

**Mémoire présenté le 12 juillet 2021
en vue de l'obtention du titre d'Actuaire de l'Institut des Actuaire
par
Mr MARTIN DECULTIEUX**

suite à son stage effectué dans le cadre de la **filière Actuariat ESSEC-ISUP**

"Réduction du nombre de méthodes de provisionnement PSAP en assurances collectives à l'international"

Confidentialité NON OUI (durée : 1 an 2 ans)

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus Membres présents du jury de l'Institut des Actuaire :

Membres présents du jury de la filière :

-

*Entreprise : AXA France
Directeur du mémoire en entreprise :
Maria VEGA PAEZ*

Membres présents du jury de l'Institut des Actuaire

-

-

-

Invité :

Signature :

Signature du candidat :

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

Signature du responsable entreprise :



MÉMOIRE D'ACTUARIAT

MASTER IN MANAGEMENT - FILIÈRE RISQUE ET ACTUARIAT

Réduction du nombre de méthodes de provisionnement PSAP en assurances collectives à l'international

Auteur :
Martin DECULTIEUX

Tutrice entreprise :
Maria VEGA PAEZ

Tutrice école :
Marie KRATZ



Juin, 2021

RÉSUMÉ

Mots clés: PSAP, CANN, assurances collectives internationale, Santé, Prévoyance

Les portefeuilles d'assurance et de réassurance pour des contrats à l'international en Santé et Prévoyance présentent bien souvent une hétérogénéité due aux différents systèmes de protection sociale locaux. Les différentes pratiques de chaque pays impliquent un nombre de méthodes et d'outils de provisionnement importants à suivre pour la société d'assurance, pouvant devenir problématique. Nous nous intéresserons particulièrement à la multiplicité des outils de provisionnement PSAP dans ce mémoire, en nous demandant s'il est possible de réduire le nombre de méthodes et d'outils de calcul de cette provision sur des garanties de Santé et de Prévoyance à l'international. Pour ce faire, deux business majeurs d'un portefeuille d'AXA France seront étudiés. Ils présentent tous deux des garanties et des caractéristiques très différentes, reflétant la complexité de la problématique. Nous présenterons tout d'abord en détail les contrats étudiés, en mettant en avant leurs spécificités liées aux pratiques locales d'assurance. Nous exposerons alors les outils actuellement utilisés pour calculer les provisions de ces contrats chez AXA France, en synthétisant leurs points forts et leurs points faibles, qui poussent l'assureur à revoir ses méthodes. Dans un second temps, nous nous attacherons à étudier les méthodes usuelles et robustes de provisionnement PSAP telles que le Chain Ladder de Mack ou le Bootstrap. Nous montrerons que les hypothèses d'homogénéité des triangles de ces méthodes, ainsi que les différentes cadences de paiement observées poussent à décomposer les outils et multiplier les méthodes. Enfin, sera présentée la méthode CANN, qui permet de lever les hypothèses fortes d'homogénéité des triangles, formulées par les méthodes classiques. Cette méthode utilisant un réseau de neurones d'une simple couche cachée, calcule les coefficients de passage du triangle en prenant en compte les différentes caractéristiques des paiements. Nous verrons qu'une telle méthode nécessite des données tête par tête massives, non présentes dans notre étude. Notre travail s'orientera alors autour de la présentation de cette méthode, avec un détail sur l'optimisation du réseau de neurones associé. La fin du mémoire est consacrée aux différentes limites de l'étude et de la méthode CANN avec les axes d'améliorations identifiés. Dans la perspective de la mise en place d'une telle méthode chez AXA France, les recommandations d'audit associées seront formulées.

ABSTRACT

Key words: IBNR, CANN, International group insurance contracts, Health, Life

Insurance and reinsurance portfolios for international Health and Life contracts are often heterogeneous due to the different local protection systems. The different local practices of each country imply an important number of provisioning methods and tools to follow up for the insurance company, which can become problematic. We will be particularly interested in the multiplicity of IBNR provisioning tools in this study, by asking ourselves whether it is possible to reduce the number of methods and tools for calculating this provision on Health and Life benefits. To do so, two major businesses of AXA France portfolio will be studied. They both present very different benefits and characteristics, reflecting the complexity of the problem. We will first present in detail the contracts studied, highlighting their specificities linked to local insurance practices. We will then introduce the tools currently used to calculate the reserves for these contracts at AXA France, summarizing their strengths and weaknesses, which are pushing the insurer to review its methods. In a second step, we will study the usual and robust methods of IBNR provisioning such as Mack's Chain Ladder or Bootstrap. We will show that the assumptions of homogeneity of the triangles of these methods, as well as the different payment rates observed, push to decompose the tools and multiply the methods. Finally, the CANN method will be presented, which allows to lift the strong hypotheses of homogeneity of the triangles, formulated by the classical methods. This method uses a neural network with a simple hidden layer, calculates the payment ratios of the triangle by taking into account the various characteristics of the payments. We will see that such a method requires massive head by head data, not present in our study. Our work will then focus on the presentation of this method, with a focus on the optimization of the associated neural network. The end of the paper is devoted to the various limits of the study and of the CANN method with the associated improvements proposed. With a view of future implementation of such a method at AXA France, the associated audit recommendations will be formulated.

REMERCIEMENTS

Je remercie tout d'abord AXA France pour la confiance accordée tout au long de mon apprentissage de 24 mois durant lequel j'ai réalisé ce mémoire. En particulier, ma manager Katia Girere pour toutes ses qualités humaines et techniques qui m'ont permis de grandir. Je remercie également Maria Vega Paez, ma tutrice de mémoire, qui m'a toujours accordé son temps et ses conseils ainsi que Maryne Sans Roudaire pour son accompagnement de qualité et sa disponibilité. Aussi, j'ai une pensée amicale pour mes collègues de la DAAP internationale qui m'ont grandement encouragé et avec qui j'ai toujours eu plaisir à travailler.

Je n'oublierai pas l'aide importante reçue de la part d'autres départements d'AXA France, qui m'ont permis de réaliser cette étude. Merci à Tatiana, Cédric, et Abderrahman de ALHIS. Merci Guillaume pour ton aide et tes conseils sur Python.

Je tiens également à remercier sincèrement Marie Kratz, responsable de la filière ESSEC-ISUP, pour m'avoir fait découvrir et aimer l'actuariat, pour son suivi régulier sur ce mémoire ainsi que sur les deux années d'apprentissage. Je remercie plus largement, l'ESSEC Business School ainsi que l'ISUP pour m'avoir délivré des cours de qualité durant près de trois années.

Enfin, d'un point de vue personnel, je suis reconnaissant envers ma famille et mes amis pour leur soutien de tout temps. En particulier, mes parents et ma sœur qui ont toujours cru en moi.

TABLE DES MATIÈRES

| | |
|--|-----------|
| RESUME | 2 |
| ABSTRACT..... | 3 |
| REMERCIEMENTS..... | 4 |
| PREAMBULE | 7 |
| ACRONYMES ET ABBREVIATIONS..... | 9 |
| INTRODUCTION | 10 |
| 1. CONTEXTUALISATION ET STRATEGIE..... | 12 |
| 1.1 CONTEXTUALISATION DES ENTITES..... | 12 |
| 1.1.1 AXA ET AXA FRANCE..... | 12 |
| 1.1.2 DIRECTION ACTUARIAT ASSURANCE DES PERSONNES (DAAP)..... | 13 |
| 1.1.3 AXA LIFE & HEALTH INTERNATIONAL SOLUTIONS (ALHIS)..... | 14 |
| 1.2 CONTEXTUALISATION REGLEMENTAIRE..... | 17 |
| 1.2.1 CODE DES ASSURANCES ET AUTORITE DES NORMES COMPTABLES..... | 17 |
| 1.2.2 SOLVABILITE II ET BEST ESTIMATE..... | 18 |
| 1.3 PRESENTATION DES CONTRATS ETUDIES..... | 19 |
| 1.3.1 PRESENTATION DES CONTRATS OSCAR..... | 19 |
| 1.3.2 PRESENTATION DES CONTRATS AUSTRALIENS..... | 26 |
| 2. LES PROPOSITIONS DE PROVISIONNEMENT | 31 |
| 2.1 L'EXISTANT | 31 |
| 2.1.1 LE PROVISIONNEMENT EXISTANT DE LA DAAP | 31 |
| 2.1.2 LE PROCESS DE PROVISIONNEMENT LORS D'UN ARRETE | 34 |
| 2.1.3 LES OUTILS DE PROVISIONNEMENT DU MARCHE | 35 |
| 2.1.4 LA DONNEE DISPONIBLE POUR L'ETUDE | 40 |
| 2.2 LES PROPOSITIONS CLASSIQUES DE PROVISIONNEMENT | 46 |
| 2.2.1 LE CHAIN LADDER CLASSIQUE..... | 47 |
| 2.2.2 LE CHAIN LADDER DE MACK (CHAIN LADDER STOCHASTIQUE)..... | 60 |
| 2.2.3 LA METHODE BOOTSTRAP..... | 66 |
| 2.2.4 CONCLUSION SUR LES METHODES CLASSIQUES | 71 |

| | |
|---|-------------------|
| 2.3 LA PROPOSITION MIXTE : LA METHODE CANN..... | 72 |
| 2.3.1 PRESENTATION DE LA METHODE..... | 72 |
| 2.3.2 SPECTRE D'ETUDE ET QUALITE DE DONNEES | 74 |
| 2.3.3 MISE EN PLACE DU RESEAU DE NEURONES | 75 |
| 2.3.4 OPTIMISATION DU RESEAU DE NEURONES | 76 |
| 2.3.5 RESULTATS DE LA METHODE CANN | 80 |
| 2.4 COMPARAISON DES RESULTATS..... | 82 |
| 2.4.1 COMPARAISON DES METHODES SUR LE SCOPE 2019 - BACKTESTING..... | 82 |
| 2.4.1 COMPARAISON DES METHODES SUR LE SCOPE 2020 - PREVISIONS..... | 84 |
| | |
| <u>3. PERSPECTIVES.....</u> | <u>89</u> |
| | |
| 3.1 LIMITES DE L'ETUDE ET PISTES D'AMELIORATION..... | 89 |
| 3.1.1 LIMITES DE L'ETUDE..... | 89 |
| 3.1.2 AMELIORATIONS POSSIBLES DE L'ETUDE..... | 90 |
| 3.2 ÉVALUATION DU MODELE CANN ET PISTES D'AUDIT | 92 |
| | |
| <u>CONCLUSION</u> | <u>95</u> |
| | |
| <u>BIBLIOGRAPHIE.....</u> | <u>97</u> |
| | |
| <u>ANNEXES</u> | <u>99</u> |
| | |
| <u>TABLE DES FIGURES.....</u> | <u>131</u> |

PRÉAMBULE

Au cours de mon Master 2 ESSEC-ISUP et afin de devenir actuaire j'ai effectué un apprentissage de 24 mois chez AXA France durant lequel 6 mois ont été consacrés à la réalisation de ce mémoire.

J'ai effectué cet apprentissage sur le site de Nanterre (92) dans la Direction Financière Stratégique et Juridique (DFSJ), plus précisément dans les équipes de la Direction Actuariat Assurances de Personnes (DAAP) sur le marché de la Santé & Collectives, branche internationale.

Les équipes de la DAAP sont garantes des provisions du marché pour lequel elles travaillent. En contact avec différents départements (Direction Technique, Souscription, Marché, Comptables), la DAAP doit construire sa vision sur les comptes en termes de provisionnement pour chaque arrêté comptable. Ces équipes sont donc constituées d'actuaires et d'autres spécialistes des assurances apportant une connaissance marché essentielle lors du provisionnement des contrats.

Arrivé en avril 2019, j'ai consacré ma première année au développement d'un nouvel outil de visualisation et de reporting des comptes du marché AXA Life & Health International Solutions (ALHIS). Différentes difficultés ont été rencontrées notamment en terme de qualité de données. Des connaissances sur Power Query (Add In Excel) et Power BI, acquises lors de précédents stages m'ont permis de standardiser les données et de proposer des visualisations répondant aux besoins exprimés par la DAAP. En outre, cette mission m'a permis de me familiariser avec les principales sources de données, de développer des compétences métier et d'avoir une vue globale de ce marché.

En début d'année 2020, j'ai commencé à travailler en binôme avec ma tutrice Maria Vega Paez (Actuaire à la DAAP) sur des sujets plus actuariels tels que la rédaction de notes techniques sur certains portefeuilles de réassurance majeurs. L'objectif de ces études est d'avoir une connaissance approfondie de certains groupes de contrats afin de comprendre les propositions de provisionnement faites par notre marché et in fine émettre notre propre avis sur ces provisions. Les contextes locaux des contrats sont parfois difficiles à comprendre de bout en bout. L'expertise actuarielle de Maria m'a régulièrement permis de surmonter ces difficultés. Différentes aides, notamment d'autres département d'AXA France plus proches de la souscription, nous ont permis de réussir cette mission.

Ce travail m'a permis d'avoir une approche « business » essentielle dans la compréhension du provisionnement. Il m'a permis de prendre du recul sur certains résultats et d'avoir des axes d'analyse supplémentaires très enrichissants pour la suite de mes travaux.

Cette mission m'a également permis d'élucider le sujet de mon mémoire, notamment lorsque j'ai constaté la diversité très importante des méthodes de provisionnement des cédantes en fonction des pays et des garanties.

ACRONYMES ET ABRÉVIATIONS

| | |
|----------------|---|
| ABS | Australian Bureau of Statistics |
| ACPR | Autorité de Contrôle Prudentiel et de Résolution |
| ADAM | Adaptative Moment Optimization |
| AGH | AXA Global Healthcare |
| ALHIS | AXA Life & Health International Solutions (Marché suivi par la DAAP – pôle international, périmètre du mémoire) |
| API | Application Programming Interface |
| APRA | Australian Prudential Regulation Authority (régulateur Australien) |
| ATO | Australian Taxation Office |
| BEL | Best Estimate Liabilities |
| CANN | Combined Actuarial Neural Network |
| CLP | Credit Life Protection (business suivi par la DAAP – pôle international, périmètre exclu du mémoire) |
| CoR | Combined Ratio (Ratio Combiné) |
| DAAP | Direction Actuariat Assurance des Personnes |
| DFSJ | Direction Finance, Stratégie, et Juridique |
| GAAP | Generally accepted accounting principles |
| GST | Good and Services Tax |
| IA | Intelligence Artificielle |
| IARD | Incendie Accidents et Risques Divers |
| IBNR | Incurred but not reported, équivalent de PSAP |
| ICD | Invalidité Courte Durée |
| IFRS | International Financial Reporting Standards |
| ILD | Invalidité Longue Durée |
| IP | Income Protection |
| IPMI | International Private Medical Insurance (une des 4 business units de ALHIS) |
| LEB | Local Solutions (une des 4 business units de ALHIS) |
| L/R | Loss Ratio |
| MAE | Mean Absolute Error (l'erreur absolue moyenne) |
| MSE | Mean Squarre Error (l'erreur quadratique moyenne) |
| MVM | Market Value Margin |
| NRWT | Non-Resident Withholding Tax |
| PSAP | Provision pour sinistre à payer, équivalent de IBNR |
| PYS | Protecting Your Super |
| RGPD | Règlement Général sur la Protection des Données |
| RMSPROP | Root Mean Square Propagation |
| SGD | Stochastic Gradient Descent |
| TPD | Total and Permanent Disability |
| UE | Union Européenne |

INTRODUCTION

AXA Life and Health International Solutions (ALHIS) est une entité d'AXA France qui propose des solutions de prévoyance et de santé à l'international en tant qu'assureur direct ou réassureur dans 61 pays. La partie réassurance du portefeuille est largement majoritaire depuis plusieurs années. Un des enjeux majeurs est de bien comprendre comment les cédantes provisionnent leurs contrats afin d'éventuellement réviser cette vision dans les comptes d'AXA France ou d'améliorer le Best Estimate (Solvency II) pour le groupe.

En réassurance, par manque de connaissance de certaines cédantes, lors de l'arrêté comptable, la Direction Actuariat Assurances de Personnes (DAAP) suit la vision Marché (ALHIS) sur le provisionnement des contrats majeurs. Pour le reste des contrats de réassurance, un outil généraliste d'estimation des provisions est utilisé à la DAAP mais reste parfois approximatif.

Chaque pays où AXA France est présent en tant que réassureur possède des particularités et implique une multiplication des méthodes provisionnements si l'on veut suivre très précisément chaque contrat, chose qui est en partie faite par le Marché. De manière caricaturale, il existe un outil « spécialisé » pour chaque type de provision (Provision Mathématique (PM), Provision pour Sinistre A Payer (PSAP), Provision pour Risque Croissant (PRC), etc.) qui est différent pour chaque pays (parfois pour chaque contrat) et qui diffère également en fonction de la garantie du contrat. Par exemple, la PSAP (aussi appelée Incurred But Not Reported en anglais - IBNR) n'est pas calculée de la même manière sur les contrats australiens en garantie décès, que sur les contrats américains en garantie Santé.

L'enjeu est donc de proposer une vision de provisionnement DAAP pour les contrats de réassurance majeurs à l'international sans pour autant multiplier les méthodes de provisionnement à suivre pour les équipes. Le portefeuille étant composé de garanties santé et prévoyance, le calcul des PSAP est crucial. On se demande alors s'il est possible de réduire le nombre de méthodes pour calculer les PSAP tout en conservant des performances satisfaisantes (c'est à dire un provisionnement précis), contrat par contrat.

Afin de présenter le contexte de l'étude, une revue détaillée des contrats sera déroulée en première partie, s'appuyant notamment sur les travaux des Notes Actuarielles évoquées dans le Préambule.

Nous explorerons ensuite certaines méthodes classiques de calcul des PSAP permettant d'avoir des repères quantitatifs, en présentant les raisons pour lesquelles ces méthodes poussent à multiplier le nombre d'outils de provisionnement.

Enfin, la méthode Combined Actuarial Neural Network (CANN), prenant en compte des caractéristiques supplémentaires sur les portefeuilles par rapport aux méthodes classiques, sera présentée dans le but de proposer un outil de provisionnement PSAP universel. En revanche,

les données disponibles à la DAAP en date de ce mémoire ne permettent pas d'espérer un provisionnement meilleur que les méthodes classiques. Cette dernière étude devrait être néanmoins utile pour les entités d'AXA France ayant accès aux données massives, ce qui sera peut-être le cas prochainement à la DAAP.

Note: Tous les montants issus de documents internes présentés par la suite ont reçu un coefficient multiplicateur afin de conserver la notion de confidentialité tout en présentant des résultats cohérents.

1. CONTEXTUALISATION ET STRATÉGIE

1.1 CONTEXTUALISATION DES ENTITÉS

1.1.1 AXA et AXA France

Depuis sa création en 1985, la marque AXA se développe à travers le monde et se donne pour mission de donner une vie meilleure à ses 108 millions de clients. Le groupe AXA possède un bilan robuste en 2020 avec 96.7 milliards d'euros de chiffre d'affaires dont 4.3 milliards d'euros de résultat net avec un ratio de solvabilité de 200% étant assis sur 71.6 milliards d'euros de capitaux propres.

AXA France est une société du Groupe AXA contribuant environ à un quart du chiffre d'affaires Groupe (25.0 Md€ en 2020). Cette société s'inscrit dans la stratégie du groupe en développant son implantation à l'international.

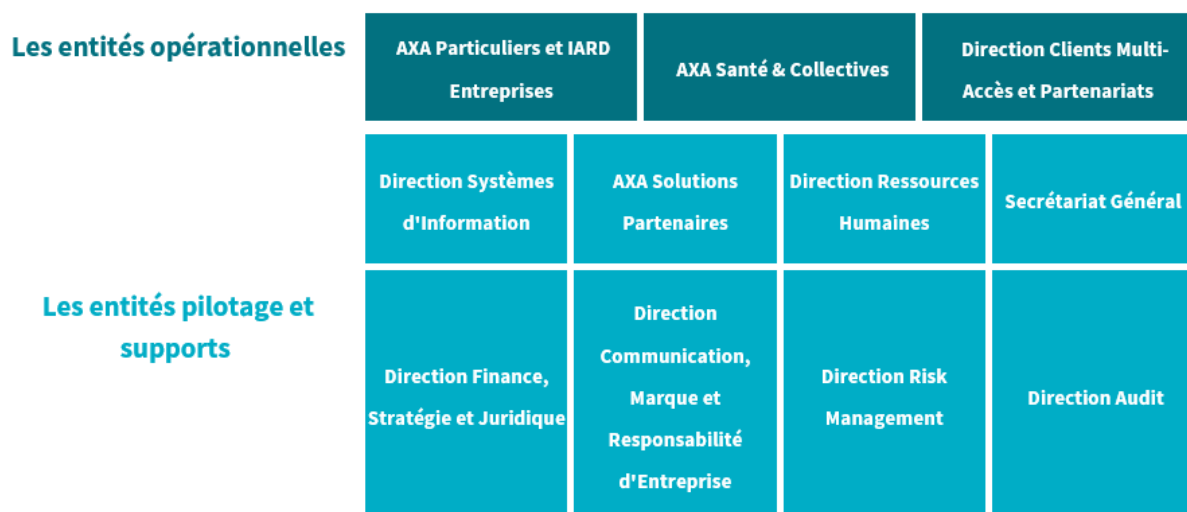


Figure 1 - Les différentes entités chez AXA France (AXA France, One AXA)

La structure d'AXA France se décompose en entités opérationnelles (au nombre de 3), et en entités de pilotage et supports (au nombre de 8). Le mémoire sera développé dans l'entité pilotage et support « Direction Finance, Stratégie et Juridique » (DFSJ), plus particulièrement au cœur de la Direction Actuariat Assurance des Personnes (DAAP).

1.1.2 Direction Actuariat Assurance des Personnes (DAAP)

La Direction Actuariat Assurance des Personnes (DAAP) appartient à l'entité de pilotage et supports DFSJ (Direction Finance, Stratégie et Juridique). La DAAP a pour mission d'être garante du provisionnement des comptes en assurance des personnes, sur les produits de santé et de prévoyance.

Deux fonctions principales guident les activités de la DAAP :

1. La fonction comptable

- Production à chaque arrêté trimestriel une vision globale du Compte de résultat afin de déterminer la rentabilité du business : loss ratios (L/R), taux de commissions, ratio combiné (CoR) en French GAAP et IFRS
- Régularisation des estimations des exercices antérieurs (estimations N-1 vs. données finales réelles), suivi et analyse des primes, des sinistres, des commissions et détermination des montants de provision
- Suivi des nouveaux business (croissance active)

2. La fonction actuarielle

- Élaboration du rapport actuariel annuel
- Suivi des process techniques avec les équipes marché
- Organisation du comité des réserves avec le risque management d'AXA France
- Suivi et accompagnement de la seconde opinion produite par le risque management
- Contribution au modèle interne d'AXA France

Cette direction se décompose en 4 pôles, décrivant des business différents :

- Épargne individuelle
- Prévoyance individuelle
- Retraite collective
- Prévoyance collective et Santé / Santé individuelle (pôle correspondant au mémoire)

Le dernier pôle Prévoyance collectives et Santé / Santé individuelle se décompose en une équipe « France » et une équipe « internationale ».

Cette dernière équipe est elle-même structurée en deux parties avec :

- Un business Creditor Lifestyle Protection (CLP)
- Un business AXA Life & Health International Solutions (ALHIS) appartenant à l'entité opérationnelle (cf. figure) AXA Santé & Collectives

C'est dans cette dernière équipe (DAAP internationale – business ALHIS) que ce mémoire a été réalisé.

1.1.3 AXA Life & Health International Solutions (ALHIS)

ALHIS correspond à la partie internationale de l'entité opérationnelle (cf. figure) AXA Santé & Collectives. Son but est d'accélérer le développement de nouveaux marchés de protection sociale principalement via des solutions de réassurance. Le portefeuille de réassurance pèse plus de 80% (Final Year 2020), avec comme marchés principaux les États-Unis, l'Inde et l'Australie.

La Stratégie ALHIS

ALHIS est un accélérateur d'entreprises et d'innovations à travers le monde ayant pour but de stimuler la croissance dans les domaines de la santé, de la protection sociale, de la prévoyance ainsi que de la retraite. ALHIS s'étend année après année, en développant des nouveaux marchés, mettant en place des succursales (Inde par exemple, Australie bientôt) et travaille tant avec des petites entreprises que des grandes multinationales. Quelle que soit la taille du client ou la localisation, l'objectif reste le même : développer la prévoyance et la santé à travers le monde.

Description du portefeuille global ALHIS

ALHIS est un regroupement de 4 business units (environ 20% du chiffre d'affaire d'AXA Santé et Collectives en 2020) :

- International Private Medical Insurance (IPMI) se compose de AXA Global Healthcare (AGH) et d'une architecture ouverte IPMI
- Réassurance internationale locale (« Local Solutions », aussi appelé LEB) qui correspond à la réassurance reçue de cédantes domiciliées à l'étranger. Ce business unit sert également d'incubateur pour les nouveaux business
- MAXIS Global Benefit Network : la joint-venture entre AXA et MetLife
- Retraite internationale qui offre des solutions d'épargne-retraite collective sur-mesure IGP

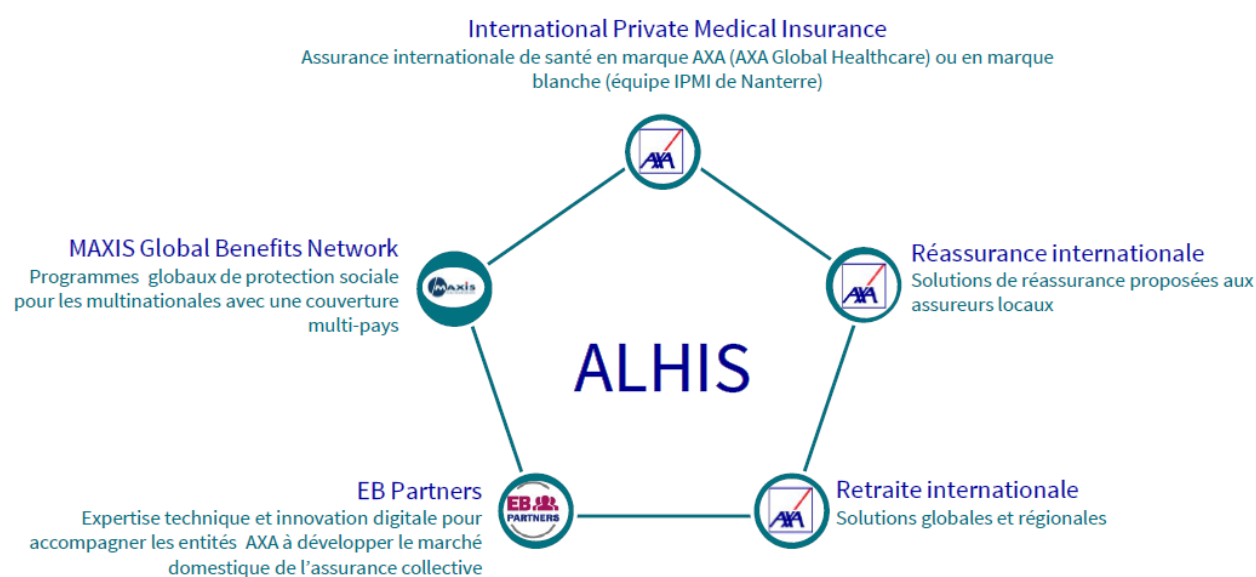


Figure 2 - Pentagone des business ALHIS

Enfin, EB Partners crée des solutions digitales pour développer l'assurance collective d'AXA à l'étranger. C'est une activité dite activité auxiliaire n'entrant pas dans le portefeuille.

Définition du périmètre de l'étude

Ce mémoire va porter sur le périmètre réassurance du business ALHIS, plus précisément sur les « Local Solutions », solutions de réassurances internationales proposées aux assureurs locaux.

Afin de pouvoir définir les contrats « Local Solutions » entrants dans l'étude, plusieurs critères ont été pris en compte :

- Importance du business en terme de chiffre d'affaires
- Volume des variations de provisions
- Croissance du business

De ce fait, le choix s'est porté sur le business Oscar, groupement de contrats de réassurance santé aux États-Unis, ainsi que sur le business Australien, groupement de contrats de prévoyance en Australie.

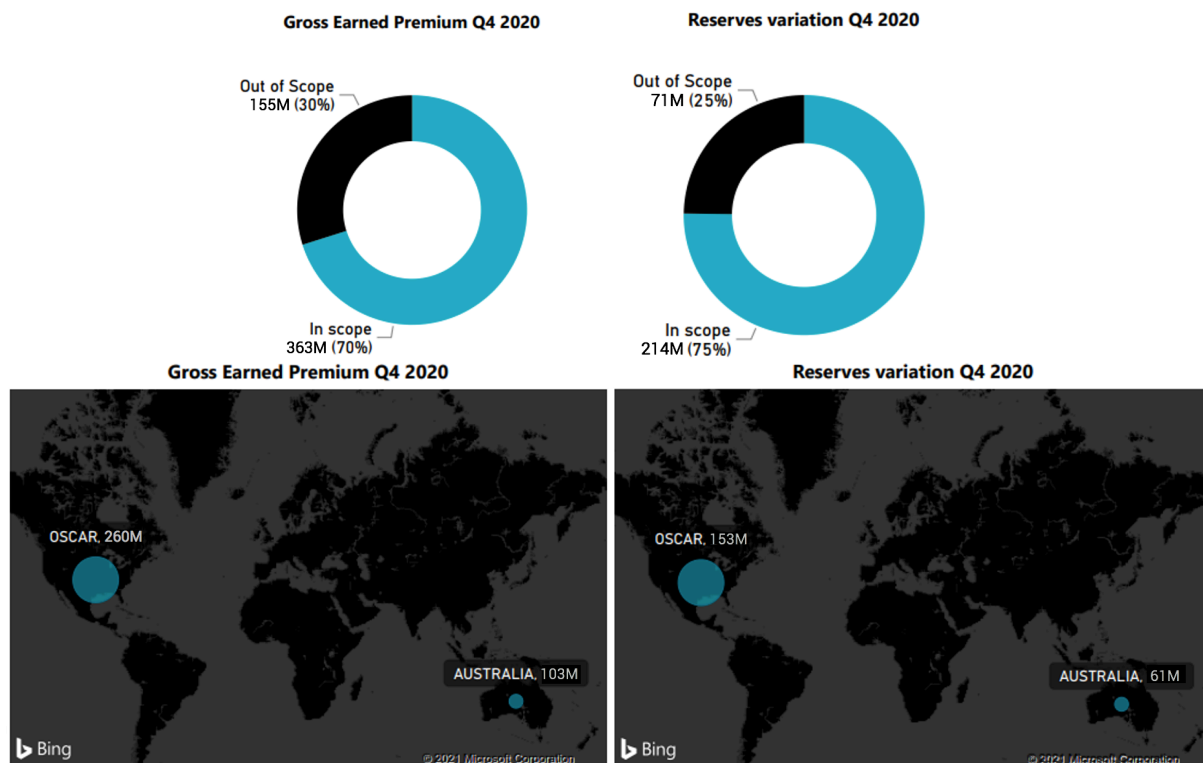


Figure 3 - Visualisation des portefeuilles étudiés

Ce choix respecte les critères de sélection mentionnés précédemment. En effet, lorsqu'on considère les contrats Oscar et les contrats Australiens, ils permettent de couvrir 70% des primes perçues en 2020 dans les contrats Local Solutions. Ils couvrent plus des trois quarts des variations de réserves constatées cette même année. Ces deux business sont par ailleurs en croissance ce qui appuie fortement leur présence dans l'étude.

1.2 CONTEXTUALISATION RÉGLEMENTAIRE

Les différentes normes et règlements applicables en France et en Union Européenne vont orienter l'étude et les méthodes utilisées pour le calcul des provisions de ce mémoire.

1.2.1 Code des Assurances et Autorité des Normes Comptables

Dans un contexte de réassurance reçue avec une cédante domiciliée à l'étranger, il est possible de rester en normes locales lors de l'intégration dans les comptes d'AXA France de l'ensemble des activités liées à ces cédantes.

Par conséquent pour la réassurance acceptée, on utilise le référentiel de la cédante sans transformer les provisions avec les règles françaises.

C'est ce qui est inscrit dans le règlement ANC 2015-11 au Chapitre II - Spécificités des opérations de réassurance acceptées, Art. 152-1 :

« En ce qui concerne les acceptations en réassurance, les entreprises enregistrent immédiatement en comptabilité tous les éléments reçus de leurs cédantes. En l'absence d'informations suffisantes, elles estiment les comptes non reçus des cédantes à la clôture de l'exercice avec pour contrepartie des comptes de régularisation qui seront soldés à l'ouverture de l'exercice suivant ou à réception des comptes des cédantes, ou elles compensent provisoirement les soldes de tous les comptes incomplets d'un même exercice par une écriture d'attente qui sera contrepassée à l'ouverture de l'exercice suivant.

En tout état de cause et quel que soit le mode de comptabilisation retenu, lorsque le réassureur connaît l'existence d'une perte, celle-ci est provisionnée pour son montant prévisible.»

Dans le business des Local Solutions, les informations reçues de la part des cédantes sont souvent insuffisantes et il est nécessaire pour AXA France d'estimer les provisions de ces contrats, malgré cette législation.

1.2.2 Solvabilité II et Best Estimate

Le Bilan sous Solvabilité II se présente sous forme de Bilan économique.

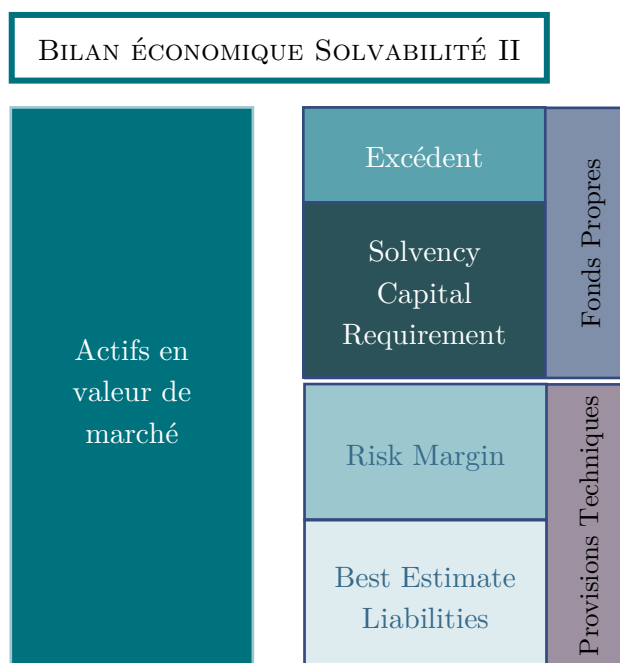


Figure 4 - Schéma du Bilan sous Solvabilité II

Les provisions techniques se calculent en valeur de marché. C'est la somme du « Best Estimate Liabilities » et de la « Market Value Margin » (ou Risk Margin en formule standard).

Le calcul du Best Estimate Liabilities correspond à la meilleure vision économique des engagements de l'assureur (y compris Provisions Techniques des réassureurs). Avec trois points de vigilance :

- En assurance vie: projection actualisée des cash-flows des passifs d'assurance
- En assurance IARD: méthode actuarielle à partir de l'analyse des triangles (déroulé des provisions, charge 50/50) sur les exercices passés

Le calcul de la Marge de risque (Market Value Margin) correspond au coût du capital et se déduit de la projection du capital requis sous Solvabilité 2.

La somme du BEL et de la MVM correspond au prix qu'exigerait une tierce partie pour honorer les engagements à la charge de l'assureur : c'est la valeur de marché.

Garante des provisions, la DAAP doit donc être capable de fournir les provisions en vision Best-Estimate. En particulier, pour les provisions pour sinistres à payer (scope du mémoire), l'estimation de la distribution peut donner accès à des quantiles, ce qui permet de connaître l'écart entre la vision prudentielle (demandée par les équipes comptables) et la vision Best-Estimate (nécessaire au Risk Management d'AXA France).

1.3 PRÉSENTATION DES CONTRATS ÉTUDIÉS

1.3.1 Présentation des contrats Oscar

Oscar est une InsurTech américaine créée en 2012 à New-York proposant des contrats de santé dans 15 états aux États-Unis. Ces contrats s'adressent aux particuliers mais également aux entreprises qualifiées de « small businesses ». Cet assureur mise sur la télémédecine et l'accompagnement des patients pour se développer et devenir leader dans l'assurance santé « tech-minded » de demain aux États-Unis.

Scope étudié

Le scope étudié dans ce mémoire sur le portefeuille américain Oscar correspond aux 4 traités historiques (les plus anciens, ayant le plus de données) : la Californie, New-York, le Texas et le New-Jersey.

Ces états ont été sélectionnés parmi les 13 états réassurés par AXA car ils sont suffisamment anciens pour pouvoir mener une étude avec du recul sur les survenances mensuelles de sinistres.

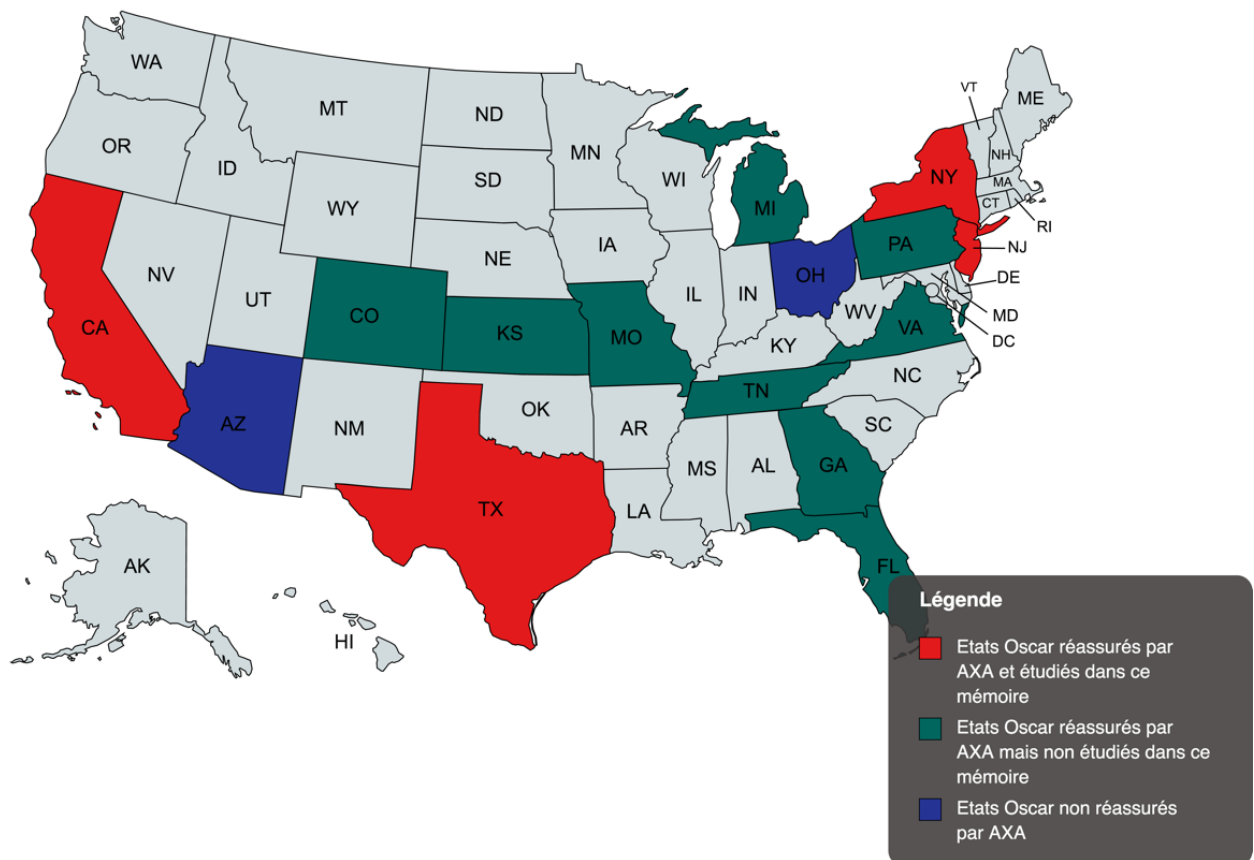


Figure 5 - Visualisation du scope du mémoire, du scope AXA, et du scope hors AXA

| TREATY LABEL | CEDING COMPANY | CURRENCY REPORTING |
|------------------|--|--------------------|
| OSCAR NEW YORK | 00600 - OSCAR INSURANCE CORPORATION | USD |
| OSCAR CALIFORNIA | 00603 - OSCAR HEALTH PLAN OF CALIFORNIA | USD |
| OSCAR NEW JERSEY | 00602 - OSCAR GARDEN STATE INSURANCE CORPORATION | USD |
| OSCAR TEXAS | 00601 - OSCAR INSURANCE CIE OF TEXAS | USD |

Figure 6 - Contrats Oscar étudiés dans ce mémoire

Description du système local

En octobre 2013, l'Obama Care (également appelé Affordable Care Act - ACA) est entrée en vigueur.

Elle comporte 2 axes principaux :

- La loi sur la protection des patients et les soins abordables
- La loi sur la réconciliation des soins de santé et de l'accessibilité à l'éducation

Cette loi a dynamisé le business des assurances santé. En effet, il n'existe pas de caisse d'assurance maladie publique aux États-Unis et tous les risques sont pris par les assureurs privés. Cela oblige les Américains à souscrire à une assurance santé privée individuelle ou collective. Bien que cette obligation ait été annulée par le gouvernement Trump, en janvier 2019 au niveau fédéral, les Américains n'ont pas résilié massivement leurs contrats de couverture santé.

Chaque assurance américaine propose 4 niveaux de couverture : Bronze, Argent, Or et Platine. Ces niveaux sont directement liés au pourcentage moyen des frais de santé pris en charge par l'assureur. Par exemple, Bronze couvre seulement 60% des dépenses alors que Platinum en couvre 90%.

Chaque assurance doit également garantir 10 services de santé « essentiels » :

- Les services aux patients ambulatoires (les soins ambulatoires que vous recevez sans être admis dans un hôpital)
- Les services d'urgence
- Hospitalisation (comme la chirurgie et les séjours de nuit)
- Les soins de grossesse, de maternité et de nouveau-né (avant et après la naissance)
- Services de santé mentale et de traitement de la toxicomanie, y compris les traitements de santé comportementale (y compris le conseil et la psychothérapie)
- Médicaments sur ordonnance
- Services et appareils de réadaptation et d'habilitation (services et appareils destinés à aider les personnes souffrant de blessures, de handicaps ou d'affections chroniques à acquérir ou à récupérer des aptitudes mentales et physiques)
- Services de laboratoire
- Services de prévention et de bien-être et gestion des maladies chroniques

- Les services pédiatriques, y compris les soins bucco-dentaires et les soins de la vue (mais la couverture des soins dentaires et de la vue pour adultes ne sont pas des prestations de santé essentielles)

Ainsi, les assureurs sont soumis aux règles américaines, ils ne peuvent pas :

- Refuser de souscrire avec quelqu'un
- Adopter une tarification plus élevée aux personnes considérées comme ayant une santé fragile
- Résilier des contrats avec des assurés malades
- Plafonner les dépenses sur les 10 services de santé « essentiels » mentionnés précédemment

Pour régler et compenser ces obligations, les États-Unis ont introduit deux mécanismes :

- L'ajustement des risques : qui donne un score à chaque assureur pour mesurer la qualité du risque qu'il prend à travers la population assurée (plus le score est faible, plus le risque est élevé).
Étant donné que tous les assureurs américains sont obligés d'accepter la demande de contrat d'assurance de toute personne, le mécanisme compensera les primes des assureurs ayant un score plus faible (prenant un risque plus élevé) par les primes des assureurs ayant un score plus élevé (car ils prennent un « bon » risque). Ce mécanisme est toujours en vigueur aujourd'hui
- (Corridor de risque : si le ratio de sinistres est trop élevé, la compagnie reçoit une compensation du pays. Ce mécanisme a été arrêté en 2016)

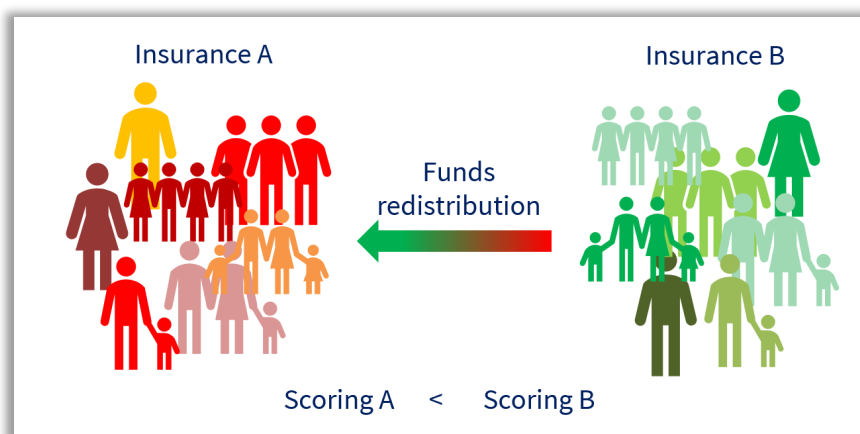


Figure 7 - Illustration de l'ajustement des risques entre différents assureurs américains

Un des points les plus importants dans le contexte américain, et qui sera remarqué dans la suite de l'étude est l'existence d'un montant Maximum Out of Pocket.

Ce montant est défini comme la limite maximale qu'un assuré doit payer pour les garanties souscrites au cours d'une année de couverture. Après avoir payé les franchises et contribué aux paiements des services de soins (frais de soins partiellement pris en charge par l'assureur, en fonction du métal souscrit), l'assuré ne contribue plus à partir d'un certain seuil. La couverture santé de l'assureur prendra en charge 100% des frais au-delà du Maximum Out of Pocket (seuil défini).

Par exemple, dans l'hypothèse où le montant de Maximum Out of Pocket est de 1 000 USD. Un assuré a des frais d'hospitalisation au cours du mois de janvier de 2.000 USD (60% payés par Oscar, 40% payés par l'assuré soit 800 USD à sa charge). Des frais d'hospitalisation au mois d'août sont de 1.000 USD (60% payés par Oscar, 40% payés par l'assuré soit 400 USD à sa charge). Mais, en août, l'assuré a atteint le montant de 1.000 USD de Maximum Out of Pocket, ce qui signifie que s'il doit payer d'autres frais médicaux (pour un sinistre sur une garantie prévue dans son contrat), Oscar paiera 100% des prochains frais jusqu'à la fin de l'année.

| CLAIM | AMOUNT TO PAY | DATE | INSURED | INSURER |
|-------|---------------|------------|-------------------------------------|-------------------------------------|
| #1 | \$2000 | 26/01/2018 | \$800 | \$1200 |
| #2 | \$1000 | 02/08/2018 | \$400 | \$600 |
| | | | \$200 | \$800 |
| | | | (because max out of pocket reached) | (because max out of pocket reached) |
| #3 | \$3000 | 01/10/2018 | \$1200 | \$1800 |
| | | | \$0 | \$3000 |
| | | | (because max out of pocket reached) | (because max out of pocket reached) |

Figure 8 - Exemple du maximum out of pocket observé

Ce mécanisme crée un effet de saisonnalité dans les demandes de remboursement de soins de santé aux États-Unis, puisque la fin de l'année est plus susceptible de présenter des montants de sinistres plus élevés.

Le montant Maximum Out Of Pocket est fixé pour tous les individus et clairement défini sur chaque contrat lors de la souscription.

Garanties proposées

Oscar propose les 4 niveaux classiques des produits d'assurance maladie américains, plus un plus basique appelé « Simple » et un autre appelé « CSR » (type de produit à prime Silver, niveau hybride entre Gold et Silver).

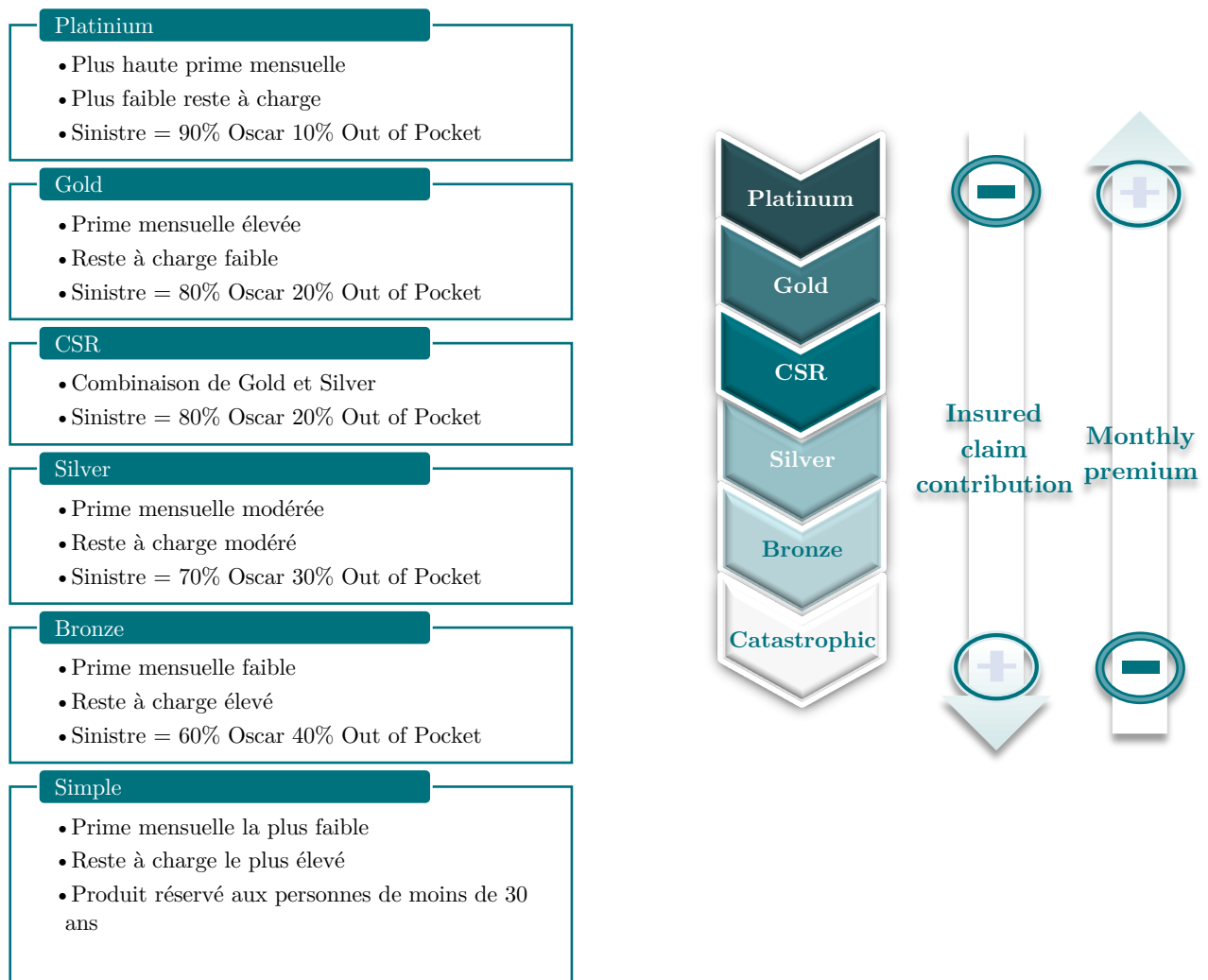


Figure 9 - Illustration des « Official Medical Packages »

Tous ces contrats sont enregistrés en garantie Santé dans les systèmes de données d'AXA France.

Population couverte

Comme mentionné précédemment, Oscar couvre des contrats individuels et groupes. Une étude approfondie du portefeuille Oscar réalisée par AXA France avait révélé une proportion de 85% de contrats individuels et 15% de contrats groupe. Le scope d'AXA France se cantonne aux contrats individuels.

Quelques détails décrivant les comportements des assurés Oscar sur l'année 2019 sont présentés dans la figure ci-dessous:

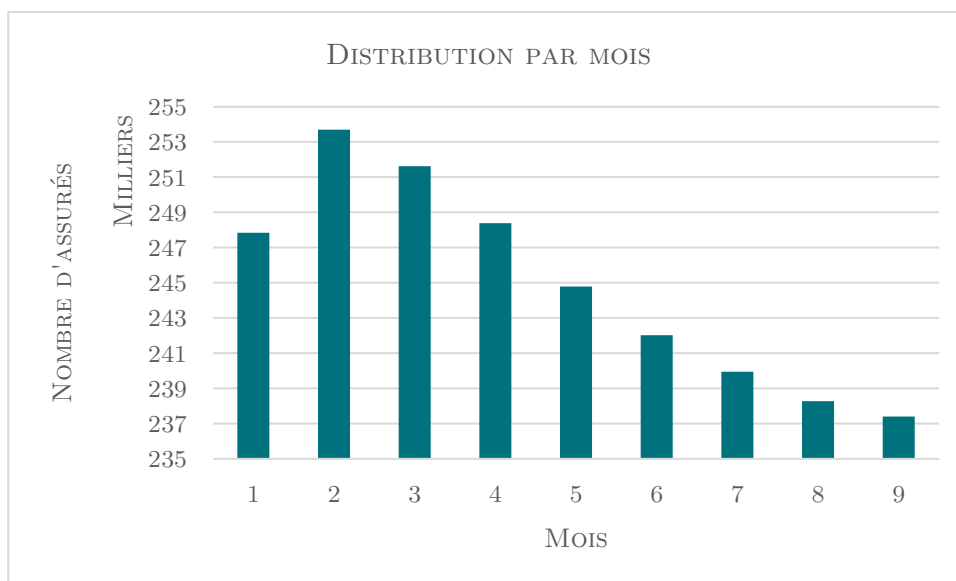


Figure 10 - Nombre d'assurés Oscar sur les 9 premiers mois de 2019, tous états confondus

La souscription se fait généralement en début d'année ce qui explique le pic atteint en février. Par la suite, certains assurés résilient leurs contrats pour diverses raisons, ce qui explique la décroissance linéaire à partir du mois de janvier jusqu'en fin d'année.

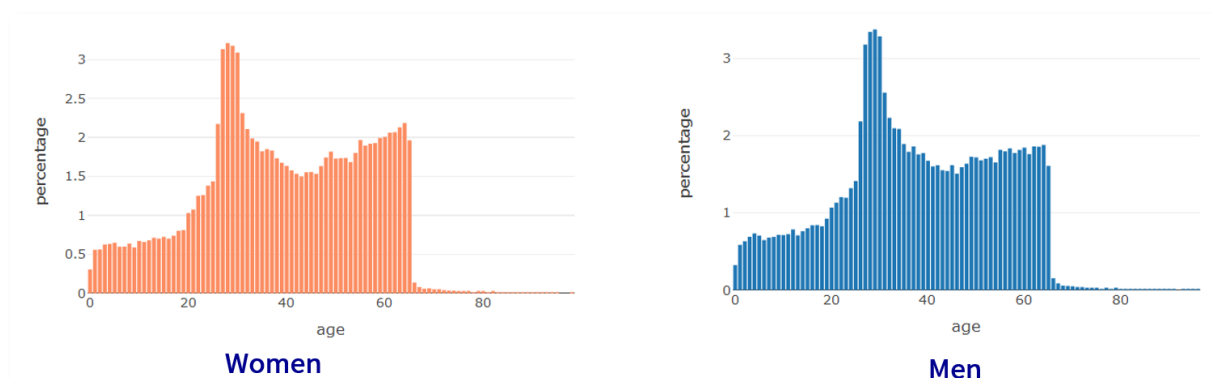


Figure 11 - Distribution par âge et par sexe pour l'année 2019

L'absence d'assurés à l'âge de 65 ans est due au système Medicare. C'est un programme national de remboursement de frais de santé s'adressant aux personnes de plus de 65 ans. En conséquence, cette catégorie de personnes ne souscrit pas de contrats santé.

La population la plus représentée est celle des Millenials (25-30 ans), probablement due à l'aspect digital de Oscar qui plaît majoritairement aux populations jeunes.

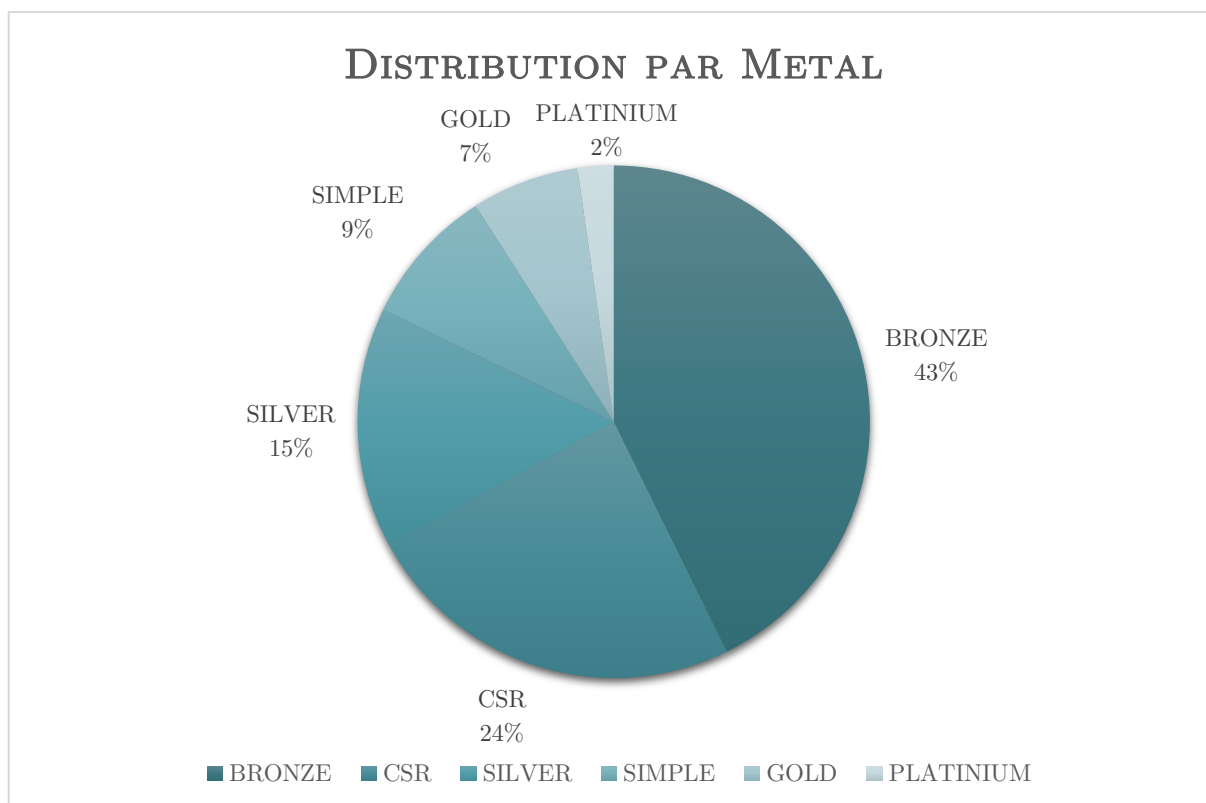


Figure 12 - Répartition des contrats par métal pour l'année 2019

Aucune information particulière concernant l'appétence au risque des assurés n'est notable. Les contrats Bronze sont largement majoritaires et principalement souscrits car ils sont les plus standards et bons marché.

Reinsurance Scheme

AXA France est réassureur d'Oscar sous forme de Quota Share depuis 2017. Oscar est également protégé à l'aide d'une réassurance dite « Excess of Loss », négociée avec un autre réassureur. La Quota Share s'applique après la couverture « Excess of Loss ».

1.3.2 Présentation des contrats australiens

AXA France est réassureur off-shore en Australie, sur des contrats de prévoyance. Le business d'AXA France en Australie se concentre sur l'un des 3 principaux piliers du système de retraite du pays: le fonds professionnel. Ce pilier est géré par des fonds de pension et est financé par l'employeur. Ces fonds de pension peuvent proposer aux salariés d'être couverts en cas d'incapacité ou de décès (Garanties réassurées par AXA France), en plus de la gestion de leur épargne. Une fois à la retraite, les assurés peuvent avoir accès à leur fonds de pension, qui peut ensuite être versé sous la forme d'une somme forfaitaire ou d'une rente définie pour leur retraite.

Scope étudié

Différents fonds de pension (appelés également « superannuation funds ») opèrent en Australie et AXA France est réassureur pour une dizaine d'entre eux. Les contrats de réassurance étudiés dans ce mémoire concernent les fonds de pension AON MasterTrust (AMT), CARE SUPER et IOOF Frontbook.

| TREATY LABEL | CEDING COMPANY | CURRENCY REPORTING |
|------------------------------|---------------------------------|--------------------|
| AMT - GROUP LIFE | AIA AUSTRALIA LIMITED | AUD |
| AMT - GSC - LT 5 YEARS | AIA AUSTRALIA LIMITED | AUD |
| AMT - GSC - ST 2 YEARS | AIA AUSTRALIA LIMITED | AUD |
| CARE SUPER AFFAIRES HORS CIA | METLIFE INSURANCE LTD AUSTRALIA | AUD |
| CARE SUPER AFFAIRES CIA | METLIFE INSURANCE LTD AUSTRALIA | AUD |
| IOOF GTL / IP | TAL LIFE LIMITED AUSTRALIA | AUD |

Figure 13 - Contrats Australiens étudiés dans ce mémoire

Description du système local

Le système de retraite en Australie repose sur les 3 piliers suivants :

- **Pensions publiques** : composé d'une pension de vieillesse sous condition de ressources, financée par l'impôt, qui fournit des prestations de base.
- **Pensions professionnelles** : système privé à cotisations définies composé de comptes individuels de retraite par capitalisation fournis par des fonds de pension. Financé par les cotisations obligatoires de l'employeur. Des garanties Décès ou Incapacité sont proposées en plus, et réassurées par AXA France.
- **Pension personnelle** : Financée par des cotisations supplémentaires du salarié en plus des cotisations obligatoires, sous réserve d'avantages fiscaux pour le salarié.

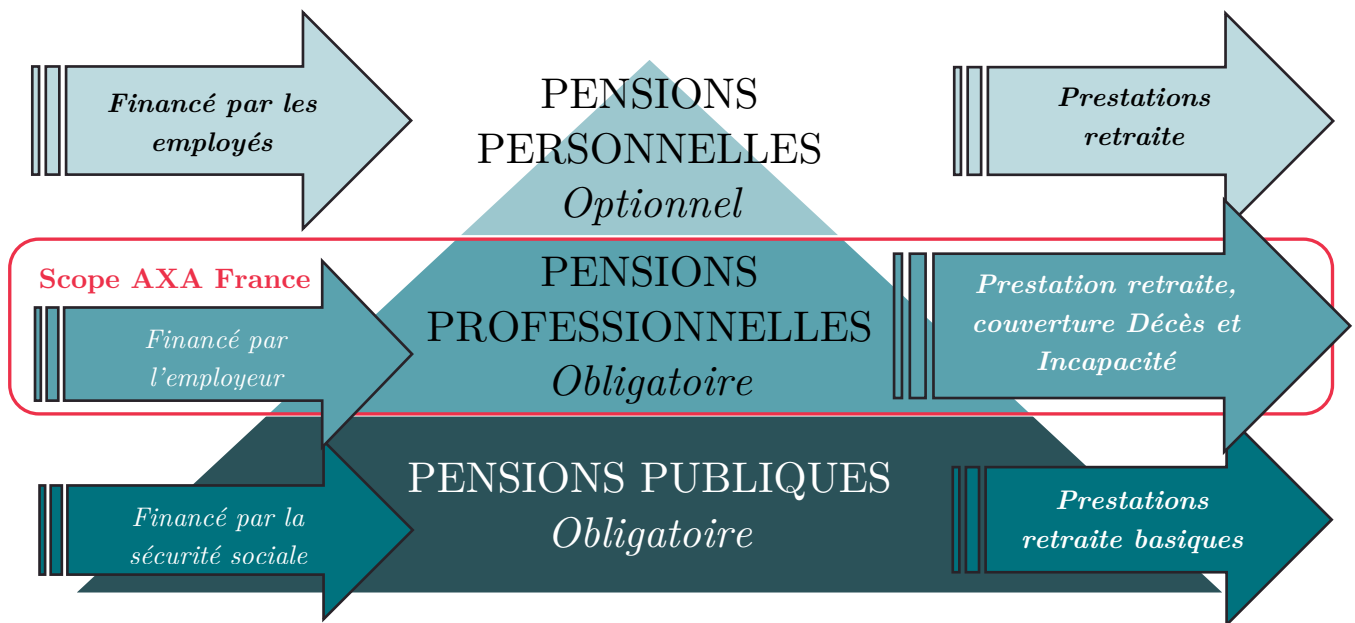


Figure 14 - Pyramide des pensions de retraite en Australie

Les fonds de pension de retraite sont mis en place par le gouvernement australien. Le but est d'encourager les actifs à placer leur épargne dans des fonds qui leur assurent un flux de revenus lorsqu'ils prennent leur retraite. Ces fonds fonctionnent sous la forme d'un trust, c'est-à-dire d'accords entre le bénéficiaire, le fonds de pension, et une société fiduciaire, cette dernière étant responsable du fonctionnement prudentiel de ses fonds, de la formulation et de la mise en œuvre d'une stratégie d'investissement.

Le gouvernement a fixé des normes minimales pour les contributions des employés ainsi que pour la gestion des fonds de pension. Les employeurs sont tenus de verser des cotisations de retraite pour leurs employés en plus de leurs salaires. Le taux de cotisation des employeurs est de 9,5 % depuis le 1er juillet 2014.

Les fonds de pension ont différents régimes définis par la loi sur la sécurité sociale, et la plupart d'entre eux sont règlementés par l'APRA (Australian Prudential Regulation Authority).

Market share by fund size

| Fund size | No. of members | No. of funds | Total AUM | Change in No. funds* | Change in AUM* | Growth in AUM |
|--------------|----------------|--------------|-----------|----------------------|----------------|---------------|
| >25B | 19.4M | 17 | \$1,313B | ↑ | ↑ | 13% |
| \$1 to \$25B | 5.5M | 55 | \$397B | ↓ | ↓ | -2% |
| < \$1B | 0.1M | 22 | \$10B | ↓ | ↑ | -19% |
| Total | 25M | 94 | \$1,720B | ↓ | ↑ | 9% |

* Change from prior year

Figure 15 - Situation des fonds de pensions Australiens en 2020

Garanties

Les fonds de pension peuvent inclure une couverture d'assurance proposant à leurs clients les garanties suivantes, par le biais de compagnies d'assurance locales :

○ **Polices d'assurance-vie de groupe**

Généralement un multiple de la prestation annuelle, versée sous forme de capital.

- Décès, quelle qu'en soit la cause
- Maladie terminale (conjointement avec le décès, pas toujours inclus)
- Invalidité totale et permanente (TPD)

○ **Polices collectives de maintien de salaire**

Pourcentage du salaire régulier versé sous forme de prestation mensuelle pendant que le travailleur est incapable de travailler sur une durée limitée.

- Protection du salaire (IP), suite à un accident ou une maladie

Les conditions relatives à la période d'indemnisation (durée maximale de la demande par sinistre), permettent de distinguer les prestations d'invalidité de courte durée (ICD), qui peuvent être versées durant 2 à 5 ans, et d'invalidité de longue durée (ILD), qui peuvent être versées jusqu'à 65 ou 70 ans.

Toutefois, la prestation d'IP (protection du salaire) prend fin à tout moment pour des raisons de fin de couverture, de rétablissement (retour au travail) ou de décès.

La prestation d'IP comporte généralement une franchise appelée période de carence, définie comme la période à partir de l'apparition de la maladie/blessure pendant laquelle une personne doit être invalide avant d'avoir droit à la prestation. Les délais d'attente disponibles sont les suivants : 14, 30, 60, 90, 180 jours, 1 an ou 2 ans.

Les assureurs proposent généralement différents niveaux de polices avec des inclusions variables et, selon les options choisies, les primes peuvent varier considérablement. Ces niveaux de polices sont définis lors de la souscription, à la maille tête par tête.

Population couverte et entités intermédiaires

Ces contrats couvrent des salariés Australiens de 18 ans jusqu'à 65 ans (certains cas allant jusqu'à 75 ans), percevant un salaire minimum de 450 AUD par mois en 2019.

AXA France réassure les compagnies locales de la manière suivante:

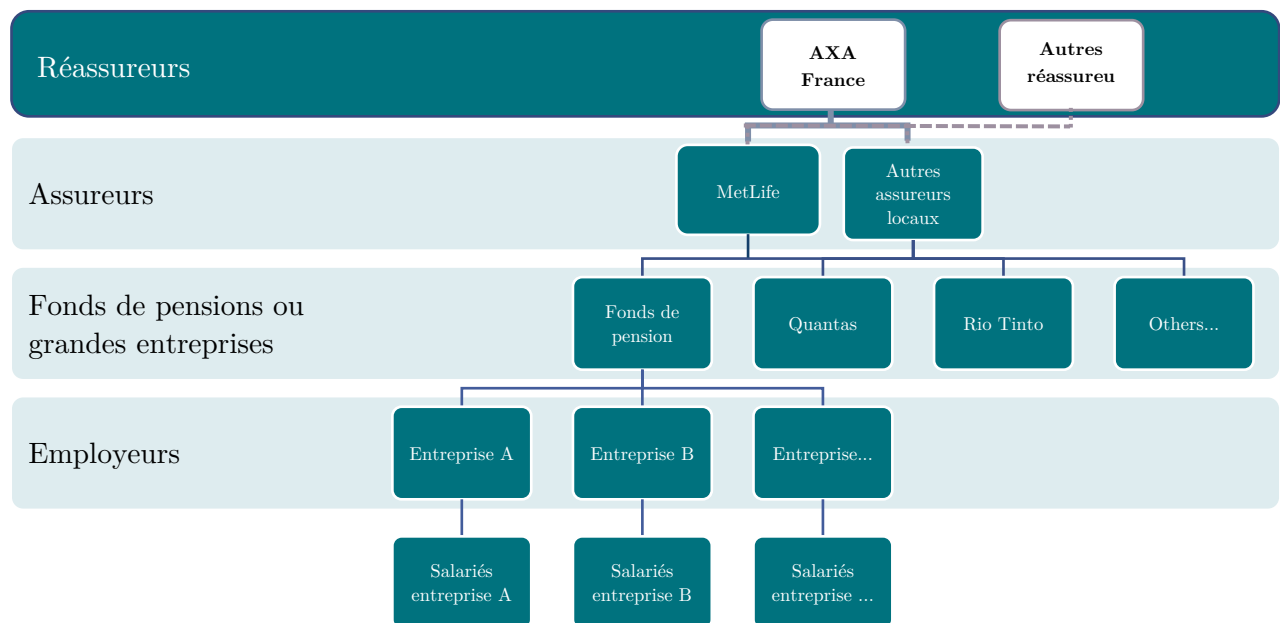


Figure 16 - Lien entre AXA France et la population couverte

Les entreprises faisant appel à des assureurs pour gérer leurs contrats de garantie décès, d'incapacité ou de protection de salaire sont soit des fonds de pension, soit des grandes entreprises australiennes comme Rio Tinto, société spécialisée dans l'extraction minière.

Pour rappel, nous étudierons les contrats Australiens de AMT, CARE SUPER et IOOF Frontbook qui sont des fonds de pension. Par la suite, la DAAP pourra étendre l'étude aux autres contrats, si les performances de provisionnement sont satisfaisantes.

La répartition de la population australienne par âge (tous les fonds de pension) est indiquée ci-dessous, selon la documentation statistique de l'APRA :

| | Number of member accounts ('000) | | | |
|------------------------|----------------------------------|---------------|------------|---------------|
| | Female | Male | Other | Total |
| By age brackets | | | | |
| <25 | 1,394 | 1,545 | 16 | 2,955 |
| 25 to 34 | 2,825 | 3,256 | 32 | 6,113 |
| 35 to 44 | 2,768 | 3,202 | 13 | 5,983 |
| 45 to 49 | 1,375 | 1,574 | 3 | 2,952 |
| 50 to 54 | 1,151 | 1,331 | 2 | 2,484 |
| 55 to 59 | 1,031 | 1,218 | 2 | 2,251 |
| 60 to 64 | 800 | 958 | 1 | 1,759 |
| 65 to 69 | 498 | 598 | 1 | 1,096 |
| 70 to 74 | 284 | 350 | - | 633 |
| 75 to 84 | 193 | 250 | - | 443 |
| 85+ | 47 | 48 | - | 94 |
| Age not available | 5 | 8 | 69 | 82 |
| Total | 12,368 | 14,338 | 140 | 26,846 |

Figure 17 - Détail des assurés fonds de pension par âge et par genre en Australie

La population australienne est estimée par l'ABS (Australian Bureau of Statistics) à 24,6 millions d'habitants en juin 2017. L'âge médian de la population australienne a augmenté de 3 ans au cours des deux dernières décennies, passant de 34 ans au 30 juin 1997 à 37 ans au 30 juin 2017.

Comme la plupart des pays développés, la population australienne vieillit en raison d'une faible natalité combinée à l'augmentation de l'espérance de vie. Cela a entraîné une diminution du nombre d'enfants (moins de 15 ans) dans la population et une augmentation plus importante des personnes âgées de 65 ans et plus.

Ces évolutions ont pour conséquence une hausse de la sinistralité sur les garanties étudiées.

Reinsurance scheme

AXA France réassure le marché australien via une quota share (majorité des contrats) ou de quota share avec excédent de plein (minorité de contrats).

La quota share est définie avec la cédante par garantie, et le montant retenu par l'assureur est soumis à une limite définie entre l'assureur et le réassureur. Au-delà de la limite définie, le réassureur prend 100% du risque (régime excédent de plein). Les différents types de réassurance sont rappelés en Annexe en page 99.

2. LES PROPOSITIONS DE PROVISIONNEMENT

Dans ce mémoire, les propositions de provisionnement seront développées à l'aide du package *ChainLadder v0.8.3* disponible sur Python. La méthode CANN sera également développée sur Python via Keras (API).

2.1 L'EXISTANT

2.1.1 Le provisionnement existant de la DAAP

L'outil existant à la DAAP est un outil d'estimation des provisions généraliste. La méthode retenue consiste à estimer poste à poste les nouveaux montants de l'année en cours, pour chaque contrat du portefeuille.

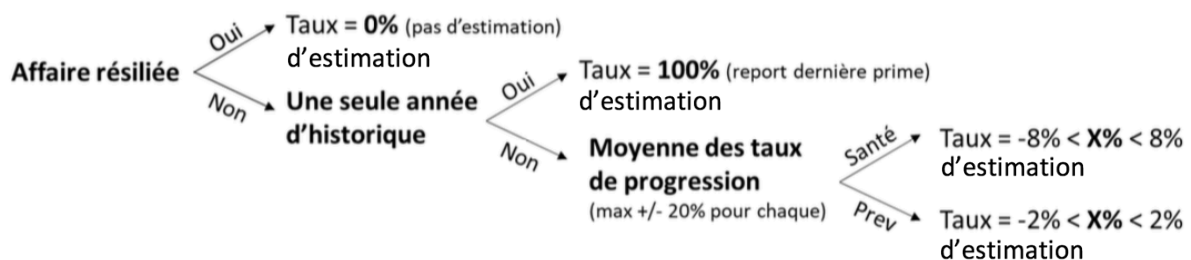
L'outil estime d'abord les primes. Le résultat obtenu est alors utilisé pour calculer les estimations de sinistres. Enfin, les estimations de sinistres et de primes sont utilisées pour calculer « la charge », ce qui permet de calculer l'estimation de provision (qui inclut la PSAP).

Estimation des primes

Pour l'estimation des primes, un taux d'estimation est calculé pour chaque contrat (appelé parfois « affaire ») non résilié. Il est fonction des taux de progression des 5 années précédentes, capés au seuil inscrit en paramètres (différent selon les garanties, +/- 8% pour la santé et +/- 2% pour les garanties de prévoyance). Ces taux minimums et maximums sont déterminés grâce à l'expérience acquise sur les contrats du portefeuille ALHIS années après années.

S'il n'y a qu'une année d'historique, le montant de la dernière survenance est reporté. Un prorata est calculé en cas d'année incomplète.

Estimation des primes



$$Prime_{\text{année } N} = Prime_{\text{année } N-1} \cdot \text{Taux d'estimation}$$

Estimation des sinistres

Pour chaque contrat non-résilié, l'outil calcule le rapport entre les sinistres observés non-nuls et les primes estimées. Ce taux estimé est reporté sur toutes les survenances vides après les derniers réels et jusqu'à l'année N. On peut alors calculer le $S/P_{estimé}$ (taux sinistres sur primes estimé).

$$S/P_{estimé} = \frac{Moyenne(Sinistres\ réels\ jusqu'à\ N-1)}{Moyenne(Primes_{jusqu'à\ N})}$$

Pour chaque contrat, le montant $Moyenne(Primes_{jusqu'à\ N})$ prend en compte les primes réelles ainsi que la prime estimée de l'année en cours (précédemment calculée).

L'estimation des sinistres de l'année en cours est donc égale à :

$$Sinistres_{estimés\ année\ N} = S/P_{estimé} \cdot Primes_{estimées\ année\ N}$$

Estimation de la charge

La charge est égale à la somme des Sinistres réels et Provisions réelles. Pour chaque contrat non-résilié, l'outil calcule le rapport entre les charges réelles non-nulles et les primes estimées. Ce taux estimé est reporté sur toutes les survenances vides après les derniers réels et jusqu'à l'année N.

$$Ratio\ Combiné_{estimé} = \frac{Moyenne(Charge\ réelle\ jusqu'à\ N-1)}{Moyenne(Primes_{jusqu'à\ N})}$$

Pour chaque contrat, le montant $Moyenne(Primes_{jusqu'à\ N})$ prend en compte les primes réelles ainsi que la prime estimée de l'année en cours (précédemment calculée).

L'estimation de la charge de l'année en cours est donc égale à :

$$Charge_{estimée\ année\ N} = Ratio\ Combiné_{estimé} \cdot Primes_{estimées\ année\ N}$$

Estimation des provisions

Le calcul des provisions correspond à la différence entre la charge estimée pour l'année N et les sinistres estimés pour l'année N.

$$Provision_{estimée\ année\ N} = Charge_{estimée\ année\ N} - Sinistres_{estimés\ année\ N}$$

Conclusion sur l'approche généraliste

L'outil utilisé est « simple » et universel. En conséquence, les temps de calcul sont acceptables malgré la multitude de contrats suivis. Il permet également des ajustements en cas d'indications particulières de la part du marché. Par exemple, certaines informations capitales vont être transmises par le marché à la DAAP lors de l'arrêté comptable (précision sur la croissance d'un contrat, résiliations, sinistre exceptionnel, etc.), ce qui permet de revoir certaines estimations ligne à ligne.

En revanche, cet outil présente un certain nombre d'inconvénients. En effet, son aspect universel ne permet pas de suivre un provisionnement précis sur chaque contrat. Le portefeuille ALHIS est un groupement de contrats assez hétérogènes en particulier sur les garanties ainsi que sur les pays.

D'une part, il existe des garanties de prévoyance (décès, incapacités, rentes, etc.) ainsi que des garanties santé. Ces garanties n'ont pas les mêmes cadences de paiement. Par exemple, une garantie santé sera généralement réglée au plus tard dans les 6 mois suivant la survenance du sinistre, tandis qu'une garantie prévoyance peut être réglée après plusieurs années suivant la déclaration du sinistre (expertises longues, procédures judiciaires, etc.).

Or, un provisionnement général ne permet pas d'intégrer cet aspect de cadence de paiement pourtant primordial dans un calcul précis de provision.

D'autre part, l'implantation mondiale des contrats ajoute de l'hétérogénéité car chaque pays possède son propre système de sécurité sociale (développé ou non), sa législation, ses coutumes ainsi que son système de santé. Par conséquent, un provisionnement sur une même garantie entre un pays comme les États-Unis et un pays d'Amérique du Sud ne sera pas le même. Par exemple, les États-Unis étant un pays développé, dans le cadre d'une garantie Santé, l'accès aux soins est rapide et le règlement du sinistre l'est aussi. Cela pour différentes raisons, en particulier la connaissance de l'accès aux soins et l'utilisation de nouvelles technologies.

En général, dans un pays développé, les assurés connaissent leur couverture santé et ont l'habitude de l'utiliser. Cela réduit le temps de déclaration de sinistre et augmente le nombre de sinistres déclarés : cadence de paiement rapide et fréquence de sinistres haute.

Les nouvelles technologies permettent également un accès aux soins rapide et un règlement de sinistre dans un laps de temps réduit. En particulier, les assureurs comme Oscar, précédemment présenté dans ce mémoire, utilisent des moyens de préventions digitaux via des applications mobiles pour orienter leurs assurés le plus rapidement possible vers le bon spécialiste dans le cadre d'une assurance santé. La facilité de compréhension des garanties via l'application combinée aux technologies de gestions développées par l'Insurtech permettent un règlement rapide des sinistres.

Dans le cas d'un pays en voie de développement, le système de santé (qu'il soit une couverture nationale ou personnelle) est moins bien connu par l'assuré et les technologies mises en place sont plus rares. De ce fait, la connaissance de l'accès aux soins est plus basse et les cadences

de paiements plus lentes. En revanche, ces deux paramètres ont tendance à évoluer rapidement dans ce type de pays tandis que dans un pays développé ces paramètres sont stables.

Par conséquent, les méthodes de calcul des provisions de ces pays, notamment en ce qui concerne les provisions pour sinistres à payer, dépendent largement des deux paramètres mentionnés précédemment.

| Avantages | Inconvénients |
|---|--|
| <ul style="list-style-type: none"> ○ Généraliste : est capable de calculer des provisions sur l'ensemble du portefeuille ○ Unique : un seul outil à maintenir pour les équipes ○ Efficace : temps de calcul acceptable ○ Corrigible : possibilité de passer des ajustements lignes à lignes en fonction des informations reçues via le marché ○ Auditable : l'outil est documenté et audité | <ul style="list-style-type: none"> ○ Précision : certains contrats majeurs nécessitent régulièrement des ajustements lignes à lignes |

Figure 18 - Avantages et inconvénient de l'outil existant

2.1.2 Le process de provisionnement lors d'un arrêté

Lors d'un arrêté, la DAAP utilise donc l'outil généraliste de provisionnement précédemment présenté. Pour chaque arrêté, le nombre de contrats se compte en milliers et par conséquent un outil universel est bienvenu. En revanche, pour certains contrats majeurs, l'outil propose une précision de provisionnement trop faible par rapport aux attentes des différentes entités travaillant avec la DAAP (les équipes comptables, le Risk Management ou encore le Marché).

Étant proche de la souscription et des entités locales, le Marché adopte une approche terrain et possède certains outils de provisionnement (créés en interne, ou récupérés via les cédantes dans le cadre des contrats de réassurance) ayant une précision acceptable sur les contrats de large envergure.

Par conséquent, la vision Marché est régulièrement suivie lors des arrêts sur les contrats majeurs.

2.1.3 Les outils de provisionnement du Marché

Le Marché ALHIS, aussi appelé Portfolio Monitoring, suit de près les contrats, de la souscription à la rentabilité. L'équipe marché est en contact rapproché avec les entités locales et possède des informations cruciales permettant le pilotage du portefeuille, avec notamment un suivi accentué des contrats d'envergure.

En conséquence, le marché possède ses propres outils de provisionnement. En particulier, sur les deux groupes de contrats étudiés dans ce mémoire : Oscar et l'Australie.

Outil de provisionnement Oscar

Pour commencer, il est à noter que le provisionnement Oscar (garantie Santé) s'effectue uniquement à travers le calcul d'une provision pour sinistre à payer (PSAP). Il n'y a pas de provision mathématique sur ce business.

Bien que le marché reçoive le détail des provisionnements PSAP effectués par la cédante Oscar, l'équipe a développé un outil permettant de challenger la vision d'Oscar sur les provisions établies.

En effet, Oscar travaille sur un provisionnement en Chain Ladder sur l'ensemble des montants de sinistres (ligne à ligne) inférieurs à \$40.000. Les sinistres supérieurs à ce montant sont étudiés au cas par cas.

A l'inverse, l'équipe marché a décidé de travailler sur un Chain Ladder prenant en compte l'ensemble des montants. Cela est dû au fait que l'étude au cas par cas devient impossible en absence d'information détaillée sur les sinistres de pointe.

Le provisionnement du portefeuille Oscar par les équipes Marché d'AXA France se fait via des « 12-month sliding Chain Ladder » état par état. C'est à dire que chaque état possède son propre triangle, sur lequel le coefficient de passage d'une période de développement à l'autre nécessite 12 mois de survénances. Ces notions seront présentées et revues en détail dans les propositions de provisionnement classiques étudiées dans la suite de ce mémoire.

Le Marché compare alors sa vision avec celle de Oscar sur les provisions. Des discussions sont entamées en cas d'écart significatif de provisions (>5% d'écart) état par état.

Une fois le provisionnement revu, ces montants sont souvent ceux que la DAAP adopte dans les comptes d'AXA France, lorsque son outil généraliste donne une estimation trop lointaine de celle d'Oscar.

Outil de provisionnement Australie

Pour commencer, il est à noter que le provisionnement sur les contrats australiens (garantie Prévoyance) s'effectue à travers le calcul d'une provision pour sinistre à payer (PSAP) ainsi qu'à travers une provision mathématique (réalisée tête par tête, une fois le sinistre déclaré – hors scope pour ce mémoire).

L'outil de provisionnement PSAP utilisé par les équipes marché vient des entités locales et se base sur les pratiques usuelles appliquées en Australie pour ce business précis.

Les triangles de calculs se font sur des nombres de sinistres (et non des paiements) ce qui permet de calculer un nombre de sinistres ultime qui sera multiplié par une charge moyenne (contenant les sinistres payés et une estimation de la provision mathématique) afin de déterminer une charge ultime.

Le calcul de la provision PSAP se faisait jusqu'alors via un Chain Ladder classique, portant sur des triangles de périodicité trimestrielle.

Une nouvelle législation est entrée en vigueur au 1^{er} juillet 2019 en Australie, portant le nom de Protecting Your Super (PYS), ce qui a entraîné une révision méthodologique pour le calcul de la PSAP sur les contrats Australiens.

Les mesures PYS sont destinées à protéger les épargnes des Australiens (Superannuation Funds) contre une érosion due aux frais et aux commissions d'assurance.

La législation comprend :

- Une modification des frais
- Le transfert des comptes inactifs à faible solde (moins de 6 000 AUD) vers l'Australian Taxation Office (ATO)
- L'annulation de l'assurance pour les membres inactifs (pas de cotisations pendant plus de 16 mois)

En conséquence de ces mesures, les dernières périodes observées présentent une accélération très forte du nombre de sinistres et un Chain Ladder classique ne permet plus de suivre avec précision les provisions de ce business. En effet, lorsqu'on applique le Chain Ladder classique, la dernière valeur du nombre de sinistres observée sur la dernière survenance a un impact très fort sur l'IBNR.

| Incrémental | Nbre Sinistres j=0 | Nbre Sinistres j=1 | Nbre Sinistres j=2 | Nbre Sinistres j=3 |
|-------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Survenance i=0 | $I_{0,0}$ | $I_{0,1}$ | $I_{0,2}$ | $I_{0,3}$ |
| Survenance i=1 | $I_{1,0}$ | $I_{1,1}$ | $I_{1,2}$ | |
| Survenance i=2 | $I_{2,0}$ | $I_{2,1}$ | | |
| Survenance i=3 | $I_{3,0}$ | | | |

Figure 19 - Triangle incrémental théorique

| Cumulatif | Nbre Sinistres j=0 | Nbre Sinistres j=1 | Nbre Sinistres j=2 | Nbre Sinistres j=3 |
|-------------------|-----------------------|-------------------------------|-------------------------------|-------------------------------|
| Survenance i=0 | $C_{0,0} = I_{0,0}$ | $C_{0,1} = I_{0,1} + C_{0,0}$ | $C_{0,2} = I_{0,2} + C_{0,1}$ | $C_{0,3} = I_{0,3} + C_{0,2}$ |
| Survenance i=1 | $C_{1,0} = I_{1,0}$ | $C_{1,1} = I_{1,1} + C_{1,0}$ | $C_{1,2} = I_{1,2} + C_{1,1}$ | |
| Survenance i=2 | $C_{2,0} = I_{2,0}$ | $C_{2,1} = I_{2,1} + C_{2,0}$ | | |
| Survenance i=3 | $C_{3,0} = I_{3,0}$ | | | |

Figure 20 - Triangle cumulé théorique

Par exemple, sur le triangle cumulé précédent, les estimations de nombre de sinistres cumulés futurs sont calculés de la manière suivante :

| Cumulatif | Nbre Sinistres j=0 | Nbre Sinistres j=1 | Nbre Sinistres j=2 | Nbre Sinistres j=3 |
|-------------------|-----------------------|---|---|---|
| Survenance i=0 | $C_{0,0}$ | $C_{0,1}$ | $C_{0,2}$ | $C_{0,3}$ |
| Survenance i=1 | $C_{1,0}$ | $C_{1,1}$ | $C_{1,2}$ | $C_{1,3} = f_3 \cdot C_{1,2}$ |
| Survenance i=2 | $C_{2,0}$ | $C_{2,1}$ | $C_{2,2} = f_2 \cdot C_{2,1}$ | $C_{2,3} = f_2 \cdot f_3 \cdot C_{2,1}$ |
| Survenance i=3 | $C_{3,0}$ | $C_{3,1} = f_1 \cdot C_{3,0}$ | $C_{3,2} = f_1 \cdot f_2 \cdot C_{3,0}$ | $C_{3,3} = f_1 \cdot f_2 \cdot f_3 \cdot C_{3,0}$ |
| Coef. | | $f_1 = \frac{\sum_{i=0}^{i=2} C_{i,1}}{\sum_{i=0}^{i=2} C_{i,0}}$ | $f_2 = \frac{\sum_{i=0}^{i=1} C_{i,2}}{\sum_{i=0}^{i=1} C_{i,1}}$ | $f_3 = \frac{\sum_{i=0}^{i=0} C_{i,3}}{\sum_{i=0}^{i=0} C_{i,2}}$ |

Figure 21 - Triangle cumulé théorique, Chain Ladder classique appliqué

On voit que si le dernier nombre de sinistres (diagonale du triangle), subit soudainement une accélération ou un ralentissement de déclaration, les estimations de la même ligne vont être impactées. En particulier, le dernier nombre de sinistres cumulé de la colonne $j = 0$ où le montant $C_{3,0}$ correspond exactement au dernier nombre de sinistres constaté $I_{3,0}$.

En effet, le poids accordé aux données observées par le passé n'est pas assez important, et une accélération ou un ralentissement brutal du nombre de sinistres dus à un changement de législation vont faire varier considérablement les réserves.

La méthode additive permet de renforcer le poids accordé aux valeurs observées par le passé. Elle consiste tout d'abord à repartir d'un triangle cumulé classique, puis d'appliquer un facteur d'échelle sur chaque année de survenance.

| Incrémental | Nbre Sinistres j=0 | Nbre Sinistres j=1 | Nbre Sinistres j=2 | Nbre Sinistres j=3 |
|-------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Survenance i=0 | $I_{0,0}$ | $I_{0,1}$ | $I_{0,2}$ | $I_{0,3}$ |
| Survenance i=1 | $I_{1,0}$ | $I_{1,1}$ | $I_{1,2}$ | |
| Survenance i=2 | $I_{2,0}$ | $I_{2,1}$ | | |
| Survenance i=3 | $I_{3,0}$ | | | |

Figure 22 - Triangle incrémental théorique

Cette normalisation est nécessaire pour la méthode additive, afin d'enlever des biais dus à la croissance du portefeuille, au changement des prestations, au changement du niveau des primes, etc.

| Facteur d'échelle | Cumulatif | Nbre Sinistres j=0 | Nbre Sinistres j=1 | Nbre Sinistres j=2 | Nbre Sinistres j=3 |
|----------------------|-------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| S_0 | Survenance i=0 | $I_{0,0}^* = \frac{I_{0,0}}{S_0}$ | $I_{0,1}^* = \frac{I_{0,1}}{S_0}$ | $I_{0,2}^* = \frac{I_{0,2}}{S_0}$ | $I_{0,3}^* = \frac{I_{0,3}}{S_0}$ |
| S_1 | Survenance i=1 | $I_{1,0}^* = \frac{I_{1,0}}{S_1}$ | $I_{1,1}^* = \frac{I_{1,1}}{S_1}$ | $I_{1,2}^* = \frac{I_{1,2}}{S_1}$ | |
| S_2 | Survenance i=2 | $I_{2,0}^* = \frac{I_{2,0}}{S_2}$ | $I_{2,1}^* = \frac{I_{2,1}}{S_2}$ | | |
| S_3 | Survenance i=3 | $I_{3,0}^* = \frac{I_{3,0}}{S_3}$ | | | |

Figure 23 - Triangle incrémental corrigé théorique

Enfin, la méthode additive calcule pour chaque période le nombre moyen de sinistres que nous avons observé, comme si le triangle était basé sur le même historique (processus de remise à l'échelle). L'IBNR est alors le nombre "additionnel" de sinistres que nous devons ajouter à l'observé pour obtenir l'ultime attendu.

| Facteur d'échelle | Cumulatif | Nbre Sinistres j=0 | Nbre Sinistres j=1 | Nbre Sinistres j=2 | Nbre Sinistres j=3 |
|-------------------|-------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| S_0 | Survenance i=0 | $I_{0,0}^* = \frac{I_{0,0}}{S_0}$ | $I_{0,1}^* = \frac{I_{0,1}}{S_0}$ | $I_{0,2}^* = \frac{I_{0,2}}{S_0}$ | $I_{0,3}^* = \frac{I_{0,3}}{S_0}$ |
| S_1 | Survenance i=1 | $I_{1,0}^* = \frac{I_{1,0}}{S_1}$ | $I_{1,1}^* = \frac{I_{1,1}}{S_1}$ | $I_{1,2}^* = \frac{I_{1,2}}{S_1}$ | $I_{0,3}^*$ |
| S_2 | Survenance i=2 | $I_{2,0}^* = \frac{I_{2,0}}{S_2}$ | $I_{2,1}^* = \frac{I_{2,1}}{S_2}$ | $\frac{1}{2} \sum I_{i,2}$ | $I_{0,3}^*$ |
| S_3 | Survenance i=3 | $I_{3,0}^* = \frac{I_{3,0}}{S_3}$ | $\frac{1}{3} \sum I_{i,1}$ | $\frac{1}{2} \sum I_{i,2}$ | $I_{0,3}^*$ |

Figure 24 - Triangle incrémental corrigé théorique, méthode additive appliquée

L'avantage de cette méthode est qu'elle ne prend pas en considération l'aspect volatile d'une accélération/ralentissement soudain(e) des dernières valeurs observées, et donne plus de poids aux données du passé. En revanche, une telle méthode tend à diluer les tendances réelles « moyen/long terme » telles que l'accélération de la cadence de déclarations de sinistres durable. Aussi, cette méthode, au même titre que le Chain Ladder Classique, ne permet pas d'évaluer une erreur d'estimation des provisions.

2.1.4 La donnée disponible pour l'étude

Différentes sources de données peuvent être choisies à la DAAP lorsqu'une étude doit être menée. Les données peuvent provenir :

- **Du DataLake (tables SAS)** : Base de données standardisée recueillant les montants des différents postes comptables classiques (Sinistres, Primes, Commissions, Provisions, etc.), avec une agrégation assez forte (aucun détail sur les sinistres).

Cette base n'a pas été retenue de par son agrégation trop forte, qui enlèverait tout intérêt à la seconde partie de l'étude dédiée à la présentation de la méthode CANN utilisant certaines caractéristiques des sinistres pour réaliser le provisionnement des PSAP.

- **Des équipes marché (fichiers Excel)** : Bases de données standardisées par portefeuille (non standardisée au global), avec une agrégation raisonnable des sinistres (présence de certaines caractéristiques chez Oscar par exemple comme la ligne de business, l'état américain où est déclaré le sinistre)

Cette base a été retenue pour sa facilité d'accès dans les serveurs, la connaissance et l'utilisation régulière des données par les équipes Marché apportant une expertise supplémentaire en cas de besoin, sa taille raisonnable permettant des calculs rapides.

Note sur la disponibilité des données tête par tête : il existe des données SharePoint (fichiers .parquets) collectées par la gouvernance de la donnée du Marché ALHIS. Cette base de données standardisée recueillant les portefeuilles majeurs (Oscar, Australie – en cours de construction, etc.), comporte une agrégation très faible (beaucoup de détails sur les sinistres tête par tête).

Lors de la réalisation du mémoire, cette base n'a pas été retenue pour différentes raisons. D'une part, la DAAP n'a pas un accès direct à ces données, et la construction d'un outil de réserve doit se baser sur une alimentation en données stable et pérenne en vue de la production trimestrielle. D'autre part, la mise en place de cette base est très récente et certains portefeuilles sont en cours de construction.

Description des données de sinistres disponibles sur le business Oscar

Pour Oscar, les données des équipes Marché ALHIS proviennent du reporting semestriel de la cédante. C'est sur ces données de sinistres que l'InsurTech Oscar calcule ses provisions.

Les montants de sinistres sont agrégés mensuellement (par survenance et par paiement), par état américain (New-York, New-Jersey, etc.) et par ligne de business.

Les différentes variables sont les suivantes :

- **state** : l'état américain où est réglé le sinistre – NJ pour New-Jersey, NY pour New-York, CA pour California et TX pour Texas
- **lob** : « Line Of Business » – IND pour les contrats individuels (les contrats groupes sont exclus du scope d'AXA France)
- **paid** : montant agrégé de sinistres (= *low_paid* + *high_paid*)
- **low_paid** : montant agrégé de sinistres inférieurs à \$40.000 (ligne à ligne)
- **high_paid** : montant agrégé de sinistres supérieurs à \$40.000 (ligne à ligne)
- **incmo** : le mois de survenance des sinistres
- **paidmo** : le mois de paiement des sinistres

Voici un extrait de la table de données à disposition :

| | state | lob | paid | low_paid | high_paid | incmo | paidmo |
|------|-------|-----|--------------|------------|-----------|-----------|----------|
| 0 | NJ | IND | 8.069611e+02 | 1452.53 | 0.0 | 5/1/2016 | 9/1/2017 |
| 1 | NJ | IND | 7.763117e+03 | 13973.61 | 0.0 | 11/1/2016 | 9/1/2017 |
| 2 | NJ | IND | 1.156189e+04 | 20811.41 | 0.0 | 10/1/2016 | 9/1/2017 |
| 3 | NJ | IND | 4.110286e+04 | 73985.15 | 0.0 | 12/1/2016 | 9/1/2017 |
| 4 | NJ | IND | 1.425655e+04 | 25661.79 | 0.0 | 9/1/2016 | 9/1/2017 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 7368 | TX | IND | 1.752740e+06 | 3154931.24 | 0.0 | 3/1/2016 | 4/1/2016 |
| 7369 | CA | IND | 1.305497e+05 | 234989.43 | 0.0 | 3/1/2016 | 4/1/2016 |
| 7370 | CA | IND | 8.145754e+04 | 146623.58 | 0.0 | 4/1/2016 | 4/1/2016 |
| 7371 | CA | IND | 7.407969e+04 | 133343.45 | 0.0 | 2/1/2016 | 4/1/2016 |
| 7372 | CA | IND | 2.846929e+04 | 51244.73 | 0.0 | 1/1/2016 | 4/1/2016 |

7373 rows x 7 columns

Figure 25 - Données disponibles pour l'étude du provisionnement PSAP sur Oscar

La prise en main des données se fait d'abord à travers l'identification des valeurs manquantes et leur traitement si elles existent. Dans cette table de données, aucune valeur manquante n'a été détectée.

Ensuite, il est souhaitable d'étudier la variable « **state** ». Il n'est pas intéressant de visualiser cette variable par montant de par l'agrégation de la table qui enlève du sens à cette analyse. En revanche, une analyse par nombre d'observations (agrégées), demeure pertinente.

Afin de savoir quel volume de données est attribué à chaque état, une visualisation du nombre d'observations de la variable « **state** » est effectuée. Ce nombre d'observations par état varie fortement pour différentes raisons :

- certains états sont entrés dans le portefeuille récemment et possèdent des données uniquement sur les derniers mois d'observation
- certains états ne proposent pas encore de garantie Santé sur la Line of Business « Groupe »

Le volume de données très variable a permis de choisir facilement les 4 états étudiés dans ce mémoire à savoir **New-York, la Californie, le Texas et le New Jersey**.

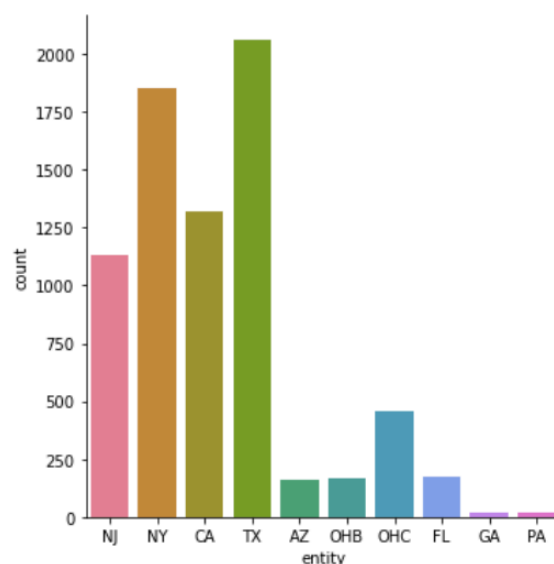


Figure 26 - Nombre d'observations de la variable «state» dans les données - Oscar

Note sur la qualité des données :

Les données sur le business Oscar sont de bonne qualité pour des méthodes de provisionnement classiques. Aucune valeur manquante n'a été identifiée et quelques vérifications ont permis de tester la cohérence des données (exemple : voir si toutes les dates de survenances sont antérieures ou égales aux dates de paiement). Un point négatif est cependant à noter : les données reçues pour l'étude s'arrêtent en date de Juin 2020. En conséquence, l'étude de l'année 2020 sera partielle. Cela pourrait être anodin sur certains portefeuilles, mais la présence du phénomène saisonnier « Maximum Out Of Pocket » évoqué précédemment implique qu'une étude complète se fait sur une année civile complète sur cette cédante.

Description des données de sinistres disponibles sur le business australien

Les données des équipes marché ALHIS proviennent du reporting effectué par l'entité locale AXA Singapour, qui suit de près le business Australien. C'est sur ces données de sinistres que le Marché a dernièrement construit son outil de provisionnement PSAP utilisant la méthode additive.

Les montants de sinistres ont été agrégés par sinistres, par date de survenance, par date de paiement, par cédante (IOOF Frontbook, AMT, etc.) et par garantie.

Les différentes variables sont les suivantes :

- **Count_of_claims** : nombre de sinistres considérés
- **Final_claim_amount** : Montant de sinistre final payé
- **GCT** : Garantie (« IP », « TPD » ou « Death »)
- **Ceding** : Compagnie cédante, dans le mémoire le scope est réduit à l'étude de « CareSuper », « IOOF Frontbook » et « AMT »
- **Incurred_date** : le jour de survenance du sinistre
- **Notification_date** : le jour de paiement du sinistre

Voici un extrait de la table de données à disposition :

| | Count_of_claims | Final_claim_amount | GCT | Ceding | Incurred_date | Notification_date |
|------|-----------------|--------------------|-------|----------------|---------------|-------------------|
| 0 | 1 | 14228.955556 | IP | IOOF Frontbook | 2015-07-11 | 2015-08-10 |
| 1 | 1 | 57458.905556 | IP | IOOF Frontbook | 2015-07-23 | 2015-08-26 |
| 2 | 1 | 744.383333 | IP | IOOF Frontbook | 2015-07-27 | 2015-09-15 |
| 3 | 1 | 8932.483333 | IP | IOOF Frontbook | 2015-08-17 | 2015-09-22 |
| 4 | 1 | 27630.700000 | IP | IOOF Frontbook | 2015-07-24 | 2015-09-23 |
| ... | ... | ... | ... | ... | ... | ... |
| 7581 | 1 | 45233.333333 | Death | CareSuper | 2013-08-29 | 2013-10-03 |
| 7582 | 1 | 82766.666667 | Death | CareSuper | 2012-12-06 | 2013-03-22 |
| 7583 | 1 | 334266.666667 | Death | CareSuper | 2014-03-27 | 2014-06-16 |
| 7584 | 1 | 147643.298611 | Death | CareSuper | 2015-08-25 | 2015-09-11 |
| 7585 | 1 | 140488.888889 | Death | CareSuper | 2014-12-31 | 2015-01-28 |

7586 rows x 6 columns

Figure 27 - Données disponibles pour l'étude du provisionnement PSAP sur l'Australie

La prise en main des données se fait d'abord à travers l'identification des valeurs manquantes et leur traitement si elles existent. Dans cette table de données, aucune valeur manquante n'a été détectée.

Ensuite, il est souhaitable d'étudier la variable « **GCT** » et « **Ceding** ».

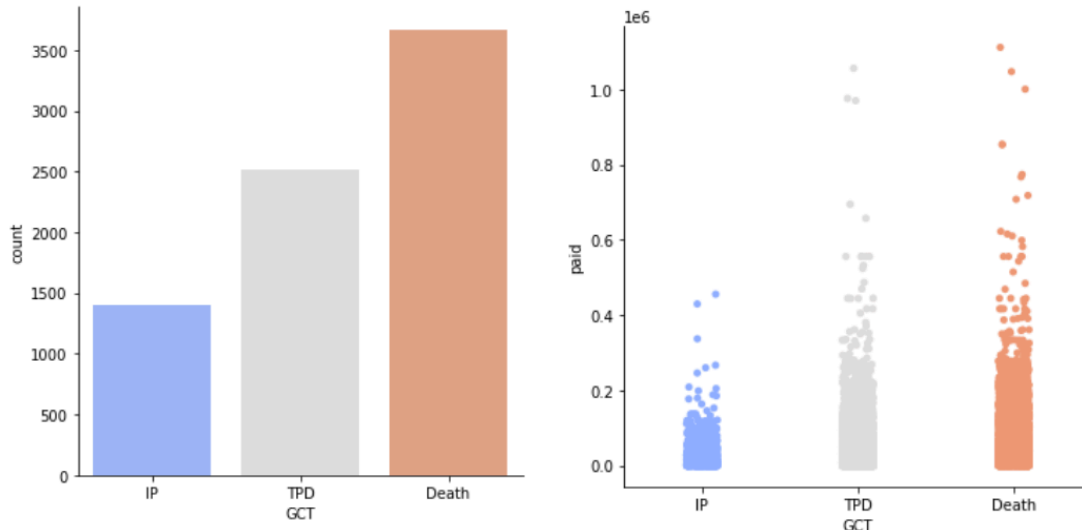


Figure 28 - Nombre d'observations et montants par «GCT» dans les données - Australie

Comme expliqué précédemment, trois garanties sont présentes sur le portefeuille australien : IP correspondant à un produit d'incapacité, ainsi que TPD et Death correspondant à un produit Décès. Il est notable que la garantie Décès est majoritaire dans le scope étudié. Egalement, et sans surprise, les valeurs de sinistres les plus élevés se situent sur les garanties Décès. Effectivement, nous avons précédemment évoqué le fait que les sinistres étaient payés en annuité pour les incapacités alors que les garanties Décès étaient réglées en une seule fois.

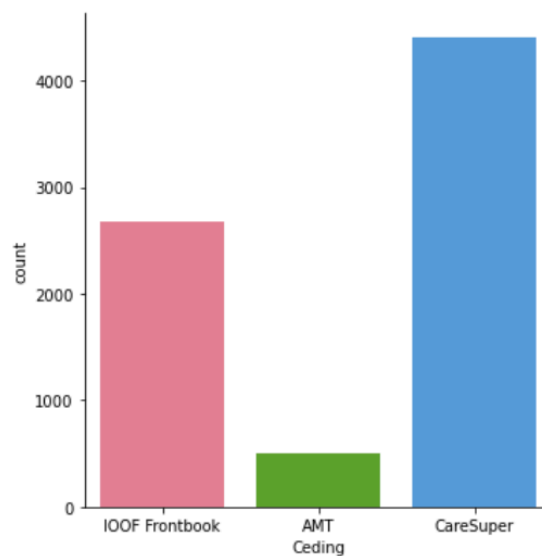


Figure 29 - Nombre d'observations par «Ceding» dans les données - Australie

En ce qui concerne les cédantes, CareSuper correspond au business le plus important en Australie. C'est pourquoi cette cédante a été retenue dans notre scope d'étude, tout comme IOOF Frontbook. En revanche, AMT présente un nombre de sinistres nettement en dessous des autres cédantes précédemment mentionnées, ce qui est dû à une entrée plus tardive dans le portefeuille.

Note sur la qualité des données :

Sur le portefeuille australien, les données observées par le passé présentent un certain nombre de biais pouvant avoir un impact notable dans la suite de l'étude.

En effet, AXA France est entré sur le marché de la Prévoyance en Australie en 2016. Or, les données transmises par le marché débutent en 2012. Les données de 2012 à 2016 sont en réalité des données acquises par AXA France lors de la tarification des contrats en 2016.

Cependant, l'observation des sinistres survenus entre 2012 et 2016 a été arrêtée en 2018. Or, les sinistres de Prévoyance se règlent en moyenne sur une durée de 3 à 5 ans. Il y a donc un phénomène de censure à droite non négligeable qui induit un nombre de sinistre sous-estimé lorsque nous effectuerons un back-testing par la suite, sur le scope 2012-2016.

2.2 LES PROPOSITIONS CLASSIQUES DE PROVISIONNEMENT

Dans cette partie présentant les méthodes classiques de provisionnement PSAP, ainsi que dans la suivante présentant la méthode Combined Actuarial Neural Network (CANN), les applications sont présentées uniquement sur le Business Oscar. Afin d'éviter des répétitions, les résultats du business Australien sont placés en Annexe. Ils sont néanmoins commentés au fur et à mesure du mémoire, à la suite des commentaires concernant Oscar.

Aussi, pour chaque business (Oscar et Australie) l'ensemble des méthodes seront appliquées sur deux scopes temporels distincts. Le premier scope sera suffisamment ancien pour permettre un back-testing. Les résultats présentés par la suite correspondent aux résultats obtenus sur le premier scope temporel : année 2019 pour Oscar. Le second scope sera le scope des dernières données disponibles afin de réaliser des prévisions.

Pour Oscar, les règlements sont rapides (voir ci-dessous en page 52). Il faut remonter à environ 6 mois pour avoir une visibilité quasi-totale des sinistres :

- Scope back-testing : étude des PSAP à fin 2019 (présenté au cours du mémoire)
- Scope récent : étude des PSAP à juin 2020

Pour l'Australie, les règlements sont plus lents (voir en Annexe en page 99). Il faut remonter à environ 4 ans pour avoir une visibilité quasi-totale des sinistres :

- Scope back-testing : étude des PSAP à fin 2016 (présenté en annexe du mémoire)
- Scope récent : étude des PSAP à juin 2020

2.2.1 Le Chain Ladder Classique

Le Chain Ladder classique est une des méthodes de provisionnement PSAP les plus communes. Ses avantages principaux sont sa facilité de mise en œuvre ainsi que sa simplicité de compréhension. Une maille agrégée de données peut être utilisée ce qui est un avantage lorsque la qualité de données est assez faible.

En revanche, cette méthode repose sur une grande stabilité des facteurs de développement. En d'autres termes, les données du triangle doivent correspondre à un périmètre homogène. Si ces contraintes ne sont pas respectées, les prévisions de charges finales (PSAP) peuvent être très instables, notamment en ce qui concerne les années récentes de survenance. Cette méthode ne permettant pas de quantifier les erreurs d'estimation, une instabilité n'est pas souhaitable.

Ce manque de quantification d'erreur d'estimation est pénalisant, notamment dans le cadre du Best-Estimate de la directive européenne Solvabilité II. C'est pourquoi nous explorerons d'autres méthodes dites « classiques » telles que le Chain Ladder de Mack ou encore la méthode Bootstrap afin de mesurer les erreurs de prédiction.

Dimensionnement du triangle

Avant de vérifier les hypothèses du Chain Ladder et d'appliquer la méthode, il est important de dimensionner correctement son triangle.

Différents éléments sont à déterminer, en particulier :

- La durée d'une période dans le triangle
- Le nombre de périodes utiles

Durée d'une période dans le triangle

Classiquement cette période peut être mensuelle, trimestrielle ou annuelle. Le choix va être orienté d'abord par la qualité des données. En effet, il faut que la période d'observation des sinistres soit plus courte que la période souhaitée dans le triangle. En d'autres termes, si les sinistres sont observés tous les jours, alors une agrégation de la donnée est possible pour construire des périodes mensuelles, trimestrielles ou annuelles. En revanche, si les sinistres sont observés trimestriellement, un triangle de période mensuelle est impossible à construire.

D'autre part, la période choisie va également dépendre de la garantie. D'après l'expérience de l'ensemble du portefeuille ALHIS en Santé, les règlements sont dits « rapides », c'est à dire que l'ensemble des sinistres survenus à une date donnée seront réglés en large majorité (plus de 90% des sinistres) à un horizon allant de 6 mois à 1 an, selon les business. En revanche, en Prévoyance, les règlements sont plus longs, c'est à dire que l'ensemble des sinistres survenus à une date donnée seront réglés en large majorité (plus de 90% des sinistres) à un horizon allant de 3 à 5 ans (expérience du portefeuille ALHIS).

En conséquence, on adopte en général une maille mensuelle pour les garanties Santé, et une maille trimestrielle pour les garanties Prévoyance, à condition d'avoir une qualité de données suffisante. Pour les méthodes classiques, les données sont en quantité et qualité suffisantes sur les périmètres étudiés dans ce mémoire. En conséquence, nous adopterons une maille mensuelle pour le portefeuille Oscar et une maille trimestrielle pour le portefeuille Australien.

Nombre de périodes utiles

Une fois la périodicité fixée, il est nécessaire de déterminer le nombre de périodes utiles afin de dimensionner correctement son triangle. Deux paramètres sont à prendre en compte avant de réaliser ce choix. Le premier est la cohérence des données. En effet, si certaines données de sinistres passés ne sont plus cohérentes avec les données de sinistres actuels (changement de cadence de paiements, nouvelles réglementations, élargissement du portefeuille, etc.), alors l'application du Chain Ladder va prendre en compte ces données obsolètes dans le calcul des coefficients de passage, ce qui va polluer les prédictions.

A noter qu'une normalisation du triangle peut enlever certains biais et résoudre ce problème dans certaines situations.

Le deuxième paramètre à prendre en compte pour ce choix est la lisibilité du triangle. Effectivement, certains triangles ayant un nombre de périodes trop important peut nuire à sa lisibilité.

Enfin, il faut également veiller à ce que la durée totale du triangle (*période . nombre de périodes utiles*) soit suffisamment longue afin de réaliser une estimation de l'ultime qui correspond bien à la fin du règlement pour une survenance donnée. En reprenant les temps de règlement mentionnés à titre d'exemple précédemment.

- En Santé, nécessité d'environ 1 an de période totale (durée d'une période 1 mois – nombre de périodes : 12)
- En Prévoyance, nécessité d'environ 4 années de période totale (durée d'une période 3 mois – nombre de périodes : 16)

Ces dimensionnements sont à vérifier avant d'appliquer les triangles. C'est ce qui a été fait sur Oscar et sur l'Australie.

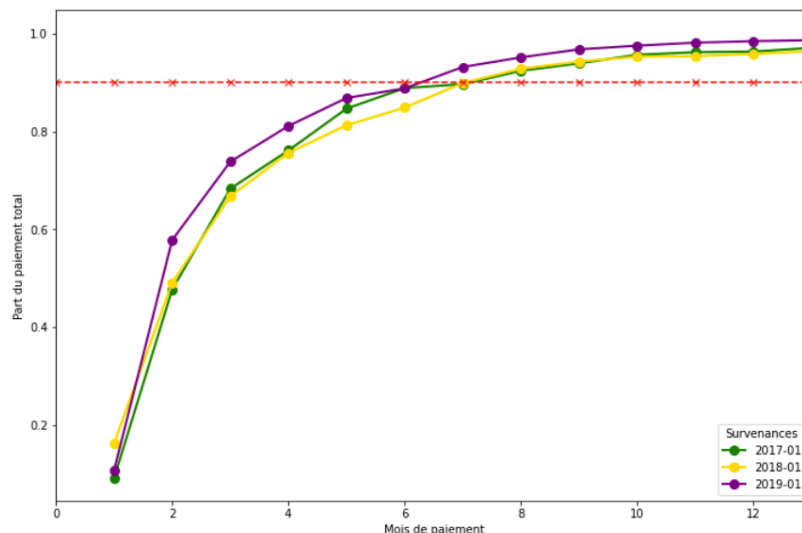


Figure 30 - Cadences réelles des paiements pour les survenances de janvier des années 2017, 2018, 2019 sur Oscar

La cadence de paiement se calcule de la manière suivante : pour chaque mois de paiement observé sur les survenances considérées (Janvier 2017, 2018, et 2019), le montant cumulé de paiement observé est divisé par l'ultime (montant de paiements cumulés total dernièrement observés, dans le cas d'Oscar, en date de juin 2020).

Notons que 90% des paiements versés (indemnisation des sinistres) sont effectués 7 mois après la survenance de l'accident, ce qui correspond à une cadence assez rapide. Effectivement, l'aspect digital de l'InsurTech Oscar permet des règlements rapides.

Pour conclure, une maille mensuelle est choisie pour Oscar, avec 12 périodes, soit un an en totalité. En ce qui concerne l'Australie, un choix de maille trimestrielle a été effectué avec 16 périodes, soit 4 ans en totalité.

Chain Ladder Classique

Soit $C_{i,j}$, le montant total cumulé des sinistres de l'année d'accident i , $1 < i < I$, qu'ils soient payés ou encourus jusqu'à l'année de développement j , $1 < j < J$. On considère $C_{i,j}$ comme une variable aléatoire dont on a une observation si $i + j < I + 1$ (triangle de liquidation). L'objectif est d'estimer le montant des sinistres ultimes et la réserve pour sinistres en suspens notée :

$$R_i = C_{i,J} - C_{i,I-i}$$

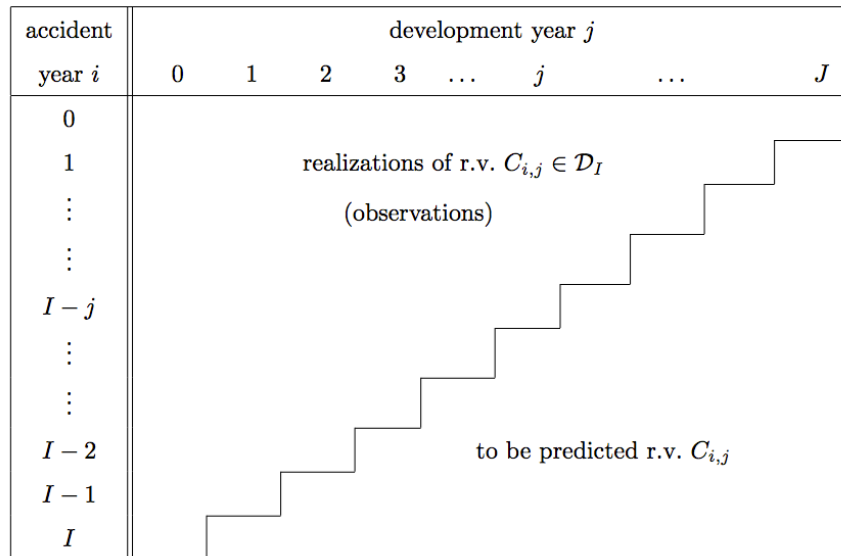


Figure 31 - Triangle cumulé théorique

Le Chain Ladder classique repose sur deux hypothèses notées H_1 et H_2 .

Tout d'abord, étant donné que l'algorithme du Chain Ladder ne tient pas compte des corrélations entre les années d'accident les variables $C_{i,j}$ des différentes années d'accident i , on note :

H_1 : Les paiements cumulés $C_{i,j}$ sont indépendants entre les années d'accident i , i.e. :

$$\forall i, j \{C_{i,1}, \dots, C_{i,j}\}, \{C_{j,1}, \dots, C_{j,j}\} \text{ sont indépendants.}$$

Dans la pratique, l'indépendance des années d'accident peut être faussée par certains effets durant l'année, comme des changements majeurs dans le traitement des sinistres.

Aussi, on note :

H_2 : Pour l'année d'accident i , il existe des facteurs de développement f_j tels que

$$E(C_{i,j+1} | C_{i,0}, \dots, C_{i,j}) = f_j \cdot C_{i,j}$$

La méthode du Chain Ladder consiste à estimer ces f_j par :

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

L'estimation des \hat{f}_j permettra de remonter de proche en proche au $C_{i,j}$ à partir des $C_{i,I-i}$, et in-fine d'effectuer le calcul de la réserve ultime, la PSAP.

Introduction du Sliding Chain Ladder

Un sliding Chain Ladder correspond à un Chain Ladder Classique qui a la particularité de calculer ses coefficients de passage via une fenêtre de dimension fixe reculant/se déplaçant (« sliding ») dans les données du passé à chaque estimation de \hat{f}_j .

Exemple d'un sliding Chain Ladder de 2 périodes effectué sur triangle de 3 périodes :

| Cumulatif | Paieiment j | Paieiment j+1 | Paieiment j+2 | Paieiment j+3 |
|----------------|-------------|---------------|---------------|---------------|
| Survenance i | $C_{i,j}$ | $C_{i,j+1}$ | $C_{i,j+2}$ | $C_{i,j+3}$ |
| Survenance i+1 | $C_{i+1,j}$ | $C_{i+1,j+1}$ | $C_{i+1,j+2}$ | |
| Survenance i+2 | $C_{i+2,j}$ | $C_{i+2,j+1}$ | | |
| Survenance i+3 | $C_{i+3,j}$ | | | |

Figure 32 - Calcul de \hat{f}_j pour le sliding Chain Ladder

Ici est calculé un « 2-month sliding Chain Ladder », ce qui correspond donc à une fenêtre fixe (en rouge) de 2 mois de profondeur de survénance. Le coefficient de passage \hat{f}_j sera calculé en faisant la division de la somme des éléments la fenêtre A_1 par la somme des éléments de la fenêtre B_1 .

Dimensionnement de la fenêtre du Sliding Chain Ladder

Tout d'abord, il faut dimensionner la profondeur de survénance des deux fenêtres fixes permettant le calcul des coefficients de passage. Le critère de choix majeur est la cohérence des données prises dans ces fenêtres. Par exemple, si les survénances anciennes présentent une cadence lente (flèches oranges), alors que les survénances récentes présentent une cadence plus rapide (flèches vertes), il faut adapter la taille des fenêtres afin de calculer un coefficient de passage cohérent avec les dernières données.

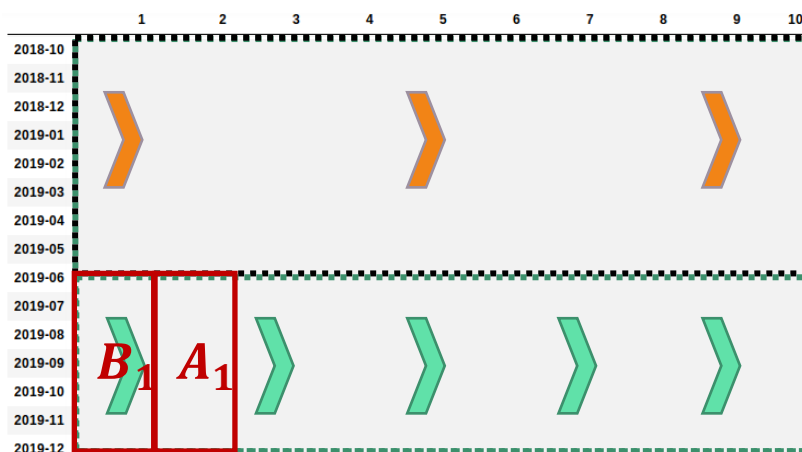


Figure 33 - Dimensionnement des fenêtres sur un triangle ayant des cadences hétérogènes

Dans notre étude, pour Oscar, certaines technologies utilisées dans la gestion des sinistres ont grandement modifié les cadences de paiement. En conséquence, la prise en compte de données « trop » antérieures ne correspond plus aux cadences de paiements actuelles. En revanche, le calcul du coefficient de passage avec une fenêtre « courte » (prenant en compte seulement quelques mois récents de survenance) peut être faussé par la volatilité des mois considérés. En effet, nous avons vu sur Oscar un phénomène de sinistralité accentué sur les derniers mois de l'année, ce qui ne reflète pas le comportement global du business.

Ce choix de profondeur de fenêtre s'est fait graphiquement sur Oscar ainsi que sur l'Australie.

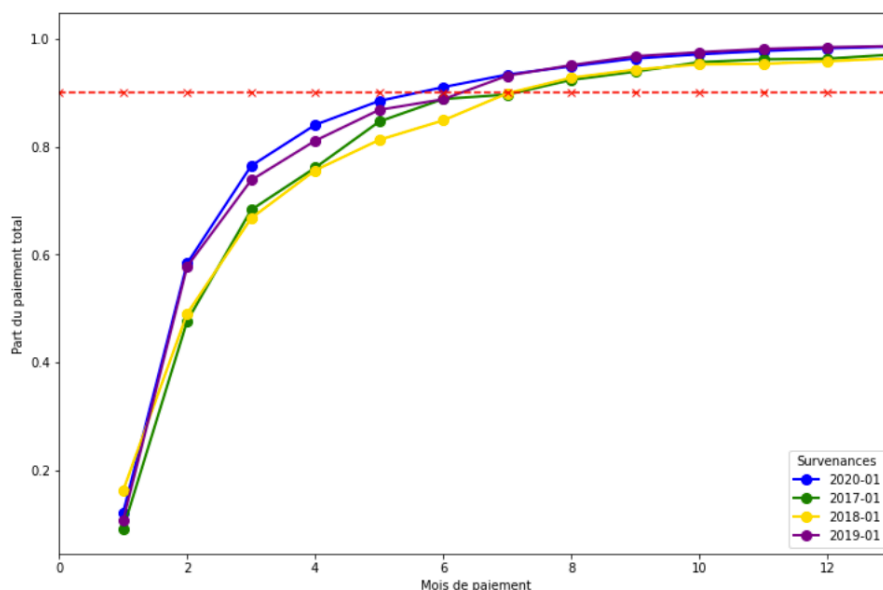


Figure 34 - Cadences de paiement pour les survenances de Janvier 2017, 2018, 2019 et 2020

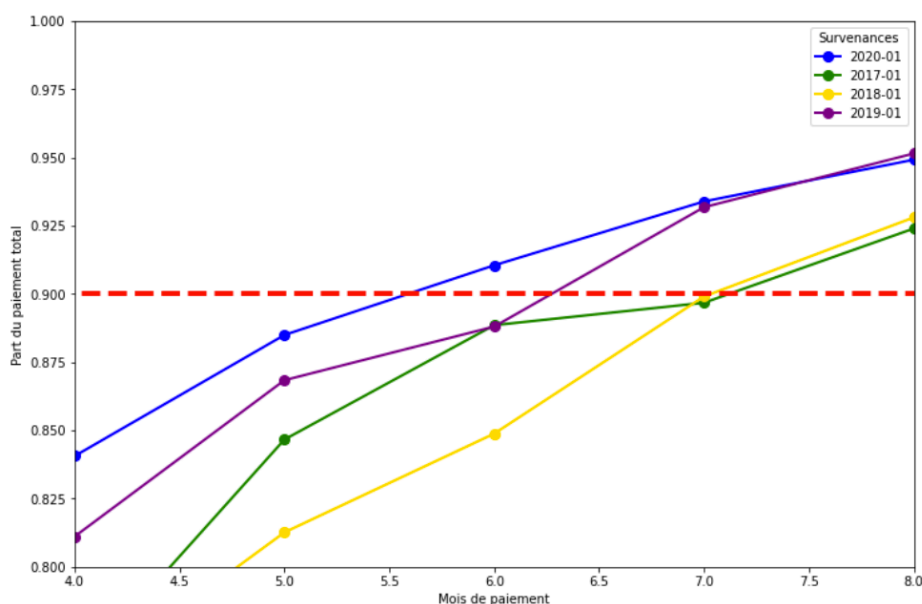


Figure 35 - Cadences de paiement pour les survenances de Janvier 2017, 2018, 2019 et 2020 (zoom)

Différentes fréquences de survenances ont été étudiées (tous les 6 mois, tous les 12 mois, tous les 18 mois) afin de constater au bout de combien de mois de paiement le seuil des 90% de règlement total était atteint.

Les survenances étudiées tous les 6 mois ne montraient pas une différence de cadence notable. Les survenances étudiées tous les 18 mois présentaient des écarts de cadence trop importants (> 1 mois de paiement de gagné).

Pour des survenances étudiées une fois par an (12 mois de profondeur de fenêtre, voir ci-dessus), une accélération de la cadence raisonnable est observée (environ 0.5 mois gagné par an). La profondeur de 12 mois sera donc adoptée sur Oscar.

En ce qui concerne l'Australie, une profondeur de 12 trimestres sera adoptée.

Revue du dimensionnement du triangle du Sliding Chain Ladder

Un problème de dimension du triangle peut être rencontré sur les lags de paiement supérieurs (j grand). Il faut alors revoir la taille de celui-ci avant de tester les hypothèses du Chain Ladder et d'appliquer la méthode du Sliding Chain Ladder.

Reprenons l'exemple du « 2-month sliding Chain Ladder », le coefficient de passage \widehat{f}_{j+2} sera calculé en faisant la division de la somme des paiements cumulés la fenêtre A_3 par la somme des paiements cumulés de la fenêtre B_3 .

Ce dernier calcul pose un problème de dimension du triangle. En effet, il manque dans chaque fenêtre A_3 et B_3 , le paiement cumulé correspondant à l'année de survenance $i - 1$. Il faut donc revoir le dimensionnement du triangle afin que le calcul du dernier coefficient de passage puisse se faire par rapport à la profondeur de survenance choisie dans la fenêtre du sliding Chain Ladder.

| Cumulatif | Paiement j | Paiement j+1 | Paiement j+2 | Paiement j+3 |
|-------------------|-------------|---------------|---------------|--------------|
| Survenance i | $C_{i,j}$ | $C_{i,j+1}$ | $C_{i,j+2}$ | $C_{i,j+3}$ |
| Survenance i+1 | $C_{i+1,j}$ | $C_{i+1,j+1}$ | $C_{i+1,j+2}$ | |
| Survenance i+2 | $C_{i+2,j}$ | $C_{i+2,j+1}$ | | |
| Survenance i+3 | $C_{i+3,j}$ | | | |

Figure 36 - Calcul de \widehat{f}_{j+2} pour le sliding Chain Ladder

Sur Oscar, un « 12-months sliding » Chain Ladder sera appliqué. Il est donc nécessaire de revoir la profondeur de survenance du triangle initialement dimensionnée à 12 périodes de 1 mois. Le calcul du dernier coefficient f_{11} nécessite donc 12 périodes absentes du triangle, en conséquence, il faut rajouter 12 périodes de survenances à ce dernier. Les dimensions du triangle Oscar seront alors les suivantes : 12 mois de paiement et 24 mois de survenance.

La révision du triangle australien ne se fait pas car les derniers mois de paiement sont calculés à l'aide d'une fonction LOGIT sur Excel.

Cet ajustement LOGIT suit la directive demandée par les équipes techniques d'AXA France en raison du manque d'expérience des sinistres de longue queue non encore observés dans le portefeuille.

L'objectif de cet ajustement est de ne pas sous-estimer le nombre final de sinistres en négligeant les sinistres potentiels. En particulier pour les prestations IP qui peuvent être en « run-off » après plusieurs années, ou encore les sinistres rouverts qui apparaissent dans certains portefeuilles.

Cette méthodologie d'ajustement est encore en cours de test pour valider son adéquation aux données australiennes. Elle ne sera pas appliquée dans ce mémoire.

Validation des hypothèses du Chain Ladder

Pour rappel, les deux hypothèses du Chain Ladder sont les suivantes :

H_1 : Les paiements cumulés $C_{i,j}$ sont indépendants entre les années d'accident i , i.e. :
 $\forall i, j \{C_{i,1}, \dots, C_{i,j}\}, \{C_{j,1}, \dots, C_{j,j}\}$ sont indépendants.

Cette hypothèse est testée via un test de corrélation de Spearman dans la pratique car il n'implique aucune distribution sous-jacente. Ce test présenté par Thomas Mack est contenu dans le package *ChainLadder* et permet de vérifier cette hypothèse.

```
t.development_correlation(p_critical=.05).t_critical.values
```

H_2 : Pour l'année d'accident i , il existe des facteurs de développement f_j tels que

$$E(C_{i,j+1} | C_{i,0}, \dots, C_{i,j}) = f_j \cdot C_{i,j}$$

Cette hypothèse est souvent vérifiée graphiquement, en affichant les $C_{i,j+1} = fonction(C_{i,j})$. La validation de cette hypothèse implique l'existence d'une droite passant par l'ensemble des points générés par les couples $\{C_{i,j}; C_{i,j+1}\}$ ainsi que par l'origine du repère.

Validation des hypothèses en maille agrégée (Oscar - tout états confondus – triangle fin 2019):

Une validation d'hypothèses sur un unique triangle (agrégation des données pour les 4 états étudiés : New-York, Californie, Texas, et New-Jersey) est d'abord effectuée.

H_1 : *Hypothèse validée (test de corrélation négatif)*

H_2 : *Graphiquement sur triangle agrégé*

Graphiquement, l'ensemble des points générés par les couples $\{C_{i,j}; C_{i,j+1}\}$ semblent être alignés sur une droite passant par l'origine du repère, sauf pour $j = 1$ où il semblerait y avoir l'existence d'un biais noté b_1 tel que $C_{i,2} = f_1 \cdot C_{i,1} + b_1$. Cette relation correspond à l'hypothèse de la méthode London-Chain, variante du Chain Ladder, ayant des hypothèses moins contraignantes.

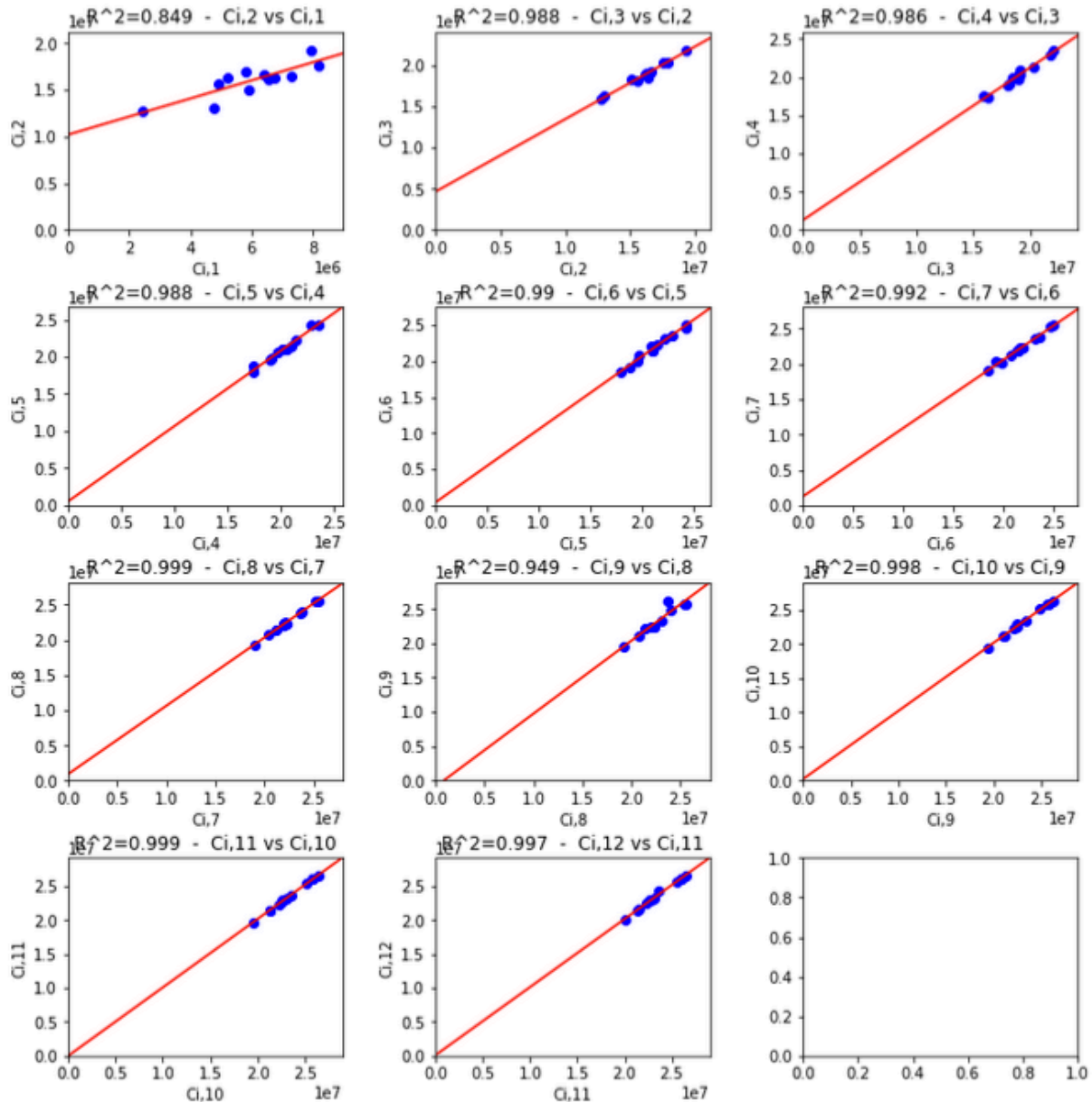


Figure 37 - Ensemble des points $\{C_{i,j}; C_{i,j+1}\}$ (Oscar - tout états confondus - triangle fin 2019)

Validation des hypothèses en maille territoriale (Oscar - états par états – triangle fin 2019):

Afin d'être plus confortable avec l'hypothèse H_2 , nous allons décomposer le triangle en triangles territoriaux. C'est à dire, un triangle pour l'état de New-York, un second pour l'état de Californie, un troisième pour l'état du Texas et un dernier pour l'état du New-Jersey.

Pour ce faire, il faut vérifier à nouveau pour chaque état les deux hypothèses H_1 et H_2 mentionnées précédemment.

H_1 : *Hypothèse validée pour chaque état (test de corrélation négatif)*

H_2 : *Graphiquement pour chaque état*

Ici est présenté le résultat de la vérification de l'hypothèse H_2 , pour l'état de New-York sur le triangle correspondant à l'étude de l'année 2019. Les résultats concernant la Californie, le Texas ainsi que le New-Jersey sont disponibles en Annexe. Il semblerait que nous sommes toujours en présence d'un biais pour $j = 1$.

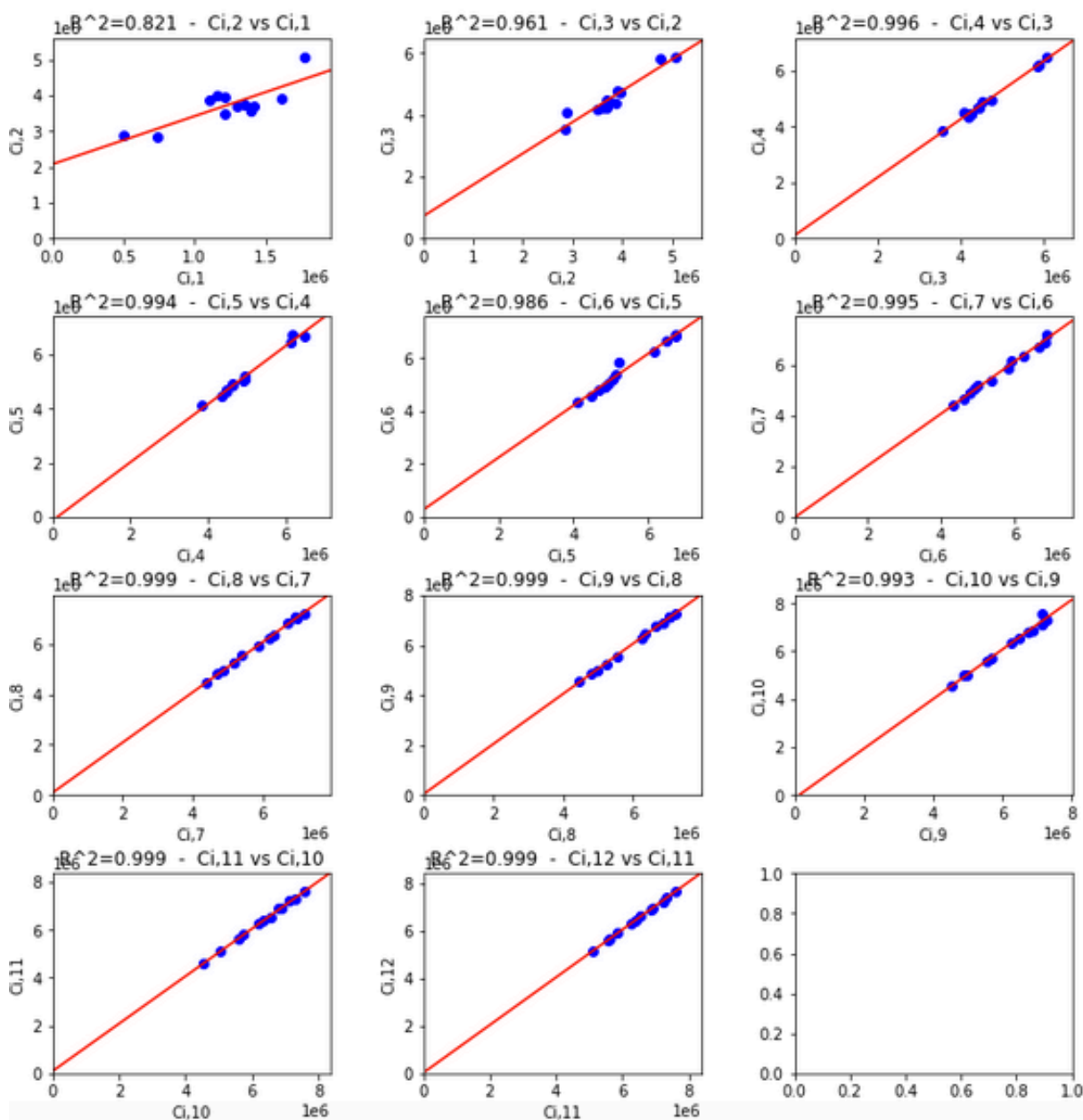


Figure 38 - Ensemble des points $\{C_{i,j}; C_{i,j+1}\}$ (Oscar - état de New-York - triangle fin 2019)

Afin de pouvoir comparer l'importance du biais pour $j = 1$, une visualisation des différents points $\{C_{i,j}; C_{i,j+1}\}$ est effectuée. Elle permet de mettre en regard les résultats observés sur le triangle agrégé ainsi que les résultats observés état par état.

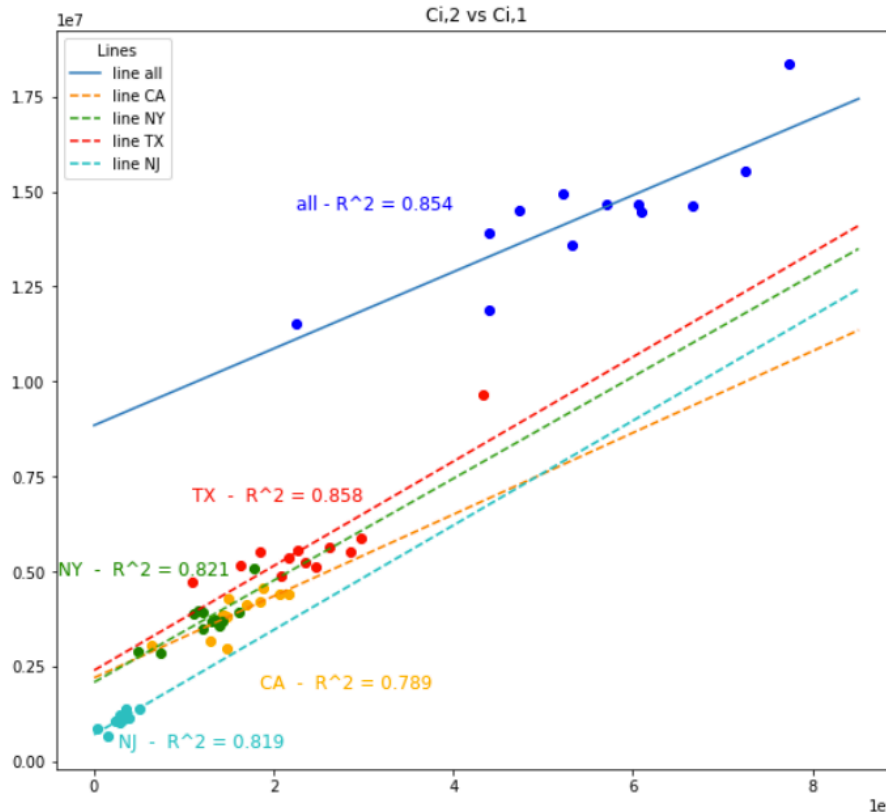


Figure 39 - Comparaison de l'ensemble des points $\{C_{i,j}; C_{i,j+1}\}$ (Oscar - triangle fin 2019)

Malgré un coefficient R^2 légèrement moins favorable par état, nous constatons une meilleure convergence vers l'origine du repère en maille territoriale (décomposition du triangle état par état). L'hypothèse H_2 du Chain Ladder Classique est plus confortable en décomposition territoriale. Cela permet de travailler sur des triangles de données plus homogènes.

Pour savoir si un travail sur une maille agrégée est tout de même possible, il faut se placer dans le cas de la méthode London Chain, qui autorise l'écriture affine de H_2 de la manière suivante :

$$C_{i,j} = f_j \cdot C_{i,j} + b_j$$

Un test statistique est alors à effectuer pour vérifier la significativité du facteur b_j . La maille territoriale étant satisfaisante, ce point ne sera pas développé.

Pour la suite, nous travaillerons donc sur des triangles décomposés par états pour Oscar.

Application du 12-months Sliding Chain Ladder et résultats sur Oscar

Voici les résultats bruts de la PSAP calculée pour chaque état Oscar sur l'année 2019.

| Méthode | New-York | Californie | Texas | New-Jersey |
|--------------------------------|--------------|--------------|--------------|-------------|
| 12-months Sliding Chain Ladder | \$15.805.570 | \$13.077.553 | \$14.542.715 | \$5.886.911 |

Figure 40 - Tableau des résultats bruts sur Oscar (12-months sliding Chain Ladder)

Sans point de comparaison, il est difficile de commenter ces résultats bruts. En revanche, il est intéressant de regarder la tendance générale qui se dégage à l'issue de l'application du Chain Ladder, lorsqu'on consolide sinistres payés et PSAP.

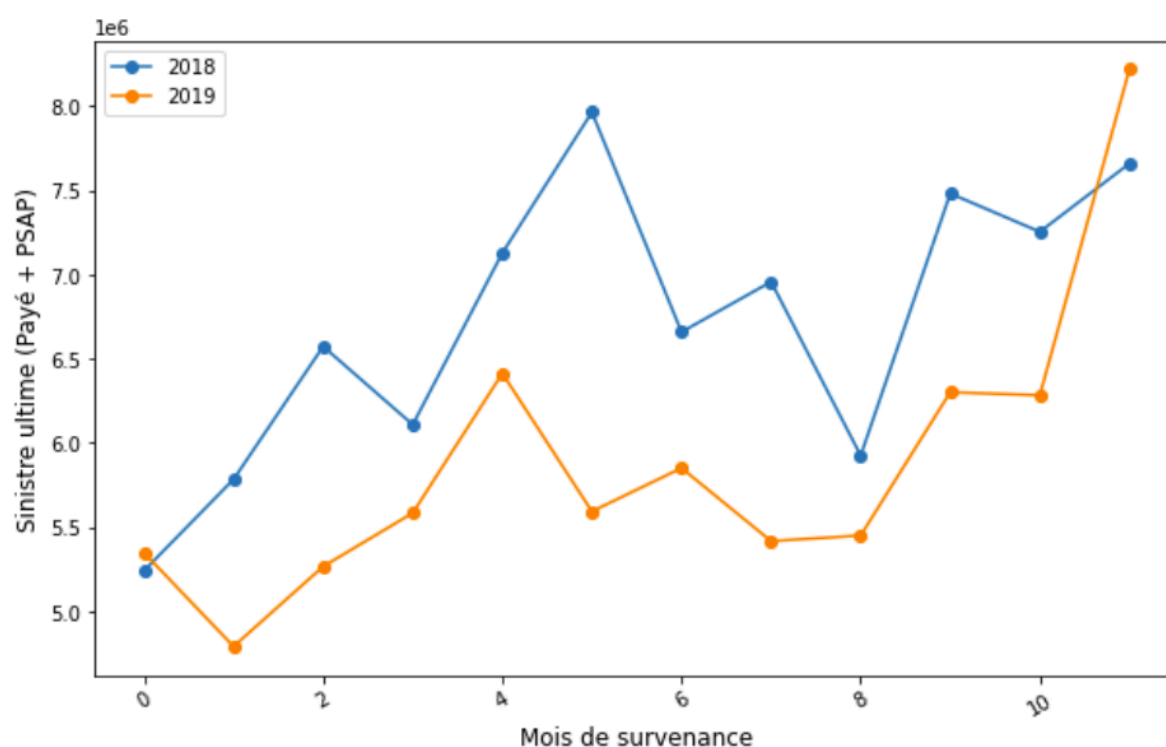


Figure 41 - Sinistres ultimes 2018 et 2019 - 12-months sliding Chain Ladder (Oscar – état de New-York – triangle fin 2019)

Lorsqu'on compare les tendances de sinistres, il est notable qu'une dégradation du portefeuille se fait au cours de l'année. Nous constatons les effets du « Maximum Out Of Pocket », largement présenté en partie 1.3.1. en page 22. Nous notons également une bosse de sinistralité entre les mois de survenance de mars et de juillet. Ce pic de sinistralité peut s'expliquer par la décroissance du nombre d'assurés au cours de l'année à partir du mois de février, présenté en page 24. Cette décroissance semble être plus impactante que le phénomène « Maximum Out Of Pocket » à partir d'un certain mois (mai pour les survenances de 2018, et avril pour les survenances de 2019). Puis, le « Maximum Out Of Pocket » reprend un poids majoritaire en fin d'année en terme de sinistralité.

Chain Ladder Classique - Commentaires sur l'Australie

L'estimation des réserves sur le périmètre australien ne se fait pas directement en montant, tel que cela peut être fait sur Oscar. Il faut tout d'abord estimer le nombre de sinistres survenus mais non encore déclarés avant de calculer un coût moyen de sinistre. Le produit du nombre de sinistres survenus mais non encore déclarés par le coût moyen de ces derniers permet alors de calculer la PSAP.

Ce process particulier (versus Oscar) vient du fait que le calcul du coût moyen varie fortement selon les années. Effectivement, le coût moyen dépend d'une multitude de facteurs sur les garanties prévoyance. Il ne dépend pas uniquement des coûts historiques. Certains facteurs sont à intégrer comme l'inflation médicale ainsi que des possibles nouvelles législations locales impactant parfois certains sinistres en cours de règlement.

Ce calcul de coût moyen de sinistres est tenu par les équipes marché ou les entités locales. Les calculs menés dans ce mémoire sur le périmètre Australien ont pour objectif de déterminer précisément le nombre de sinistres. Le calcul du coût moyen sera considéré hors scope.

Pour l'Australie, l'agrégation des données (nombre de sinistres) sur un unique triangle est tout d'abord effectuée. Cette agrégation permet de valider les hypothèses du Chain Ladder. En revanche, d'un point de vue pratique les calculs de coûts moyens sont effectués par garantie. En effet, les process de règlement des garanties incapacité diffèrent grandement des garanties décès et leur mélange n'est pas souhaitable.

Nous retiendrons donc pour la suite de l'étude Australienne une décomposition de triangle par garantie, pour la raison technique précédemment expliquée, malgré le fait que les hypothèses du Chain Ladder soient validées en maille agrégée.

| Méthode | TPD | Death | IP |
|--------------|-----|-------|-----|
| Chain Ladder | 385 | 260 | 186 |

Figure 42 - Tableau des résultats bruts sur l'Australie (Chain Ladder)

2.2.2 Le Chain Ladder de Mack (Chain Ladder Stochastique)

L'intérêt des méthodes stochastiques est de pouvoir quantifier la volatilité des réserves estimées, et d'obtenir une marge d'erreur sur le montant des provisions. La méthode du Chain Ladder de Mack, tout comme la méthode Bootstrap construisent les réserves sous une approche probabiliste en considérant leur distribution. L'intérêt de ces méthodes s'est renforcé suite à la mise en place de la réglementation Solvabilité II.

Cependant, le Chain Ladder de Mack est excessivement sensible aux facteurs de développement du triangle. De la même manière que pour le Chain Ladder Classique, une stabilité de ces facteurs est nécessaire afin de ne pas perturber l'estimation des réserves et de la quantification de leur volatilité.

Bien que cette méthode permette d'avoir une estimation de l'erreur de prédiction, elle impose à l'assureur de fixer lui-même la loi de distribution. Ce choix est souvent fait à travers une loi normale, qui n'est pas forcément représentative de la distribution, notamment lorsque des sinistres « extrêmes » sont enregistrés dans le portefeuille.

Chain Ladder de Mack

Le Chain Ladder de Mack repose sur trois hypothèses. Deux d'entre elles ont déjà été présentées et vérifiées précédemment avec le Chain Ladder Classique en page 54 (H_1 et H_2) :

H_1 : Les paiements cumulés $C_{i,j}$ sont indépendants entre les années d'accident i , i.e. :
 $\forall i, j \{C_{i,1}, \dots, C_{i,j}\}, \{C_{j,1}, \dots, C_{j,j}\}$ sont indépendants.

H_2 : Pour l'année d'accident i , il existe des facteurs de développement f_j tels que

$$E(C_{i,j+1} | C_{i,0}, \dots, C_{i,j}) = f_j \cdot C_{i,j}$$

H_3 : Il existe $\hat{\sigma}_1^2, \dots, \hat{\sigma}_{n-1}^2$ paramètres de variance tels que pour tous $i = 1, \dots, n$ et pour tous $j = 1, \dots, n - 1$, nous avons :

$$\text{Var}(C_{i,j+1} | C_{i,1}, \dots, C_{i,j}) = \text{Var}(C_{i,j+1} | C_{i,j}) = \hat{\sigma}_j^2 \cdot C_{i,j}$$

L'hypothèse H_3 introduit un estimateur de la variance $\hat{\sigma}_j$.

Sous ces trois hypothèses, il est possible d'appliquer le théorème de Thomas Mack qui permet d'estimer la volatilité des estimations de provisions (mean square error, notée mse), pour chaque année de survenance :

$$\widehat{mse}(\widehat{R}_i) = \hat{C}_{i,I}^2 \cdot \sum_{j=i+1}^{I-1} \frac{\hat{\sigma}_j^2}{\hat{f}_j^2} \cdot \left(\frac{1}{\hat{C}_{i,j}} + \frac{1}{\sum_{k=1}^{I-j} C_{k,j}} \right)$$

Avec possibilité de définir un estimateur sans biais de $\hat{\sigma}_j^2$:

$$\hat{\sigma}_j^2 = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)$$

On peut alors remonter à l'erreur totale d'estimation de réserve :

$$\widehat{mse}(\widehat{R}) = \sum_{i=1}^I \left[\widehat{mse}(\widehat{R}_i) + \hat{C}_{i,I} \left(\sum_{j=i+1}^I \hat{C}_{i,j} \right) \right]$$

Enfin, il faut choisir une fonction de distribution des provisions et il est alors possible de travailler en quantile de réserves.

Validation des hypothèses du Chain Ladder de Mack

Validation des hypothèses en maille territoriale (Oscar - états par états – triangle fin 2019):

H₁ : Hypothèse validée pour chaque état – voir en page 54

H₂ : Hypothèse validée pour chaque état – voir en page 54

H₃ : Hypothèse de l'existence d'un estimateur de la variance $\hat{\sigma}_j^2$

Cette hypothèse **H₃** est souvent vérifiée graphiquement, en affichant les résidus $\frac{C_{i,j+1} - \hat{f}_j \cdot C_{i,j}}{\sqrt{C_{i,j}}}$, avec $i = 1, \dots, n - j$, pour j fixé. Ces résidus ne doivent présenter aucune structure non aléatoire.

Ici est présenté le résultat de la vérification de l'hypothèse **H₃**, pour l'état de New-York sur le triangle correspondant à l'étude de l'année 2019. Les résultats concernant la Californie, le Texas ainsi que le New-Jersey sont disponibles en Annexe en page 112. Il semblerait que nous sommes en présence d'un biais pour les j de 1 à 4 (présence d'une structure non aléatoire).

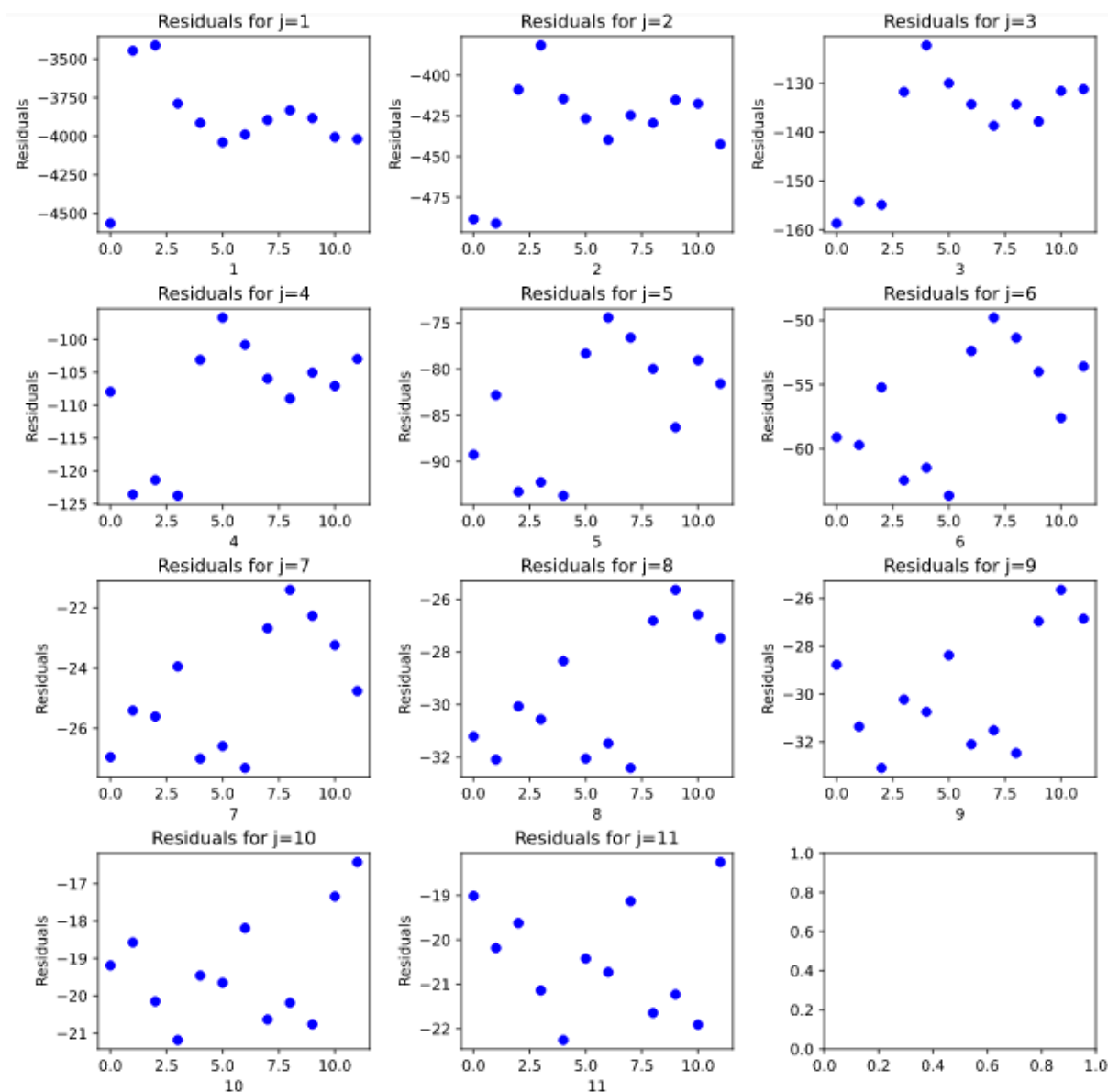


Figure 43 - Ensemble des résidus en fonction de i (Oscar – état de New-York – triangle fin 2019)

Nous considèrerons cette hypothèse H_3 comme validée, tout en gardant à l'esprit que l'estimation de l'erreur de Mack puisse présenter un biais. En effet, les lags correspondant aux j supérieurs à 4 semblent valider l'hypothèse. En revanche, pour les j inférieurs ou égaux à 4, l'hypothèse H_3 est moins confortable.

Application du Chain Ladder de Mack et résultats sur Oscar

Voici les résultats bruts de la PSAP calculée pour chaque état Oscar sur l'année 2019 ainsi que des erreurs associées.

| Méthode | New-York | Californie | Texas | New-Jersey |
|-------------------|----------------------------------|----------------------------------|----------------------------------|---------------------------------|
| Mack Chain Ladder | \$15.805.570 Err. \$2.013.853 | \$13.077.553 Err. \$2.480.865 | \$14.542.715 Err. \$3.043.340 | \$5.886.911 Err. \$4.618.650 |

Figure 44 - Tableau des résultats bruts sur Oscar (12-months sliding Mack Chain Ladder)

Les résultats correspondent logiquement aux PSAP déjà relevées sur la méthode du 12-months sliding Chain Ladder. En effet, c'est la même méthode de calcul de coefficients f_j . En revanche, nous avons accès à l'erreur de Mack ce qui permet d'avoir une première approche de la volatilité de l'estimation des réserves.

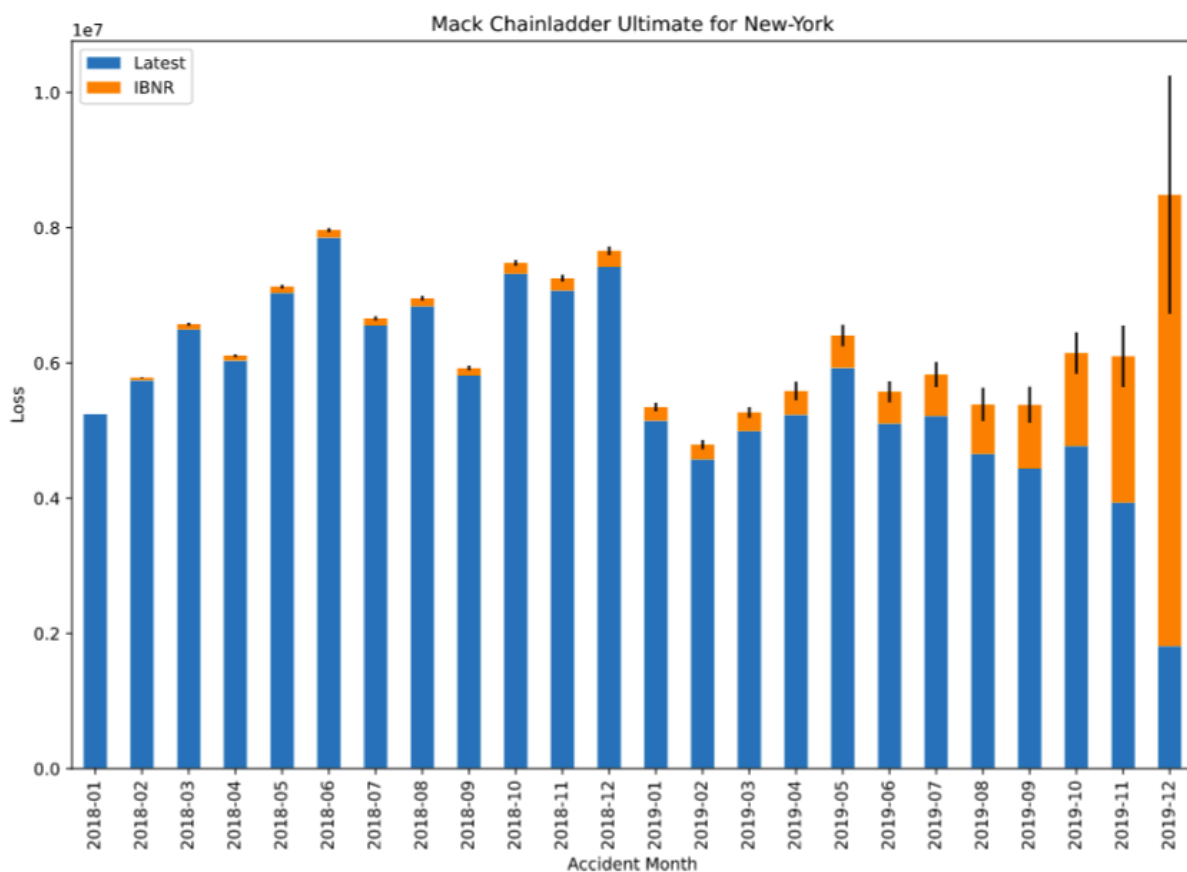


Figure 45 - Résultats par mois de survenance (état de New-York - Mack Chain Ladder)

Nous constatons donc que l'erreur de Mack est assez importante sur les mois les plus récents, puis elle devient insignifiante assez rapidement. En d'autres termes, la volatilité d'estimation se situe lorsque peu de données sont observées, ce qui est cohérent.

Avec la PSAP (moyenne) et l'erreur de Mack (variance), la méthode de Mack suppose que la distribution des réserves suit une loi normale et donne accès à des quantiles.

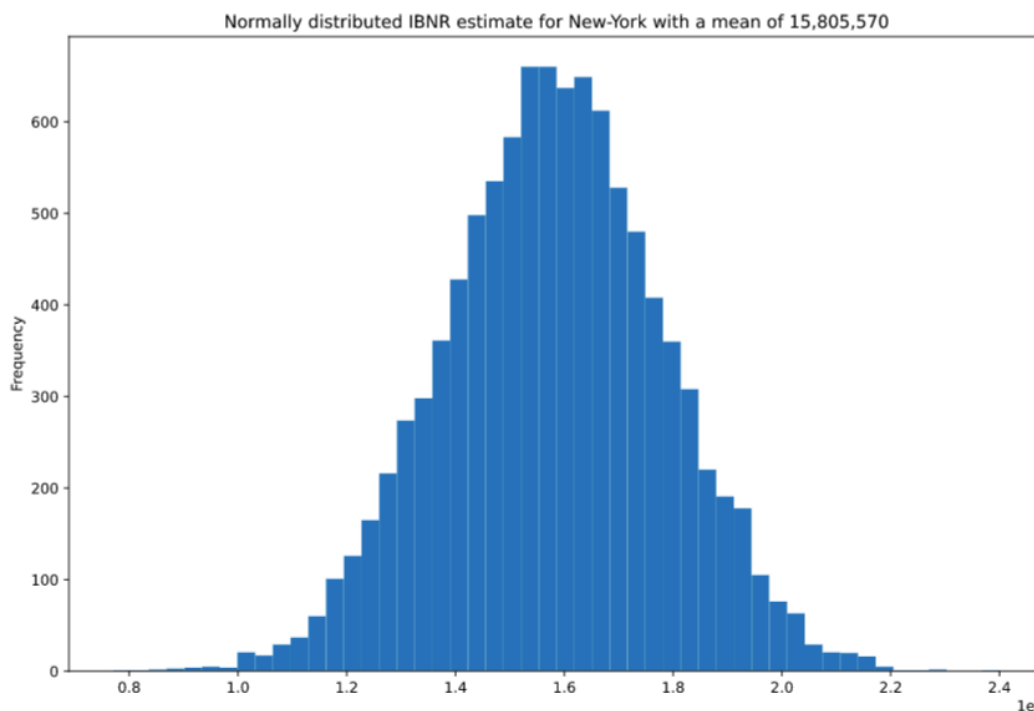


Figure 46 – PSAP totale de distribution supposée normale (état de New-York - Mack Chain Ladder)

Classiquement, le montant de PSAP *estimation (Best Estimate) + variance* est retenu. C'est un montant prudent qui correspond à une estimation Best-Estimate à laquelle est ajoutée une prudence. Dans le cas d'une loi Normale, ce montant correspond au quantile %. Ce montant est retenu pour chaque état et sera comparé aux montants proposés par les autres méthodes, dans une prochaine partie.

Chain Ladder de Mack - Commentaires sur l'Australie

Tout d'abord, l'hypothèse H_3 est bien plus confortable en Australie. En effet, sur les trois triangles (garanties TPD, Death et IP), les résidus présentent un nuage de point pour tous les j considérés, représentatifs d'une structure aléatoire.

Ensuite, les résultats sont comme mentionnés pour Oscar, identiques au Chain Ladder Classique en ce qui concerne l'estimation de la PSAP. En revanche, nous avons accès à l'erreur d'estimation de Mack. Cette erreur demeure importante sur plusieurs trimestres de survenance, notamment sur la garantie Décès. Cela est cohérent par rapport aux types de garanties présentes sur ce portefeuille : garanties qualifiées de branches longues, ce qui signifie que certains sinistres sont déclarés tardivement, après plusieurs trimestres d'observation de sinistres. Ces déclarations très tardives apportent de la volatilité sur les survenances antérieures.

| Méthode | TPD | Death | IP |
|------------|---------|---------|----------|
| Mack Chain | 385 | 260 | 186 |
| Ladder | Err. 55 | Err. 60 | Err. 127 |

Figure 47 - Tableau des résultats bruts sur l'Australie (Mack Chain Ladder)

Note: Pour rappel, les résultats du périmètres Australien sont donnés en nombre de sinistres, comme expliqué en page 59.

Aussi, une volatilité forte est notée sur la garantie IP. Cela est certainement dû au fait que cette garantie s'est ouverte en 2014 avec très peu de sinistres jusqu'à la mi-2015, puis au second semestre de cette même année est survenu près de 10 fois plus de sinistres (constaté au Q3 2015). Cette augmentation soudaine du business est due à l'arrivée de la cédante IOOF Frontbook sur cette garantie. L'exposition au risque étant dès lors plus grande, les sinistres ont augmenté.

| | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 27 | 30 |
|---------------|---------|---------|---------|---------|--------|--------|--------|--------|--------|--------|
| 2014Q1 | 1.0000 | 5.0000 | 6.0000 | 6.0000 | 7.0000 | 7.0000 | 7.0000 | 7.0000 | 7.0000 | 7.0000 |
| 2014Q2 | | 3.0000 | 3.0000 | 3.0000 | 4.0000 | 4.0000 | 4.0000 | 4.0000 | 6.0000 | |
| 2014Q3 | 1.0000 | 4.0000 | 5.0000 | 5.0000 | 5.0000 | 6.0000 | 6.0000 | 7.0000 | | |
| 2014Q4 | 1.0000 | 4.0000 | 5.0000 | 5.0000 | 5.0000 | 5.0000 | 5.0000 | | | |
| 2015Q1 | 1.0000 | 2.0000 | 3.0000 | 4.0000 | 4.0000 | 4.0000 | | | | |
| 2015Q2 | 3.0000 | 4.0000 | 4.0000 | 4.0000 | 4.0000 | | | | | |
| 2015Q3 | 8.0000 | 37.0000 | 42.0000 | 52.0000 | | | | | | |
| 2015Q4 | 15.0000 | 52.0000 | 64.0000 | | | | | | | |
| 2016Q1 | 9.0000 | 41.0000 | | | | | | | | |
| 2016Q2 | 13.0000 | | | | | | | | | |

Figure 48 - Triangle observé en garantie TPD

2.2.3 La méthode Bootstrap

La méthode Bootstrap, largement appréciée pour sa robustesse, est une méthode qui consiste à « fabriquer » de l'information à partir des sinistres connus et observés ainsi que ceux prédits par la méthode du Chain Ladder. C'est le principe du ré-échantillonnage.

Cette méthode est considérée comme plus flexible que d'autres méthodes stochastiques de reserving. En particulier, aucune hypothèse concernant la distribution sous-jacente des réserves n'est demandée.

Dans ce mémoire, nous appliquerons le Bootstrap sous les trois hypothèses du Chain Ladder de Mack, vérifiées précédemment en page 61.

Il faut cependant que les variables $C_{i,j}$ (observées et estimées) du triangle complet initial soient indépendantes et identiquement distribuées.

Cette dernière hypothèse ne se vérifie pas sur les paiements du triangle (non mutuellement indépendants). Il faut donc travailler avec les résidus de Pearson.

Les étapes clés du Bootstrap

Le Bootstrap appliqué via la méthode du Chain Ladder Classique peut se résumer en 7 étapes clés. En voici la description:

1. Détermination des \hat{f}_j via la méthode du Chain Ladder Classique

| Cumulatif | Paiement j | Paiement j+1 | Paiement j+2 | Paiement j+3 |
|----------------|-------------|---|---|---|
| Survénance i | $C_{i,j}$ | $C_{i,j+1}$ | $C_{i,j+2}$ | $C_{i,j+3}$ |
| Survénance i+1 | $C_{i+1,j}$ | $C_{i+1,j+1}$ | $C_{i+1,j+2}$ | |
| Survénance i+2 | $C_{i+2,j}$ | $C_{i+2,j+1}$ | | |
| Survénance i+3 | $C_{i+3,j}$ | | | |
| Facteurs | - | $\hat{f}_j = \sum_{i=1}^{l-j} C_{i,j+1} / \sum_{i=1}^{l-j} C_{i,j}$ | $\hat{f}_{j+1} = \sum_{i=1}^{l-j} C_{i,j+2} / \sum_{i=1}^{l-j} C_{i,j+1}$ | $\hat{f}_{j+2} = \sum_{i=1}^{l-j} C_{i,j+3} / \sum_{i=1}^{l-j} C_{i,j+2}$ |

Figure 49 – Calcul des \hat{f}_j via la méthode du Chain Ladder Classique pour le Bootstrap

2. Calcul des estimateurs de $C_{i,j}$ pour le triangle supérieur, à partir de la diagonale

| Cumulatif | Paieiment j | Paieiment j+1 | Paieiment j+2 | Paieiment j+3 |
|----------------|---|---|---|-----------------------------------|
| Survenance i | $\widehat{C}_{i,j} = \frac{\widehat{C}_{i,j+1}}{\widehat{f}_j}$ | $\widehat{C}_{i,j+1} = \frac{\widehat{C}_{i,j+2}}{\widehat{f}_{j+1}}$ | $\widehat{C}_{i,j+2} = \frac{C_{i,j+3}}{\widehat{f}_{j+2}}$ | $\widehat{C}_{i,j+3} = C_{i,j+3}$ |
| Survenance i+1 | $\widehat{C}_{i+1,j} = \frac{\widehat{C}_{i+1,j+1}}{\widehat{f}_j}$ | $\widehat{C}_{i+1,j+1} = \frac{C_{i+1,j+2}}{\widehat{f}_{j+1}}$ | $\widehat{C}_{i+1,j+2} = C_{i+1,j+2}$ | |
| Survenance i+2 | $\widehat{C}_{i+2,j} = \frac{C_{i+2,j+1}}{\widehat{f}_j}$ | $\widehat{C}_{i+2,j+1} = C_{i+2,j+1}$ | | |
| Survenance i+3 | $\widehat{C}_{i+3,j} = C_{i+3,j}$ | | | |

Figure 50 – Calcul des $\widehat{C}_{i,j}$ via la méthode du Chain Ladder Classique pour le Bootstrap

3. Retour à un triangle incrémental

| Incrémental | Paieiment j | Paieiment j+1 | Paieiment j+2 | Paieiment j+3 |
|----------------|---|---|---|---|
| Survenance i | $\widehat{X}_{i,j} = \widehat{C}_{i,j}$ | $\widehat{X}_{i,j+1} = \widehat{C}_{i,j+1} - \widehat{C}_{i,j}$ | $\widehat{X}_{i,j+2} = \widehat{C}_{i,j+2} - \widehat{C}_{i,j+1}$ | $\widehat{X}_{i,j+3} = \widehat{C}_{i,j+3} - \widehat{C}_{i,j+2}$ |
| Survenance i+1 | $\widehat{X}_{i+1,j} = \widehat{C}_{i+1,j}$ | $\widehat{X}_{i+1,j+1} = \widehat{C}_{i+1,j+1} - \widehat{C}_{i+1,j}$ | $\widehat{X}_{i+1,j+2} = \widehat{C}_{i+1,j+2} - \widehat{C}_{i+1,j+1}$ | |
| Survenance i+2 | $\widehat{X}_{i+2,j} = \widehat{C}_{i+2,j}$ | $\widehat{X}_{i+2,j+1} = \widehat{C}_{i+2,j+1} - \widehat{C}_{i+2,j}$ | | |
| Survenance i+3 | $\widehat{X}_{i+3,j} = \widehat{C}_{i+3,j}$ | | | |

Figure 51 – Calcul des $\widehat{X}_{i,j}$ via la méthode du Chain Ladder Classique pour le Bootstrap

4. Calcul des résidus de Pearson. $\forall i + j < I$, on a:

$$r_{i,j} = \frac{X_{i,j} - \widehat{X}_{i,j}}{\sqrt{\widehat{X}_{i,j}}}$$

5. N itérations du Bootstrap

Le Bootstrap comporte N itérations (paramètre renseigné par l'utilisateur). Pour chaque itération, chaque cellule du triangle supérieur prend une valeur parmi les résidus précédemment calculés (tirage aléatoire avec remise). Un échantillon peut alors comporter plusieurs fois la même valeur de résidus, dans ses différentes cellules.

6. Calcul des N triangles incrémentaux

A l'aide de la formule présentée en 4. il est possible de repasser en triangle incrémental. $\forall i + j < I$, on a:

$$X_{i,j}^* = r_{i,j}^* \cdot \sqrt{\widehat{X}_{i,j} + \widehat{X}_{i,j}}$$

7. Calcul des N triangles cumulatifs supérieurs et estimation des réserves pour les N triangles

De la même manière que cela a été présenté pour le Chain Ladder Classique précédemment, les réserves sont calculées à partir des triangles cumulatifs supérieurs.

Ces calculs génèrent donc N estimations de réserves ce qui permet d'avoir accès à des quantiles de réserves, ne reposant sur aucune distribution préalablement déterminée. Lorsque N est grand, la moyenne de ces réserves doit converger vers celle proposée par la méthode de Mack.

Application du Bootstrap Chain Ladder et résultats sur Oscar

Voici les résultats bruts de la PSAP calculée pour chaque état Oscar sur l'année 2019 avec la méthode Bootstrap (Quantile 50% et 70%). Nous avons choisi N=10.000 simulations.

| Méthode | | New-York | Californie | Texas | New-Jersey |
|--------------|--------------|--------------|--------------|--------------|-------------|
| Bootstrap | Quantile 50% | \$15.691.523 | \$12.858.585 | \$14.315.412 | \$5.810.993 |
| Chain Ladder | Quantile 70% | \$16.280.774 | \$13.328.873 | \$14.898.043 | \$6.054.747 |

Figure 52 - Tableau des résultats bruts sur Oscar (Bootstrap Chain Ladder)

Le choix du quantile 70% est un choix usuel chez AXA France, qui permet d'avoir une vision Best-Estimate (quantile 50%) avec une marge de prudence assez importante. Ce quantile peut être modulé selon les besoins spécifiques.

Comme mentionné précédemment, le quantile 50% du Bootstrap doit être très proche de l'estimation moyenne proposée par le Mack Chain Ladder. C'est pourquoi, nous visualisons ici les résultats du Bootstrap au regard de ceux obtenus avec la méthode de Mack.

Nous constatons que le Bootstrap réduit la volatilité de la distribution de réserves PSAP par rapport au résultat obtenu avec la méthode de Mack. Nous obtenons une distribution légèrement moins symétrique que celle proposée par la loi Normale (de paramètres : Moyenne = PSAP estimée, avec Variance = Erreur de Mack). En conséquence, dans le cas étudié de New-York, le montant *Estimation(Mack Chain Ladder) + Erreur de Mack* est plus prudent qu'un Bootstrap 70%. Ce point sera revu et détaillé dans la partie dédiée à la comparaison des méthodes.

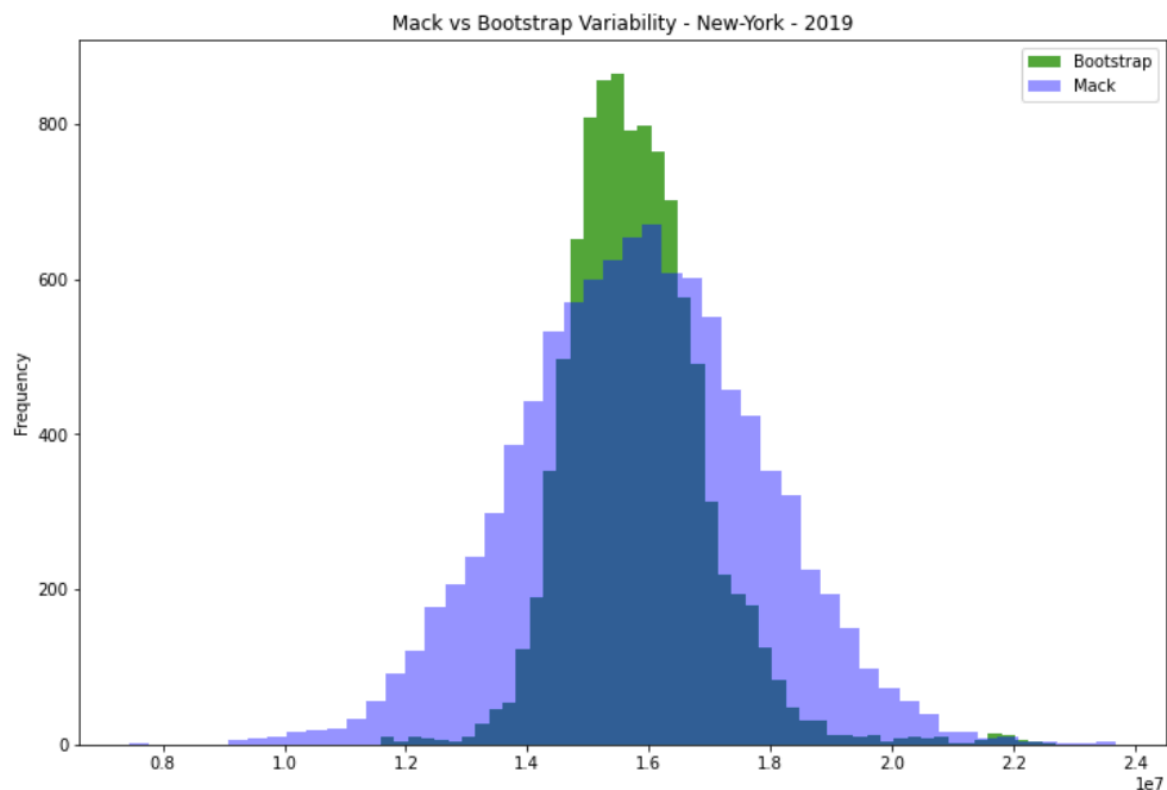


Figure 53 – Résultats Bootstrap mis au regard du Mack Chain Ladder (état de New-York – 2019)

Pour les autres états composants le portefeuille Oscar, les résultats renforcent les analyses effectuées sur New-York. La réduction de la volatilité est assez forte avec une asymétrie de la distribution des réserves proposées par le Bootstrap (coefficient de Skewness positif). La distribution des réserves selon une loi Normale est certainement un choix qui ne reflète pas la réalité.

Bootstrap Chain Ladder - Commentaires sur l'Australie

En ce qui concerne l'Australie, nous obtenons également une réduction de la volatilité par rapport au Mack Chain Ladder pour les garanties Death et IP.

En revanche, le résultat obtenu par la méthode Bootstrap pour la garantie TPD est quasiment identique à celui de la distribution normale proposé par la méthode de Mack. Une distribution normale des réserves semble être adaptée pour cette garantie.

| Méthode | | TPD | Death | IP |
|--------------|--------------|-----|-------|-----|
| Bootstrap | Quantile 50% | 383 | 259 | 186 |
| Chain Ladder | Quantile 70% | 409 | 269 | 195 |

Figure 54 - Tableau des résultats bruts sur l'Australie (Bootstrap Chain Ladder)

Note: Pour rappel, les résultats du périmètres Australien sont donnés en nombre de sinistres, comme expliqué en page 59.

Comme mentionné précédemment, le quantile 50% du Bootstrap doit être très proche de l'estimation moyenne proposée par le Mack Chain Ladder. C'est une égalité pour la garantie TPD (383) ainsi que pour la garantie IP (186), et quasiment égale pour la garantie Décès (260 pour Mack vs. 259 pour le Bootstrap).

2.2.4 Conclusion sur les méthodes classiques

Nous avons appliqué des méthodes classiques parmi les plus usuelles à savoir le Chain Ladder Classique (12-months Sliding Chain Ladder pour le business Oscar), le Chain Ladder de Mack et enfin le Chain Ladder Bootstrap.

Il est à noter que les méthodes stochastiques permettent une quantification d'erreur. Dans le cas du Chain Ladder de Mack, la loi de distribution des réserves doit être renseignée, alors que le Bootstrap ne nécessite pas ce paramètre, ce qui le rend plus flexible.

D'autres méthodes classiques auraient pu être traitées dans cette partie, mais l'objet de ce mémoire n'est pas de faire une revue de toutes ces méthodes. L'intérêt de cette partie est plutôt d'avoir un point de comparaison avec le provisionnement du Marché ALHIS qui utilise majoritairement des Chain Ladder, tout en présentant la multiplication des outils dans le cas d'utilisation de méthodes classiques.

En effet, nous avons vu à travers cette partie que pour des portefeuilles volumineux comme Oscar ou l'Australie, les méthodes usuelles utilisant les triangles de règlement reposent bien souvent sur des hypothèses d'homogénéité des paiements cumulés assez fortes. C'est pourquoi nous avons adopté une décomposition territoriale sur Oscar par exemple. Cette multiplication de triangles peut être lourde à traiter.

Deuxièmement, bien souvent, les triangles varient en taille selon les garanties. Aussi, les méthodes utilisées, bien que regroupées sous le nom de Chain Ladder, fluctuent. Un 12-month Sliding Chain Ladder est utilisé sur Oscar tandis qu'un Chain Ladder standard est appliqué sur l'Australie. Dernièrement, les entités locales en Australie sont passées à une méthode dite additive, suite au changement de législation PYS présenté en page 38. Ce changement au local va probablement amener une révision de méthode de provisionnement PSAP sur ce périmètre dans les prochains mois à la DAAP.

Pour l'ensemble des particularités de provisionnement PSAP liées à chaque business qui viennent d'être mentionnées, il paraît difficile d'utiliser une méthode dite classique comme méthode universelle.

La méthode CANN, présentée au chapitre suivant, devient alors intéressante à étudier.

2.3 LA PROPOSITION MIXTE : LA MÉTHODE CANN

Cette partie s'appuie sur l'article *Neural Networks Applied to Chain-Ladder Reserving* de Mario Wüthrich (ASTIN Bulletin, 2018). La méthode a été initiée et discutée dans un ASTIN Working Party. Une telle méthode doit s'appuyer sur un jeu de données tête par tête suffisamment volumineux afin d'obtenir des performances satisfaisantes. Durant la réalisation de ce mémoire, la disponibilité de la donnée tête par tête sur les portefeuilles étudiés n'était pas effective pour la DAAP. Seules des tables agrégées, possédant quelques caractéristiques supplémentaires des montants de paiement ont pu être recueillies. Cependant, un accès à des données individuelles devrait être possible prochainement.

L'objectif de cette partie est de présenter la méthode CANN, et d'exposer l'ensemble de la démarche nécessaire au développement d'un outil universel de reserving PSAP. Compte tenu du type de données disponibles et utilisées pour ce mémoire, à savoir des données agrégées, nous n'attendons pas de bons résultats dans le cas présent. Nous espérons que cette méthode pourra être testée dès lors qu'un accès aux données individuelles sera possible.

Cette méthode CANN a été développée sur R dans l'article *Neural Networks Applied to Chain-Ladder Reserving*. Pour ce mémoire, l'application sera effectuée sur Python. En effet, l'expertise de Data Science et plus particulièrement du Deep Learning chez AXA France est principalement présente du Python. Ce choix était alors logique pour des questions de développement de l'outil ainsi que pour sa pérennité.

2.3.1 Présentation de la méthode

Cette méthode, appelée Combined Actuarial Neural Network (CANN), repose sur la prise en compte de variables explicatives des sinistres dans le calcul des coefficients de passage du Chain Ladder. Le calcul de ces coefficients permettra ensuite de calculer l'ultime de la même manière que ce qui a été présenté précédemment dans la méthode du Chain Ladder en page 49, pour ensuite calculer la PSAP.

$$R_i = C_{i,J} - C_{i,I-i}$$

avec $C_{i,J}$, la charge ultime et $C_{i,I-i}$ le dernier paiement cumulé observé pour l'année de survenance i .

Considérons une table de données ayant des caractéristiques $x \in X$ avec $C_{i,j}(x) > 0$ pour une période de développement $1 \leq j \leq J$ et une période de survenance $1 \leq i \leq I - j$. Nous étudions alors la relation suivante énoncée dans le Chain Ladder Classique, en considérant les caractéristiques de paiements $x \in X$:

$$E(C_{i,j+1} | C_{i,0}(x), \dots, C_{i,j}(x)) = f_j \cdot C_{i,j}(x)$$

Ces caractéristiques de paiement $x \in X$ sont des caractéristiques ligne à ligne pouvant être agrégées dans une table. Par exemple, l'âge de l'assuré, son genre, la partie du corps ayant été blessée, etc.

Les facteurs de Chain Ladder f_j sont trouvés en minimisant la fonction de perte suivante :

$$\mathcal{L}_j = \sum_{i=1}^{I-j} \sum_{C_{i,j}(x) > 0} \frac{(C_{i,j+1}(x) - f_j(x) C_{i,j}(x))^2}{\sigma_j^2 C_{i,j}(x)}$$

qui s'écrit alors :

$$\mathcal{L}_j = \frac{1}{\sigma_j^2} \sum_{i=1}^{I-j} \sum_{C_{i,j}(x) > 0} C_{i,j}(x) \left(\frac{C_{i,j+1}(x)}{C_{i,j}(x)} - f_j(x) \right)^2 \quad (*)$$

La modélisation des coefficients f_j se fait à travers un réseau de neurones comportant une seule couche cachée (peu profond), composée de q neurones. Si la dimension de X , se note d , alors on peut écrire le modèle de régression associé à la couche cachée définie par:

$$x \rightarrow z_k(x) = \tanh \left(\omega_{k,0} + \sum_{l=1}^d \omega_{k,l} x_l \right)$$

avec $k = 1, \dots, q$ et $\omega_{k,l}$ les poids associés aux neurones

La propagation de la fonction d'activation tangente hyperbolique se fait alors de la manière « feed-forward » jusqu'à la couche de sortie :

$$x \rightarrow \log f_j(x) = \beta_0 + \sum_{k=1}^q \beta_k z_k(x)$$

Pour un nombre de neurones $q \in \mathbb{N}$, on cherche à minimiser la fonction perte définie en (*):

$$\alpha \rightarrow \mathcal{L}_j = \mathcal{L}_j(\alpha)$$

avec α , les paramètres de réseaux :

$$\alpha = (\beta_0, \dots, \beta_q, \omega_{1,0}, \dots, \omega_{q,d}) \in \mathbb{R}^{q+1+q(d+1)}$$

La minimisation de cette fonction \mathcal{L}_j aura pour objectif de trouver les meilleurs paramètres de réseaux α (pour un nombre de neurones q donné). On utilise alors un algorithme d'optimisation (à choisir) afin de minimiser la perte pas à pas.

La fonction de perte utilisée est la fonction de perte au carré ('mse' mean squared error du package Keras). La perte peut alors s'écrire :

$$\mathcal{L}_{j+1}^0 = \sigma_j^2 \mathcal{L}_j = \sum_{i=1}^{I-j} \sum_{C_{i,j}(x) > 0} \left(\frac{C_{i,j+1}(x)}{\sqrt{C_{i,j}(x)}} - f_j(x) \cdot \sqrt{C_{i,j}(x)} \right)^2$$

On remarque que $\sigma_j^2 > 0$ n'influence pas la calibration du modèle, c'est pourquoi il est possible de l'omettre. Désormais, la fonction de perte est une fonction de perte non pondérée et il est alors possible de travailler avec la fonction de perte 'mse' de Keras.

La variable de sortie est définie comme $Y_{i,j+1}(x) = C_{i,j+1}(x) \cdot \sqrt{C_{i,j}(x)}$

La fonction de régression s'écrit alors :

$$f_j^0(x) = f_j(x) \cdot \sqrt{C_{i,j}(x)} = \exp \left\{ \beta_0 + \sum_{k=1}^q \beta_k z_k(x) + \frac{1}{2} \log C_{i,j}(x) \right\}$$

L'estimation de ces coefficients $f_j(x)$ permet de calculer l'ultime de la même manière que ce qui a été présenté précédemment dans la méthode du Chain Ladder. On en déduit alors la PSAP pour une survénance i par différence entre l'ultime (estimé) et le dernier paiement cumulé observé $C_{i,j}$ (réel). La somme des PSAP par année de survénance constitue alors la provision PSAP totale.

2.3.2 Spectre d'étude et qualité de données

L'étude réalisée dans l'article de Wüthrich porte sur une table de sinistres santé détaillée tête par tête, comportant une quantité de données importante (environ 5 millions de lignes). L'approche ne considère pas les sinistres individuels (tête par tête) en tant que tels, mais plutôt des sous-portefeuilles (homogènes) de différents types de sinistres pour lesquels une hypothèse de Chain Ladder est faite et des réserves sont calculées.

Bien que la donnée disponible pour ce mémoire ne soit pas en quantité importante (voir en page 40), nous observons cette notion de sous-portefeuilles homogènes pour lesquels une hypothèse de Chain Ladder est faite.

De plus, ce travail (documentation ainsi que code Python associé) pourra être réutilisé lorsque la donnée massive des portefeuilles étudiés sera disponible pour la DAAP, notamment à travers les fichiers .parquet évoqués en page 40 « *Note sur la disponibilité des données tête par tête* ».

2.3.3 Mise en place du réseau de neurones

La mise en place du réseau de neurones est faite à l'aide de Keras sur Python. Le code appliqué reprend l'Annexe B, intitulée *Neural network architecture and calibration*, de l'article de Wüthrich.

Il faut tout d'abord définir les variables explicatives à utiliser. Dans notre cas, sur Oscar, nous utiliserons la maille territoriale (état américain dans lequel se trouve l'assuré), la ligne de business (contrats individuels ou contrats groupe) et enfin le mois de survenance (pour capter l'effet du Maximum Out of Pocket).

A noter que seuls les contrats individuels appartiennent au scope d'AXA France sur les données de 2019, utilisées dans ce mémoire. Néanmoins, nous intégrons cette variable explicative, sans filtre sur les contrats individuels, pour deux raisons. Tout d'abord, elle peut permettre un apprentissage supplémentaire pour le modèle. Ensuite, les contrats groupes sont réassurés depuis le début de l'année 2021 et correspondent donc au scope actuel et futur.

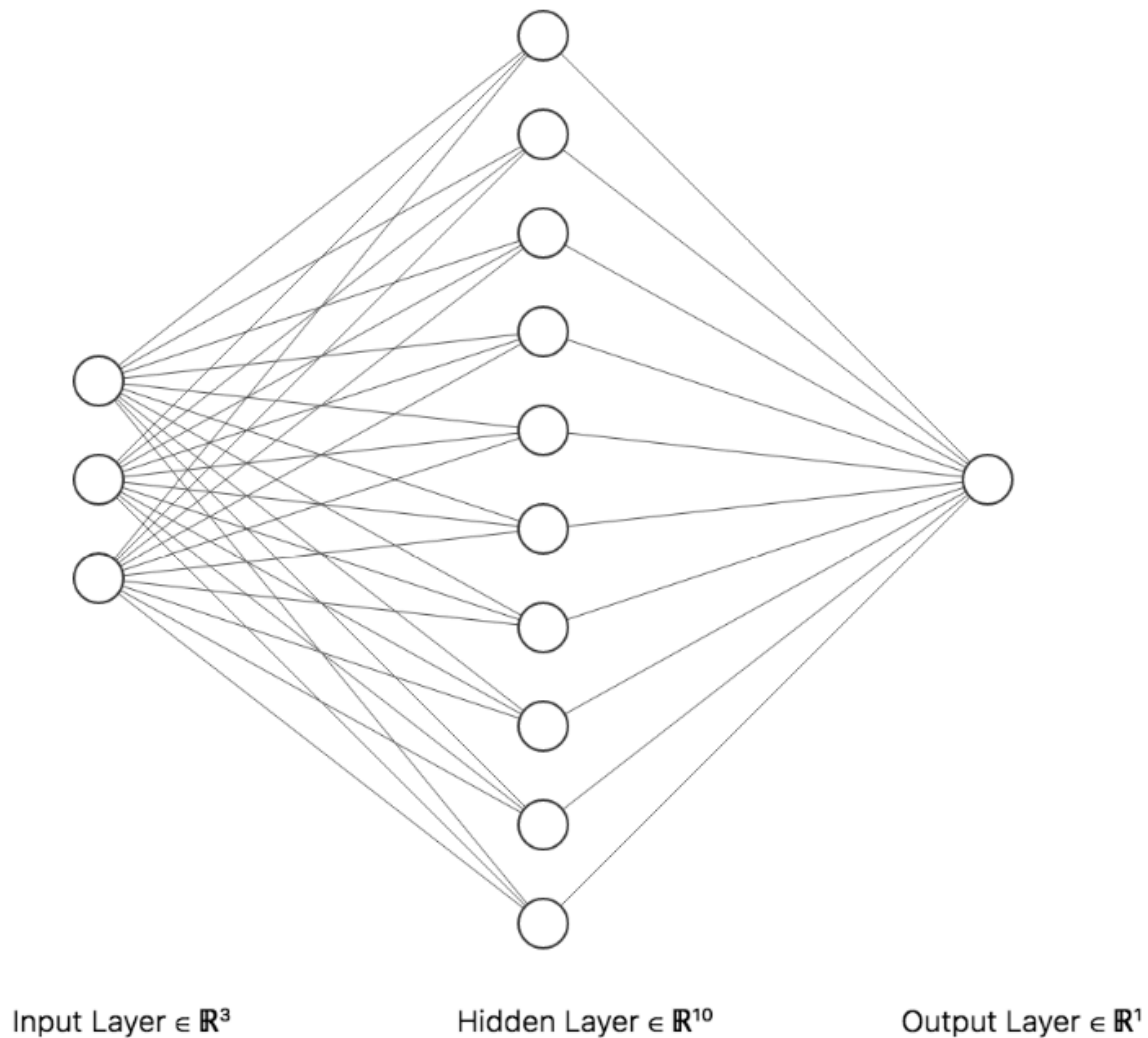


Figure 55 – Réseau de neurone simple couche appliqué sur Oscar

Une fois la table de données établie, il faut encoder les variables discrètes avant de les donner à la couche d'entrée du réseau de neurones. Cela est fait à l'aide de la technique *One Hot Encoding*, qui est disponible sur la librairie *Pandas* de Python.

Afin de pouvoir travailler sur des variables à échelle comparable, il faut ensuite appliquer une mise à l'échelle des variables continues via la fonction *MinMaxScaler* disponible sur la librairie *Scikitlearn*.

On peut alors entraîner le réseau de neurones avec le jeu de données disponible. Il faut alors trouver les hyperparamètres optimaux permettant les meilleures performances pour le réseau de neurones.

2.3.4 Optimisation du réseau de neurones

Les hyperparamètres sont les paramètres dont la valeur est utilisée pour contrôler la structure du réseau de neurones ainsi que son processus d'apprentissage. Pour en citer quelques-uns : le nombre de neurones, l'optimiseur, etc. En revanche, les valeurs des autres paramètres

mentionnés précédemment dans la présentation du modèle (les $\alpha = (\beta_0, \dots, \beta_q, \omega_{1,0}, \dots, \omega_{q,d}) \in \mathbb{R}^{q+1+q(d+1)}$) sont obtenues par apprentissage et mises à jour par le modèle lui-même afin de minimiser la fonction de perte définie.

Méthode Grid Search

La méthode d'optimisation Grid Search consiste à établir un ensemble de valeurs possibles pour chaque hyperparamètre. Par exemple le nombre de neurones q de la couche cachée peut prendre 8 valeurs dans notre étude. L'optimiseur peut être soit « rmsprop » ou « adam ». La méthode consiste alors à entraîner le modèle pour chaque combinaison de paramètres avec pour but ultime de sélectionner la combinaison ayant les meilleures performances sur le set de validation.

Voici les hyperparamètres qui vont varier dans notre optimisation :

- **Le nombre de neurones dans la couche cachée**

Le nombre de neurones q de la couche cachée doit être calibré de manière appropriée. Un nombre q trop faible entraîne une perte d'information tandis qu'un q trop grand entraînera de l'overfitting. Le nombre de neurones est de 20 sur l'article. Ayant un jeu de données bien moins conséquent pour cette étude, des nombres inférieurs sont testés. Il est généralement recommandé de prendre q égal à une à trois fois le nombre de variables explicatives.

- 8 valeurs possibles choisies : [5, 8, 10, 12, 14, 16, 18, 20]

- **L'optimiseur**

L'optimiseur correspond à la méthode de mise à jour des paramètres du réseau de neurones $\alpha = (\beta_0, \dots, \beta_q, \omega_{1,0}, \dots, \omega_{q,d}) \in \mathbb{R}^{q+1+q(d+1)}$. L'objectif est d'avoir un trajet vers le minimum de la fonction de perte pas trop long (temps de calcul) et suffisamment précis. Wüthrich utilise l'optimiseur Root Mean Square Propagation (noté rmsprop). L'optimiseur Adam (Adaptive Moment Optimization) sera également testé. Il utilise la méthode de descente de gradient stochastique (Stochastic Gradient Descend, SGD) pour effectuer le processus d'optimisation. Il est efficace à utiliser et consomme très peu de mémoire.

- 2 valeurs possibles [rmsprop, adam]

Soit 16 combinaisons possibles (8×2).

D'autres choix ont été effectués. En particulier, une seule fonction de perte a été utilisée. La fonction de perte correspond à la manière de quantifier les erreurs pour le réseau de neurones. Celle choisie est la même que l'article, à savoir la fonction d'erreur quadratique moyenne (mean square error, notée mse). Cette fonction pénalise les erreurs de surestimation ainsi que les erreurs de sous-estimation.

Aussi, un early-stopping a été ajouté. Le mécanisme d'early-stopping permet d'arrêter l'apprentissage du réseau de neurones lorsque la perte du set de validation ne diminue plus ou augmente. Cela permet d'éviter le sur-apprentissage (overfitting) sur le set de données d'entraînement.

Cette technique d'early-stopping permet également de ne pas tester des modèles faisant varier le nombre d'epochs. Le nombre d'epochs est un hyperparamètre qui définit le nombre de fois que l'algorithme d'apprentissage va travailler sur l'ensemble des données d'apprentissage. S'il est trop grand, il y a un risque de sur-apprentissage. Il est fixé à 100 dans l'article. Dans notre cas, l'early-stopping arrêtera l'apprentissage avant l'overfitting. Il sera également fixé à 100 (et sera en réalité inférieur ou égal, car stoppé par l'early-stopping)

Enfin, un mécanisme de régularisation (pénalisation L1 par exemple) peut parfois être implémenté pour limiter également l'overfitting. Il encourage certains poids du réseau de neurones à prendre une valeur nulle. Ce mécanisme a été testé mais n'a pas eu d'influence sur les résultats. Il est principalement adapté à des architectures de réseaux complexes, ce qui n'est pas le cas dans notre étude.

Sélection de la meilleure combinaison de paramètres

Pour chaque combinaison de paramètres précédemment définie, sont collectées les erreurs d'apprentissage (mse sur le set d'apprentissage), les erreurs de validation (mse sur le set de validation) ainsi que l'erreur de back-testing (distance entre la valeur PSAP prédite et la valeur réellement observée).

Parmi les 16 combinaisons de paramètres possibles décrites dans la Méthode Grid Search, nous allons choisir la combinaison de paramètres qui minimise les métriques suivants:

- L'erreur de validation :
 - o Correspond à l'erreur que le modèle génère sur un set de données non encore connu (sur lequel il n'a pas appris)
- L'écart entre l'erreur de validation et l'erreur d'apprentissage :
 - o Correspond à la notion de sur-apprentissage. Si l'erreur d'apprentissage est faible et que l'erreur de validation est grande alors le modèle a sur-appris (overfitting). Si l'erreur de validation est plus faible que l'erreur d'entraînement alors il est possible que le modèle n'ait pas assez appris (underfitting)
- L'erreur de back-testing
 - o Correspond à la distance entre le montant de PSAP prédictif du modèle CANN et le montant de sinistres non payés à date mais constatés et payés réellement à posteriori (back-testing)

Le classement des 16 modèles se fait de la manière suivante : un rang est associé à chaque modèle pour chaque variable. Par exemple, le modèle minimisant le plus l'erreur de validation prendra le rang 1 de la variable « rang_erreur_validation ». De la même manière, le modèle obtenant l'erreur de backtest la plus faible se verra associer le rang 1 sur la variable « rang_erreur_backtest ».

Une fonction scoring est alors définie. Elle fait la moyenne des 3 rangs observés (rang de l'erreur de validation, rang de l'écart entre l'erreur de validation et l'erreur d'apprentissage, et rang de l'erreur de back-testing) pour chaque modèle. On peut alors afficher cette fonction et choisir le modèle optimum, minimisant la fonction scoring.

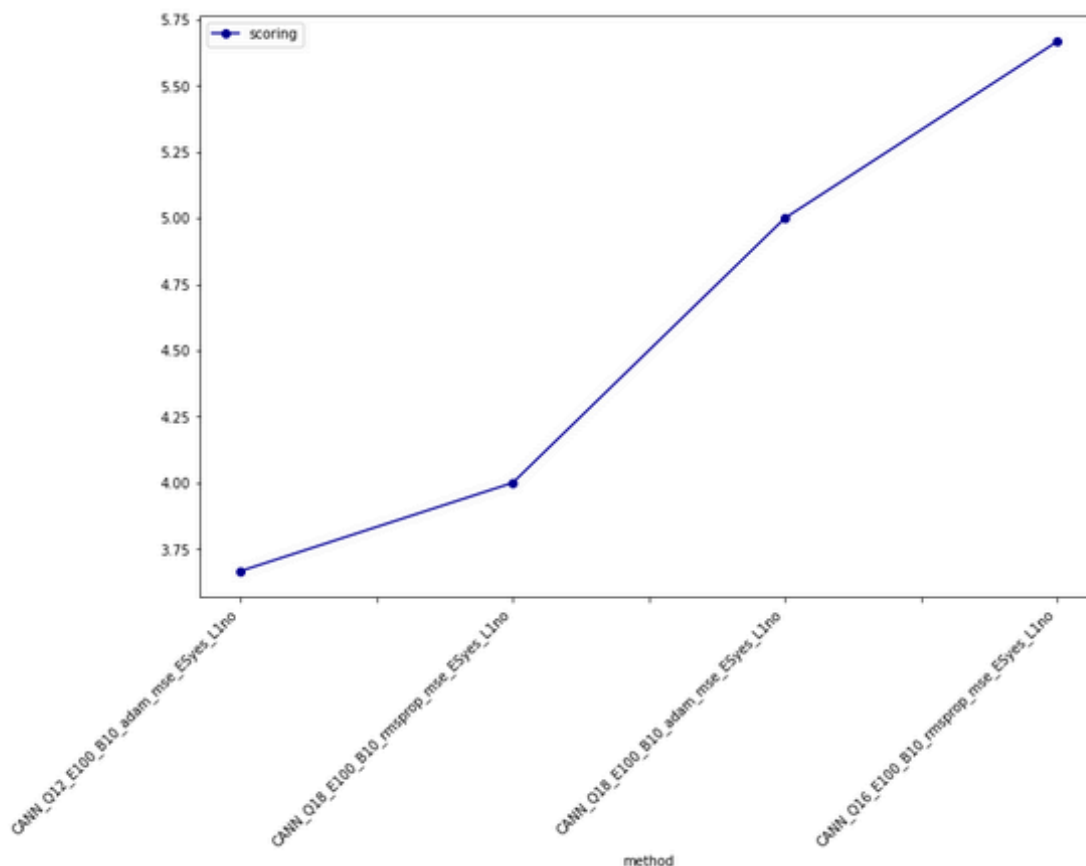


Figure 56 – Fonction scoring pour les 4 meilleurs modèles

Deux modèles semblent se distinguer des autres si l'on suit la métrique précédemment construite. Il s'agit des modèles présentant un rang moyen de 3.7 et 4.0. On sélectionne alors la fonction qui minimise le plus l'erreur de back-testing, ce qui nous mène au modèle CANN_Q18_E100_B10_rmsprop_mse_ESyes_L1no. Ce nom de modèle traduit un nombre de neurones égal à 18 pour un optimiseur « rmsprop ». Le reste des paramètres étant fixés et déjà énumérés précédemment (Epoch = 100, Batch size = 10, Fonction de perte = mean square error, Application d'un Early-Stopping, Pas de pénalisation L1).

2.3.5 Résultats de la méthode CANN

Le réseau de neurones CANN_Q18_E100_B10_rmsprop_mse_ESyes_L1no nous donne les résultats suivants :

| Méthode | New-York | Californie | Texas | New-Jersey |
|---------|-------------|-------------|-------------|-------------|
| CANN | \$5.572.755 | \$4.631.700 | \$4.866.923 | \$2.659.798 |

Figure 57 - Tableau des résultats bruts sur Oscar (CANN)

Ces résultats sont obtenus avec l'intervention du early-Stopping pour un nombre d'époques relativement faible (33 itérations). Cette stagnation de l'apprentissage est due au jeu de données étudié trop restreint ainsi qu'à l'architecture assez simpliste du réseau mis en place.

Aussi, il est noté ici qu'aucune estimation de l'erreur n'est disponible pour cette méthode à cette date.

CANN - Commentaires sur l'Australie

Le réseau de neurones CANN_Q10_E100_B10_adam_ESyes_L1no nous donne les résultats suivants :

| Méthode | TPD | Death | IP |
|---------|-----|-------|-----|
| CANN | 174 | 292 | 108 |

Figure 58 - Tableau des résultats bruts sur l'Australie (CANN)

Note: Pour rappel, les résultats du périmètres Australien sont donnés en nombre de sinistres, comme expliqué en page 59.

Ces résultats sont obtenus avec l'intervention du Early-Stopping pour un nombre d'Epochs relativement faible (26 itérations). De la même manière que sur Oscar, cette stagnation de l'apprentissage est due au jeu de données restreint étudié ainsi qu'à l'architecture assez simpliste du réseau mis en place.

2.4 COMPARAISON DES RÉSULTATS

Cette partie est consacrée à la comparaison des résultats entre les différentes méthodes abordées au cours de ce mémoire. Il est rappelé ici que compte tenu des données disponibles et utilisées pour cette étude, aucun résultat meilleur qu'une méthode classique précédemment présentée n'est attendu pour la méthode CANN.

2.4.1 Comparaison des méthodes sur le scope 2019 - Backtesting

Tout d'abord, une comparaison sur le scope temporel correspondant à l'année 2019 est effectuée. C'est sur ce scope que l'ensemble des méthodes ont été présentées précédemment. Autrement dit, on se place à l'arrêté comptable du dernier trimestre de l'année 2019 (Q4 2019).

Pour le business Oscar, une comparaison en maille territoriale est alors effectuée. Voici les différents montants présentés :

- « **REAL** » **en jaune** : le montant de PSAP réellement payé, constaté après plusieurs mois de recul. En d'autres termes, c'est l'ensemble des paiements réellement effectués depuis fin décembre 2019 (observés sur l'année 2020) sur les survivances de 2019. Ce montant est un montant permettant de mesurer les performances de nos modèles puisque c'est ce qu'on cherchait à estimer le plus précisément en fin d'année 2019.
- « **MACK** » **en noir** : le montant de PSAP « Best-Estimate » de la méthode de Mack. Cela correspond également au montant PSAP du Chain Ladder classique, puisque nous avons effectué le même calcul de coefficients de passage f_j pour ces deux méthodes.
- « **MACK+StdErr** » **en gris** : le montant de PSAP « Best-Estimate » de la méthode de Mack plus son erreur standard, sous hypothèse d'une distribution normale des réserves.
- « **BOOTSTRAP 50%** » **en rouge foncé** : le quantile 50% de PSAP donné par la méthode Bootstrap. Autrement dit, le montant PSAP « Best-Estimate » de cette méthode.
- « **BOOTSTRAP 70%** » **en rouge clair** : le quantile 70% de PSAP donné par la méthode Bootstrap.
- « **C.A.N.N._Q18_E100_B10_rmsprop_mse_yes** » **en vert pomme** : le montant « Best-Estimate » de la méthode CANN pour le modèle choisi après le processus d'optimisation.

- « AXA reserves » en bleu : le montant de provisions PSAP placé par la DAAP dans les comptes d'AXA France au Q4 2019.

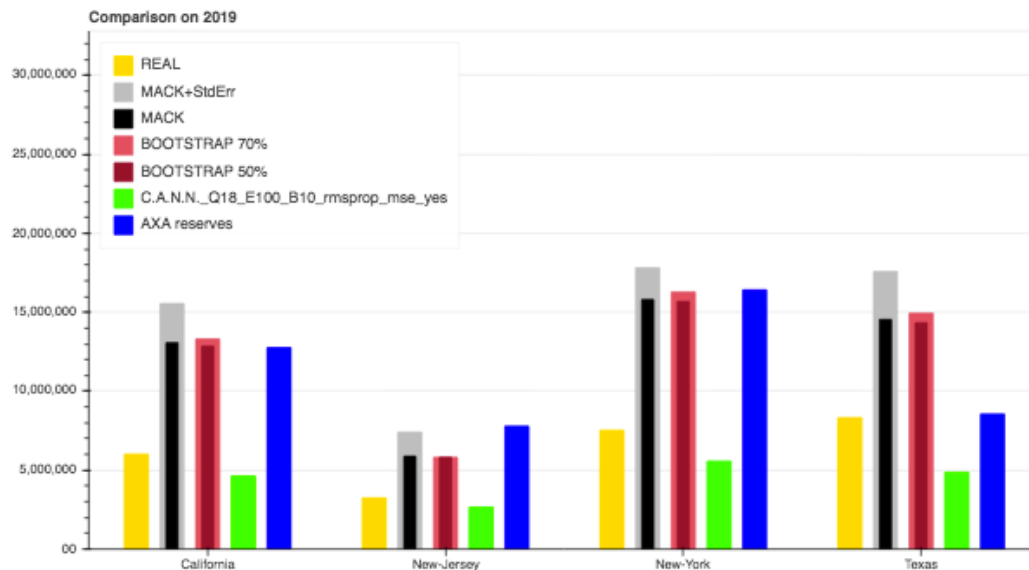


Figure 59 – Comparaison des différents modèles (back-testing sur Oscar)– Vision Final Year 2019

Le but est d'approcher le plus possible le montant « REAL » présenté en jaune. La méthode s'approchant le plus de cette valeur est la méthode CANN. En revanche, c'est une approche « par le bas ». C'est à dire que le montant réel est sous-estimé sur l'ensemble des états. Bien qu'aucun résultat ne soit particulièrement attendu de par les données disponibles pour une telle méthode, statuons tout de même qu'une méthode sous-estimant les réserves est très dangereuse d'un point de vue solvabilité pour une société d'assurance.

Ensuite, nous retrouvons le constat fait en page 69 lorsque nous comparons les deux méthodes classiques stochastiques. La moyenne de ces deux méthodes (leur montant « Best-Estimate ») converge. En revanche un écart est à noter entre les montants « MACK+StdErr » et « BOOTSTRAP 70% ». Cet écart est lié à deux effets qui se cumulent. D'une part, comme constaté en page 69, le Bootstrap permet de réduire la volatilité concernant la distribution des réserves PSAP. D'autre part, pour une distribution normale des réserves (Mack Chain Ladder), la somme de son espérance et de son écart type correspond au quantile 84% des réserves. Ce montant en gris, est systématiquement au-dessus du montant du quantile 70% lié à la méthode Bootstrap, en rouge clair.

Enfin, le montant « AXA reserves », mis dans les comptes d'AXA France au Q4 2019 par la DAAP est un montant issu de la vision marché ALHIS. Ce montant n'est pas issu de l'outil généraliste utilisé par la DAAP. En effet, comme expliqué précédemment, les prévisions liées à des contrats majeurs tels que Oscar sont corrigées par la vision marché qui utilise des outils spécialisés. Ce montant n'est pas exactement celui du « 12-months Chain Ladder Classique » (montant « MACK » en noir), malgré le fait que cette méthode soit appliquée par le marché. En effet, lors de l'arrêté, les équipes marché peuvent ajuster leur vision en fonction d'informations clés reçues de la part de la cédante Oscar. Ces ajustements entraînent des écarts

avec les prévisions brutes données par l'outil « 12-months Chain Ladder Classique ». Cet écart est notamment constaté sur le Texas, où une approche de la valeur réelle est mieux réussie par rapport aux autres méthodes.

Plus généralement, les méthodes permettant une quantification d'erreur sont mieux appréciées dans le contexte de vision « Best-Estimate » des comptes. Il est important de savoir la prudence prise lors de l'estimation des réserves. La quantification d'erreur est uniquement possible via les deux méthodes classiques stochastiques. Une préférence est attribuée à celle qui réduit la volatilité des réserves, à savoir le Bootstrap.

2.4.1 Comparaison des méthodes sur le scope 2020 - Prévisions

L'ensemble de l'étude présentée précédemment a également été effectuée sur le scope du Half-Year 2020 sur les contrats Oscar. Pour chaque méthode, l'ensemble des hypothèses ont été à nouveau vérifiées.

Une comparaison en maille territoriale est à nouveau effectuée avec les mêmes catégories de montants que ceux décrits précédemment. Néanmoins, deux changements sont à noter :

- **L'absence de « REAL » en jaune** : ce montant n'est pas disponible car le recul temporel n'est pas assez important pour réaliser un back-testing.
- **« OSCAR » en orange** : ce montant correspond aux dernières estimations de réserves effectuées par la cédante.

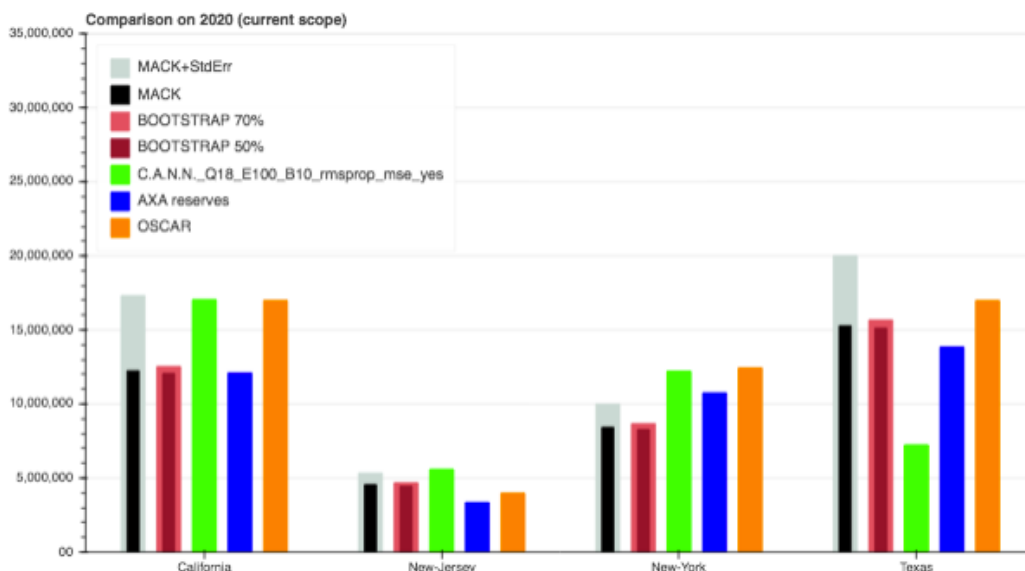


Figure 60 – Comparaison des différents modèles (prévisions sur Oscar) – Vision Half Year 2020

Ces montants sont donc des montants de prévisions et aucun repère ne peut nous permettre de distinguer une méthode comme plus performante que les autres (absence de back-testing).

En revanche, certains commentaires effectués précédemment sont toujours d'actualité. En particulier, le Bootstrap permet de réduire la volatilité par rapport au Mack Chain Ladder.

De plus, nous appliquons le réseau de neurones sans ré-apprentissage sur le scope 2020, afin de voir comment se comporte le modèle ayant appris sur des données antérieures (scope 2019). L'apprentissage sur l'année 2019, décrite comme une année classique par le marché, présente un biais lorsqu'on souhaite prédire la PSAP 2020 puisque cette dernière est fortement impactée par le Covid-19. En particulier, une baisse de plus de 50% de la sinistralité sur la survenance d'Avril entre l'année 2019 et 2020 est constatée. Au-delà du caractère limitant de la quantité de données, la méthode CANN induit un questionnement sur les scopes d'apprentissages.

Aussi, nous notons un écart entre les provisions PSAP proposées par Oscar et celles proposées par le marché ALHIS (notées « AXA reserves »). Cet écart provient du fait que les méthodes de calcul ne sont pas exactement identiques. En fait, la cédante Oscar distingue ses paiements ligne à ligne en deux catégories : les sinistres inférieurs à \$40.000 (notés « low paid ») et les sinistres supérieurs à ce seuil (notés « high paid »). Pour les sinistres « low paid », un Chain Ladder Classique est effectué, tandis qu'un traitement au cas par cas est effectué pour les sinistres de pointe dépassant les \$40.000.

En conséquence, lorsqu'un écart assez important (>5%) entre les PSAP des méthodes ALHIS et les PSAP des méthodes Oscar est constaté, une discussion entre la cédante et le réassureur a lieu afin de comprendre la différence. Un ajustement sur les provisions PSAP d'AXA peut alors avoir lieu en fonction des raisons évoquées. Cet ajustement est généralement effectué au semestre 2 de l'année, raison pour laquelle des écarts importants restent présents sur le graphique ci-dessus relatif à la vision Half-Year avec une vision Oscar plus prudente que celle d'AXA France.

Comparaison des méthodes - Commentaires sur l'Australie

De la même manière que ce qui est effectué sur Oscar, deux scopes temporels ont été étudiés sur l'Australie. Le premier correspond au scope 2012-2016, et le second au scope 2016-2020.

Pour le premier scope, dont les résultats sont commentés tout au long du mémoire, on se place au Half-Year 2016. L'objectif est de prévoir la PSAP des survenances allant de mi-2012 à mi-2016 tout en bénéficiant d'un recul temporel de 4 années, permettant de connaître les montants réellement payés in fine.

Ce processus de back-testing est pollué par la qualité des données faible sur ce scope puisqu'elles présentent une censure à droite, comme expliqué dans la partie

Pour l'Australie, une comparaison par garantie est effectuée. La description des montants est la même que celle effectuée en page 82. Deux précisions sont néanmoins apportées :

- « **REAL** » en jaune : le montant de PSAP réellement payé, constaté après plusieurs années de recul. En d'autres termes, c'est l'ensemble des paiements réellement effectués depuis mi-2016 (observés depuis juin 2016) sur les survenances correspondant au scope étudié (2012-2016).
- L'absence de « **AXA reserves** » en bleu : au Half-Year 2016, AXA France n'était pas encore réassureur de ces contrats et n'a donc pas constitué de provisions sur ces survenances. Ces données antérieures (2012-2016) correspondent à des données acquises pour la souscription des contrats.

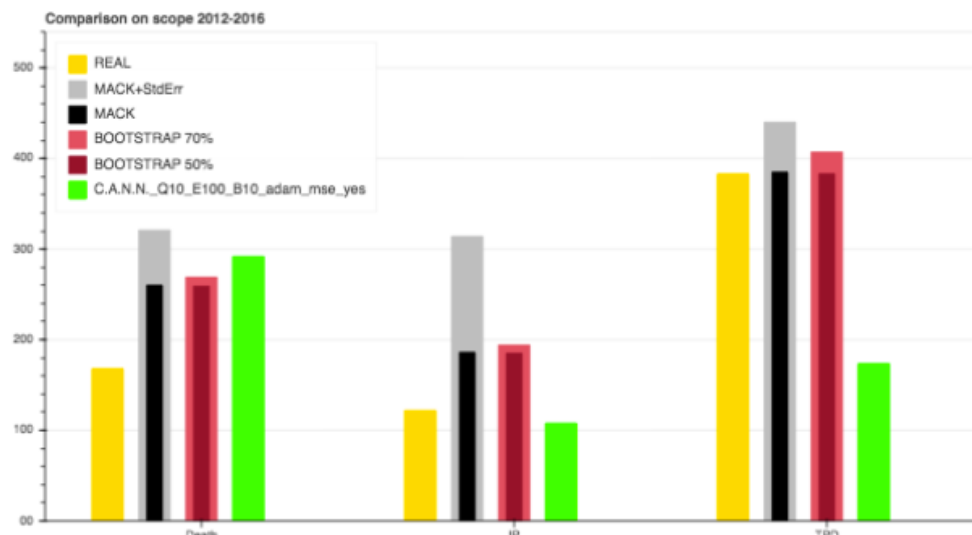


Figure 61 – Comparaison des différents modèles (back-testing sur l'Australie) – Vision Half Year 2016

De la même manière que ce que nous avons pu constater sur Oscar concernant les méthodes classiques stochastiques, une réduction de la volatilité est constatée globalement sur le business Australien pour la méthode Bootstrap par rapport à la méthode de Mack. Ces deux méthodes

surestiment modérément le montant réel (lui-même sous-estimé par la censure à droite). Elles donnent donc satisfaction.

En revanche, la méthode CANN semble ici plus agressive et l'estimation de la garantie TPD présente le même inconvénient que celui remarqué sur Oscar, à savoir une sous-estimation. Compte tenu de la quantité limitée de données pour une telle méthode, ces résultats ne sont en aucun cas des arguments permettant de rejeter la méthode CANN.

Sur le scope 2016-2020, des prévisions sont effectuées. Une comparaison par garantie est à nouveau effectuée avec les mêmes catégories de montants que ceux décrits précédemment. Néanmoins, trois changements sont à noter :

- **L'absence de « REAL » en jaune** : ce montant n'est pas disponible car le recul temporel n'est pas assez important pour réaliser un back-testing.
- **« AXA – CL modif » en bleu foncé** : ce montant correspond aux estimations d'un Chain Ladder modifié (triangle tronqué à 12 périodes) du marché ALHIS, sans fonction LOGIT appliquée pour les sinistres de longue queue.
- **« AXA – Additive » en bleu ciel** : ce montant correspond à la dernière méthode étudiée par le marché ALHIS, en lien avec les pratiques locales suite à la mise en place du PYS (voir en page 35), sans fonction LOGIT appliquée pour les sinistres de longue queue.

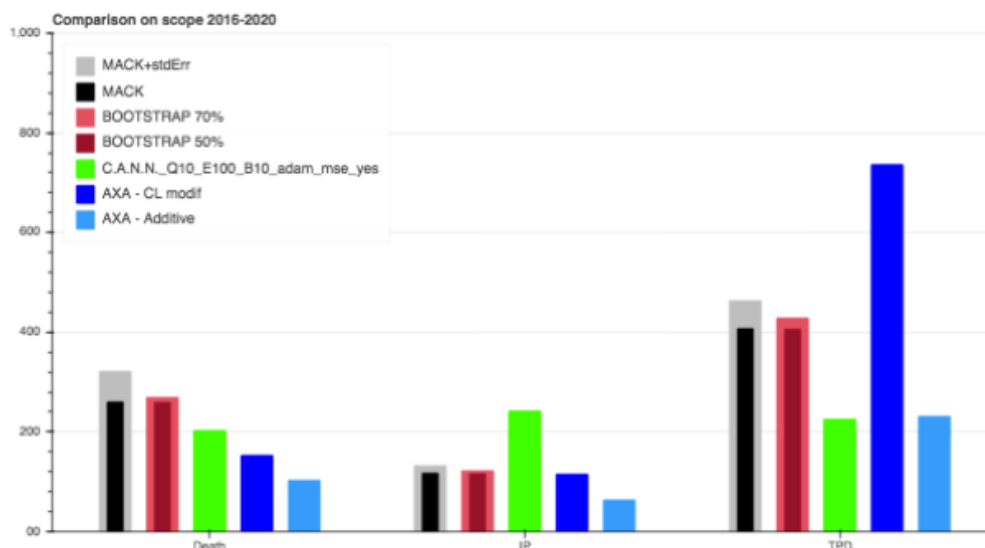


Figure 62 – Comparaison des différents modèles (prévisions sur l'Australie) – Vision Half Year 2020

Tout d'abord, un écart important entre les méthodes est constaté, notamment sur la garantie TPD. En effet, certaines données atypiques sur les derniers règlements peuvent avoir une influence considérable sur le montant de la charge ultime calculée par le Chain Ladder. C'est

pourquoi, le Marché ALHIS a décidé d'utiliser une méthode dite « additive », comme évoqué en page 35, afin de limiter l'impact de l'accélération des derniers paiements (dus à la nouvelle réglementation locale PYS)

Concernant la méthode CANN, nous notons que le choix des modèles a été effectué via la fonction scoring prenant en compte le rang de back-testing. Or, les données étant censurées à droite pour le portefeuille Australien, nous savons que ce rang de back-testing ne reflète pas forcément la réalité. En conséquence, le meilleur modèle choisi par la fonction scoring n'est probablement pas le meilleur aux yeux de la réalité.

De la même manière que ce qui a été fait sur Oscar, le modèle CANN a effectué son apprentissage sur le scope antérieur (ici 2012-2016) et nous l'appliquons sans ré-apprentissage sur le dernier scope en date. Ici aussi, la méthode CANN induit un questionnement sur les scopes d'apprentissages.

3. PERSPECTIVES

L'intérêt de cette partie est de voir comment l'étude effectuée peut apporter des points de vigilance sur les méthodes appliquées ainsi que sur la donnée utilisée. Nous proposons quelques pistes d'améliorations concernant les parties dédiées aux méthodes de provisionnement. Enfin, nous présentons les pistes d'audit concernant l'utilisation des réseaux de neurones dans la perspective d'une mise en place d'un tel outil de provisionnement à la DAAP.

3.1 LIMITES DE L'ÉTUDE ET PISTES D'AMÉLIORATION

3.1.1 Limites de l'étude

Les limites ainsi que pistes d'amélioration de cette étude sont nombreuses. En premier lieu, il est nécessaire de rappeler les limites de la donnée utilisée, avant de présenter celles de la méthode CANN.

Limites – données utilisées

Pour les méthodes classiques, la donnée utilisée reste satisfaisante. Néanmoins, pour chaque portefeuille, un biais de données majeur est à noter.

D'une part, sur le business Oscar, la table de données disponible ne nous donne pas le niveau de métal (Metal Level) souscrit. Nous supposons qu'il a une grande influence dans l'homogénéité des triangles étudiés. En d'autres termes, bien que les hypothèses du Chain Ladder soient validées en décomposition territoriale, il est possible qu'une décomposition par type de contrat soit plus satisfaisante d'un point de vue résultat et performance vis à vis du back-testing.

D'autre part, sur le business Australien, la censure à droite constatée empêche de réaliser un back-testing très précis. Autrement dit, les montants de paiements tardifs donnés dans le back-testing pour l'Australie sont une sous-estimation de la réalité. D'après les cadences de paiements, présentées en Annexe en page 99, on constate qu'au bout de deux ans sur de telles garanties, approximativement 80% de la charge est réglée. Il y aurait donc selon les survénances, environ 20% de sous-estimation de sinistres sur ces montants de back-testing.

Enfin, les quantités de données (nombre d'observations, granularité des variables, etc.) sont trop restrictives pour une méthode utilisant des réseaux de neurones telle que CANN, ce qui sera davantage développé par la suite.

Limites – méthode CANN

L'application d'une telle méthode nécessite un jeu de données tête par tête important. Il est donc nécessaire de pouvoir avoir accès à une base de données massive comportant des caractéristiques de sinistres précises telles que celles utilisées dans l'article de Wüthrich (2018).

En particulier, sur le périmètre Oscar, des variables telles que l'âge de l'assuré, le niveau de métal du contrat, la partie du corps blessée, le type de consultation (comme les 10 services essentiels présentés dans la partie Description du système local), ou encore le lieu de consultation pourraient être largement utiles pour la méthode CANN. Le provisionnement PSAP ici correspond à des sinistres non encore déclarés (Incurred But Not Yet Reported). En conséquence, l'ensemble des variables collectées pouvant expliquer la vitesse de la remontée de données à l'assureur, sont des variables à utiliser.

Enfin, bien que cette méthode nous permette de réduire le nombre d'outils de provisionnement PSAP, nous avons vu que CANN était limitante sur deux aspects. D'une part, elle ne permet pas la quantification d'erreur de prédiction, ce qui la pénalise vis à vis d'une méthode Bootstrap qui donne accès à des quantiles de PSAP. D'autre part, cette méthode pose la question du scope d'apprentissage. En effet, lorsque l'algorithme apprend (ajuste ses paramètres) sur un scope temporel passé, il faut que le scope sur lequel les prévisions sont effectuées (après apprentissage) soit homogène au scope d'apprentissage. Or, nous avons vu avec l'exemple du business Oscar que l'apprentissage sur une année ordinaire telle que 2019 pouvait probablement perturber les prévisions sur une année exceptionnelle comme 2020, notamment sur les garanties Santé qui ont été fortement impactées par le Covid-19.

3.1.2 Améliorations possibles de l'étude

Améliorations – Méthodes classiques

Comme mentionné précédemment, le but de cette étude n'est pas d'effectuer une revue des méthodes possibles pour calculer les PSAP. L'objectif est davantage de montrer que certaines méthodes usuelles multiplient le nombre d'outils et qu'une méthode universelle telle que CANN est bienvenue pour répondre à notre problématique.

Néanmoins, il aurait été possible de développer l'étude des méthodes classiques de provisionnement selon deux axes :

- **Le calcul des coefficients de passage** : d'autres méthodes usuelles existent sur les Chain Ladder, notamment la méthode dite « 3 sur 5 » assez répandue chez AXA France. L'application d'un Chain Ladder 3 sur 5 permettrait d'avoir un nouveau repère dans la comparaison des méthodes.

- **La classification en groupes homogènes des paiements** : un clustering autre que celui établi pourrait être envisagé. En particulier, peut être que certains états sur Oscar présentent des paiements homogènes, ce qui permettrait de réduire partiellement le nombre de triangles. Par exemple, un clustering mènerait peut-être à associer le New-Jersey et New-York, en laissant le Texas et la Californie indépendants. Cela réduirait le nombre de triangles à manipuler.

Aussi, concernant l'Australie, l'étude s'est limitée au calcul du nombre de sinistres. D'une part, ce calcul ne traite pas le manque d'expérience des sinistres de longue queue non encore observés dans le portefeuille. Un ajustement LOGIT tel que décrit en page 54 doit être effectué.

D'autre part, pour les sinistres Australiens, un taux de rejet (decline rates) est à prendre en compte afin de capter les sinistres en cours d'acceptation, sur lesquels une expertise est en cours. L'ensemble de l'étude s'est concentrée sur les sinistres ayant un statut « finalisé » aux yeux des assureurs.

Améliorations – Méthode CANN

En ce qui concerne la méthode CANN, au-delà de la structure du réseau de neurones, l'amélioration proposée est davantage une question de données en input. En particulier, deux axes d'amélioration peuvent être présentés (en dehors de l'obtention indispensable d'une table tête par tête) :

- **L'enrichissement de la table de données tête par tête** : ajout de variables explicatives telles que le trimestre d'accident, le nombre d'habitants dans la ville de l'assuré, catégorisation des médecins, etc.
- **La construction d'une table unique pour l'ensemble du portefeuille** : passage d'une méthode unique (CANN appliquée business par business) à un outil unique (une seule méthode une seule table)

De manière pragmatique, une telle méthode sera possible à mettre en œuvre à l'aide des tables (en cours de construction) du département gouvernance de la donnée de ALHIS. La brève description de ces tables est disponible en page 40. La mise en place de la méthode CANN nécessitera également des documentations et des explications indispensables à l'audit, ce qui est présenté en page suivante.

3.2 ÉVALUATION DU MODÈLE CANN ET PISTES D'AUDIT

L'Autorité de Contrôle Prudentiel et de Résolution (ACPR) a publié en Juin 2020 un document de réflexion intitulé *Gouvernance des algorithmes d'intelligence artificielle dans le secteur financier*. Ce document a pour objectif d'identifier les critères d'évaluation des algorithmes d'intelligence artificielle, tout en donnant des recommandations sur leur gouvernance. En particulier, la notion d'auditabilité du modèle est abordée.

Lors de la conception et de la mise en place d'un outil de reserving utilisant une méthode dite d'intelligence artificielle (IA) telle que CANN, un certain nombre de précautions sont nécessaires.

En particulier, le développement d'un algorithme d'IA doit suivre 4 grands principes :

- **Le traitement adéquat des données** : évaluation des risques réglementaires et éthiques.

L'évaluation part de la source des données (conformité aux normes sectorielles, Règlement Général sur la Protection des Données – RGPD, etc) jusqu'à leur preprocessing (traitement des données avant leur entrée dans l'algorithme). Enfin, les traitements post-algorithme entrent également dans l'évaluation.

Dans le cas du modèle CANN que nous avons appliqué, une attention particulière sera portée sur la source de donnée, ainsi que les traitements effectués notamment le One Hot Encoding lors du preprocessing. Enfin, une attention sera portée sur les traitements post-algorithmiques tels que les manipulations permettant de passer de l'output du réseau de neurones à la prédiction des réserves par sous portefeuilles.

- **La performance du modèle** : une sélection de métriques de performance est nécessaire.

La recommandation principale est que ces métriques doivent permettre une évaluation technique de l'algorithme en considérant un compromis entre simplicité et efficacité de l'algorithme.

Dans le cas du modèle CANN que nous avons mis en place, les métriques de performance utilisées sont présentées en page 78.

- **La stabilité du modèle** : identification des différentes sources d'instabilités liées au modèle et mise en place de méthodes de résolution.

Différents risques opérationnels ou de conformité proviennent des sources d'instabilités. En particulier, l'instabilité temporelle, l'instabilité liée à un défaut de robustesse ou encore l'instabilité liée aux processus de réapprentissage sont mentionnées.

Dans le cas du modèle CANN que nous avons appliqué, au fur et à mesure de l'évolution des portefeuilles ALHIS, un processus de réapprentissage est nécessaire afin de lutter contre une possible instabilité temporelle. Ce réapprentissage peut néanmoins induire

des instabilités de modèle. Si ces dernières ne peuvent être palliées, il est conseillé d'archiver les versions successives des modèles entraînés.

- **L'explicabilité du modèle** : observation, justification, approximation et réplication du modèle.

C'est ici que la notion d'auditabilité intervient. En effet, elle est intimement liée avec la notion d'explicabilité qui se décline en 4 niveaux :

- **Niveau d'explicabilité 1 : Observation** – *Que fait l'algorithme?*
Éléments explicatifs requis: Observation des résultats + fiche descriptive

⇒ *Pour la méthode CANN : visualisation des résultats sur Python avec fiche descriptive s'appuyant sur l'article de Wüthrich (2018)*
- **Niveau d'explicabilité 2 : Justification** – *Pourquoi l'algorithme donne-t-il un tel résultat?*
Éléments explicatifs requis: Présentation simplifiée d'éléments explicatifs de niveau 3 ou 4 ou génération par l'algorithme lui-même de justifications obtenues par apprentissage.

⇒ *Pour la méthode CANN : Présentation simplifiée d'éléments explicatifs de niveau 3 ou 4*
- **Niveau d'explicabilité 3 : Approximation** – *Comment fonctionne l'algorithme?*
Éléments explicatifs requis: Méthodes explicatives post-modélisation ou analyse structurelle de l'algorithme.

⇒ *Pour la méthode CANN : Analyse structurelle de l'algorithme avec explication couche par couche du réseau de neurones*
- **Niveau d'explicabilité 4 : Réplication** – *Comment prouver que l'algorithme fonctionne correctement?*
Éléments explicatifs requis: Revue ligne à ligne du code source avec examen des données utilisées ainsi qu'un examen détaillé de l'ensemble des paramètres du modèle.

⇒ *Pour la méthode CANN : Revue ligne à ligne du code Python. Examen des sources de données (conformité sectorielle, RGPD, etc.). Examen détaillé de la méthode Grid Search appliquée permettant l'obtention des hyperparamètres choisis.*

Ces différents niveaux d'explication ne sont pas tous nécessaires selon les différents cas d'usage. Néanmoins, dans le cas de la DAAP, la production des réserves est auditée. D'après le tableau des cas d'usage disponible en Annexe page 128, un niveau d'explicabilité 3 est au minimum requis.

CONCLUSION

La réduction du nombre de méthodes de provisionnement PSAP peut devenir un véritable enjeu pour un groupe d'assurance, notamment lorsque certains de ses portefeuilles sont morcelés et hétérogènes.

Durant cette étude concernant les contrats de Santé et Prévoyance collectives à l'international, nous avons traité deux portefeuilles présentant des garanties différentes, dans des pays différents. Ces business représentatifs de la diversité du portefeuille global ont permis d'identifier trois phénomènes menant à une décomposition des méthodes et des outils de provisionnement PSAP.

Tout d'abord, une décomposition par garantie doit être faite, notamment à cause des cadences de paiements très différentes menant à des triangles de règlement présentant des périodicités disparates.

D'autre part, l'évolution des législations locales peut également moduler cette cadence de paiement et avoir un impact sur la méthode de provisionnement PSAP utilisée.

Enfin, nous avons montré que certaines méthodes classiques telles que le Chain Ladder impliquaient parfois une multiplication des outils dans un même contrat (décomposition en plusieurs triangles) afin de respecter les hypothèses d'homogénéité de ces méthodes.

Une méthode universelle permettant de lever ces contraintes d'homogénéité peut alors être intéressante. Nous avons présenté et appliqué la méthode CANN sur les deux portefeuilles étudiés. Ne possédant pas plus de données que celles des méthodes très robustes telles que le Chain Ladder de Mack ou le Bootstrap, nous n'attendons pas de meilleurs résultats. Un compromis aurait pu être trouvé en conservant les méthodes classiques tout en réduisant le nombre de triangles via un regroupement de périmètres homogènes par clustering. Aussi, une régression sur la faible quantité de données existantes aurait pu donner satisfaction.

Cette méthode CANN apparaît néanmoins comme la réponse à la problématique en cas de données tête par tête massives à disposition. La mise en place d'une telle méthode implique des choix pouvant faire varier les résultats obtenus. En particulier, le choix du scope d'apprentissage est crucial, et la notion de dérive temporelle doit être suivie. Aussi, les choix d'optimisation des hyperparamètres peuvent avoir une grande influence sur les estimations finales.

Bien que cette méthode donne une estimation Best-Estimate de la PSAP, en considérant des variables discrètes et continues dans le calcul des coefficients de passage du triangle, elle ne propose pas de quantification d'erreurs d'estimation. Cela peut être pénalisant lorsqu'on la compare à une méthode stochastique classique.

Enfin, des précautions sont à prendre en terme de documentation et d'audit. Un niveau d'explicabilité élevé est recommandé par l'ACPR pour ce type de méthode, avec notamment une analyse structurelle de l'algorithme à fournir dans notre cas d'étude.

BIBLIOGRAPHIE

ACPR – pôle Fintech-Innovation, DUPONT Laurent, FLICHE Olivier, YANG Su, Gouvernance des algorithmes d'intelligence artificielle dans le secteur financier, Juin 2020, [Disponible en ligne.](#)

ACPR, Solvabilité II, MAJ le 19/03/2019, Disponible sur : <https://acpr.banque-france.fr/europe-et-international/assurances/reglementation-europeenne/solvabilite-ii>

APRA, Statistics – Annual Superannuation Bulletin, Table 12, page 21, [Disponible en ligne.](#)

APRECIALIS, Le vocabulaire technique de la Réassurance, La Tribune de l'assurance, Septembre 2009, [Disponible en ligne.](#)

CLEISS, Le régime américain de sécurité sociale, MAJ 2021, Disponible sur : https://www.cleiss.fr/docs/regimes/regime_etatsunis.html

CNAM – département informatique, Apprentissage, réseaux de neurones et modèles graphiques, pages 50 à 52 [Disponible en ligne.](#)

CHAINLADDER, P&C Loss reserving with Python, Version 0.8.4, Open Source, Disponible sur : <https://chainladder-python.readthedocs.io/en/stable/index.html>

HABIB & RIBAN, Quelle méthode de provisionnement pour des engagements non-vie dans Solvabilité 2 ?, Chapitre 2.1, page 13, [Disponible en ligne.](#)

KPMG, Super Insights 2020, Superannuation funds over 2018/19, page 19, Disponible sur : <https://assets.kpmg/content/dam/kpmg/au/pdf/2020/super-insights-report-2020.pdf>

MACK Thomas, Distribution-free calculation of the standard error of Chain Ladder reserve estimates, Astin Bulletin 23, 1993, pages 215 à 217, [Disponible en ligne.](#)

MACK Thomas, Measuring the Variability of Chain Ladder Reserve Estimates, pages 103 à 125, [Disponible en ligne.](#)

OECD, Awareness and education on risk and insurance, Disponible sur: <https://www.oecd.org/finance/insurance/38962007.pdf>

OSCAR, Health insurance made for you, Disponible sur: <https://www.hioscar.com/individuals>

WUTHRICH Mario, Neural Networks Applied to Chain-Ladder Reserving, Disponible sur :
https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3209181_code769240.pdf?abstractid=2966126&mirid=1

ANNEXES

Annexe 1.3.2 – 1. Rappel sur les différents types de réassurance

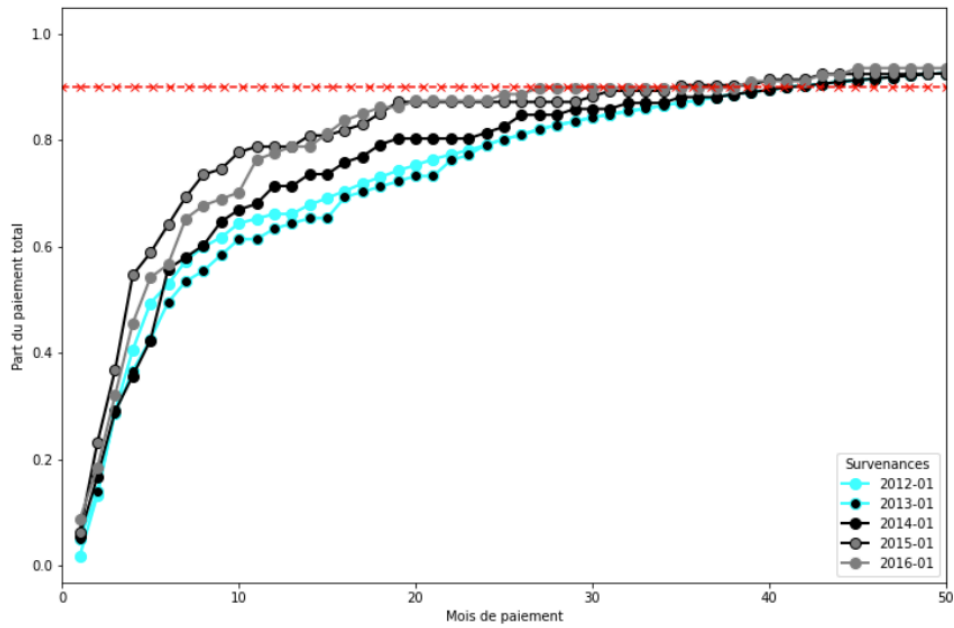
- **Les traités proportionnels**

- Quote-part: cession d'un pourcentage du risque de l'assureur direct au réassureur.
- L'excédent de plein: contrat par lequel l'assureur direct a un montant dit de rétention qui correspond à son engagement. Tous les assurés dont les capitaux garantis cumulés dépassent le montant de rétention sont cédés en réassurance pour la part du capital assuré dépassant ce plein. La part cédée en réassurance s'appelle l'excédent de plein.

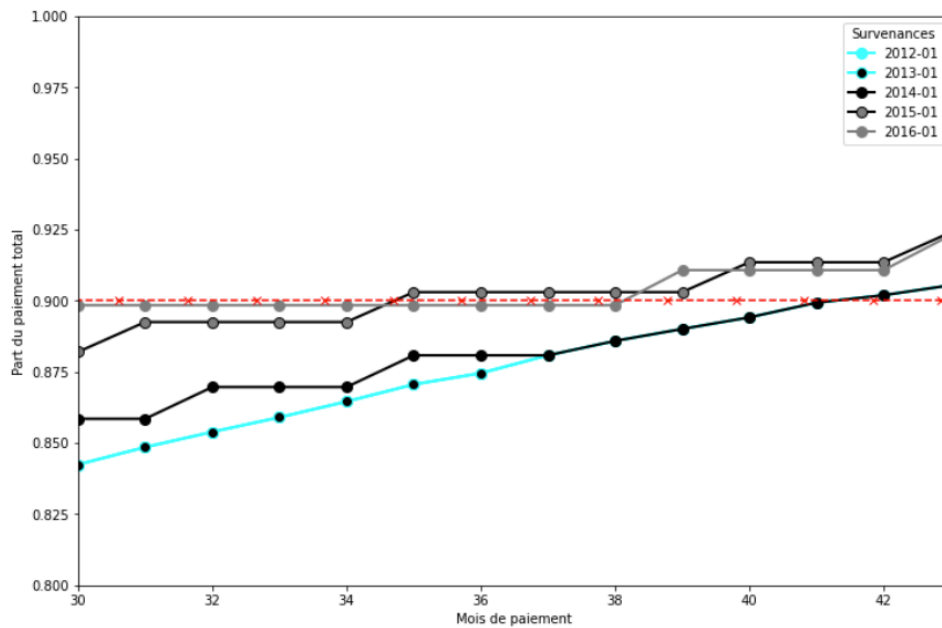
- **Les traités non proportionnels**

- L'excédent de sinistre: l'assureur direct conserve une partie du sinistre fixe (appelé priorité) et le réassureur accepte un montant de sinistre fixe dépassant la priorité (appelé portée). Si l'assureur direct veut s'engager uniquement sur le montant de priorité, la partie du sinistre supérieure à la somme de la priorité et de la portée doit être supportée par une autre réassurance.
- L'excédent de perte: type de réassurance globale lorsqu'un certain nombre d'affaires dépassent un niveau de rapport sinistre à prime fixé.

Annexe 2.2.1 – 1. Dimensionnement du triangle - AUSTRALIE

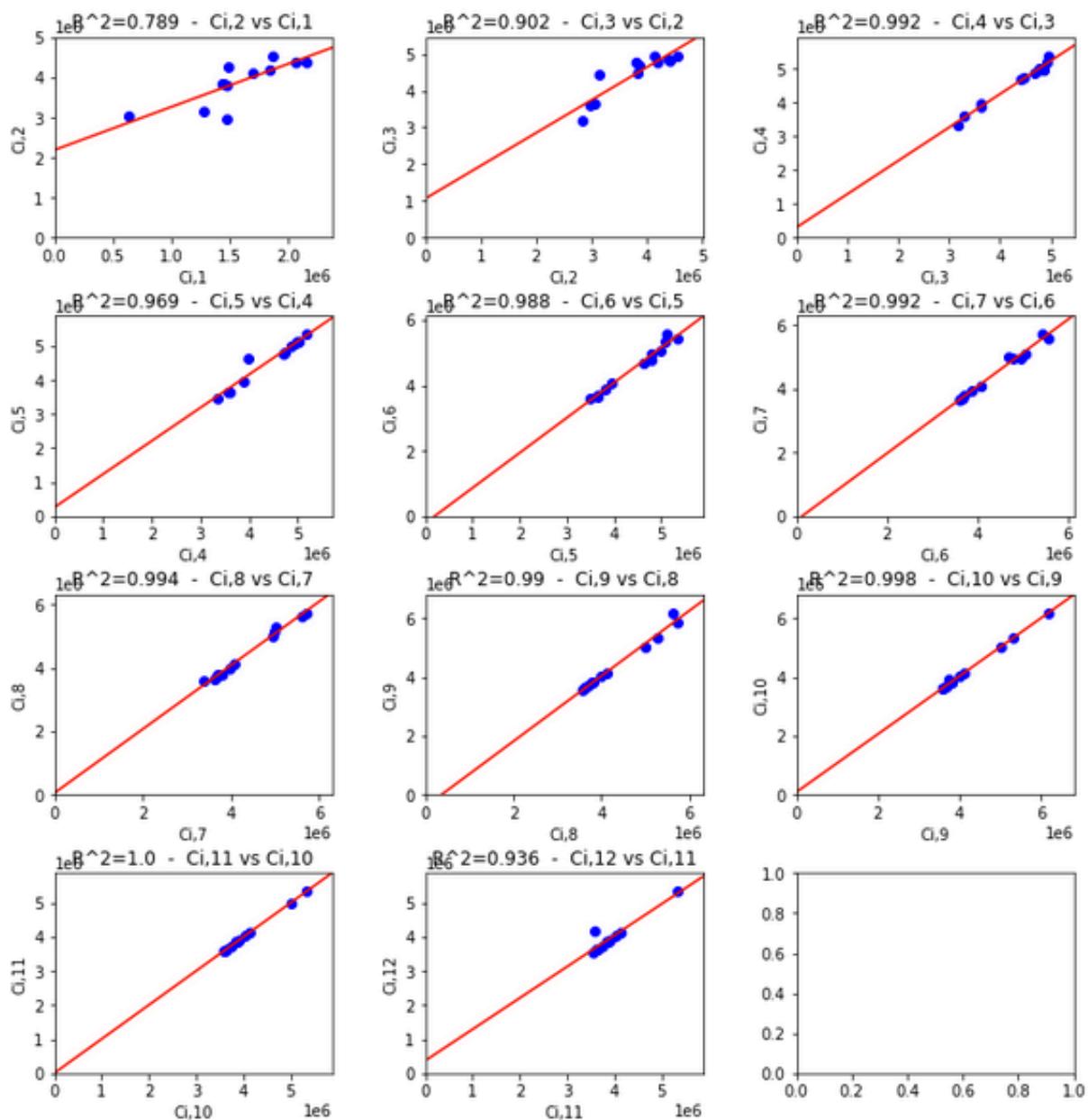


Visualisation des cadences de paiement pour les survenances de Janvier 2012, 2013, 2014, 2015 et 2016

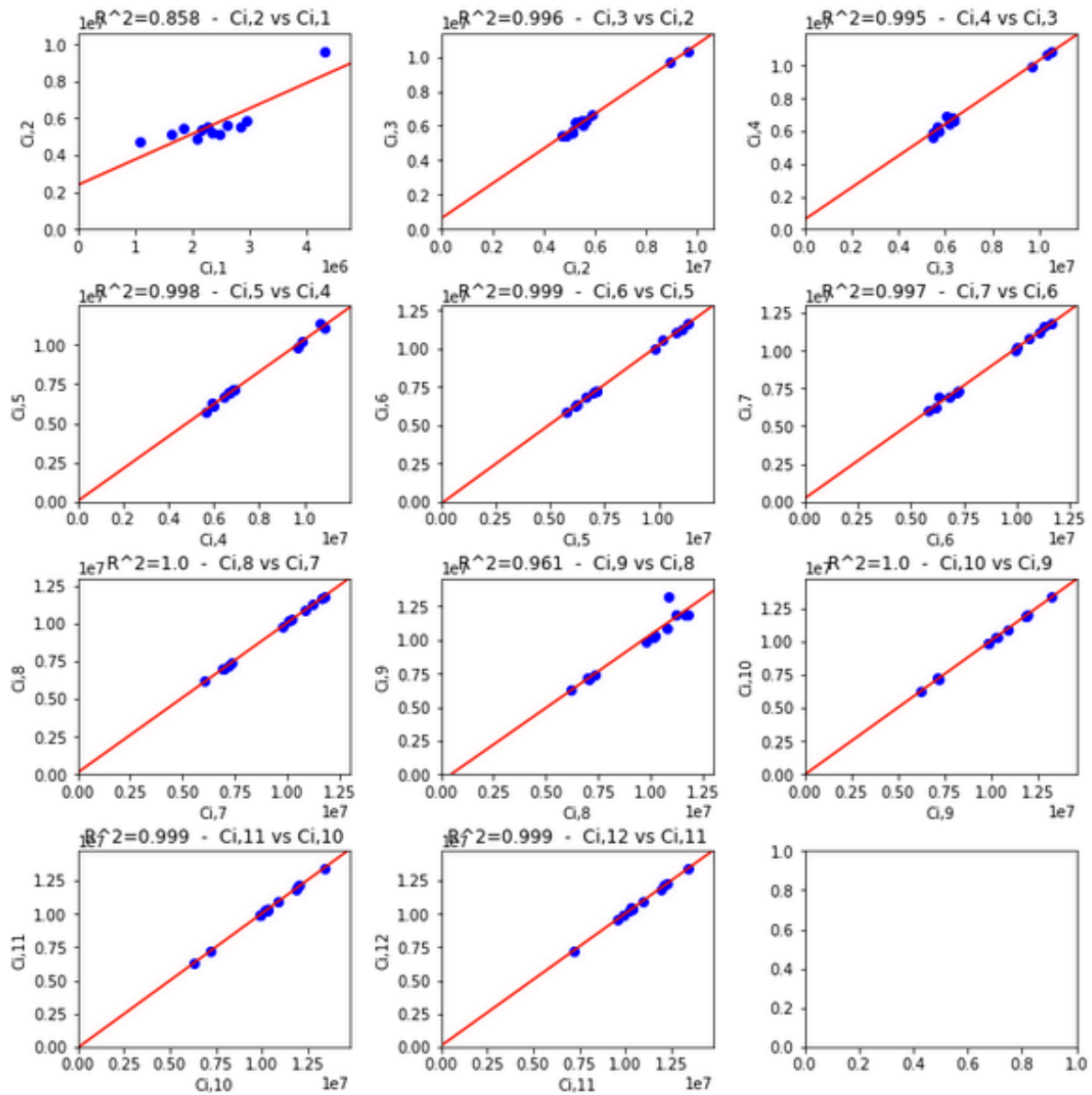


Visualisation des cadences de paiement pour les survenances de Janvier 2012, 2013, 2014, 2015 et 2016 (zoom)

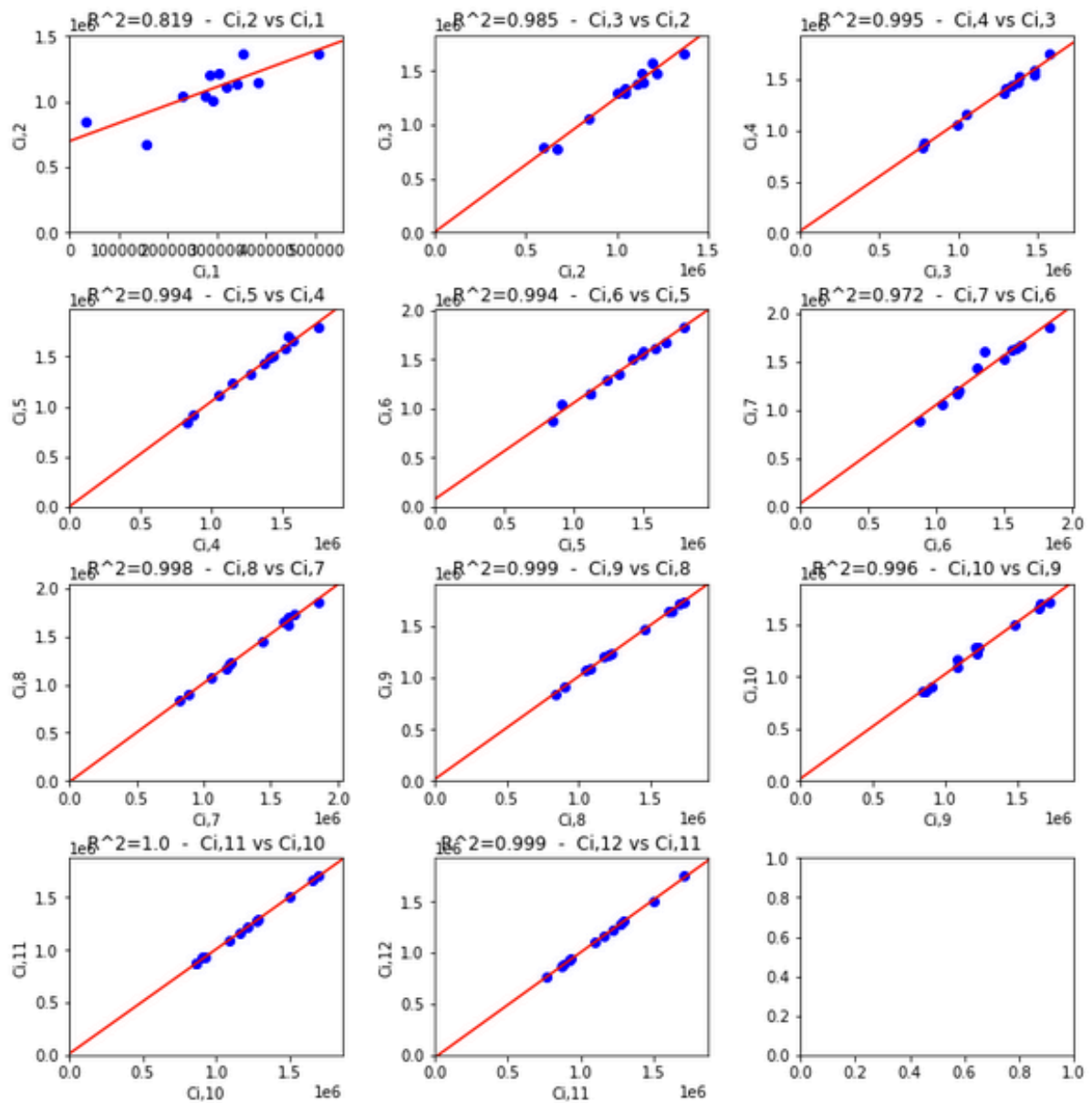
Annexe 2.2.1 – 2. Chain Ladder – Validation hypothèse H_2 - OSCAR



Ensemble des points $\{C_{i,j}; C_{i,j+1}\}$ (Oscar – état de Californie – triangle fin 2019)

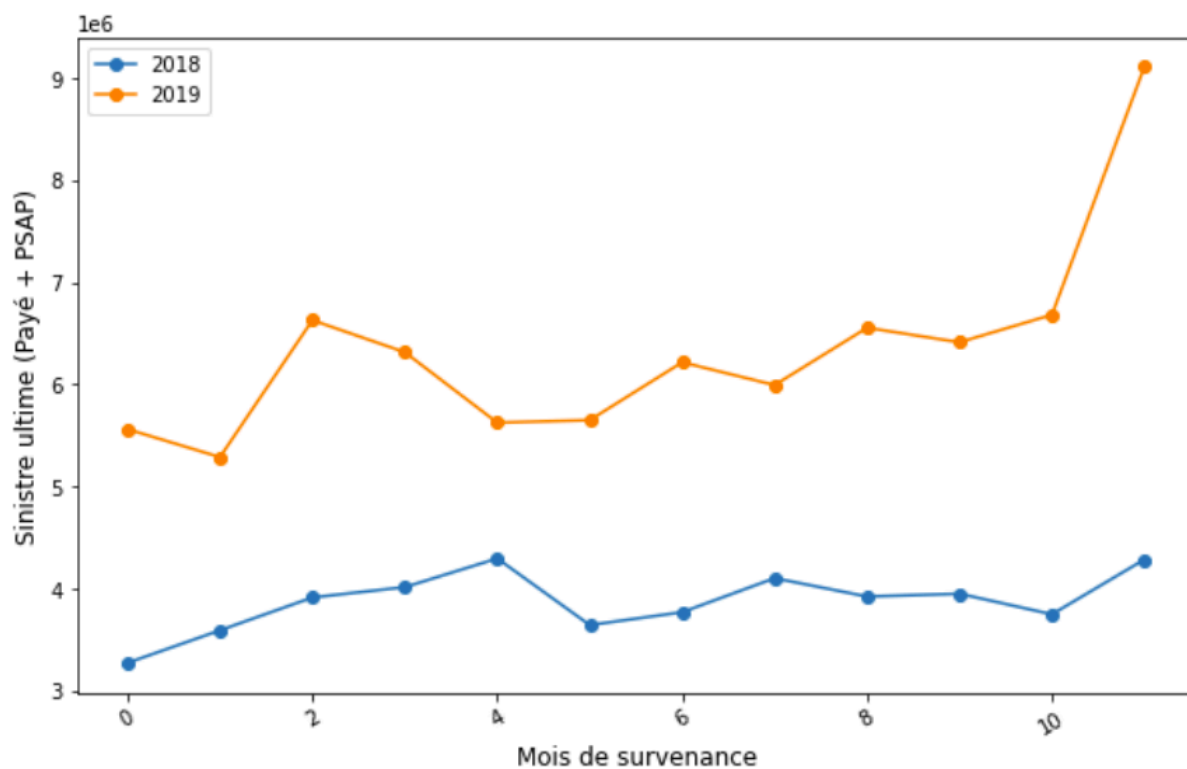


*Ensemble des points $\{C_{i,j}; C_{i,j+1}\}$ (Oscar – état du **Texas** – triangle fin 2019)*

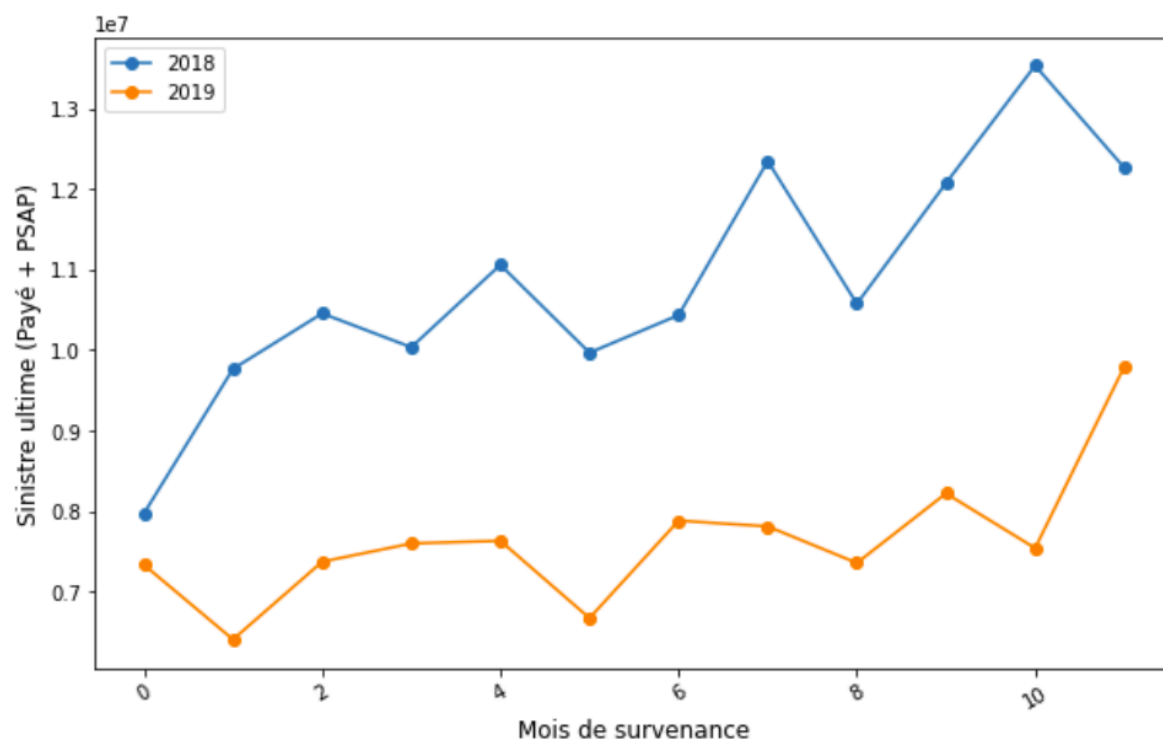


Ensemble des points $\{C_{i,j}; C_{i,j+1}\}$ (Oscar – état du New-Jersey – triangle fin 2019)

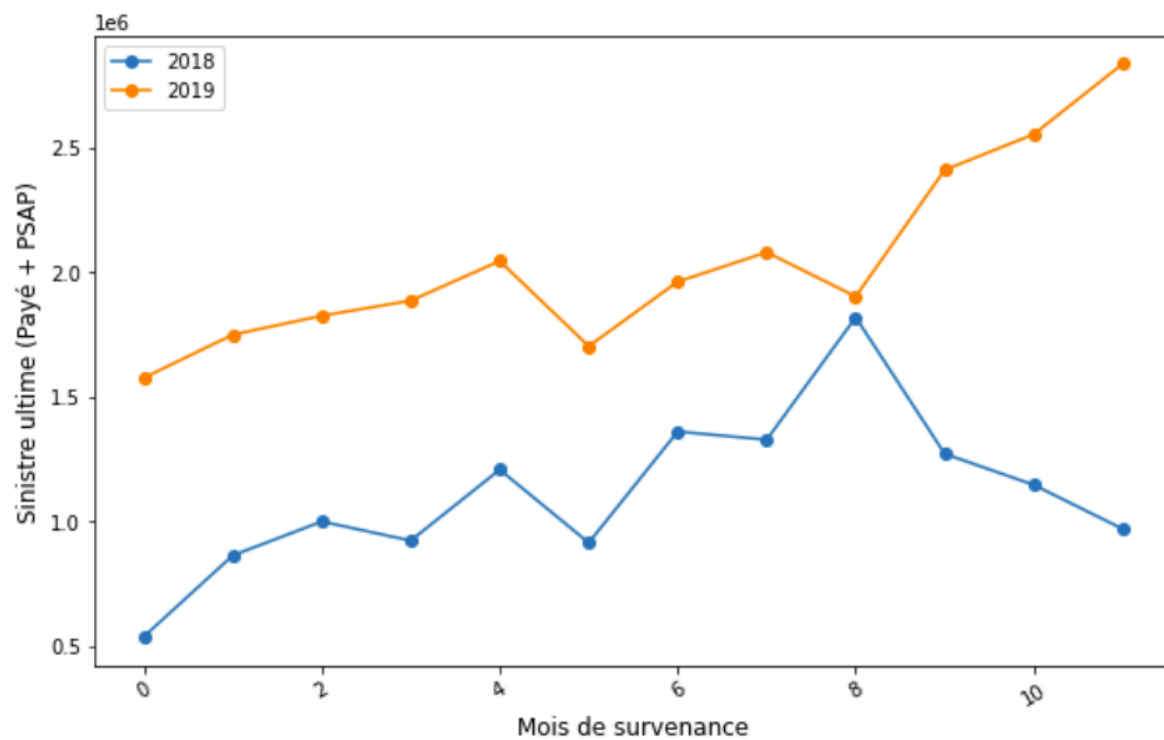
Annexe 2.2.1 – 3. Chain Ladder – Application - OSCAR



*Sinistres ultimes 2018 et 2019 après application du 12-months sliding Chain Ladder (Oscar – état de **Californie** – triangle fin 2019)*

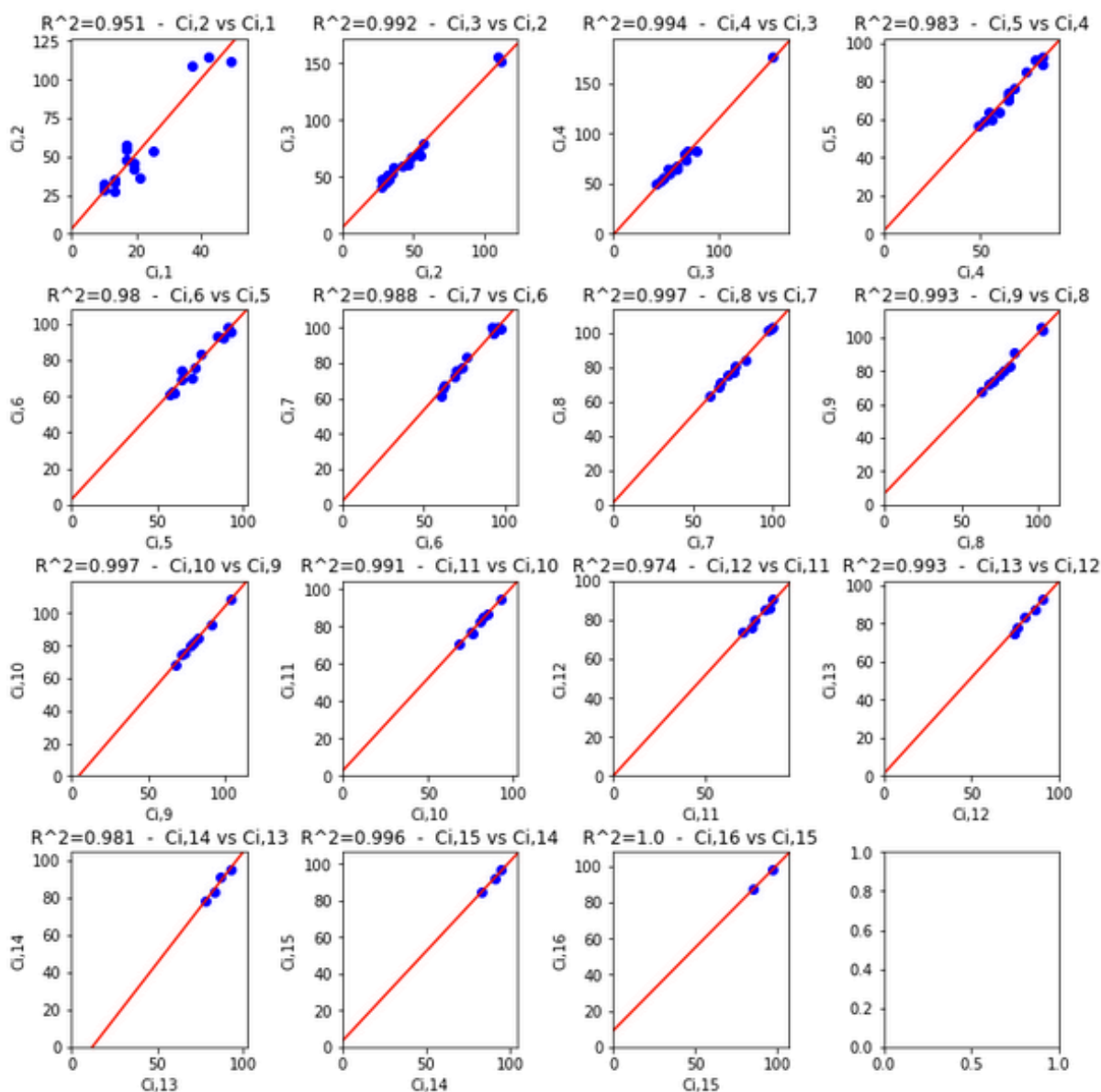


*Sinistres ultimes 2018 et 2019 après application du 12-months sliding Chain Ladder (Oscar – état du **Texas** – triangle fin 2019)*

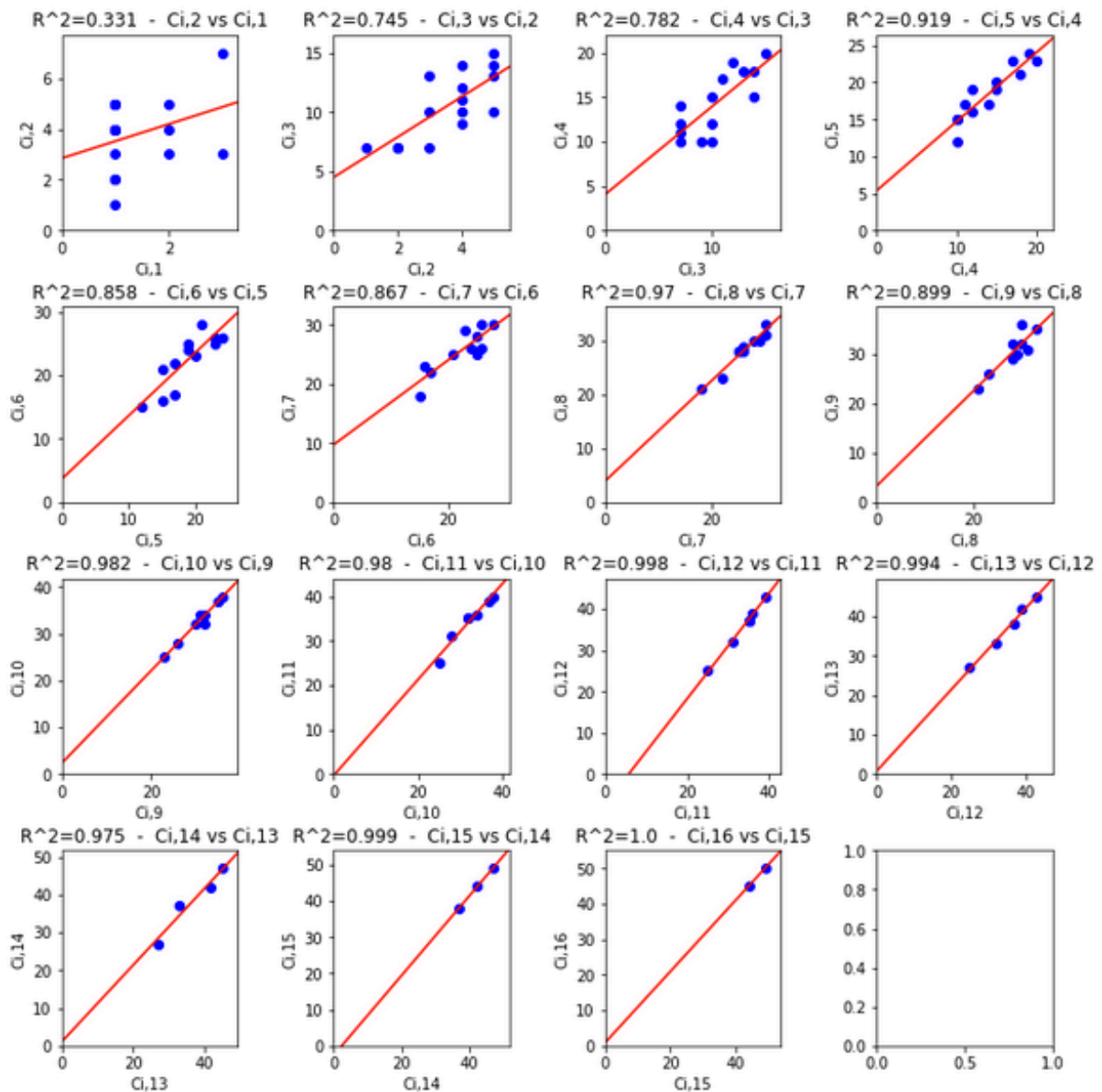


*Sinistres ultimes 2018 et 2019 après application du 12-months sliding Chain Ladder (Oscar – état du **New-Jersey** – triangle fin 2019)*

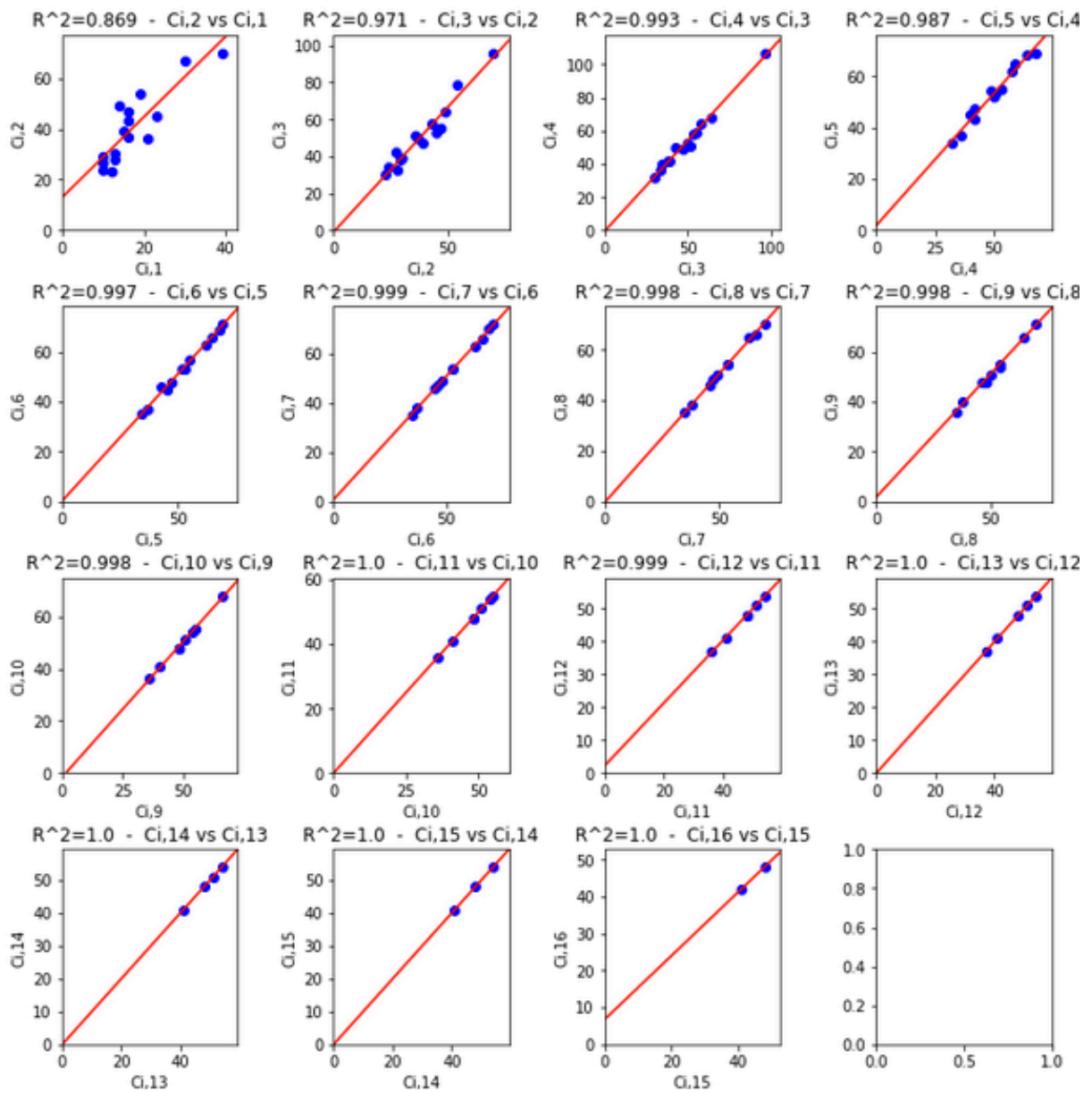
Annexe 2.2.1 – 4. Chain Ladder – Validation hypothèse H_2 - AUSTRALIE



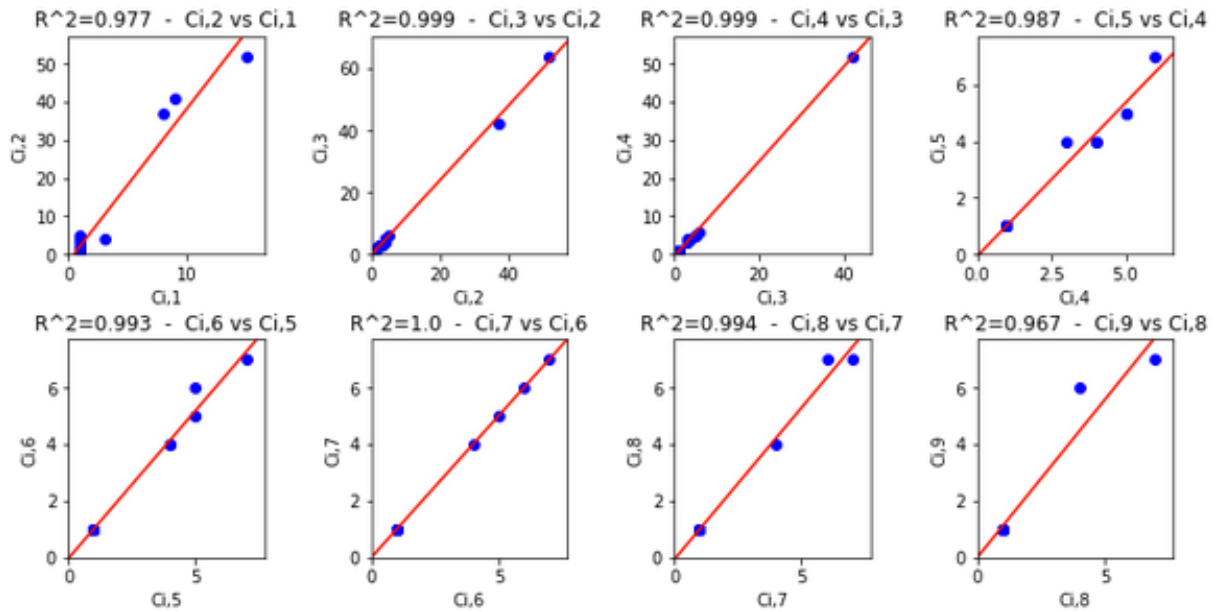
Ensemble des points $\{C_{i,j}; C_{i,j+1}\}$ (Australie – Toutes garanties agrégées – triangle au half-year 2016)



Ensemble des points $\{C_{ij}; C_{i,j+1}\}$ (Australie – Garantie TPD – triangle au half-year 2016)



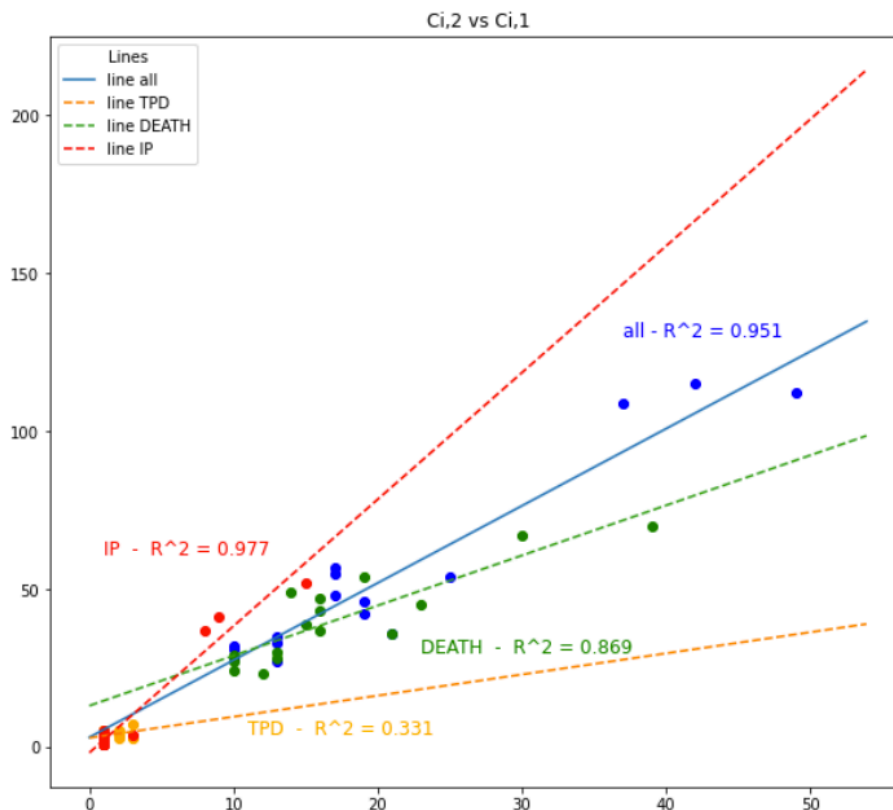
Ensemble des points $\{C_{i,j}; C_{i,j+1}\}$ (Australie - Garantie Décès - triangle au half-year 2016)



Ensemble des points $\{C_{i,j}; C_{i,j+1}\}$ (Australie – Garantie IP – triangle au half-year 2016)

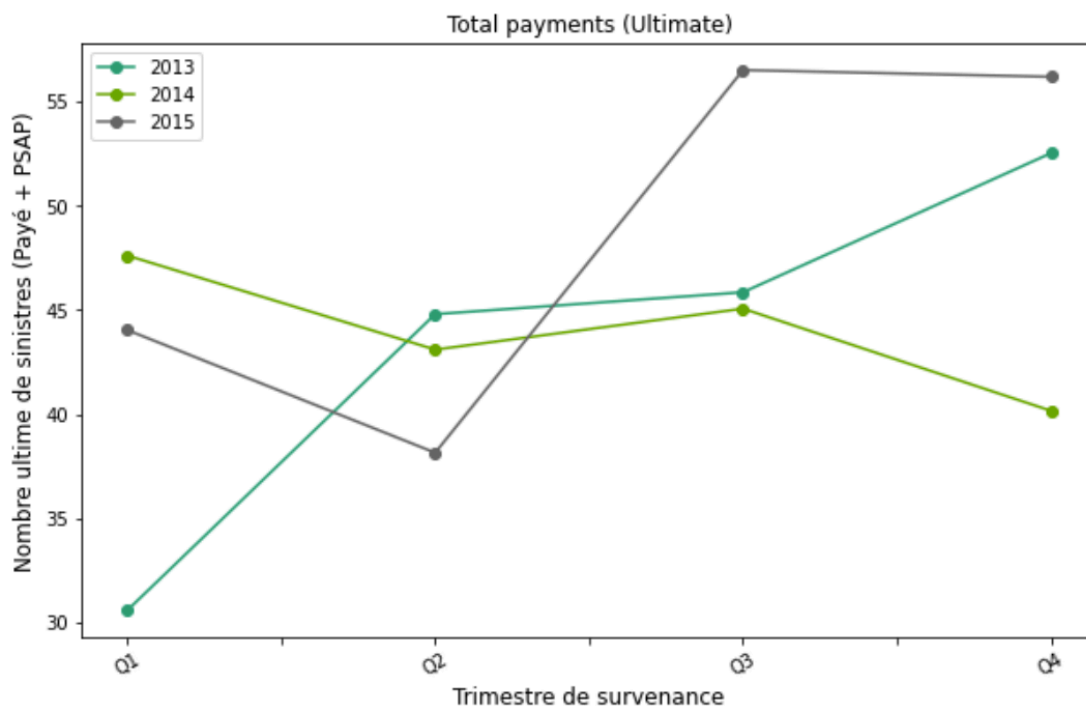
Note:

Le portefeuille de la garantie IP commence seulement en 2014, c'est pourquoi seulement 8 couples $\{C_{i,j}; C_{i,j+1}\}$ sont affichés. Ils correspondent aux données disponibles entre 2014 et 2016.

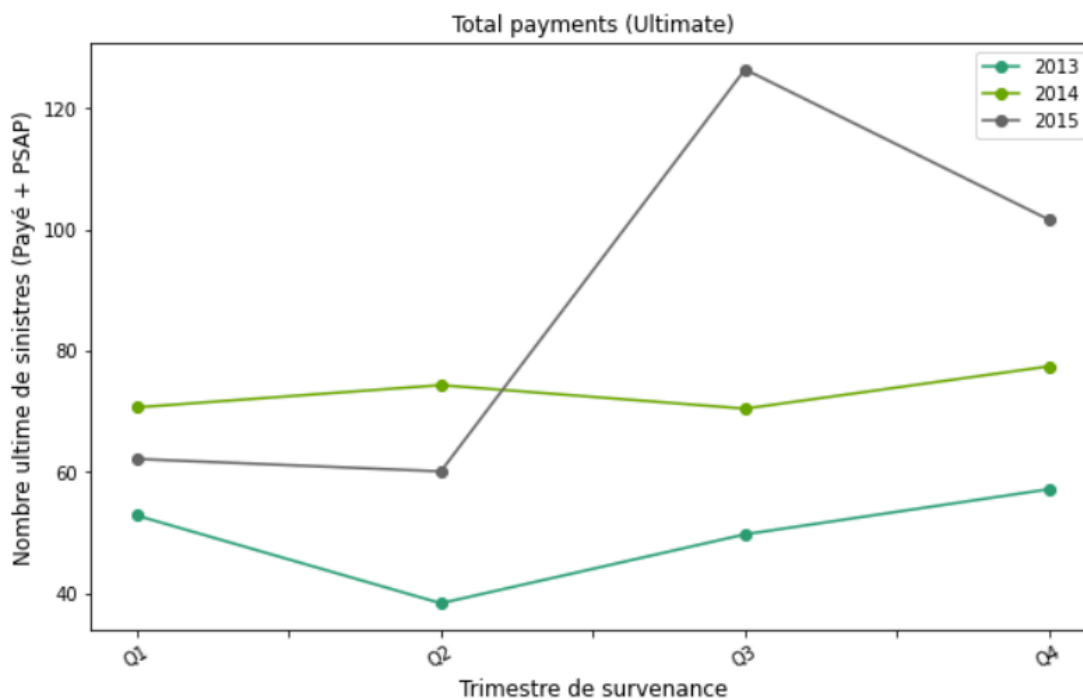


Comparaison de l'ensemble des points $\{C_{i,1}; C_{i,2}\}$

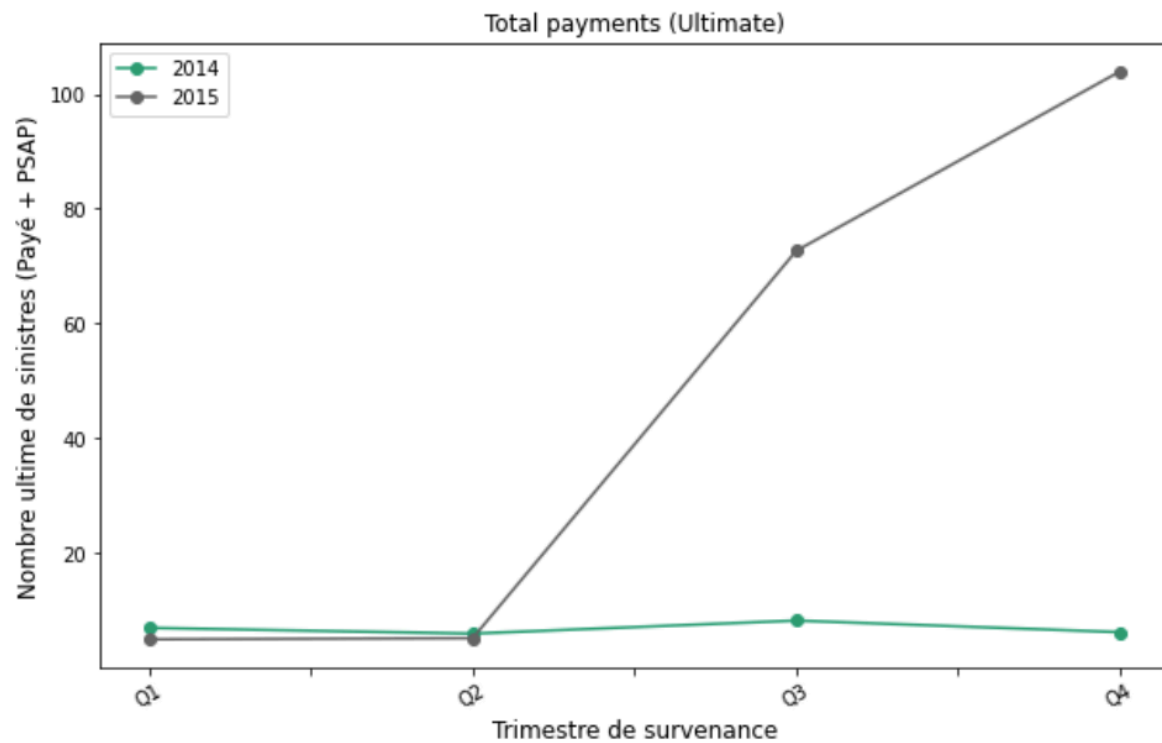
Annexe 2.2.1 – 5. Chain Ladder – Application - AUSTRALIE



Sinistres ultimes 2013, 2014 et 2015 après application du Chain Ladder (Australie – Garantie TPD – triangle au half-year 2016)



Sinistres ultimes 2013, 2014 et 2015 après application du Chain Ladder (Australie – Garantie Death – triangle au half-year 2016)

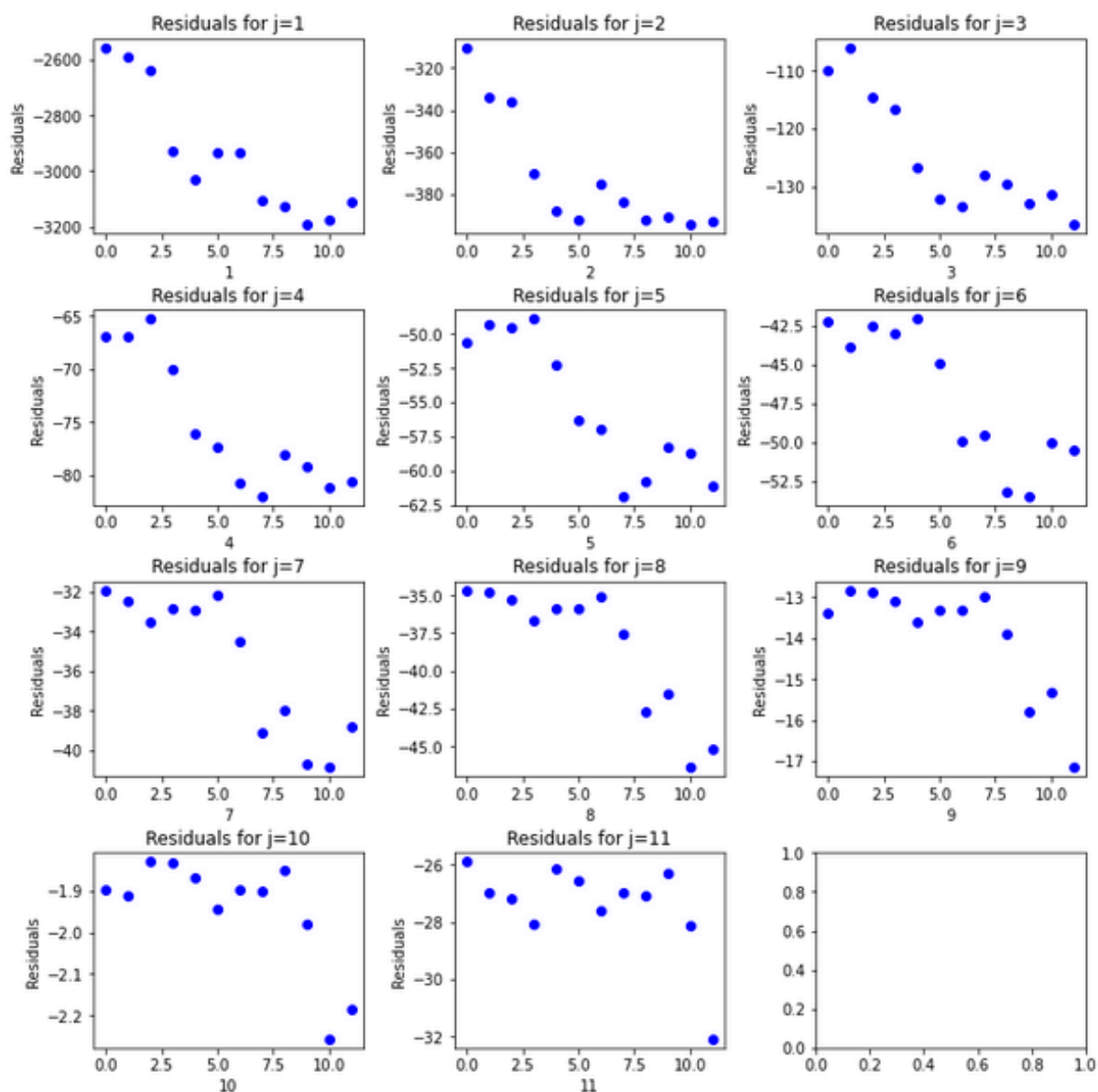


*Sinistres ultimes 2014 et 2015 après application du Chain Ladder (Australie – Garantie IP
- triangle au half-year 2016 2016)*

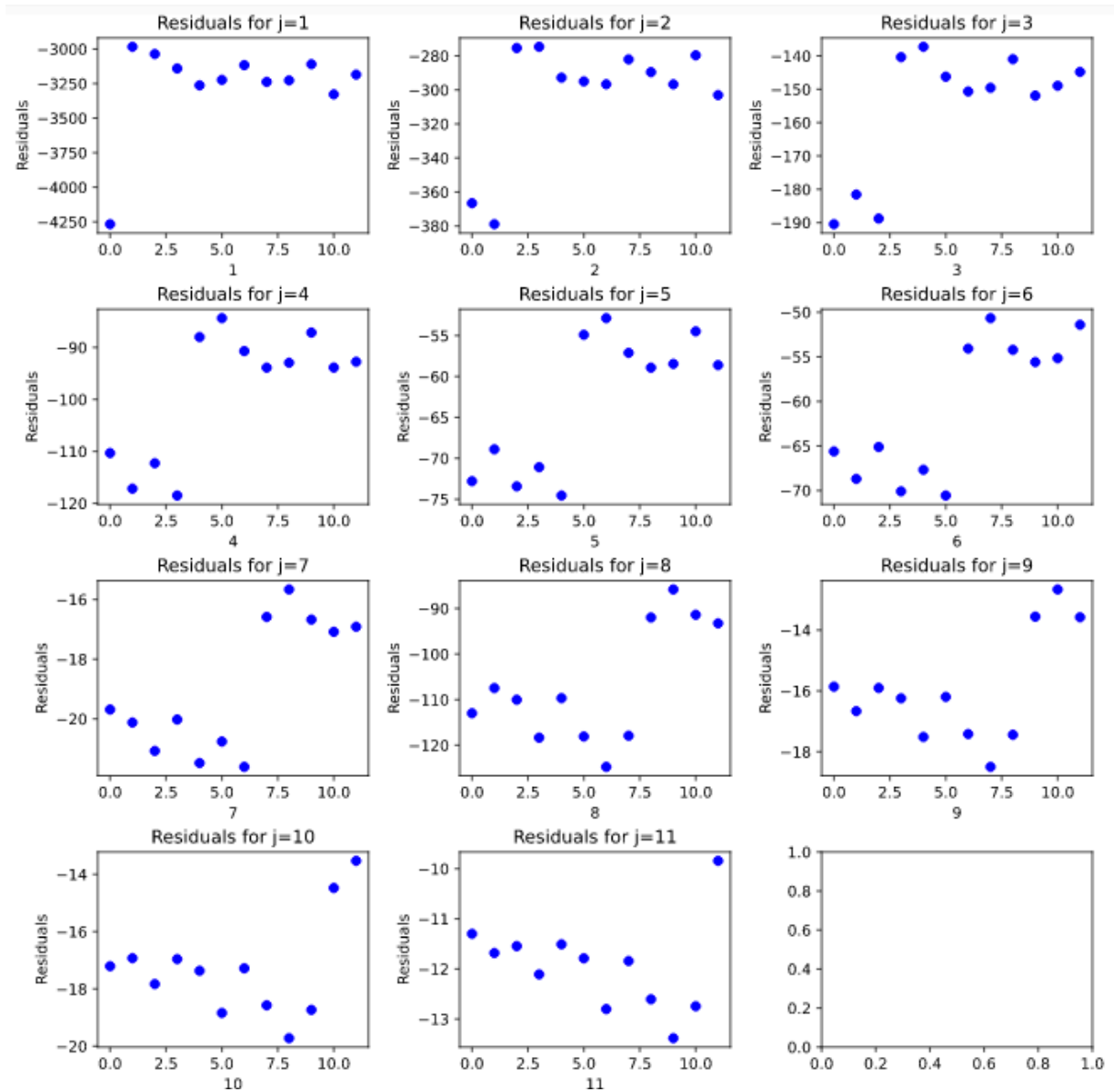
Note:

Le portefeuille de la garantie IP commence seulement en 2014, c'est pourquoi l'année 2013 n'est pas présente.

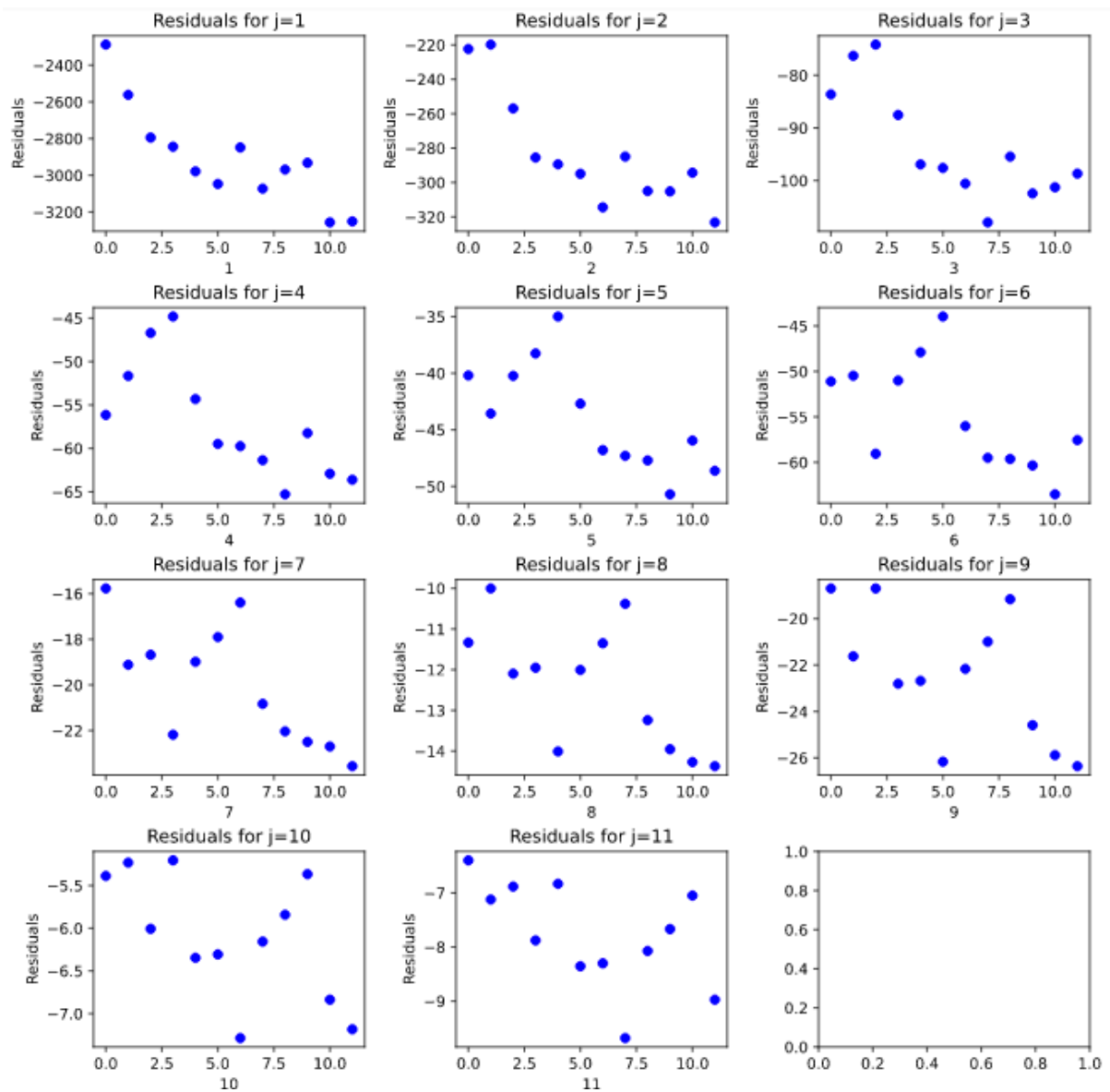
Annexe 2.2.2 – 1. Chain Ladder de Mack – Validation hypothèse H_3 - OSCAR



Ensemble des résidus en fonction de i (Oscar – état de Californie – triangle fin 2019)

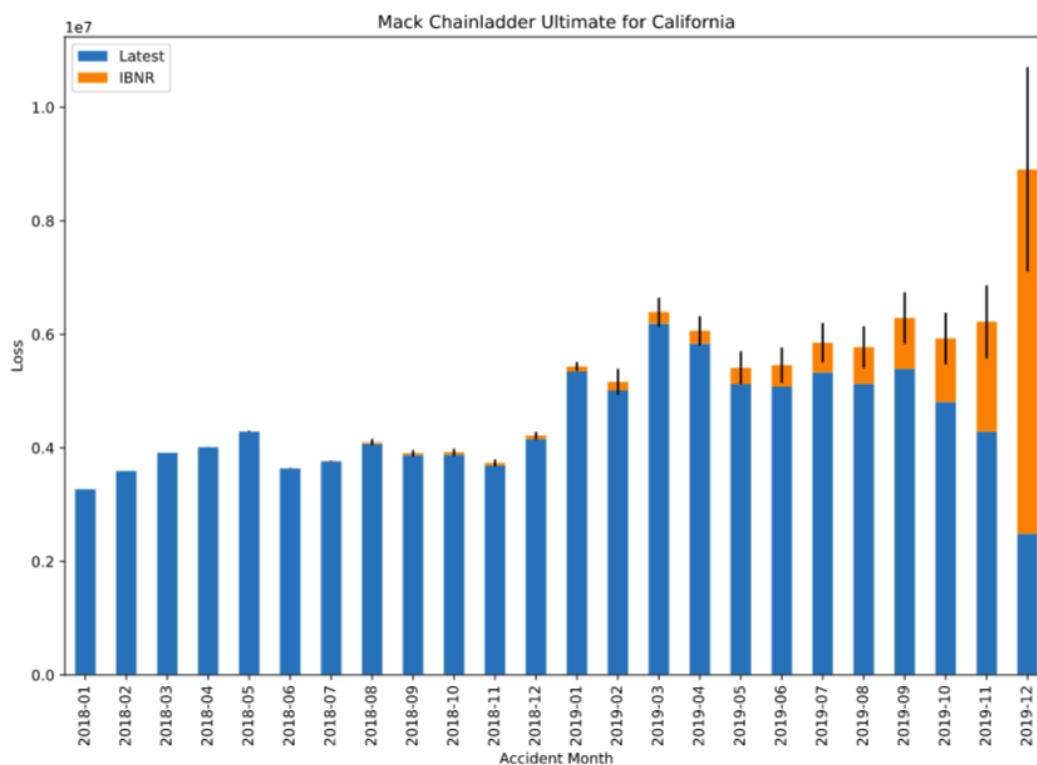


*Ensemble des résidus en fonction de i (Oscar – état du **Texas** – triangle fin 2019)*

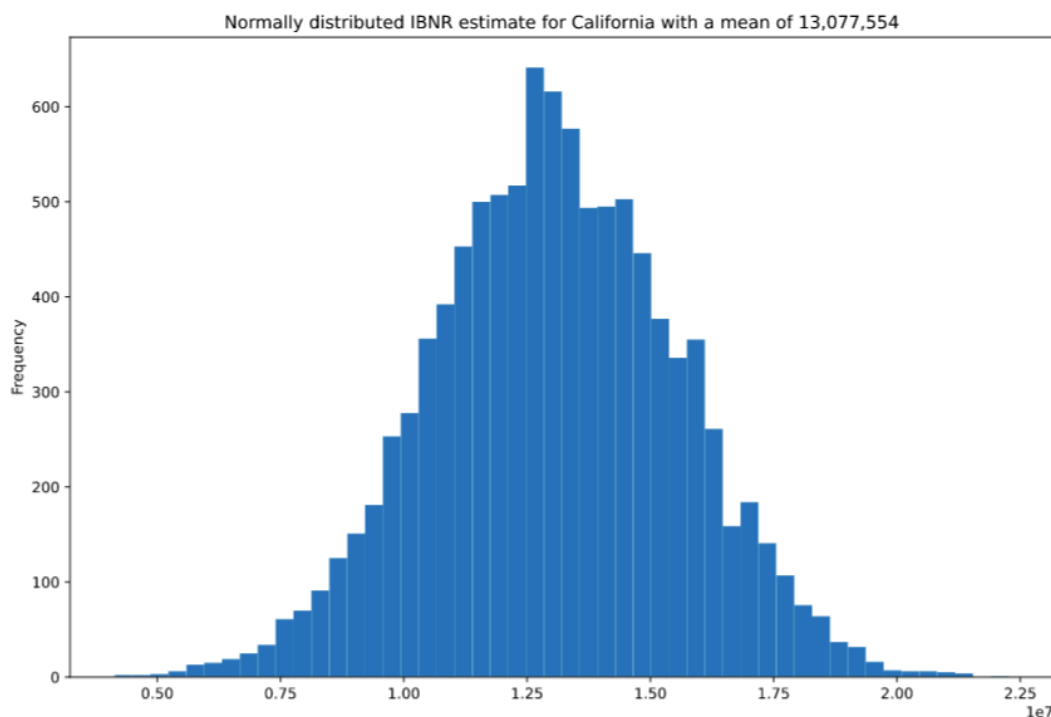


Ensemble des résidus en fonction de i (Oscar – état du New-Jersey – triangle fin 2019)

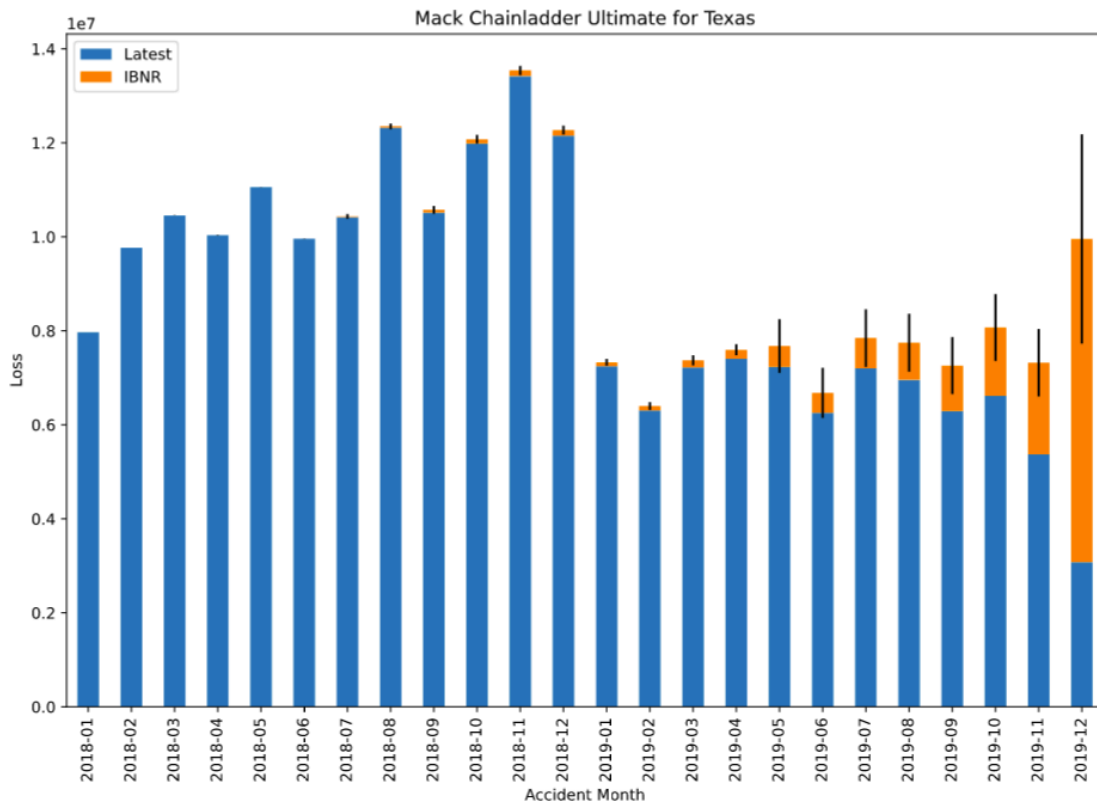
Annexe 2.2.2 – 2. Chain Ladder de Mack – Application – OSCAR



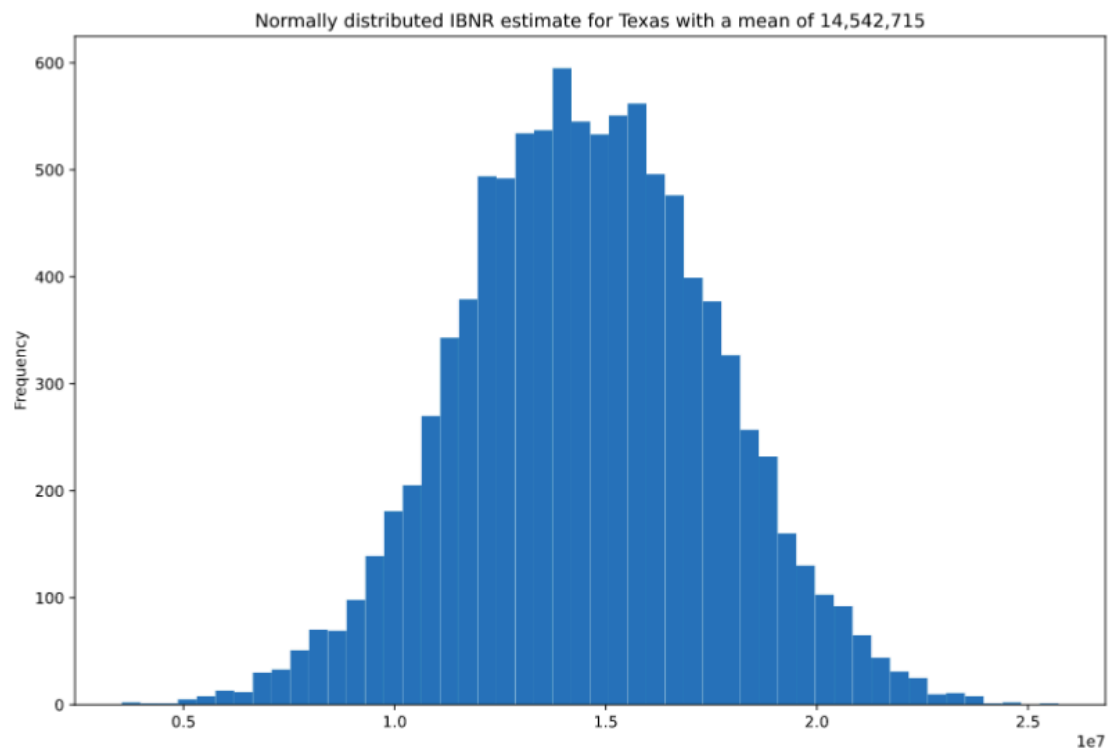
*Résultats par mois de survenance (état de **Californie** - Mack Chain Ladder – triangle fin 2019)*



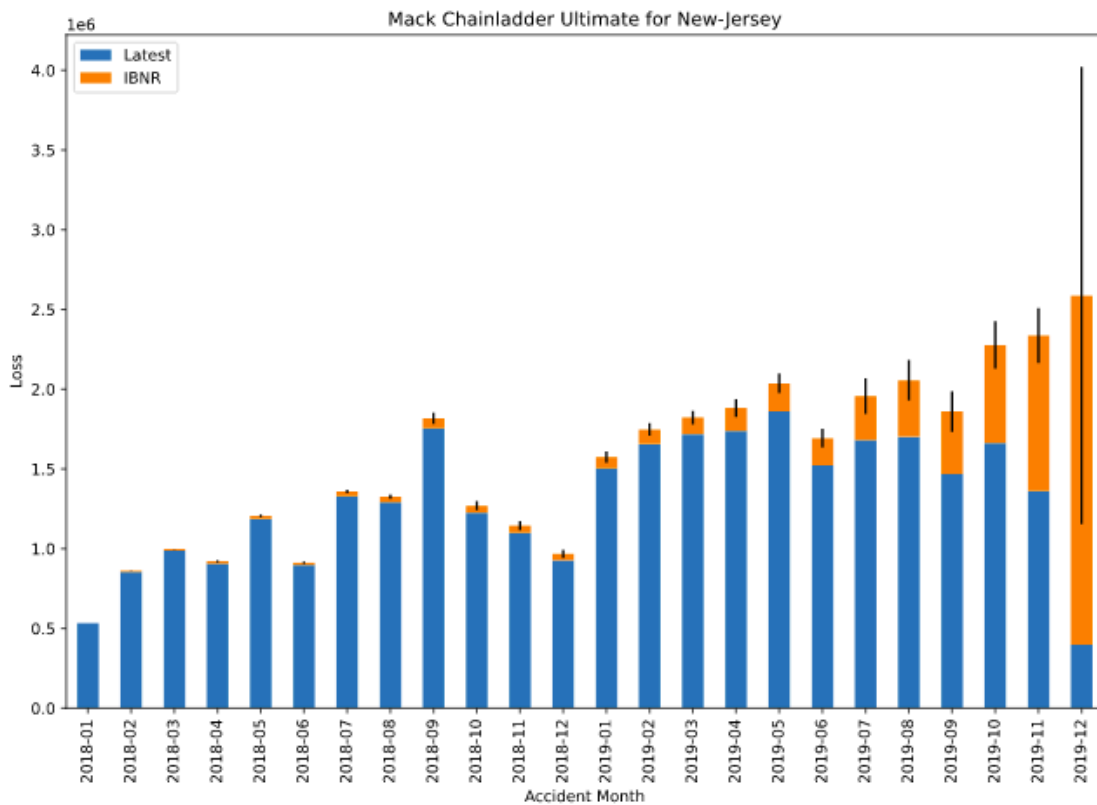
*PSAP totale de distribution supposée normale (état de **Californie** - Mack Chain Ladder – triangle fin 2019)*



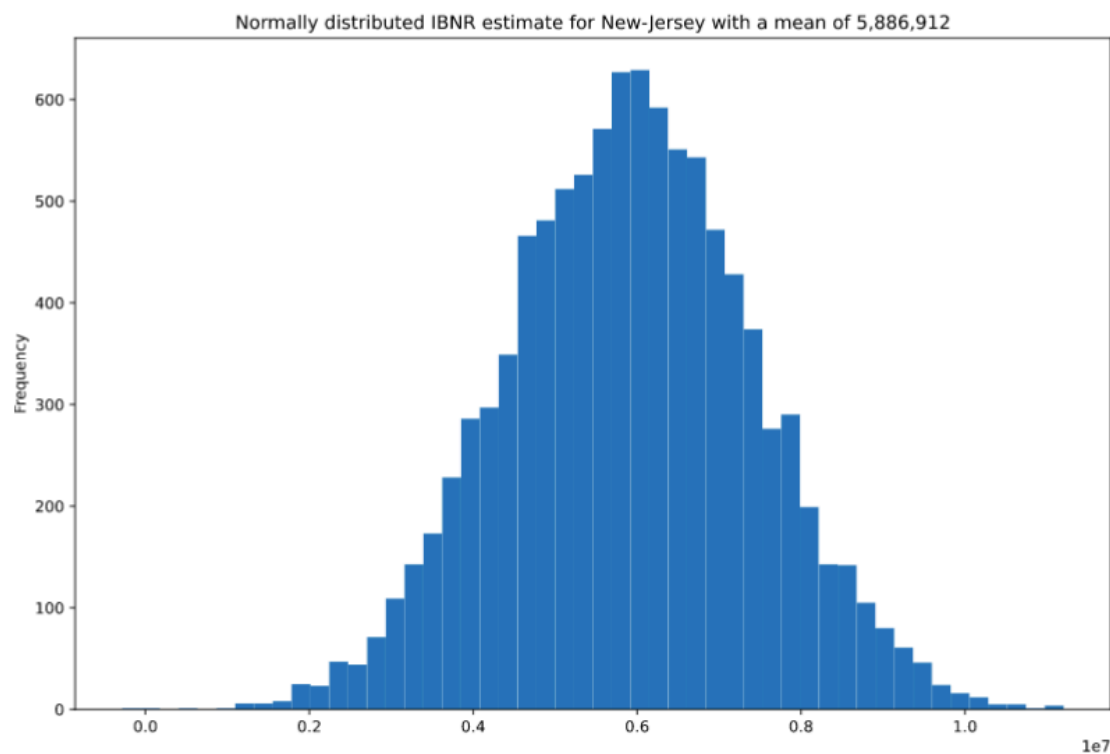
*Résultats par mois de survenance (état du **Texas** - Mack Chain Ladder – triangle fin 2019)*



*PSAP totale de distribution supposée normale (état du **Texas** - Mack Chain Ladder – triangle fin 2019)*

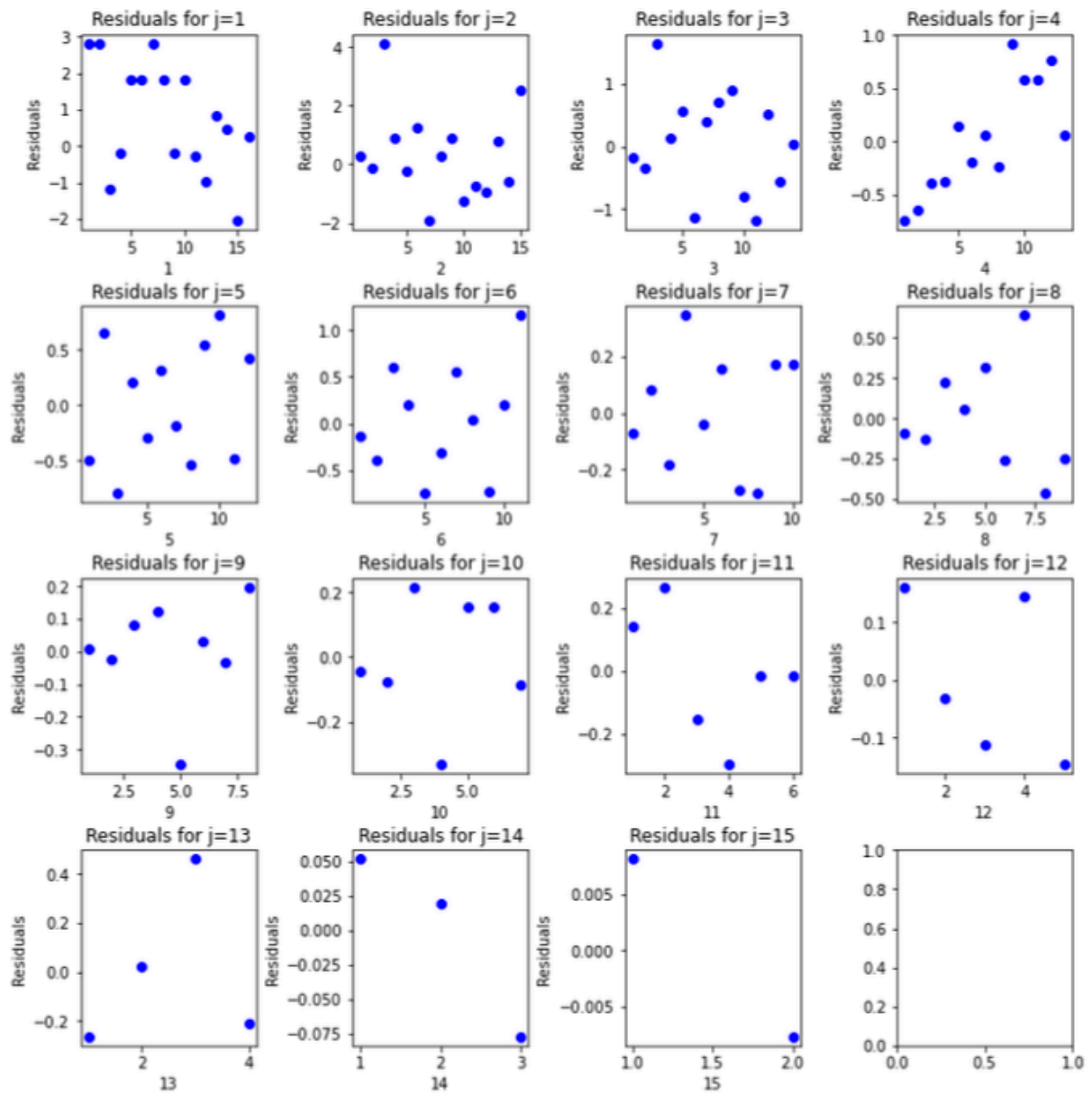


*Résultats par mois de survenance (état du **New-Jersey** - Mack Chain Ladder – triangle fin 2019)*

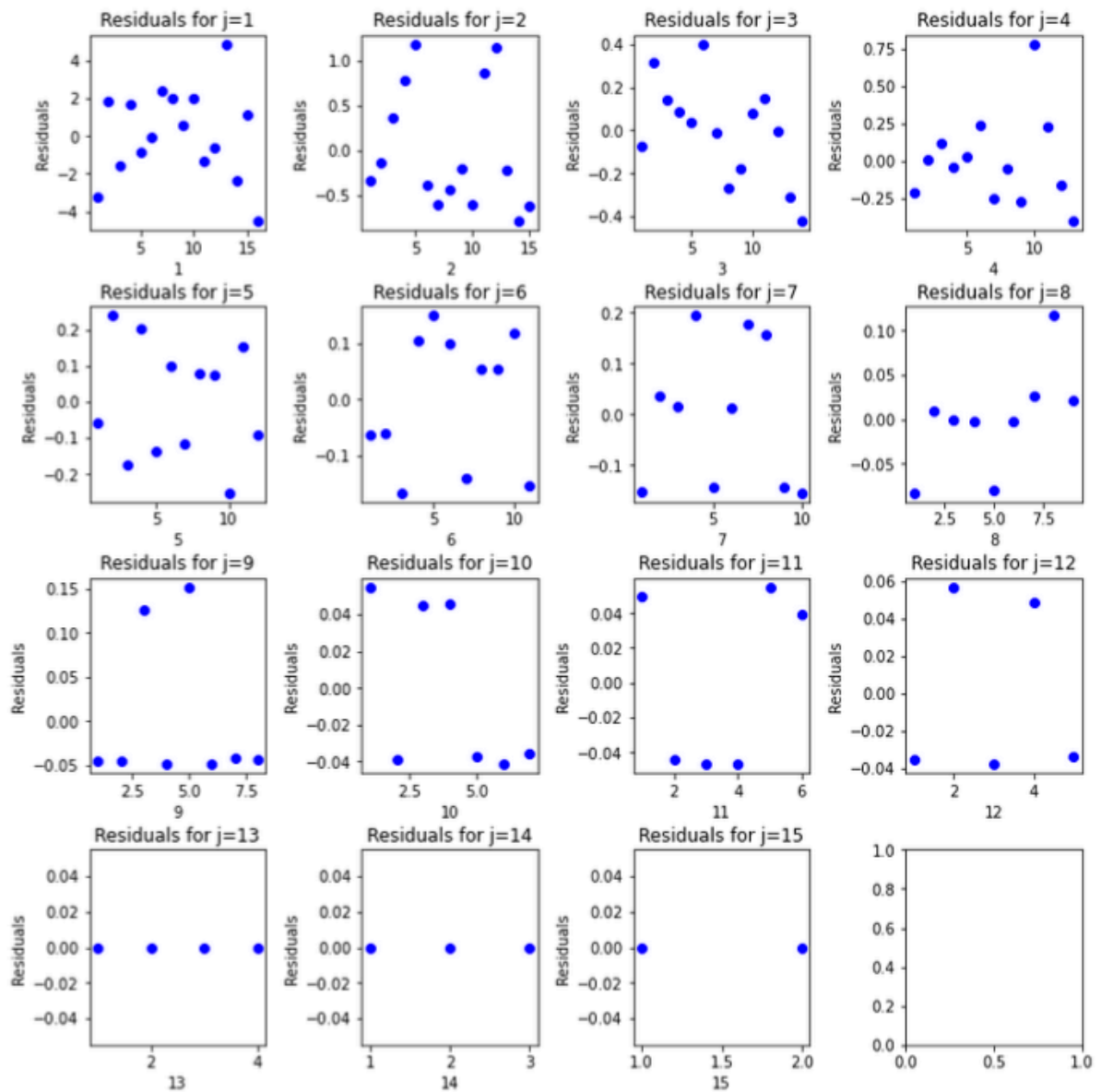


*PSAP totale de distribution supposée normale (état du **New-Jersey** - Mack Chain Ladder – triangle fin 2019)*

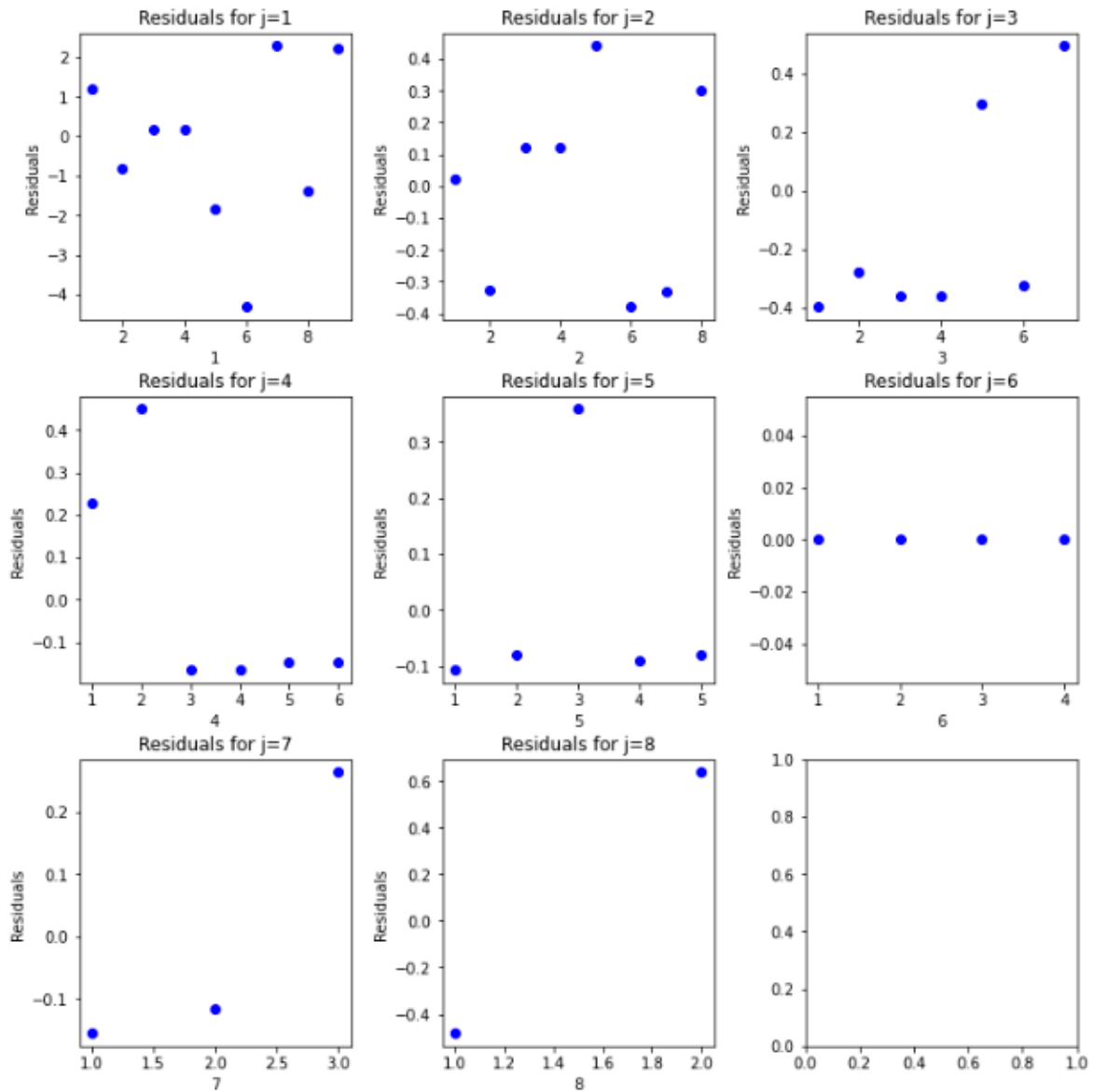
Annexe 2.2.2 – 3. Chain Ladder de Mack – Validation hypothèse H_3 - AUSTRALIE



*Ensemble des résidus en fonction de i (Australie – Garantie **TPD** – triangle au half-year 2016)*

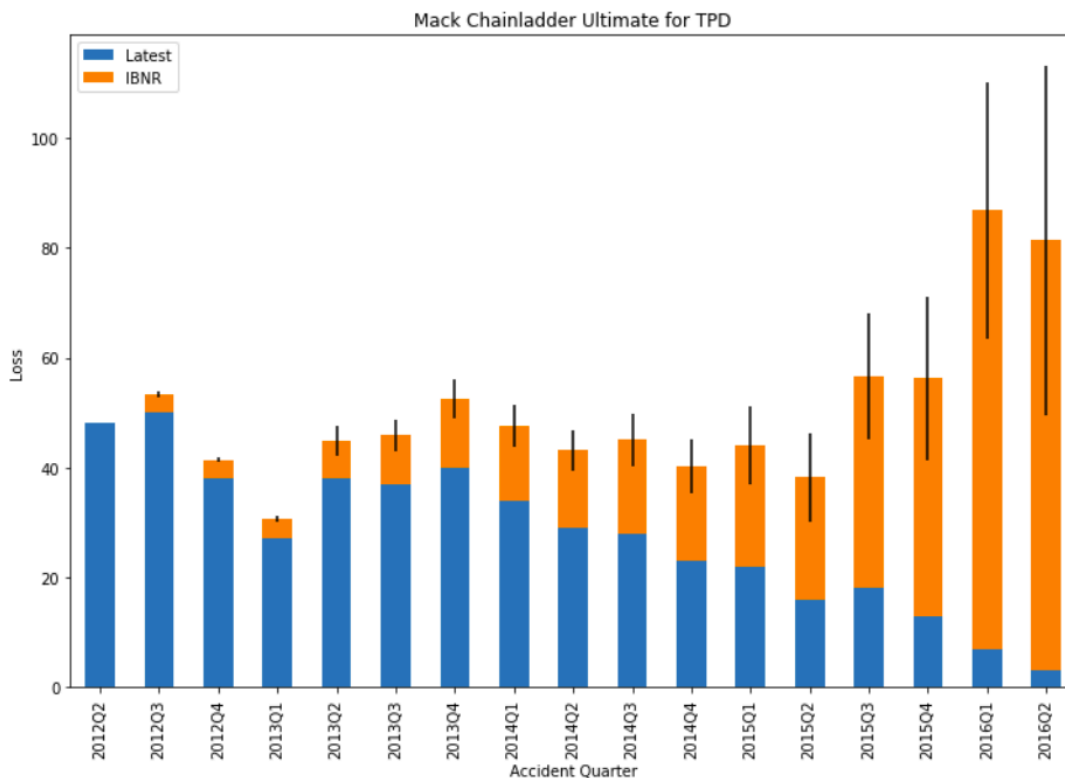


*Ensemble des résidus en fonction de i (Australie – Garantie **Death** – triangle au half-year 2016)*

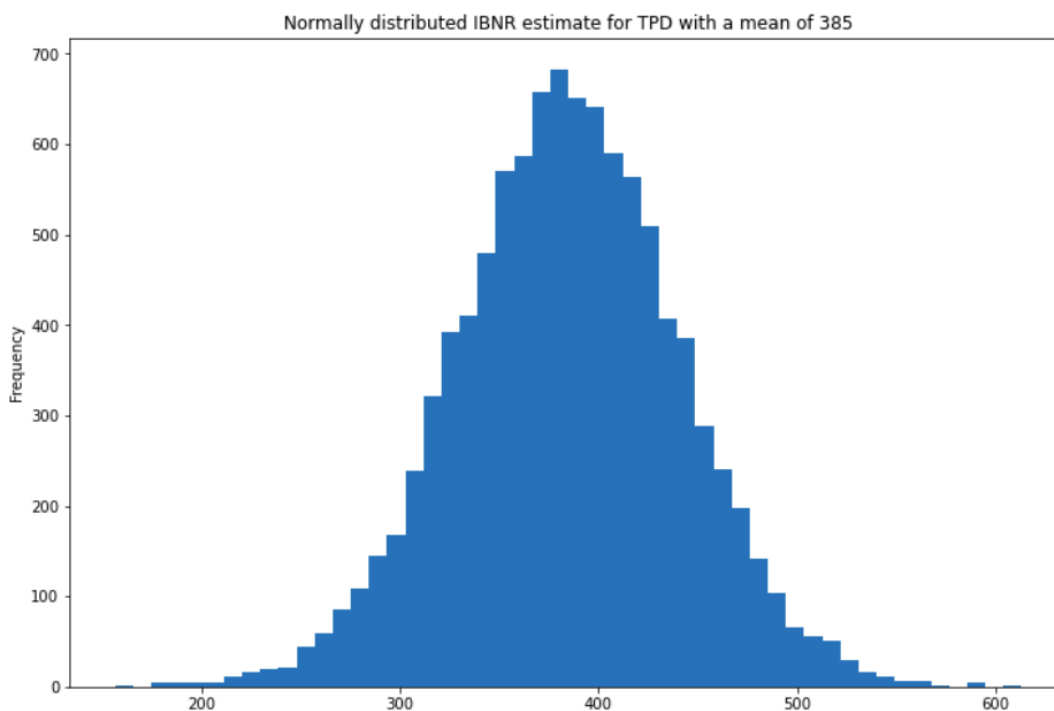


Ensemble des résidus en fonction de i (Australie – Garantie IP – triangle au half-year 2016)

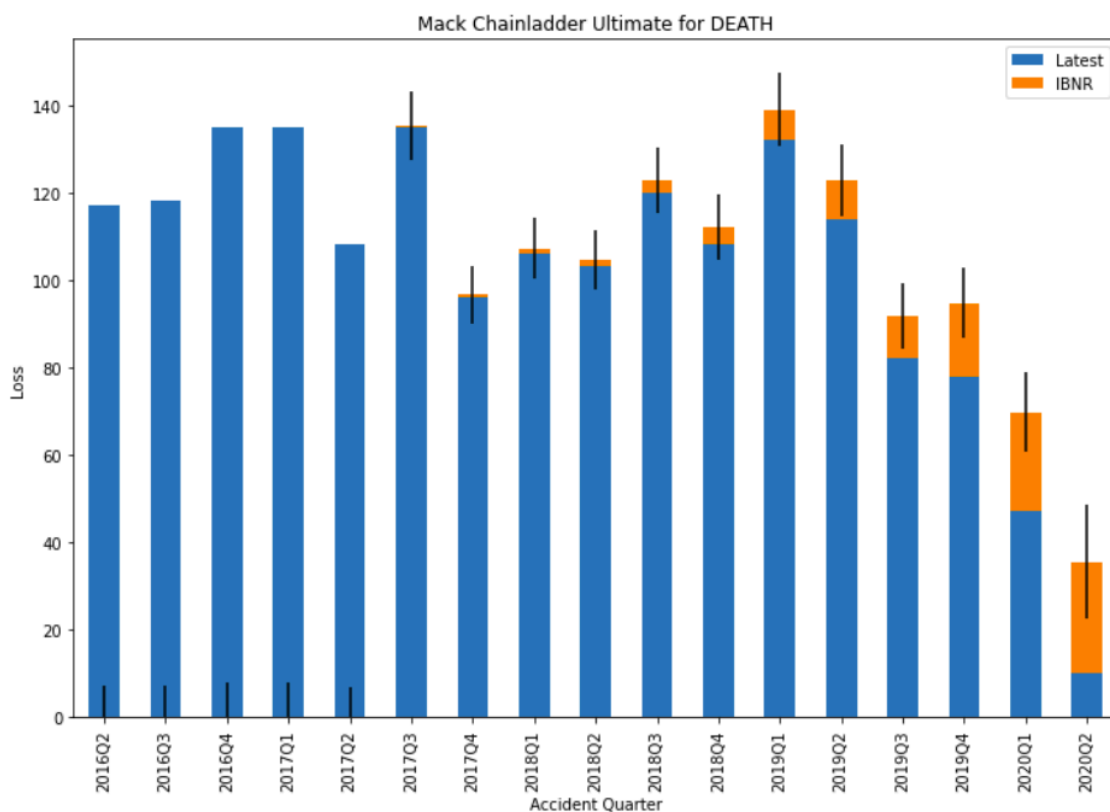
Annexe 2.2.2 – 4. Chain Ladder de Mack – Application - AUSTRALIE



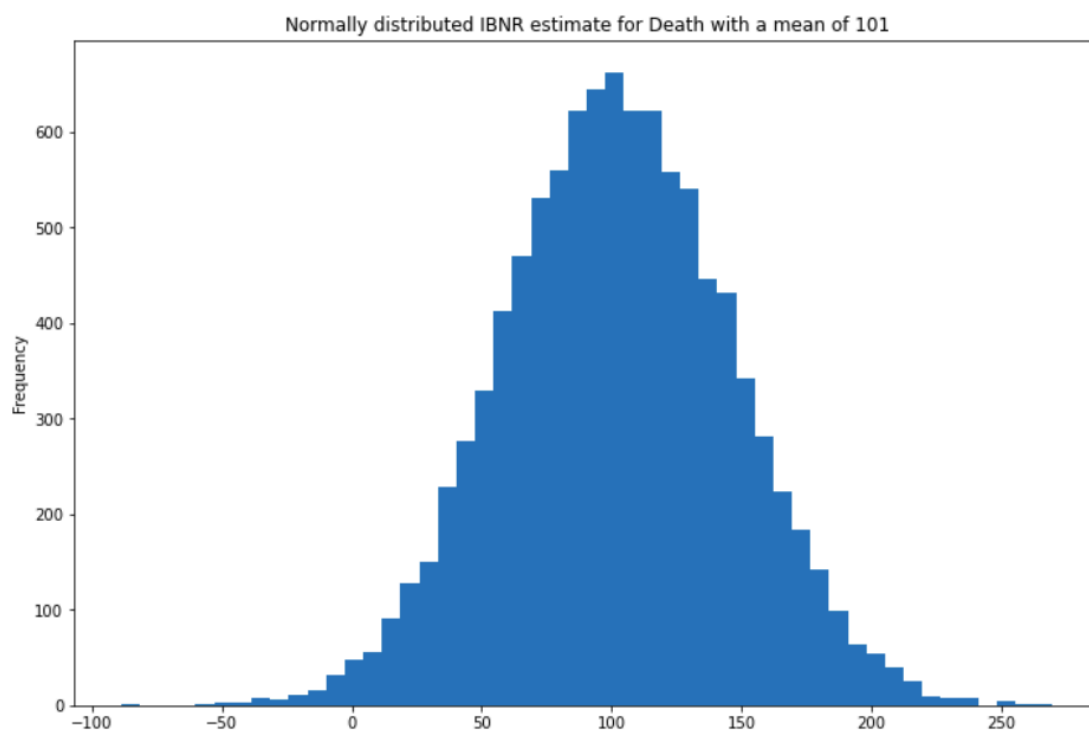
*Résultats par trimestre de survenance (Garantie **TPD** - Mack Chain Ladder – triangle au half-year 2016)*



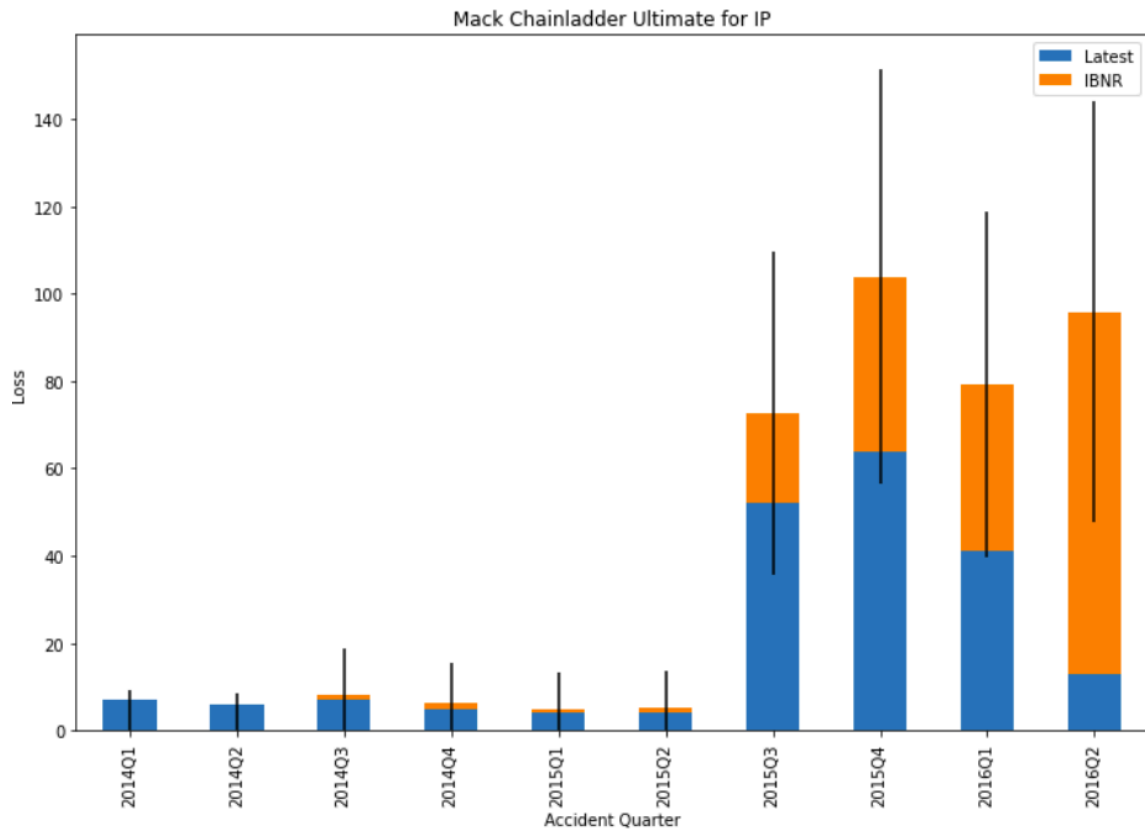
*PSAP totale de distribution supposée normale (Garantie **TPD** - Mack Chain Ladder – triangle au half-year 2016)*



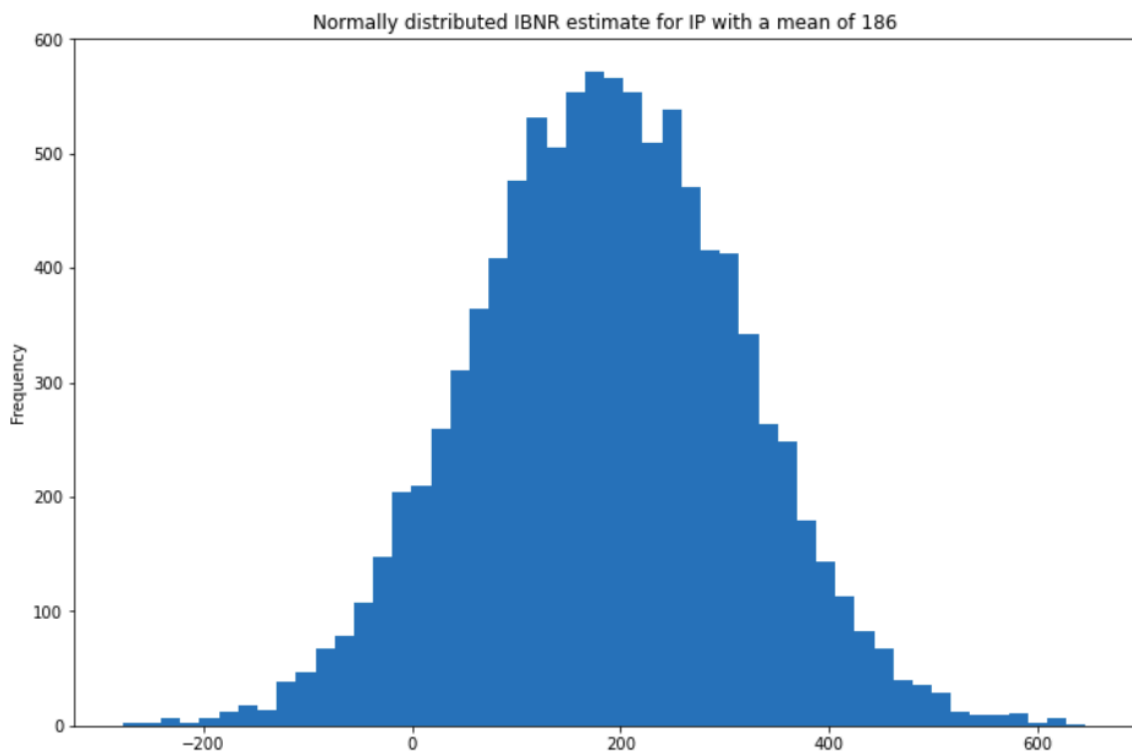
*Résultats par trimestre de survenance (Garantie **Death** - Mack Chain Ladder – triangle au half-year 2016)*



*PSAP totale de distribution supposée normale (Garantie **Death** - Mack Chain Ladder – triangle au half-year 2016)*

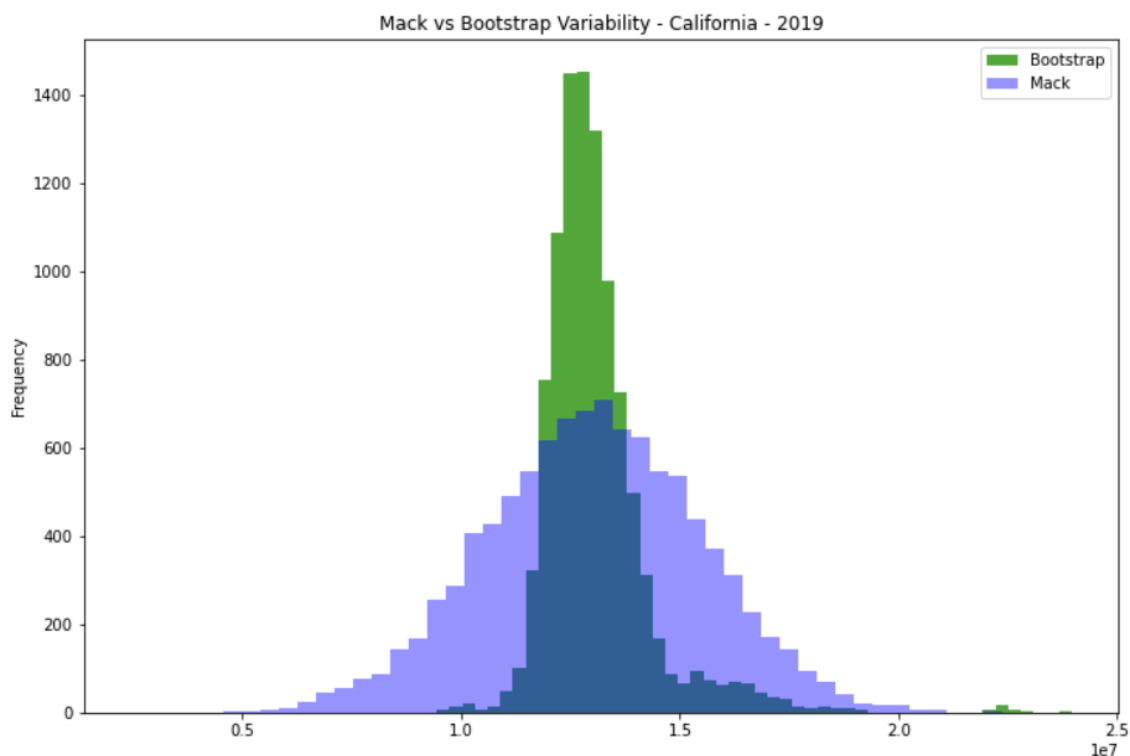


Résultats par trimestre de survenance (Garantie IP - Mack Chain Ladder – triangle au half-year 2016)

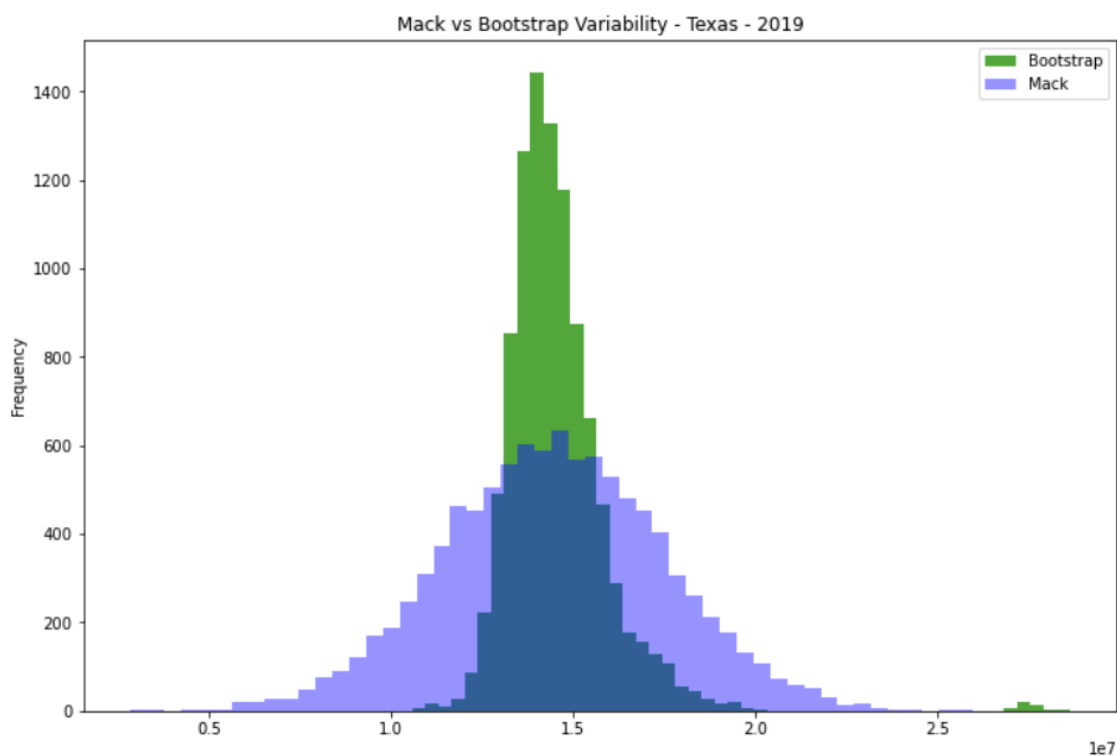


PSAP totale de distribution supposée normale (Garantie IP - Mack Chain Ladder – triangle au half-year 2016)

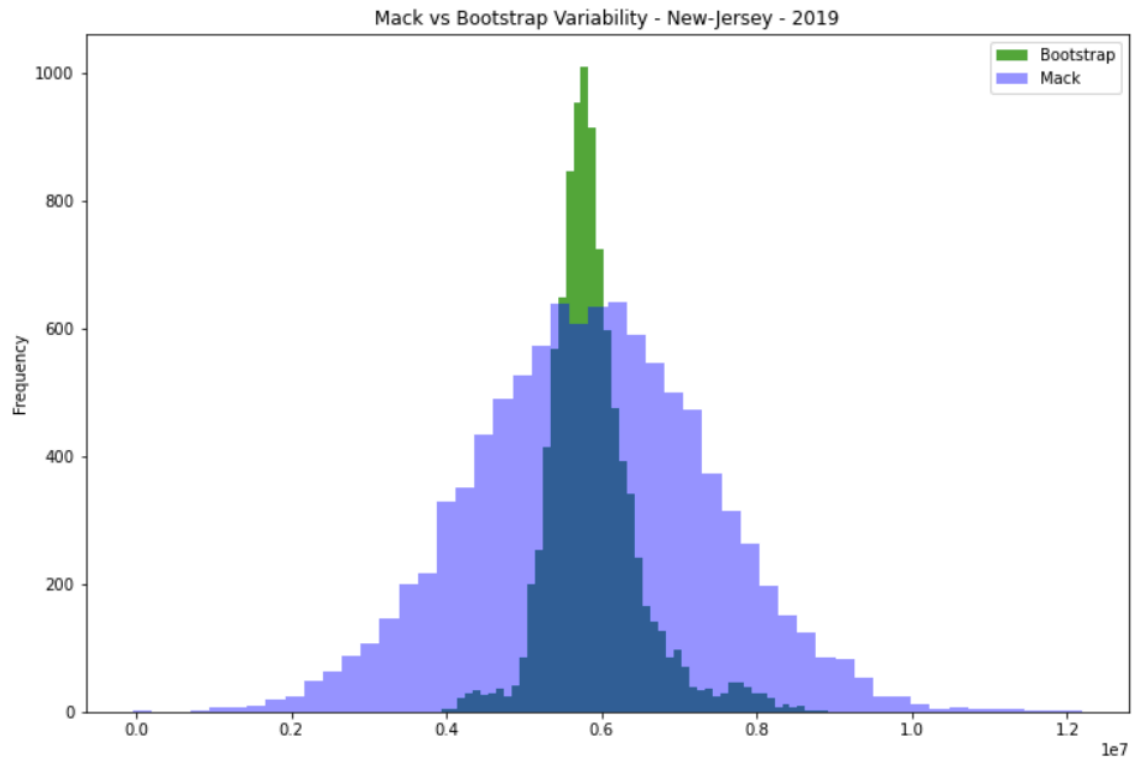
Annexe 2.2.3 – 1. Bootstrap – Application - OSCAR



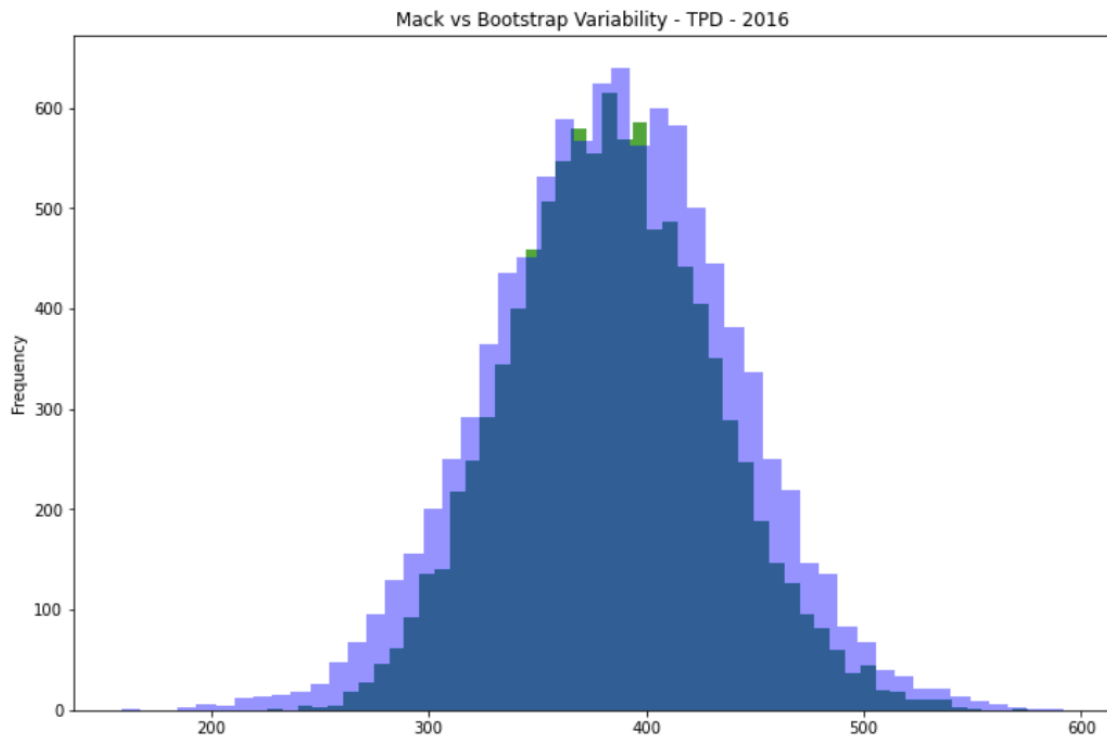
Résultats Bootstrap mis au regard du Mack Chain Ladder (état de Californie – fin 2019)



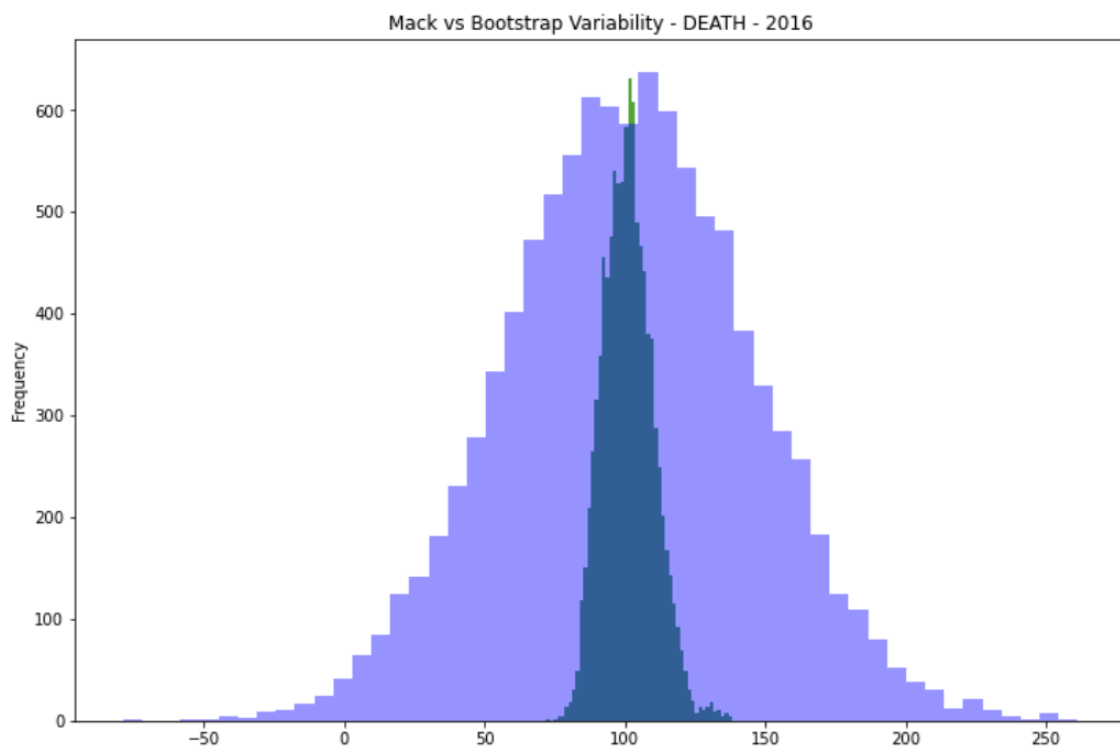
Résultats Bootstrap mis au regard du Mack Chain Ladder (état du Texas – fin 2019)



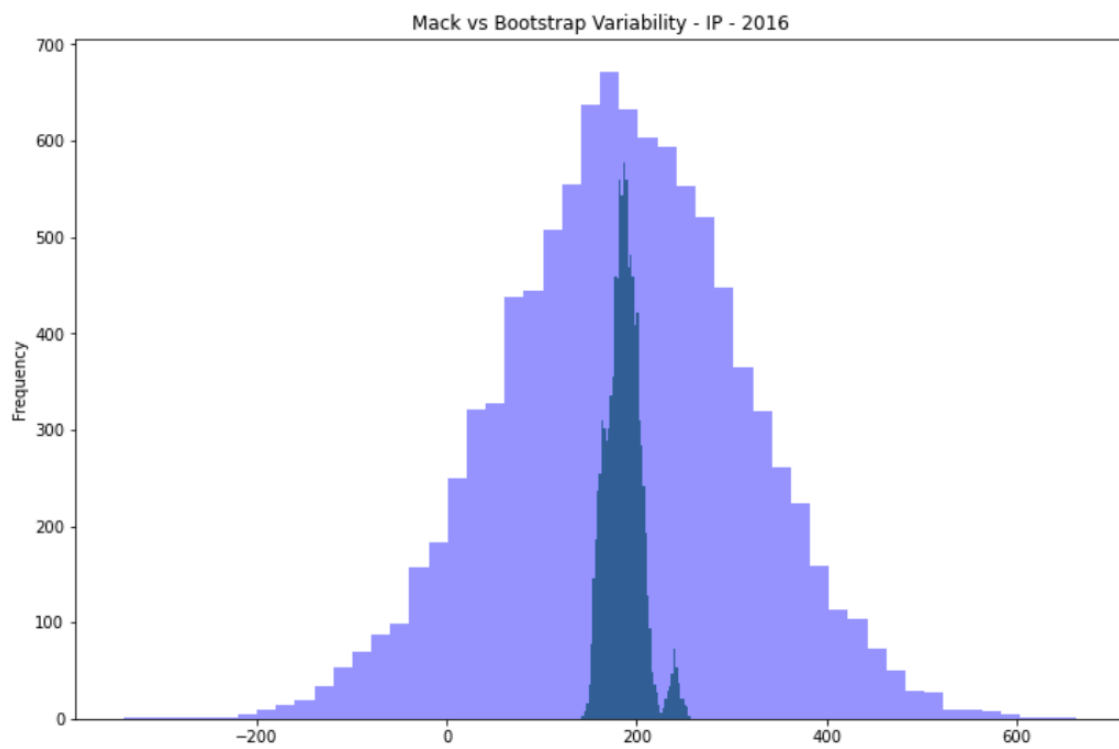
Résultats Bootstrap mis au regard du Mack Chain Ladder (état du New-Jersey – fin 2019)

Annexe 2.2.3 – 2. Bootstrap – Application - AUSTRALIE

Résultats Bootstrap mis au regard du Mack Chain Ladder (Garantie TPD – vision half-year 2016)

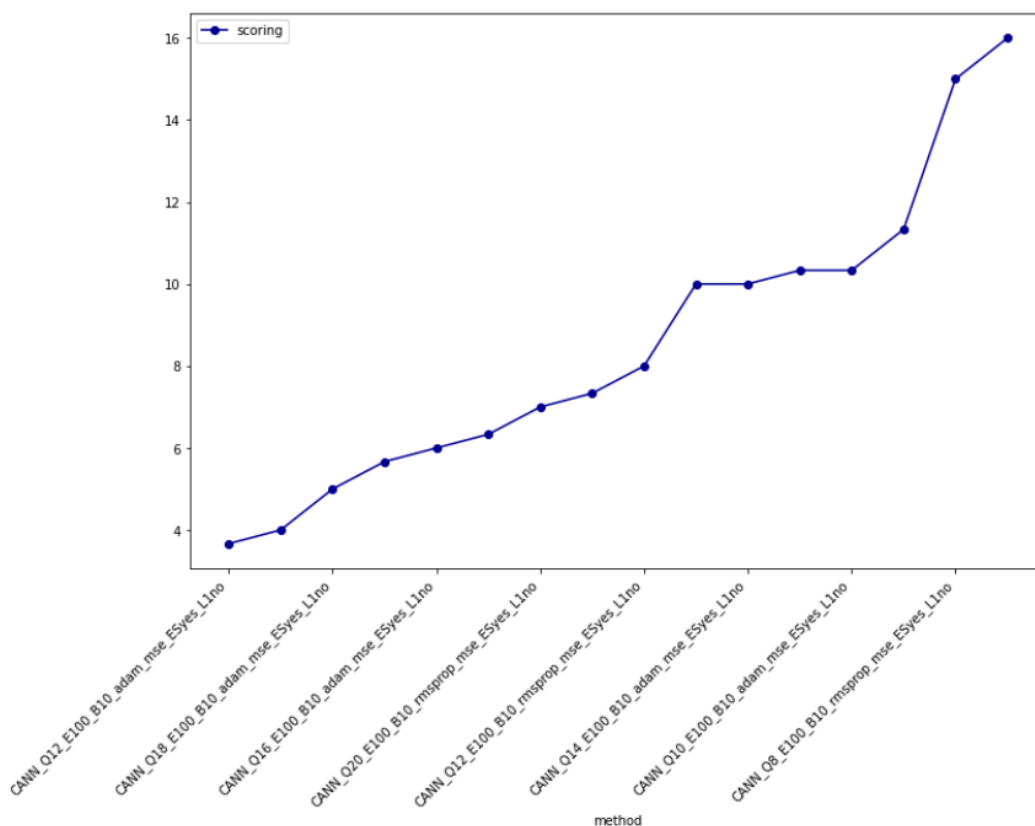


Résultats Bootstrap mis au regard du Mack Chain Ladder (Garantie Death – vision half-year 2016)

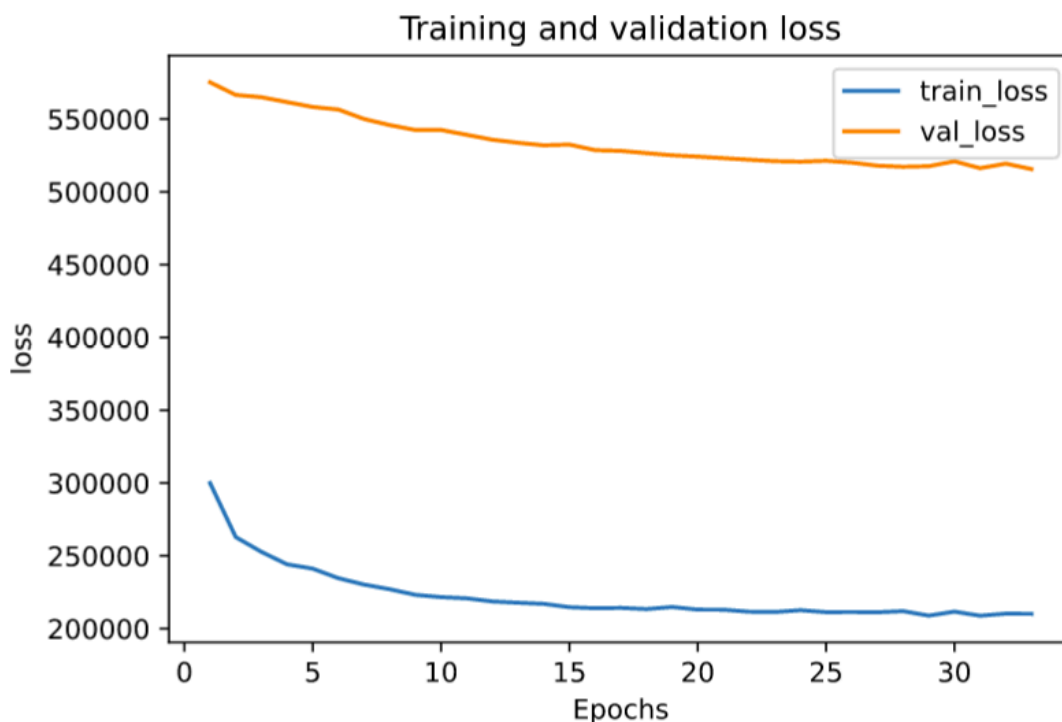


Résultats Bootstrap mis au regard du Mack Chain Ladder (Garantie IP – vision half-year 2016)

Annexe 2.3.4 – 1. CANN – Optimisation - OSCAR

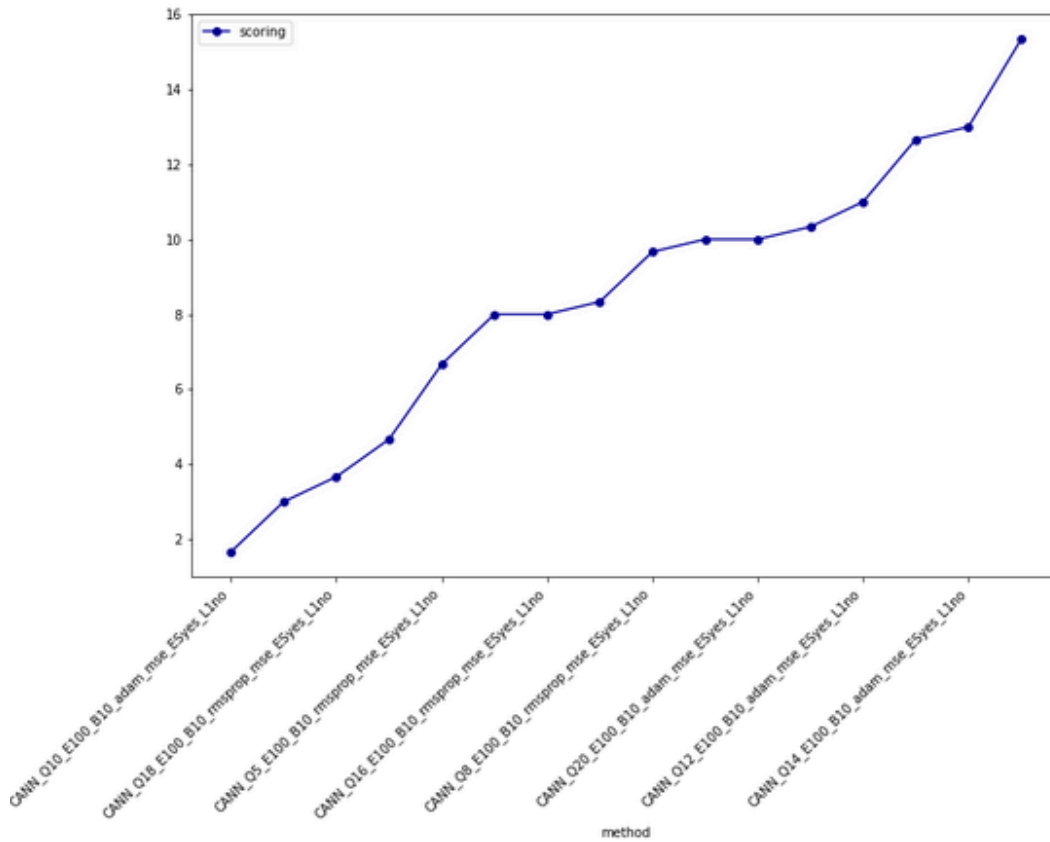


Fonction scoring pour le choix du modèle - Oscar

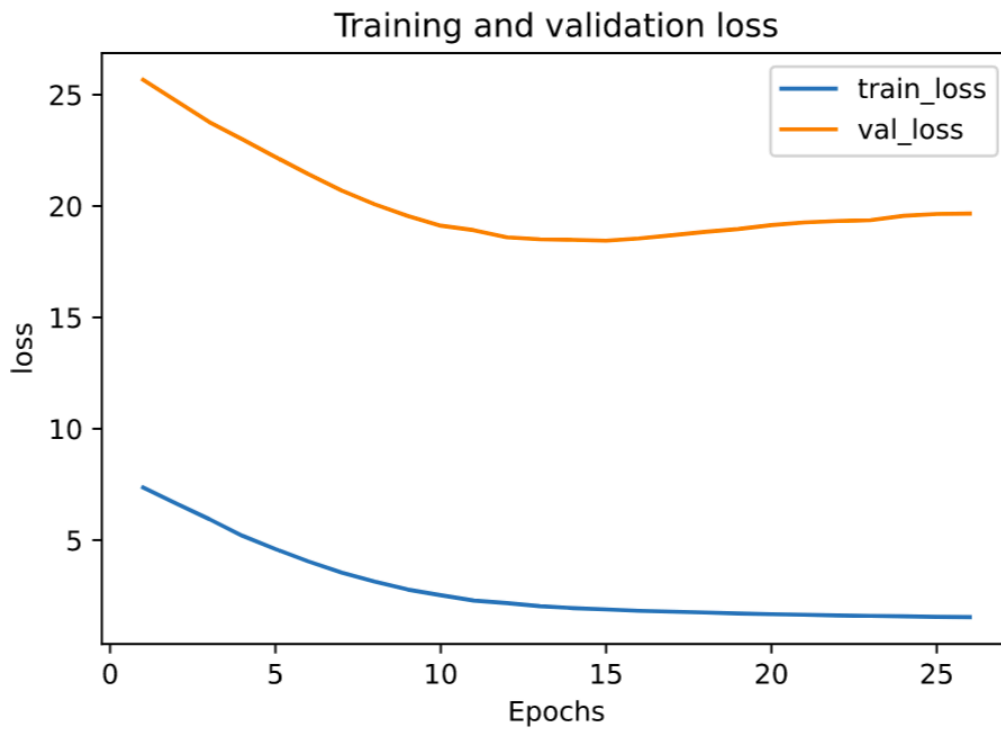


Training et validation loss pour le modèle choisi - Oscar

Annexe 2.3.4 – 2. CANN – Optimisation - AUSTRALIE



Fonction scoring pour le choix du modèle - Australie



Training et validation loss pour le modèle choisi - Australie

Annexe 3.1 – 1. Tableau des cas d'usage - ACPR

| Cas d'usage | | | Critères d'explicabilité | | | Niveau d'explication requis |
|----------------------|----------------------|----------------------------------|---|--|--|-----------------------------|
| Domaine | Processus métier | Fonctionnalité de l'IA | Audience de l'explication | Contexte | Risque associé | |
| Contrats d'assurance | Gestion de contrat | Propositions d'indemnisation | Client | Processus d'indemnisation | Risque opérationnel (insatisfaction du client) | 1 |
| | | | Contrôleur interne | Vérification au quotidien du bon fonctionnement du processus | - Risque opérationnel - Risque de conformité (respect du contrat) - Risque financier | 2 |
| | | | Auditeur | Évaluation de l'algorithme | - Risque opérationnel - Risque de conformité (respect du contrat) - Risque financier | 3 |
| | Proposition de vente | Pré-remplissage de devis | Client | Demande de devis en ligne | Risque de conformité (mauvaise information du client, manquement au devoir de conseil, présence de biais discriminatoire...) | 2 |
| | | | Contrôleur ou auditeur interne | Évaluation de la conformité | Risque de conformité (mauvaise information du client, manquement au devoir de conseil, présence de biais discriminatoire...) | 3 |
| Modèles internes | Conception du modèle | Calcul des ratios de solvabilité | Équipe de validation | Validation des modèles, et de la politique de changement de modèle | - Risque de modèle (de solvabilité) - Risque de conformité | 4 |
| | | | Organes d'administration, de gestion ou de surveillance | Approbation | - Risque de modèle (de solvabilité) - Risque de conformité | 2 |
| Sécurité financière | Gel des avoirs | Remontée d'alertes | Agent en niveau 2 | Analyse des alertes | Néant (si le comportement de l'analyste n'est pas modifié par l'existence de l'algorithme) | 1 |
| | | | Contrôleur de l'algorithme | Contrôle permanent | - Risque opérationnel (faux positifs et faux négatifs) - Risque de conformité (obligation de résultat) | 2 |
| | | | Auditeur | Contrôle périodique | - Risque opérationnel (faux positifs et faux négatifs) - Risque de conformité (obligation de résultat) | 3 |

Tableau des cas d'usages pour l'explicabilité des modèles IA (ACPR)

TABLE DES FIGURES

| | |
|---|----|
| Figure 1 - Les différentes entités chez AXA France (AXA France, One AXA) | 12 |
| Figure 2 - Pentagone des business ALHIS | 15 |
| Figure 3 - Visualisation des portefeuilles étudiés | 16 |
| Figure 4 - Schéma du Bilan sous Solvabilité II | 18 |
| Figure 5 - Visualisation du scope du mémoire, du scope AXA, et du scope hors AXA | 19 |
| Figure 6 - Contrats Oscar étudiés dans ce mémoire | 20 |
| Figure 7 - Illustration de l'ajustement des risques entre différents assureurs américains | 21 |
| Figure 8 - Exemple du maximum out of pocket observé | 22 |
| Figure 9 - Illustration des « Official Medical Packages » | 23 |
| Figure 10 - Nombre d'assurés Oscar sur les 9 premiers mois de 2019, tous états confondus .. | 24 |
| Figure 11 - Distribution par âge et par sexe pour l'année 2019 | 24 |
| Figure 12 - Répartition des contrats par métal pour l'année 2019 | 25 |
| Figure 13 - Contrats Australiens étudiés dans ce mémoire | 26 |
| Figure 14 - Pyramide des pensions de retraite en Australie | 27 |
| Figure 15 - Situation des fonds de pensions Australiens en 2020 | 27 |
| Figure 16 - Lien entre AXA France et la population couverte | 29 |
| Figure 17 - Détail des assurés fonds de pension par âge et par genre en Australie | 29 |
| Figure 18 - Avantages et inconvénient de l'outil existant | 34 |
| Figure 19 - Triangle incrémental théorique | 36 |
| Figure 20 - Triangle cumulé théorique | 37 |
| Figure 21 - Triangle cumulé théorique, Chain Ladder classique appliqué | 37 |
| Figure 22 - Triangle incrémental théorique | 38 |
| Figure 23 - Triangle incrémental corrigé théorique | 38 |
| Figure 24 - Triangle incrémental corrigé théorique, méthode additive appliquée | 39 |
| Figure 25 - Données disponibles pour l'étude du provisionnement PSAP sur Oscar | 41 |
| Figure 26 - Nombre d'observations de la variable «state» dans les données - Oscar | 42 |
| Figure 27 - Données disponibles pour l'étude du provisionnement PSAP sur l'Australie | 43 |
| Figure 28 - Nombre d'observations et montants par «GCT» dans les données - Australie | 44 |
| Figure 29 - Nombre d'observations par «Ceding» dans les données - Australie | 44 |
| Figure 30 - Cadences réelles des paiements pour les survenances de janvier des années 2017, 2018, 2019 sur Oscar | 49 |
| Figure 31 - Triangle cumulé théorique | 50 |
| Figure 32 - Calcul de fj pour le sliding Chain Ladder | 51 |
| Figure 33 - Dimensionnement des fenêtres sur un triangle ayant des cadences hétérogènes .. | 51 |
| Figure 34 - Cadences de paiement pour les survenances de Janvier 2017, 2018, 2019 et 2020 | 52 |
| Figure 35 - Cadences de paiement pour les survenances de Janvier 2017, 2018, 2019 et 2020 (zoom) | 52 |
| Figure 36 - Calcul de $fj + 2$ pour le sliding Chain Ladder | 53 |

| | |
|---|----|
| Figure 37 - Ensemble des points $C_{i,j}; C_{i,j+1}$ (Oscar - tout états confondus – triangle fin 2019)..... | 55 |
| Figure 38 - Ensemble des points $C_{i,j}; C_{i,j+1}$ (Oscar – état de New-York – triangle fin 2019)..... | 56 |
| Figure 39 - Comparaison de l'ensemble des points $C_{i,j}; C_{i,j+1}$ (Oscar – triangle fin 2019) | 57 |
| Figure 40 - Tableau des résultats bruts sur Oscar (12-months sliding Chain Ladder)..... | 58 |
| Figure 41 - Sinistres ultimes 2018 et 2019 - 12-months sliding Chain Ladder (Oscar – état de New-York – triangle fin 2019) | 58 |
| Figure 42 - Tableau des résultats bruts sur l'Australie (Chain Ladder)..... | 59 |
| Figure 43 - Ensemble des résidus en fonction de i (Oscar – état de New-York – triangle fin 2019)..... | 62 |
| Figure 44 - Tableau des résultats bruts sur Oscar (12-months sliding Mack Chain Ladder) . | 63 |
| Figure 45 – Résultats par mois de survenance (état de New-York - Mack Chain Ladder) ... | 63 |
| Figure 46 – PSAP totale de distribution supposée normale (état de New-York - Mack Chain Ladder)..... | 64 |
| Figure 47 - Tableau des résultats bruts sur l'Australie (Mack Chain Ladder)..... | 65 |
| Figure 48 – Triangle observé en garantie TPD | 65 |
| Figure 49 – Calcul des f_j via la méthode du Chain Ladder Classique pour le Bootstrap | 66 |
| Figure 50 – Calcul des $C_{i,j}$ via la méthode du Chain Ladder Classique pour le Bootstrap ... | 67 |
| Figure 51 – Calcul des $X_{i,j}$ via la méthode du Chain Ladder Classique pour le Bootstrap ... | 67 |
| Figure 52 - Tableau des résultats bruts sur Oscar (Bootstrap Chain Ladder) | 68 |
| Figure 53 – Résultats Bootstrap mis au regard du Mack Chain Ladder (état de New-York – 2019)..... | 69 |
| Figure 54 - Tableau des résultats bruts sur l'Australie (Bootstrap Chain Ladder)..... | 70 |
| Figure 55 – Réseau de neurone simple couche appliqué sur Oscar | 76 |
| Figure 56 – Fonction scoring pour les 4 meilleurs modèles..... | 79 |
| Figure 57 - Tableau des résultats bruts sur Oscar (CANN)..... | 80 |
| Figure 58 - Tableau des résultats bruts sur l'Australie (CANN) | 81 |
| Figure 59 – Comparaison des différents modèles (back-testing sur Oscar)– Vision Final Year 2019 | 83 |
| Figure 60 – Comparaison des différents modèles (prévisions sur Oscar) – Vision Half Year 2020 | 84 |
| Figure 61 – Comparaison des différents modèles (back-testing sur l'Australie) – Vision Half Year 2016 | 86 |
| Figure 62 – Comparaison des différents modèles (prévisions sur l'Australie) – Vision Half Year 2020 | 87 |