Modèle de provisionnement des sinistres graves et son allocation économique aux différentes succursales d'AXA Corporate Solutions
(Note de synthèse)

Duc Hien VU

11 octobre 2015

Note de Synthèse

Problématique et Sujet

AXA Corporate Solutions (AXA CS) est une entité du Groupe AXA qui assure les grands risques internationaux IARD. Face à la présence des sinistres graves à fréquence faible et à coût élevé sur la majorité des branches d'activité, AXA CS a besoin d'un modèle pour les provisionner et d'une allocation économique entre les différentes succursales. Dans la majorité des méthodes de provisionnement, les sinistres graves sont écartés du processus pour assurer la stabilité du triangle de liquidation. Un modèle pour les provisionner nécessite donc une étude spécifique qui fait intervenir la théorie des valeurs extrêmes et des techniques moins utilisées dans la pratique. D'ailleurs, pour assurer la stabilité du triangle de provisionnement des sinistres attritionnels, AXA CS considère un sinistre comme étant grave dès lors que sa charge touche le seuil grave, peu importe son positionnement à l'ultime par rapport à ce dernier. Par conséquent, nous devons envisager dans le modèle différents scénarios d'évolution des sinistres. La provision sera calculée par branche d'activité mais au niveau global d'AXA CS. Une étape supplémentaire sera de répartir les réserves entre les différentes succursales d'AXA CS.

Démarches

Revue littéraire

Nous avons dans un premier temps revu les méthodes de provisionnement classiques qui peuvent être regroupées dans deux classes : les méthodes agrégées et les méthodes ligne à ligne.

Les méthodes agrégées consistent toutes à développer le triangle de provisionnement qui agrège les données par année d'origine (année de survenance ou année de déclaration par exemple), avec ou sans présence d'un vecteur d'exposition comme information a priori. Nous avons analysé et comparé ces méthodes pour bien identifier leurs avantages et défauts. Les deux visions extrêmes sont Chain-ladder qui ne repose que sur le triangle pour estimer la charge ultime et la méthode des sinistres espérés qui prévoit l'ultime directement par un avis expert (produit du montant de primes avec le ratio S/P par exemple). Toutes les autres méthodes tendent à crédibiliser les deux visions extrêmes en donnant différents poids aux estimateurs de l'ultime selon deux approches extrêmes.

Les méthodes ligne à ligne cherchent à développer des sinistres non-clos individuellement pour pouvoir utiliser le maximum d'information disponible. Elles proposent différentes approches, tant paramétriques que non-paramétriques, en travaillant sur : soit des montants cumulés, soit des montants incrémentaux et faisant intervenir divers domaines tels que Processus poissonnien, Chaîne de Markov, Apprentissage Statistique, Econométrie, Séries Temporelles et Modélisation stochastique. Elles sont cependant non adaptées au contexte d'AXA CS où nous disposons de peu de données et peu de variables explicatives (comparé aux réassureurs), avec des contraintes très particulières.

L'ensemble des méthodes présentées dans cette revue littéraire a constitué une base de réflexion avec beaucoup d'idées intéressantes qui nous ont servi pour la suite de cette étude.

Modèles de référence

Nous sommes ensuite retournés vers les travaux de recherche menés chez AXA CS concernant les sinistres graves. Ils ont été faits avec plus ou moins les mêmes contraintes requises pour notre modèle de provisionnement et par conséquent sont a priori plus adaptés que les méthodes trouvées dans la littérature et la pratique. Nous disposons d'un modèle de provisionnement des graves en utilisant le ratio Payé/Charge (*Paid/Incurred ou P/I*), appelé modèle 1, et d'un modèle pour modéliser les sinistres graves futurs, appelé modèle 2. Notre but n'étant pas la modélisation, nous avons tout d'abord ajusté le modèle 2 pour mieux répondre au problème de provionnement.

Nous comparons ensuite le modèle 1 et le modèle 2 ajusté. Les deux séparent les réserves pour les sinistres graves connus étant toujours en cours de développement (notés *IBNeR* pour *Incurred But Not enough Reported*) et celles pour les sinistres graves tardifs non encore connus (notés *IBNyR* pour *Incurred But Not yet Reported*). La projection d'*IBNeR* se fait par des méthodes ligne à ligne, étant donné une liste détaillée des sinistres graves à développer. La projection d'*IBNyR*, intervenant après la projection d'*IBNeR*, utilise une approche fréquence-coût. Le nombre d'*IBNyR* (nombre de sinistres graves tardifs) se déduit de la projection du triangle de nombres par des méthodes agrégées. La sévérité est calibrée sur le vecteur de charges ultimes obtenu à travers la projection d'*IBNeR*.

Modèle 1

Dans le modèle 1, nous estimons directement la charge ultime de chacun des sinistres connus en lui appliquant un facteur de développement à l'ultime. L'hypothèse sous-jacente est que ce dernier est proportionnel au ratio *P/I* qui mesure l'état d'avancement dans le règlement d'un sinistre. Nous en déduisons le montant d'*IBNeR*. Pour l'estimation d'*IBNyR*, nous envisageons deux scénarios : la charge ultime d'un sinistre est supérieure ou égale au seuil grave ou le cas contraire. Pour chacun des scénarios, nous projetons un triangle de nombre correspondant en utilisant la méthode *Additive Loss Reserving (ALR)*. Cette méthode, s'appliquant au triangle de nombres incrémentaux, estime pour chacune des années de développement un ratio "Nombre incrémental de graves par unité d'exposition". La sévérité dans le cas Ultime au-delà du seuil grave est estimée par une loi de Pareto, celle dans le cas contraire est estimée par une moyenne empirique des sinistres historiques.

Modèle 2 et Modèle 2 ajusté

Le modèle 2 tient compte de l'effet d'*IBNeR* dans la sévérité et de l'effet d'*IBNyR* dans la fréquence. Pour développer les sinistres connus et non-clos, nous appliquons une méthode dite Chain-ladder ligne à ligne qui est très similaire au Chain-ladder classique. Nous estimons des *link-ratios* communs pour tous les sinistres qui sont en fait des moyennes, pondérées par les montants de charges, des facteurs de passages individuels au-delà d'un seuil d'*IBNeR* (qui est choisi en-dessous du seuil grave pour tenir compte d'effet d'*IBNeR* tout en assurant l'homogénéité des facteurs individuels). Nous procédons ensuite comme dans Chain-ladder mais cette fois-ci sinistre par sinistre pour estimer la charge ultime. Nous obtenons donc un vecteur de charge ultime par sinistre. Nous en déduisons la sévérité des sinistres graves qui nous sert dans la projection d'*IBNyR*. L'estimation des nombres d'*IBNyR* se fait par la méthode Schnieper qui estime, année par année, le nombre de dégradations (nombre de sinistres qui dépassent le seuil grave) et le nombre d'améliorations (nombre de sinistres graves qui retombent en dessous du seuil). Le nombre de dégradations est supposé proportionnel à l'exposition (la prime) et le nombre d'améliorations est supposé proportionnel au nombre de graves cumulé.

Le modèle 2 ajusté est la version adaptée du modèle 2 à la vision des sinistres graves chez AXA CS. Comme nous provisionnons tous les sinistres qui touchent le seuil grave à n'importe quel moment, leur charge ultime ne sera pas forcément au-dessus. La base de sinistres utilisée pour la projection d'*IBNeR* devrait être étendue et le seuil d'*IBNeR* deviendrait obsolète. La méthodologie de Chain-ladder ligne à ligne reste cependant la même. Concernant la projection d'*IBNyR*, la méthode Schnieper n'est plus appropriée puisque nous voulons provisionner chez AXA CS également les améliorations. Au final, nous sommes obligés de revenir vers la méthode *ALR* appliquée à deux différents triangles de nombres suivant deux situations de la charge ultime. Par conséquent, le modèle 2 ajusté ne diffère du modèle 1 que dans l'estimation d'*IBNeR*.

Idées d'amélioration

Nous avons ensuite vérifié les hypothèses des deux modèles et les avons testées sur des données réelles. Les deux modèles souffrent de défauts. Nous avons donc cherché des idées d'amélioration, à la fois sur la projection d'*IBNeR* et la projection d'*IBNyR*.

Projection d'*IBNeR*

(i) Chain-ladder ligne à ligne et choix de pondération

Partons du modèle 2 ajusté. Nous pensons que dans un développement ligne à ligne, il vaut mieux appliquer des *link-ratios* spécifiques à chacun des sinistres, suivant ses caractéristiques. D'ailleurs, le fait de pondérer les facteurs individuels par le montant de charge individuelle pour estimer les *link-ratios* n'est pas approprié aux sinistres graves dont le développement est hétérogène. Nous proposons donc une nouvelle version de Chain-ladder ligne à ligne dans laquelle les facteurs de développement individuels de référence sont pondérés en fonction de la proximité (l'inverse de la distance) entre la charge du sinistre à développer et celle des sinistres de référence. Notre intuition est que les sinistres avec montants de charges plus ou moins proches tendent à se développer d'une façon similaire. Pour être prudent, seuls les facteurs de développement individuels après le premier dépassement du seuil grave sont pris comme référence.

Après avoir testé différentes mesures de distance, nous avons retenu la distance L^1 normalisée par la charge du sinistre à développer. Nous introduisons également une tolérance pour ne pas surpondérer les sinistres trop proches du sinistres à développer. Enfin, un paramètre β est calibré sur l'historique des sinistres pour quantifier l'impact de la proximité sur le développement des sinistres. Si ce paramètre est 0, la proximité entre sinistres n'a aucun impact sur leur développement et par conséquent les facteurs de développement individuels de référence seront équipondérés. Plus ce paramètre est grand, plus les sinistres "proches" se développent de manière similaire. Nous n'avons donc pas besoin de vérifier notre intuition ; la méthode se justifie elle-même à travers le calibrage du paramètre β.

(ii) Munich Chain-ladder ligne à ligne

Après avoir modifié la méthode Chain-ladder ligne à ligne du modèle 2 ajusté, nous avons voulu la combiner avec la méthode du ratio P/I du modèle 1. Aucune méthode ligne à ligne dans la littérature ne répond à notre besoin. Par contre, dans la classe des méthodes agrégées, il existe Munich Chain-ladder qui vise à corriger Chain-ladder classique en utilisant à la fois les Payés P et les Charges I. Nous avons développé une version ligne à ligne de Munich Chain-ladder pour notre projection d'*IBNeR*. L'idée est de développer les payés et les charges simultanément et assurer que ces deux quantités ne s'éloignent pas trop au cours des années de développement, afin que le ratio P/I converge vers 1 à l'ultime. Pour une année de développement donnée, nous disposons de différents ratios P/I pour les sinistres. Si un sinistre a un ratio P/I plus élevé que la moyenne sur tous les sinistres, sa charge I se développera plus vite que prédit Chain-ladder pour rattraper l'écart avec son payé P et/ou inversement son payé P se développera moins vite que prédit Chain-ladder pour réduire l'écart avec sa charge I.

Le test sur différentes branches d'activités a montré que la méthode Munich Chain-ladder ligne à ligne n'apportait une valeur ajoutée que sur la branche Responsabilité Civile où nous avons constaté à la fois une corrélation linéaire entre résidus(I/P) - résidus(P) et celle entre résidus(P/I) - résidus(I). Sur les autres branches, nous n'avons observé que la première. Il paraît que généralement, les payés s'ajustent pour rattraper les charges mais pas inversement. Comme nous avons choisi de développer les charges par Chain-ladder ligne à ligne (en considérant qu'elles sont plus pertinentes que les payés dans notre cas), nous ne gagnons rien en développant parallèlement les payés sachant que les charges n'apprennent pas des payés.

Nous avons décidé de conserver notre version ligne à ligne de Chain-ladder sur les charges et ne pas utiliser Munich Chain-ladder, d'une part parce que cette dernière n'était intéressante que sur la branche Responsabilité Civile, d'autre part parce que notre modèle mathématique pour Munich Chain-ladder ligne à ligne n'est pas rigoureusement développé (les facteurs de développement individuels sont équipondérés plutôt que pondérés en fonction de la proximité des sinistres).

Projection d'*IBNyR*

Les deux modèles 1 et 2 ajusté implémentent la méthode *ALR* pour projeter le triangle de nombres de sinistres graves. Nous avons testé d'autres méthodes agrégées, à savoir Chain-ladder, Bornhuetter-Ferguson mixte (avec cadence Chain-ladder et a priori *ALR*), Cape-Cod, Benktander-Hovinen. Il n'y a pas de "meilleure" méthode dans l'absolu, sachant que nous utilisons à la fois un triangle de nombre qui est instable et un vecteur de primes comme exposition qui n'est pas forcément corrélé à la fréquence des sinistres graves. Nous avons retenu Bornhuetter-Ferguson mixte qui est un bon compromis entre deux visions extrêmes : Chain-ladder et nombres de sinistres espérés. Cette méthode est aussi facile à comprendre et à implémenter.

Modèle retenu

Le modèle final comporte trois modules principaux : la sélection du seuil grave, la projection d'*IBNeR* et la projection d'*IBNyR*.

Sélection du seuil grave

La sélection du seuil grave est faite sous l'hypothèse que des dépassements de seuil suivent une loi Pareto Généralisée et donc nous pouvons utiliser les méthodes : Graphique de Hill, Fonction de dépassement moyen, Estimateur de Cheng&Peng et Bootstrap de Danielsson. Nous cherchons à travers ces méthodes un bon compromis entre un seuil suffisamment élevé pour bien approcher la loi Pareto généralisée et un nombre de points suffisamment grand pour bien estimer les paramètres de cette distribution.

La base de données utilisée pour les deux autres modules (projection d'*IBNeR* et projection d'*IBNyR*) regroupe tous les sinistres dont la charge a dépassé le seuil grave au moins une fois. La provision *IBNeR* correspond à la différence entre la charge ultime et la charge actuelle des sinistres présents dans cette base de données. La provision *IBNyR* correspond aux sinistres graves non encore déclarés ou déclarés mais n'ayant pas encore touché le seuil grave.

Projection d'*IBNeR* : Méthode Chain-ladder ligne à ligne

Notons $\mathcal{K}_n(j)$ le sous-ensemble des sinistres dont la charge à l'année de développement j est connue pour l'année courante n. Pour un sinistre k dont la dernière charge connue correspond à l'année de développement j (i.e. survenu à l'année $n-j+1$), nous estimons sa charge ultime par :

$$\widehat{C_{k,n}} = C_{k,j}\,\widehat{F_{k,j}}\,\widehat{F_{k,j+1}}\ldots\,\widehat{F_{k,n-1}}$$

avec :

$$\widehat{F_{k,j}} = \frac{1}{\sum_{k' \in \mathcal{K}_n(j+1)} w^j_{k,k'}} \sum_{k' \in \mathcal{K}_n(j+1)} w^j_{k,k'} F_{k',j}$$
$$\widehat{\widehat{F_{k,s}}} = \frac{1}{\sum_{k' \in \mathcal{K}_n(s+1)} \widehat{w^s_{k,k'}}} \sum_{k' \in \mathcal{K}_n(s+1)} \widehat{w^s_{k,k'}} F_{k',s} \quad \forall\, s = j+1, \dots, n-1$$

où

— $F_{k',j} = \dfrac{C_{k',j+1}}{C_{k',j}}$ les facteurs de développement individuels ;
— $w^j_{k,k'} = \left(max\left\{\varepsilon, \left|\dfrac{C_{k,j} - C_{k',j}}{C_{k,j}}\right|\right\}\right)^{-\beta}$ mesurant la proximité entre k et k' à l'année de développement j (dans le cas où $C_{k,j}$ est connue) ;
— $\widehat{w^j_{k,k'}} = \left(max\left\{\varepsilon, \left|\dfrac{\widehat{C_{k,j}} - C_{k',j}}{\widehat{C_{k,j}}}\right|\right\}\right)^{-\beta}$ mesurant la proximité entre k et k' à l'année de développement j (dans le cas où $C_{k,j}$ est non-connue et doit être estimée). Comme la mesure de proximité w dépend de la charge du sinistre à développer, elle n'est connue que pour l'année courante. Pour les années suivantes, nous devons estimer sa charge pour calculer sa proximité avec les sinistres de référence ;
— $\varepsilon > 0$ un paramètre, appelé tolérance, introduit pour but de ne pas sur-pondérer les sinistres trop "proches" du sinistre à développer ;
— $\beta \geq 0$ le paramètre mesurant l'impact de la proximité au niveau de la charge entre deux sinistres sur leur développement.

Les facteurs de développement individuels sont calculés comme dans Chain-ladder classique. Les poids $w^j_{k,k'}$ font la particularité de la méthode. Plus la charge d'un sinistre de référence $C_{k',j}$ est proche de celle du sinistre à développer $C_{k,j}$, plus on donne un poids important à $F_{k',j}$. La présence de la tolérance ε dit que les poids sont capés à $\varepsilon^{-\beta}$. Le poids est équi-réparti entre les sinistres dont la charge se trouve dans l'intervalle $[C_{k,j}(1-\varepsilon)\,;\, C_{k,j}(1+\varepsilon)]$. D'une telle manière, un sinistre avec un développement atypique et une charge trop proche du sinistre à développer n'influerait pas gravement notre *link-ratio* estimé. L'intervalle défini ci-dessus est appelé voisinage de $C_{k,j}$. Il est d'autant plus grand que la charge $C_{k,j}$ est élevée. En effet, les sinistres de charge très élevée sont rares, nous voulons étendre son voisinage pour avoir suffisamment de points "voisins".

Le paramètre β est calibré par une technique d'apprentissage statistique dite *leave-one-out cross-validation*. Nous faisons parcourir β dans un intervalle suffisamment grand. Pour chaque valeur de β, nous écartons successivement un sinistre de notre base de données, estimons des $link - ratios$ pour ce sinistre en utilisant les autres sinistres comme sinistres de référence. Nous comparons le vrai développement de ce sinistre au développement prédit par notre algorithme. Au final, le β optimal est celui qui minimise l'erreur moyenne de prédiction. Ce β optimal nous dit si la proximité entre sinistres a un impact important ou faible, voire nul sur leur développement.

Projection d'*IBNyR* : Méthode Bornhuetter-Ferguson avec cadence Chain-ladder et a priori *ALR*

La projection d'*IBNyR* est faite avec une approche fréquence-coût (nous avons besoin de la projection d'*IBNeR* avant ce module pour estimer la sévérité des sinistres graves sur les montants de charge ultime des sinistres présents dans la base). Nous distinguons deux cas : Charge ultime ≥ Seuil grave et Charge ultime < Seuil grave.

Comme dans les méthodes de référence, la sévérité est calibrée par des lois extrêmes dans le premier cas et simplement par une moyenne pondérée dans le deuxième cas. Ce qui diffère des méthodes de référence dans ce module est la projection des triangles de nombres. Dans les deux cas, nous utilisons la même méthode appelée Bornhuetter-Ferguson avec cadence Chain-ladder et a priori *ALR*. Cette méthode nécessite un triangle de nombres incrémentaux des sinistres graves $(X_{i,j})$ (nous en déduisons le triangle de nombres cumulés $C_{i,j} = \sum_{m=1}^{j} X_{i,m}$) [1] et un vecteur d'exposition (E_i). L'hypothèse sous-jacente est que le nombre de sinistres graves est a priori proportionnel à l'exposition.

Pour chacune des années de développement $j = 1, \ldots, n$, nous estimons le ratio "Nombre de sinistres graves survenus par unité d'exposition" : $\widehat{\alpha_j} = \frac{\sum_{i=1}^{n-j+1} X_{i,j}}{\sum_{i=1}^{n-j+1} E_i}$. Nous en déduisons un ratio ultime : $\widehat{\alpha} = \sum_{j=1}^{n} \widehat{\alpha_j}$. En re-multipliant ce dernier au vecteur d'exposition, nous obtenons un vecteur "a priori" du nombre ultime de sinistres graves par année de survenance : $\widehat{\mu_i} = E_i \widehat{\alpha}$. (Ce n'est en fait pas un vrai a priori puisque $\widehat{\alpha}$ dépend du triangle $(X_{i,j})$). Nous appliquons ensuite Bornhuetter-Ferguson pour estimer le nombre ultime de sinistres graves pour chacune des années de survenance et le nombre d'*IBNyR* :

$$\widehat{C_{i,n}}^{BF} = C_{i,n-i+1} + (1 - \widehat{\beta_{n-i+1}})\widehat{\mu_i}$$

où la cadence est estimée par Chain-ladder : $\widehat{\beta_{n-i+1}} = \frac{1}{\prod_{j=n-i+1}^{n-1} \widehat{f_j}^{CL}}$.

Risque à 1 an

Pour quantifier le risque de provisionnement à un an, nous avons utilisé Monte-Carlo non paramétrique (l'équivalence de Bootstrap non-paramétrique) pour simuler N scénarios pour le développement des sinistres dans un an et ensuite avons ré-appliqué notre modèle pour prévoir l'ultime dans chaque scénario. Nous avons obtenu une distribution de la charge ultime et ainsi une distribution du *CDR (Claim Development Result)*. Comme notre modèle est scindé en deux modules, nous avons proposé trois différents calculs de *CDR* : *CDR(IBNeR)*, *CDR(IBNyR)* et *CDR(IBNR)* pour quantifier la volatilité à 1 an de chacun des modules et la volatilité à 1 an globale.

1. Attention : nous utilisons la notation C mais ça correspond au nombre plutôt qu'au montant de charge

Allocation économique entre succursales

La dernière étape est d'allouer le montant de réserves pour chacune des branches d'activité entre différentes succursales d'AXA CS. Pour les réserves *IBNeR*, la méthode Chain-ladder ligne à ligne nous permet de connaître le montant d'*IBNeR* d'un sinistre donné et l'affecter à la succursale à laquelle il est rattaché. La répartition des *IBNyR* fait intervenir les profils de risque (engagement, parts, rétention, primes) de chacune des succursales.

Mots clés : Provisionnement, Sinistres graves, TVE, *IBNeR*, *IBNyR*, méthode ligne à ligne, proximité de charges, Chain-ladder, Bornhuetter-Ferguson, *ALR*, allocation économique

Summary

Context and Subject

AXA Corporate Solutions (AXA CS) is an entity of the AXA Group which insures International Major Risks in Property and Casualty. Given the presence of large claims with "low frequency and high severity" on most Lines of Business (LoB), AXA CS needs a Major Losses Reserving model and an economic allocation between its different branches. In the majority of reserving methods, large claims are excluded from the process to ensure the stability of the run-off triangle. A model for major losses reserving therefore requires a specific study that involves extreme value theory and techniques lesser used in practice. Moreover, to ensure the stability of attritional claims reserving triangle, AXA CS considers a claim as major loss (ML) as soon as its Incurred is above the ML threshold, regardless of its ultimate position. Therefore, we must consider in our model different scenarios of claims evolution. The reserves amount will be calculated by LoB but at AXA CS global level. The final step will be the reserves allocation between the different branches of AXA CS.

Research approaches

Litterature review

We reviewed the traditional reserving methods that can be grouped into two classes : the aggregated methods and the individual methods.

The aggregated methods consist in developing the run-off triangle of aggregated data by year of origin (year of occurrence or reporting year, for example), with or without the presence of an exposure vector as prior information. We analyzed and compared these methods to properly identify their advantages and shortcomings. The two extreme visions are Chain-ladder that only relies on the triangle to estimate the ultimate and the expected claims method that directly provides the ultimate thanks to an expert opinion (the product of the premium and the loss ratio for example). All other methods tend to credibilize the two extreme visions by giving different weights to the ultimate estimators of two extreme approaches.

The individual methods tend to develop opened claims individually in order to use the maximum available information. They offer different approaches, both parametric and non-parametric, working on either cumulative amounts or incremental amounts and involving various fields such as Poisson processes, Markov Chain, Machine Learning, Econometrics, Time Series and Stochastic Modeling.

They are however not adapted to the context of AXA CS where we have little data and few explanatory variables (compared to reinsurers), with very specific constraints.

All methods presented in this literature review constitute a great reflection basis with interesting ideas that have helped us for the remaining of our study.

Reference models

We then concentrated our analysis to the researches conducted at AXA CS on large claims. The methods were built with more or less the same constraints required in our reserving model and are consequently more suitable than the methods found in the literature and practice. We have a ML reserving model using the Paid/Incurred (or P/I) ratio, called model 1 and another model that simulate major losses for a future year, called Model 2. Since our goal is not modeling, we adjusted Model 2 to better suit our reserving subject.

We then compare the model 1 and model 2 adjusted. The two separate reserves for the known claims that have not settled yet (noted IBNeR for Incurred But Not enough Reported) and those for large claims not yet reported (noted IBNyR for Incurred But Not yet Reported). The IBNeR projection is done by individual methods, given a detailed list of large claims to be developed. The IBNyR projection, occuring after the IBNeR projection, uses a frequency-severity approach. The number of IBNyR (number of large claims not yet reported) is derived from the projection of the triangle of numbers by aggregated methods. The severity is calibrated on the ultimate obtained through the IBNeR projection.

Model 1

In the model 1, we estimate directly the ultimate for each known claim by applying an ultimate development factor. The underlying assumption is that this factor is proportional to the P/I ratio which reflects the progress in the settlement of a claim. We deduce the IBNeR reserve for each claim. For IBNyR projection, we consider 2 cases : claims with ultimate above ML threshold and otherwise. In each case, we project a corresponding triangle of numbers thanks to Additive Loss Reserving (ALR) method. This method, applied to a triangle of incremental numbers, estimate for each year of occurrence a ratio "expected incremental number of large claims per unit of exposure". The severity in the case where the ultimate is above the ML threshold is fitted by a Pareto distribution, whereas the severity in the other case is simply estimated by an empirical mean of historical claims.

Model 2 and Model 2 adjusted

In the model 2, IBNeR effect is taken into the severity modeling whereas the IBNyR effect is incorporated in the frequency modeling. Known and opened claims are developed by a method called Individual Chain-ladder inspired by the classic Chain-ladder. We estimate commun link-ratios for all claims by averaging (weighted by Incurred amounts) individual development factors for passages above the IBNeR threshold (which is chosen below the ML threshold to take into account the IBNeR effect, while maintaining the homogeneity of retained individual development factors). The procedure is similar to the classic Chain-ladder but this time we project claim per claim individually. Thus, we obtain a vector of ultimate from which the severity of large claims needed in the IBNyR projection is deduced. The IBNyR numbers are estimated by the Schnieper method that estimates for each year of occurrence the number of degradations (the number of claims that exceed the ML threshold) and the number of improvements (the number of claims that fall below the ML threshold). The latter is assumed to be proportional to the previous Incurred whereas the number of degradations is expected to be proportional to the exposure (premium).

The model 2 adjusted is the adapted version of the model 2 to the ML vision of AXA CS. As we provision all claims that have reached or will reach the ML threshold at least once, their ultimate are not necessarily above the threshold. The claims database used in IBNeR projection should be extended and the IBNeR threshold becomes now obsolete. However, the methodology of Individual Chain-ladder is retained. Concerning the IBNyR projection, the Schnieper method is no longer appropriate since we also need to calculate reserves for improvement-claims. Finally, we returned to the ALR method, applied to two different triangles of numbers corresponding to two situations of the ultimate. As a consequence, the model 2 adjusted only differs from the model 1 in the IBNeR projection.

Ideas for improvement

We then checked the assumptions of the two models and tested them on real data. Both models present shortcomings. We therefore sought new ideas for improvement, in both the IBNeR projection and the IBNyR projection.

IBNeR projection

Individual Chain-ladder and choice of weighting

We shall start with the model 2 adjusted. We believe that in an individual projection, we better apply specific link-ratios to each claim, according to its characteristics. Moreover, the weighting of individual factors by the amount of individual Incurred for estimating link-ratios is not appropriate to large claims whose development is heterogeneous. We propose a new version of Individual Chain-ladder in which individual development factors are weighted based on Incurred proximity (inverse of distance) between claim to be developed and reference claims. Our intuition is that claims with close Incurred amounts tend to develop in a similar way. To be carreful, only individual development factors after the first exceedance of the ML threshold are taken as reference factors.

After testing different measures of distance, we retained the L^1 distance, standardized by the Incurred of the claim to be developed. We also introduce a tolerance parameter in order not to overweight claims that are too close to the claim to be developed. Finally, a parameter β is calibrated on claims history to quantify the impact of claims proximity on their development. If this parameter is 0, claims proximity has no impact on their development and therefore individual development factors are equally weighted. The bigger the parameter is, the more similar the development of claims with close Incurred are. Consequently, we do not need to confirm our intuition ; the method justifies itself thanks to the calibration of β parameter.

Individual Munich Chain-ladder

After modifying the Individual Chain-ladder method found in the model 2 adjusted, we want to combine it with the P/I ratio method of the model 1. No method found the litterature review deals with this combination. Nevertheless, in the class of aggregated methods, there is Munich Chain-ladder that tends to correct the classic Chain-ladder by using both Incurred losses (I) and Paid losess (P) in claims development. We thus develop an individual version of Munich Chain-ladder for our IBNeR projection. The idea is to project the Paid losses and the Incurred losses simultaneously and to ensure that these two amounts do not stray too much during different development years. The aim is that the P/I ratio will be close to 1 at the ultimate. Given a development year, we have different P/I ratios corresponding to different claims. If a claim has its P/I ratio higher than the mean of all claims, its Incurred will develop faster than predicted Chain-ladder to reduce the gap with its Paid and/or conversely, its Paid will develop slower than predicted Chain-ladder to wait for its Incurred.

Tests on different Lines of Business (LoB) have proved that Individual Munich Chain-ladder only brought value on Liability LoB where we found a linear correlation between both residues(I/P) - residues(P) and residues(P/I)-residues(I). On other LoB, we only observe the first correlation. It seems that generally, Paid losses adjust to catch Incurred losses but not vice versa. As we have chosen to apply Individual Chain-ladder to Incurred (by considering that Incurred losses are more relevant than Paid losses in our context), we gain nothing by projecting in parallel Paid amounts, knowing that Incurred amounts do not learn from Paid amounts.

We decided to keep our Individual Chain-ladder version applied to Incurred and not to use Individual Chain-ladder, firstly because it was only interesting for Liability LoB, secondly because our mathematical model for Individual Munich Chain-ladder was not rigorously developed (individual development factors are equally weighted rather than weighted according to the claims proximity).

IBNyR projection

Both model 1 and model 2 adjusted use the ALR method to project the triangle of numbers. We haved tested other aggregated methods, namely Chain-ladder, mixed Bornhuetter-Ferguson (with a Chain-ladder pattern and a priori ALR), Cape-Cod and Benktander-Hovinen. There is no "best" method in the absolute sense, knowing that the triangles of numbers are unstable and the vector of exposure is not necessarily correlated to the frequency of large claims. We finally selected mixed Bornhuetter Ferguson which is a good compromise between the two extreme views : Chain-ladder method and expected claims method. This method is also easy to understand and to implement.

Final model

The selected model is composed of three main modules : the ML threshold selection, the IBNeR projection and the IBNyR projection.

ML threshold selection

The ML threshold is selected under the assumption that the excess over the threshold follows a Generalized Pareto Distribution (GPD) so that we can use Hill plot, Mean Excess Function, Cheng&Peng's estimator and Danielsson's bootrap. We find through these methods a good compromise between a high enough threshold to properly approach the Generalized Pareto Distribution and a sufficiently large number of points for calibrating GPD parameters.

The database used for the other two modules (IBNeR projection and IBNyR projection) includes all claims whose Incurred has once exceeded the ML threshold. The IBNeR amount is the difference between the ultimate and the actual Incurred of claims in this database. The IBNyR reserves correspond to large claims that have not yet been reported and claims that have not reached the ML threshold yet but are expected to.

IBNeR projection : Individual Chain-ladder

Let $\mathcal{K}_n(j)$ the subset of claims whose Incurred of development year j is known in current year n. For a claim k whose last known Incurred amount corresponds to the development year j (i.e. occurred in year $n - j + 1$), we estimate its Ultimate by :

$$\widehat{C_{k,n}} = C_{k,j}\,\widehat{F_{k,j}}\,\widehat{\widehat{F_{k,j+1}}}\dots\widehat{\widehat{F_{k,n-1}}}$$

with :

$$\widehat{F_{k,j}} = \frac{1}{\sum_{k'\in\mathcal{K}_n(j+1)} w^j_{k,k'}} \sum_{k'\in\mathcal{K}_n(j+1)} w^j_{k,k'} F_{k',j}$$

$$\widehat{\widehat{F_{k,s}}} = \frac{1}{\sum_{k'\in\mathcal{K}_n(s+1)} \widehat{w^s_{k,k'}}} \sum_{k'\in\mathcal{K}_n(s+1)} \widehat{w^s_{k,k'}} F_{k',s} \quad \forall\, s = j+1,\dots,n-1$$

where :

— $F_{k',j} = \dfrac{C_{k',j+1}}{C_{k',j}}$ are idividual development factors ;

— $w^j_{k,k'} = \left(max\left\{\varepsilon, \left|\dfrac{C_{k,j} - C_{k',j}}{C_{k,j}}\right|\right\}\right)^{-\beta}$ measures the proximity between k and k' in the development year j (in case $C_{k,j}$ is known) ;

— $\widehat{w^j_{k,k'}} = \left(max\left\{\varepsilon, \left|\dfrac{\widehat{C_{k,j}} - C_{k',j}}{\widehat{C_{k,j}}}\right|\right\}\right)^{-\beta}$ measures the proximity between k and k' in the development year j (in case $C_{k,j}$ is unknown and needs to be estimated). As claims proximity w depends on the Incurred of the claim to be developed, it is only known for the current year. For the next years, we have to estimate the Incurred to calculate the proximity with the reference claims ;

— $\varepsilon > 0$ a tolerance parameter introduced to not over-weight claims that are too close to the claim to be developed ;

— $\beta \geq 0$ a parameter measuring claims proximity impact on their development.

Individual development factors are calculated as in the classic Chain-ladder. The weights $w^j_{k,k'}$ make the difference in the method. The closer $C_{k,j}$ and $C_{k',j}$ are, the bigger weight given to $F_{k,j}$ is. The presence of tolerance ε means that weights are capped at $\varepsilon^{-\beta}$. Claims whose Incurred is in the interval $[C_{k,j}(1-\varepsilon)\ C_{k,j}(1+\varepsilon)]$ are equally weighted. In such a way, a claim with atypical development and Incurred too close to the claim to be developed will not seriously affect our estimated link-ratio. The interval defined above is called neighborhood of $C_{k,j}$. The bigger $C_{k,j}$ is, the wider its neigborhood is. Indeed, very large claims are rare, we want to extend their neighborhood to have enough "neighbour" points.

The β parameter is calibrated by a statistical learning technique called leave-one-out cross-validation. We vary β on a sufficiently large interval. For each value of β, we discard successively one claim from our dataset, then estimate $link - ratios$ for this one by using other claims as references. We compare the true development of this claim to the estimation in our algorithm. Finally, the optimal β is the one that minimizes the average error prediction . This optimal β tells us if the claims proximity impacts strongly or weakly or even does not affect their development.

BNyR projection : Bornhuetter-Ferguson method with Chain-ladder pattern and ALR prior

The IBNyR projection is done using a frequency-severity approach (we need the IBNeR projection before this module in order to estimate the large claims severity based on the ultimate of all claims presented in the dataset). We distinguish two cases : Ultimate $\geq$ ML threshold and Ultimate $<$ ML threshold.

As in reference methods, the severity is calibrated by extreme distributions in the first case and simply by a weighted average in the second case. What differs from the reference methods in this module is the projection of the triangles of numbers. In both cases, we use the same method called Bornhuetter-Ferguson with a Chain-ladder pattern and an ALR prior. This method requires a triangle of incremental numbers of large claims $(X_{i,j})$ (we then deduce the triangle of cumulative numbers $C_{ij} = \sum_{m=1}^{j} X_{i,m})$ [2] and an exposure vector (E_i). The underlying assumption is that the number of large claims is a priori proportional to the exposure.

For each development year $j = 1, \ldots, n$, we estimate the "average number of large claims occured per exposure unit" : $\widehat{\alpha_j} = \dfrac{\sum_{i=1}^{n-j+1} X_{i,j}}{\sum_{i=1}^{n-j+1} E_i}$. We deduce an ultimate ratio : $\widehat{\alpha} = \sum_{j=1}^{n} \widehat{\alpha_j}$. By multiplying this ratio to the exposure vector, we obtain an a priori vector of ultimate number of large claims by year of occurrence : $\widehat{\mu_i} = E_i \widehat{\alpha}$. (It is not a true a priori because $\widehat{\alpha}$ depends on the triangle $(X_{i,j})$). Then, we apply Bornhuetter-Ferguson method to estimate the ultimate number and the IBNyR number of larges claims for each year of occurrence :

$$\widehat{C_{i,n}}^{BF} = C_{i,n-i+1} + (1 - \widehat{\beta_{n-i+1}})\,\widehat{\mu_i}$$

2. Note : we use the notation C but it is the number rather than the Incurred amount

with Chain-ladder pattern : $\widehat{\beta_{n-i+1}} = \dfrac{1}{\prod_{j=n-i+1}^{n-1} \widehat{f_j}^{CL}}$.

One-year reserve risk

To quantify the one-year reserve risk, we used non-parametric Monte-Carlo method (an equivalence of Bootstrap non-parametric) to simulate N scenarios of claims development for next year and re-applied our model to predict the ultimate of each scenario. We obtained a CDR (Claim Development Result) distribution. Since our model is divided into two modules, we proposed three different calculations of *CDR* : *CDR(IBNeR)*, *CDR(IBNyR)* and *CDR(IBNR)*, in order to measure the specific one-year volatility of each module and the global one-year volatility.

Economic Allocation between AXA CS branches

The final step is to allocate the reserves amount of each LoB between the different AXA CS branches. For IBNeR reserves, the Indivual Chain-ladder method allows us to know the IBNeR amount of each claim and therefore assign it to the branch to which the claim belongs. The IBNyR allocation requires knowledge of the risk profiles (commitment, shares, retention, premiums) of each branch.

Key words : Reserving, Atypical claims, Major Losses, EVT, IBNeR, IBNyR, individual approach, incurred proximity, Chain-ladder, Bornhuetter-Ferguson, ALR, economic allocation