est un croisement entre un aléa et une vulnérabilité. Après avoir analysé la situation géographique du portefeuille, il est donc nécessaire d'analyser les caractéristiques physiques des sites assurés.

La base de données d'exposition donne des informations sur certaines caractéristiques physiques des sites du portefeuille. Il s'agit de la branche d'assurance, de l'activité et du matériau de construction. Les autres caractéristiques physiques n'ont pas été étudiées en raison du faible volume d'information disponible.

a) Analyse par branche d'assurance et par activité

Le portefeuille se compose de risques dans quatre lignes de business différentes : résidentiel, industriel, commercial et municipal. Chaque branche dispose de plusieurs activités, qui décrivent plus précisément le site assuré : hôtel, maison individuelle, etc.

Poids de chaque branche d'assurance et activité

Les deux graphes ci-dessous présentent le poids de chaque branche et activité en terme de nombre de sites, de NIV, d'AAL et de primes.

Les activités « commerce de détail » et « hôtel » relèvent de la branche commercial et l'activité « maison individuelle » de la branche résidentiel.

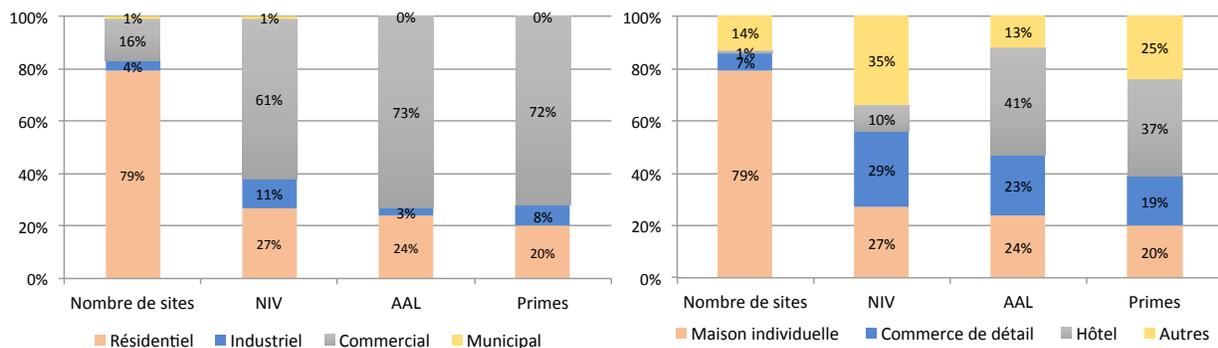


Figure 94 : poids de chaque branche d'assurance et activité

Le risque et les primes du portefeuille sont concentrés dans la branche commercial : les polices de cette LoB représentent près de 73% des sinistres et des primes. Cette LoB représente peu de sites (16%) car il s'agit principalement de grands comptes ayant des primes et sommes assurées importants. Son *Loss Ratio* est de 99% mais il n'est plus que de 89% si on exclut les sites relevant de l'activité hôtel. Cette activité représente moins de 1% des sites mais près de 41% des sinistres pour 37% des primes. Cela s'explique par le fait que les hôtels sont généralement situés en bord de mer et ont des sommes assurées importantes.

La majorité des polices est du risque résidentiel (79%). Ces polices représentent 24% des sinistres pour 20% des primes. En moyenne, ce segment est le moins profitable, avec un *Loss Ratio* de 118%. A noter que le *Loss Ratio* n'est plus que de 40% si on exclut les polices relevant de l'activité maison individuelle. Cette activité représente 24% des sinistres du portefeuille pour 20% des primes. Cela s'explique par le fait que la majorité de la population mexicaine vit près des côtes et est donc exposée au risque ouragan. Par ailleurs, les maisons individuelles sont généralement moins résistantes que les immeubles et les bâtiments commerciaux : elles sont donc plus vulnérables au risque d'ouragan.

L'activité hôtel est de loin la plus risquée, avec un taux de destruction de 0,46‰. Son taux de prime moyen est de 4,34‰, soit un *Loss Ratio* de 107%.

Avec un taux de destruction de près de 0,1‰, les activités commerce de détail et maison individuelle sont environ deux fois plus risquées que le reste du portefeuille. Alors qu'elles sont deux fois plus risquées, ces activités sont cependant tarifées au même prix que le reste du portefeuille : environ 0,8‰. De ce fait, elles présentent les plus mauvais *Loss Ratio* du portefeuille : environ 118%.

Résultats par police

Pour identifier les lignes de business et les activités qui concentrent les polices les moins rentables du portefeuille, les résultats par police ont été représentés sous la forme d'un nuage de points.

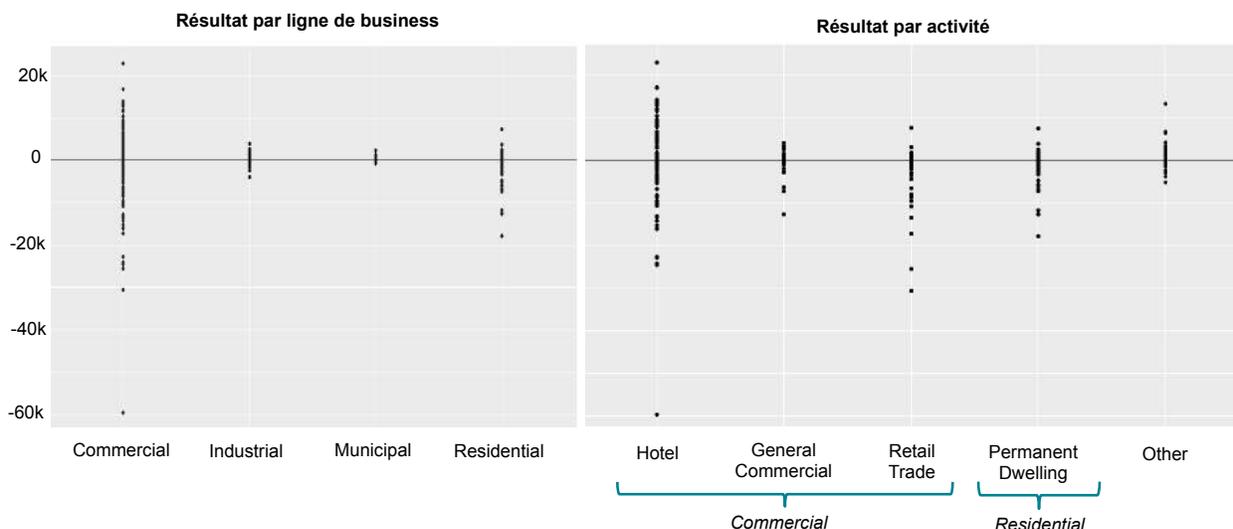


Figure 95 : résultat par branche d'assurance et activité (en EUR)

119 polices ont un résultat inférieur à -10 000 euros. Elles relèvent des lignes de business commercial et résidentiel. La branche commercial est celle dont les résultats sont les plus volatiles.

La branche commercial a 80 polices avec un résultat inférieur à -10 000 euros, représentant un déficit global de 5,8 millions. Parmi ces polices, 40 relèvent de l'activité hôtel, 28 de l'activité commerce de détail (*retail trade*) et 7 de l'activité commercial général (*general commercial*). Par ailleurs, 77% des ces polices sont localisées à Quintana Roo et 15% en Basse-Californie du Sud.

La branche résidentiel a 39 polices avec un résultat inférieur à -10 000 euros, représentant un déficit global de 1,3 million d'euros. Toutes ces polices relèvent de l'activité maison individuelle (*permanent dwelling*). 46% des polices sont localisées à Quintana Roo et 44% en Basse-Californie du Sud.

Les polices ayant les plus mauvais résultats sont donc concentrées sur quatre activités et sur deux zones CRESTA côtières. L'activité permet de mieux identifier les polices les moins rentables car elle donne des informations supplémentaires sur le risque : localisation, type de bâtiment etc.

b) Analyse par matériau de construction

Les sites du portefeuille ont des matériaux de construction variés : *concrete* (béton), *reinforced concrete* (béton armé), *reinforced masonry* (maçonnerie armée), *steel* (acier) et *steel frame superior* (acier supérieur).

Les matériaux de construction utilisés au Mexique sont hétérogènes (*AIR Worldwide, 2014 [18]*). Les maisons individuelles sont principalement construites en béton et en maçonnerie. Les bâtiments commerciaux ont des matériaux de construction très variables : on trouve par exemple des bâtiments construits en maçonnerie avec peu d'étages et des bâtiments en acier.

Poids de chaque matériau de construction

Le graphe ci-dessous présente le poids de chaque matériau de construction en terme de nombre de sites, de NIV, d'AAL et de primes.

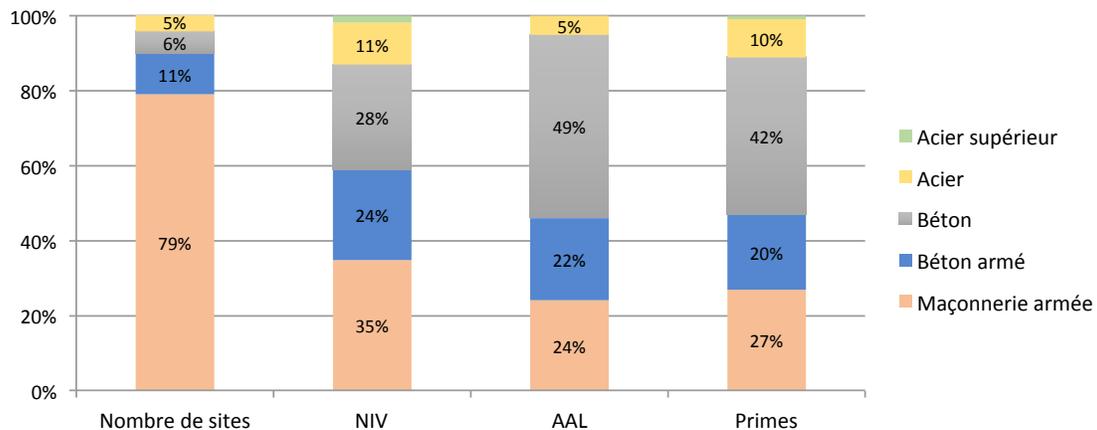


Figure 96 : poids de chaque matériau de construction

Les matériaux de construction les moins rentables sont le béton et le béton armé : leur part dans les sinistres du portefeuille est supérieure à leur part dans les primes. Ce sont également les matériaux les plus risqués : ils ont respectivement un taux de destruction moyen de 0,198‰ et de 0,1‰. Ces matériaux sont sous-tarifés par rapport à leur niveau de risque : ils ont un taux de prime moyen de 0,178‰ et 0,095‰ soit un *Loss Ratio* de 111% et 106%.

Le matériau le plus fréquemment utilisé est la maçonnerie armée : 79% des sites du portefeuille en sont composés. Ce type de structure est moins risqué que le béton et le béton armé (taux de destruction moyen est de 0,078‰) et plus rentable (taux de prime est de 0,088‰, soit un *Loss Ratio* de 88%).

L'acier et l'acier supérieur sont les matériaux de construction les moins fréquents et les plus rentables du portefeuille : leur *Loss Ratio* moyen est respectivement de 50% et 27%. Ce sont également les matériaux les moins risqués : leur taux de destruction moyen est respectivement de 0,053‰ et 0,012‰.

Résultats par matériau de construction

Pour identifier les matériaux de construction qui concentrent les polices les moins rentables du portefeuille, les résultats par polices ont été représentés sous la forme d'un nuage de points.

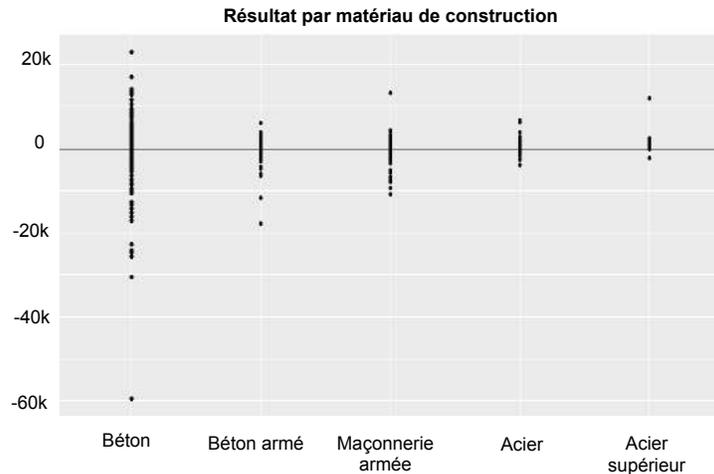


Figure 97 : répartition du résultat par matériau de construction

119 polices ont un résultat inférieur à -10 000 euros. Ces polices sont principalement construites en béton, en béton armé et en maçonnerie armée.

Parmi ces polices, 56 sont construites en béton. Il s'agit principalement d'hôtels (63% des polices). On trouve également des activités de commerce de détail (13% des polices). 79% des polices sont situées à Quintana Roo et 11% en Basse-Californie du Sud.

Les polices construites en béton armé comptent 18 polices avec un résultat inférieur à -10 000 euros. La majorité de ces polices (90%) relève de l'activité maison individuelle et sont principalement situées à Quintana Roo (83% des polices).

46 polices constituées de maçonnerie armée ont un résultat inférieur à -10 000 euros. Ces polices sont principalement des maisons individuelles (52%) et des activités de commerce de détail (39%). Elles sont situées à Quintana Roo (46%) et en Basse-Californie du Sud (52%).

Annexe 4 :

Sensibilité des résultats aux paramètres de croisement et de reproduction

L'objectif de cette partie est de présenter les résultats d'études de sensibilité aux paramètres de reproduction et de mutation de l'algorithme génétique.

Pour rappel, les paramètres suivants ont été utilisés :

- **Reproduction :**
 - Sélection des parents : par tournoi
 - Croisement : uniforme avec un taux de 50%
- **Mutation :**
 - Type de mutation : inversion de bit
 - Probabilité de mutation : 1%

Dans un premier temps, la sensibilité à la sélection des parents et aux croisements sera analysée. **Dans un second temps**, on étudiera la sensibilité des résultats aux paramètres de mutation.

a) Sensibilité à la sélection des parents et aux croisements

Les chromosomes parents ont été sélectionnés par tournoi avec un paramètre de 2. Une fois les parents sélectionnés, les chromosomes des enfants sont obtenus en effectuant un croisement uniforme avec une probabilité de 50%.

Cette partie a pour objet d'effectuer des études de sensibilité sur les méthodes de sélection et de croisement.

Sélection des parents

Les nuages de points ci-dessous présentent les résultats obtenus avec une sélection de type roulette, rang et tournoi.

Pour que les résultats soient comparables, les différentes options ont été initialisées avec la même population aléatoire.

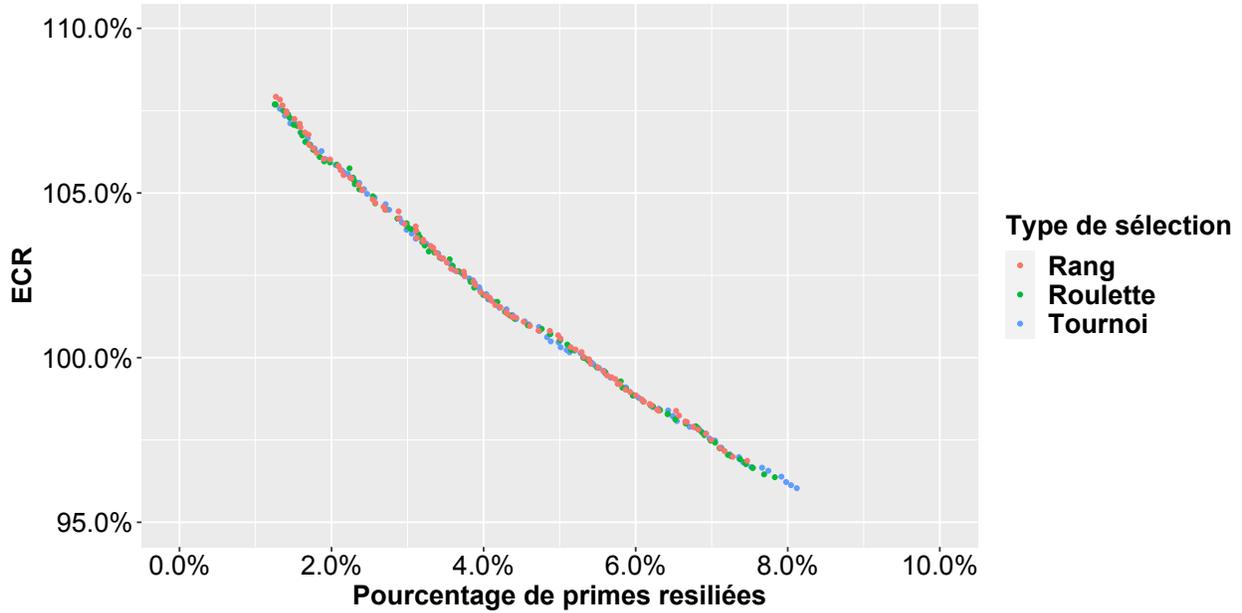


Figure 98 : sensibilité des résultats au type de sélection

Les résultats obtenus sont stables et sont peu sensibles au type de sélection choisi. La sélection par tournoi a cependant permis d’explorer davantage d’espace entre 7,5% et 8%.

Type de croisement

Le nuage de points ci-dessous présente les résultats obtenus avec des croisements en un point, des croisements uniformes et sans croisement.

Pour que les résultats soient comparables, les différentes options ont été initialisées avec la même population aléatoire.

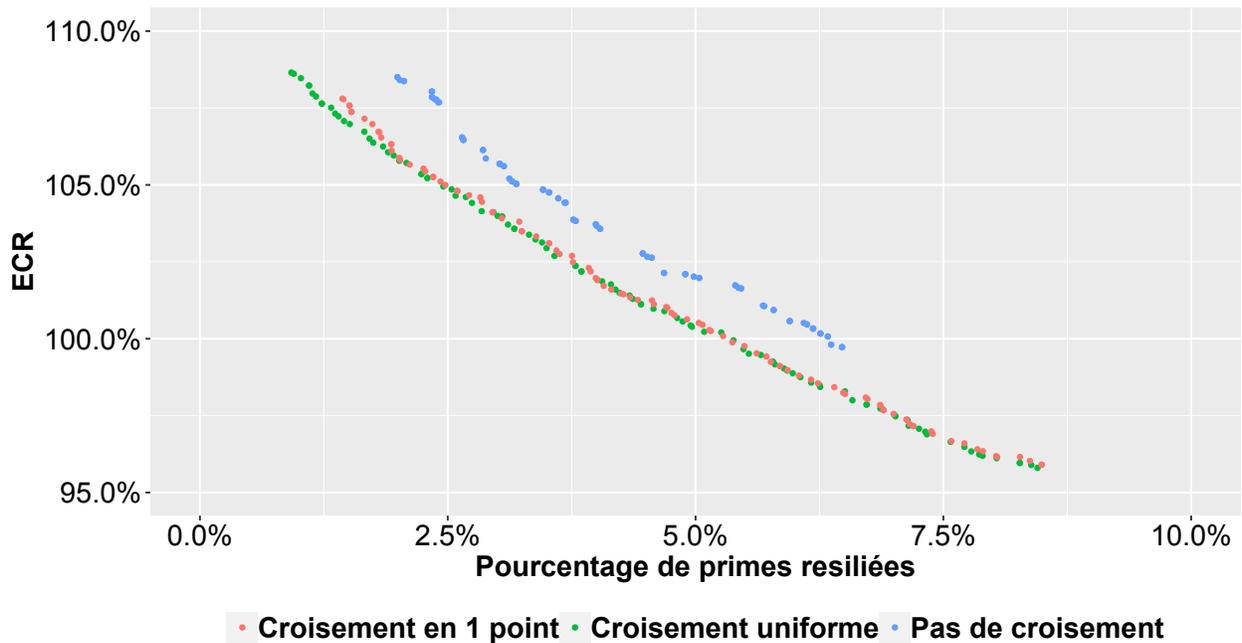


Figure 99 : sensibilité des résultats au type de croisement

On constate que les chromosomes obtenus par croisement uniforme ont permis d'explorer davantage d'espace que ceux obtenus par croisement en un point : le plus petit chromosome obtenu par croisement uniforme correspond à 1% de primes résiliées, contre 1,4% pour les croisements en un point.

Par ailleurs, jusqu'à 2% de primes résiliées, les résultats obtenus par croisement uniforme sont de meilleure qualité que ceux obtenus par croisement en un point. Au-delà de 2% de primes résiliées, les résultats sont très proches.

Les chromosomes obtenus sans croisement donnent des résultats beaucoup moins bons que ceux obtenus avec croisement uniforme et en un point : en moyenne, pour un même pourcentage de primes résiliées, les résultats obtenus sans croisement ont 1,6 point d'ECR de plus que les chromosomes obtenus avec croisement.

Par ailleurs, la population obtenue sans croisement n'a exploré qu'une faible partie de l'espace : aucun chromosome n'a de primes résiliées inférieures à 2%.

Ce constat montre l'importance des opérations de croisement dans l'évolution de la population de l'algorithme génétique.

Ces résultats sont en ligne avec la littérature : dans une étude publiée en 2010, Picek et Golub ([14]) mettent en lumière le fait que parmi les algorithmes génétiques encodés en binaire, les moins performants sont ceux qui n'effectuent pas d'opération de croisement, suivis de ceux qui effectuent des croisements en un point.

Les algorithmes génétiques appartenant à la classe des algorithmes heuristiques, les paramètres optimaux sont obtenus par expérience ou en effectuant des études de sensibilité.

Taux de croisement uniforme

Dans notre étude, un taux de croisement uniforme a été utilisé. Ce type de croisement est caractérisé par un taux α : il s'agit de la probabilité que l'enfant 1 hérite de chacun des gènes du parent 1. La probabilité que l'enfant 1 hérite des gènes du parent 2 est de $1 - \alpha$. Inversement, la probabilité que l'enfant 2 hérite des gènes du parent 1 est de $1 - \alpha$ et des gènes du parent 2 est de α .

Le nuage de points ci-dessous présente les résultats obtenus pour des taux de croisement uniforme de 10%, 25% et 50%. Aucune sensibilité n'est faite pour un taux supérieur à 50% les résultats sont symétriques : un taux de croisement de α et de $1 - \alpha$ donnent les mêmes résultats (l'enfant 2 aura les gènes de l'enfant 1 et inversement).

Pour que les résultats soient comparables, les différentes options ont été initialisées avec la même population aléatoire.

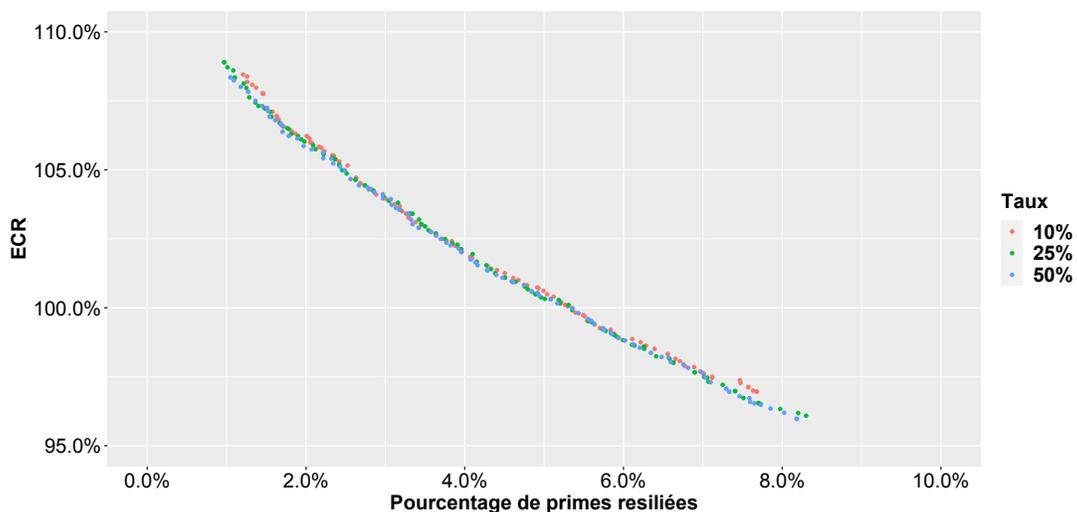


Figure 100 : sensibilité des résultats au taux de croisement uniforme

On constate que les résultats obtenus avec un taux de croisement uniforme de 50% et 25% sont légèrement meilleurs que ceux obtenus avec un taux de 10%.

Les résultats sont cependant stables et on en déduit que notre problème est peu sensible au taux de croisement uniforme choisi.

Conclusion

Les croisements sont un élément clé dans l'évolution de la population : en faisant se reproduire les meilleurs individus, les chromosomes se perfectionnent au fil des générations.

Les meilleurs résultats sont obtenus par croisement uniforme mais ils sont peu sensibles au type de sélection des parents et au taux de croisement utilisé.

b) Sensibilité aux mutations

Dans notre étude, les mutations ont été effectuées par inversion de bit. Comme suggéré dans la littérature, un pourcentage de $\frac{1}{\text{nombre de gènes}}$ a été choisi. Les chromosomes étant composés de 100 gènes, un pourcentage de 1% a donc été utilisé.

Le problème étant encodé en binaire, les mutations par inversion de bit sont l'option privilégiée et aucune étude de sensibilité ne sera réalisée sur le type de mutation. Cette partie a pour objectif d'étudier la sensibilité des résultats au taux de mutation utilisé.

Taux de mutation

Le nuage de points ci-dessous présente les résultats obtenus avec des taux de mutation de 1%, 5%, 50% et 100%. Le type de mutation est une inversion de bit et chaque option a été initialisée aléatoirement avec la même population.

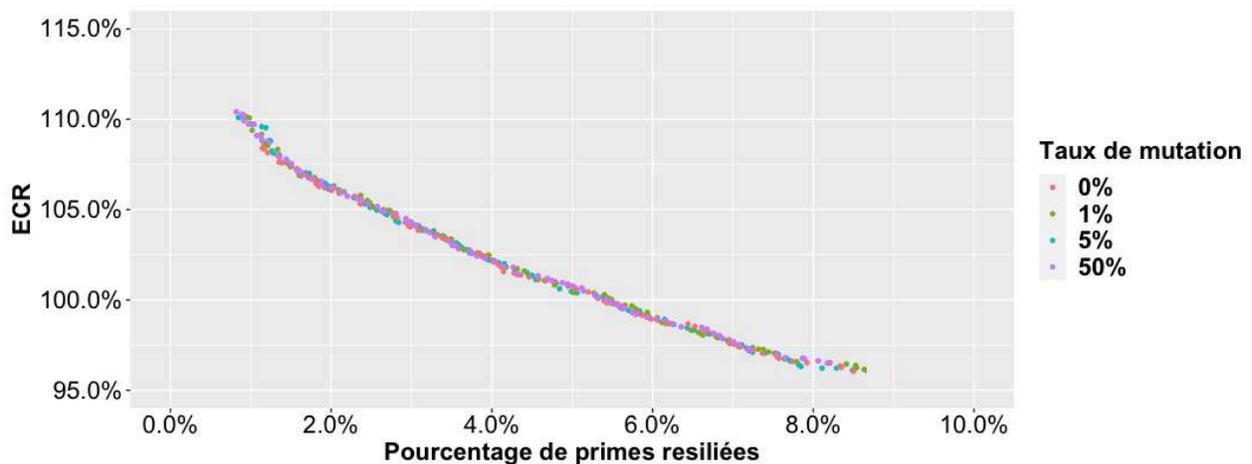


Figure 101 : sensibilité des résultats au taux de mutation

Pour des taux de mutation entre 0% et 50%, on obtient des résultats stables. Notre problème est donc peu sensible aux mutations. On remarque cependant que la population avec un taux de mutation de 0% explore moins l'espace que celle avec des taux de 1%, 5% et 50% : aucun chromosome n'a un pourcentage de primes résiliées en dessous de 1,2% et au-dessus de 7,8%. Les mutations permettent donc de mieux explorer l'espace mais ne permettent pas d'améliorer de manière matérielle les résultats.

Conclusion

Les résultats sont peu sensibles au taux de mutation : on obtient des résultats stables pour un taux de mutation entre 0% et 50%. La sensibilité de notre problème au taux de mutation est donc limitée, ce qui signifie que la population s'améliore principalement grâce aux croisements. Ce constat est en ligne avec les conclusions de Schlierkamp-Voosen (*Schlierkamp-Voosen, 1993, [19]*) : pour les populations de grande taille, les croisements sont plus efficaces que les mutations pour améliorer la population.