

Mémoire présenté devant l'ENSAE Paris  
pour l'obtention du diplôme de la filière Actuariat  
et l'admission à l'Institut des Actuaires  
le 15/03/2022

Par : **Cheick Oumar Kouyaté**

Titre : **Méthode de compression d'un portefeuille de  
contrats d'assurance vie dans une approche Full ALM**

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

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

*Membres présents du jury de la filière*

*Entreprise : **Allianz***

*Nom : Nicolas BARADEL*

*Signature :*

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

*Directeur du mémoire en entreprise :*

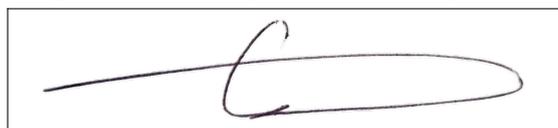
*Nom : Ghassane LAANAITI*

*Signature :*

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

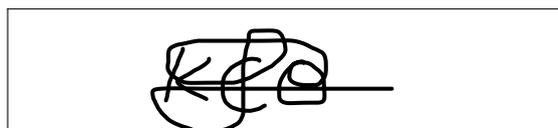
Secrétariat :

Signature du responsable entreprise



Bibliothèque :

Signature du candidat



# Remerciements

En premier lieu, je tiens à exprimer toute ma reconnaissance à toute l'équipe Modélisation vie d'Allianz France pour leur disponibilité ainsi que pour leur implication dans la réussite de ce mémoire.

Je tiens à accorder des remerciements particuliers à Ghassane LAANAITI, actuaire chez Allianz France, mon encadrant pour ce mémoire. Je le remercie pour ses innombrables idées, ses conseils avisés, sa bienveillance ainsi que son investissement tout au long de l'élaboration de ce mémoire.

Ce mémoire n'aurait également pu aboutir sans l'aide de Emmanuel MARTINEZ, actuaire manager chez Allianz France. Je le remercie pour ses conseils, son soutien, ses encouragements ainsi que pour le temps qu'il a consacré à répondre à mes nombreuses questions.

J'adresse également des remerciements à Houda AFAILAL, actuaire au sein de l'équipe Modélisation vie chez Allianz France pour son suivi et l'intérêt porté à mes travaux.

Je remercie Stéphanie PELLETIER, directrice Actuariat et Nicolas BOURE, directeur Actuariat Vie chez Allianz France pour leur accueil au sein de leur service.

J'ai, bien entendu, une pensée pour ma famille et mes proches pour leur soutien sans faille dans toutes mes entreprises.

En dernier lieu, je tiens à remercier le corps professoral de l'ENSAE ainsi que toutes les personnes ayant participé, de près ou de loin, à l'aboutissement de ce mémoire.

# Résumé

La directive Solvabilité 2 a généralisé l'utilisation des modèles ALM (Asset Liability Management) pour l'évaluation des provisions des produits d'assurance vie avec assez peu de contraintes sur la maille de projection. Cela a poussé les compagnies à opter pour des modélisations ALM avec l'approche Flexing qui permet d'avoir rapidement des résultats, mais à une large maille de calcul donc approximative. Mais, avec l'entrée en vigueur de la norme IFRS 17 et sa nouvelle définition de segmentation de portefeuilles, Allianz France a pour ambition de migrer de son modèle en Flexing vers un nouveau modèle basé sur l'approche Full ALM dont l'un des défis majeurs est le temps de calcul nécessaire pour obtenir les résultats en raison du trop grand nombre de polices dans le portefeuille. Une solution a été envisagée et reste conforme au Règlement Délégué de la commission européenne de janvier 2015 qui stipule qu'il est possible de faire des regroupements de polices dès lors que les polices considérées et les risques sous-jacents sont similaires et que l'utilisation de ces regroupements de polices (model points) ne fausse pas les résultats. L'objet de notre étude a résidé donc dans la construction de ces Model Points. En d'autres termes, il s'agit de transformer le portefeuille assurantiel géré par Allianz France en de nouvelles tables de données représentant le plus fidèlement possible ces portefeuilles. Une première approche basée sur de l'optimisation linéaire fut proposée par un cabinet de conseil. Mais celle-ci en plus d'avoir un ratio précision-compression peu satisfaisant, utilise les cashflows de BEL (meilleure estimation en vision économique des engagements futurs de l'assureur) comme variable d'agrégation. Pour pallier les nombreuses limites liées à cette approche, nous avons recouru à une méthodologie basée sur un algorithme assez répandu de classification statistique avec comme nouvelle variable d'agrégation les probabilités de sortie. Il s'agit d'une procédure en deux étapes. La première étape consiste à créer des groupes homogènes de contrats. La seconde étape permet la définition d'un contrat représentatif ou "moyen" pour chacune des classes. Grâce à cette approche, nous sommes parvenu à un taux de compression du portefeuille de 87% avec une erreur relative de moins de 9 euros pour 10 000 contre 20 euros pour 10 000 avec l'approche linéaire.

**Mot-clés :** *Assurance vie, ALM, Flexing, Full ALM, Best Estimate of Liabilities, algorithme de kmeans*

# Abstract

The Solvency 2 directive has generalized the use of ALM (Asset Liability Management) models for the valuation of reserves for life insurance products with relatively few constraints on the projection grid. This has led companies to opt for ALM modeling with the Flexing approach, which allows them to obtain results quickly, but at a large scale of calculation, which is therefore approximate. However, with the entry into force of IFRS 17 and its new definition of portfolio segmentation, Allianz France's ambition is to migrate from its Flexing model to a new model based on the Full ALM approach, one of the major challenges of which is the calculation time required to obtain the results due to the large number of policies in the portfolio. A solution has been envisaged and remains in line with the Delegated Regulation of the European Commission of January 2015, which stipulates that it is possible to group policies as long as the policies considered and the underlying risks are similar and that the use of these policy groupings (model points) does not distort the results. The purpose of our study was therefore to construct these Model Points. In other words, the aim was to transform the insurance portfolio managed by Allianz France into new data tables representing these portfolios as faithfully as possible. A first approach based on linear optimization was proposed by a consulting firm. But, in addition to having a rather poor precision-compression ratio, this approach uses the BEL cash flows (best estimate in economic vision of the insurer's future commitments) as aggregation variable. To overcome the many limitations of this approach, we have used a methodology based on a fairly common statistical classification algorithm with probabilities of exit as the aggregation variable. This is a two-step procedure. The first step consists in creating homogeneous groups of contracts. The second step allows the definition of a representative or "average" contract for each class. With this approach, we achieved a portfolio compression rate of 87% with a relative error of less than 9 euros per 10,000 compared to 20 euros per 10,000 with the linear approach.

**Keywords :** *Life insurance, ALM, Flexing, Full ALM, Best Estimate of Liabilities, kmeans algorithm*

# Sigles et Abréviations

ALIM	Asset and Liabilities Interaction Management
ALM	Asset Liability Management
BEL	Best Estimate of Liabilities
DTW	Dynamic Time Warping
IFRS	International Financial Reporting Standard
LOP	Linear OPTimization
LSIM	Liability Stochastic Interaction Model
MP	Model Points
MSRE	MultiSupport REnte
NNLS	Non-négative Least Squared
PB	Participation aux Bénéfices
PM	Provision Mathématique
RAFM	Risk Agility Financial Modeler
TMG	Taux Minimum Garanti
UC	Unités de Comptes

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Cadre de l'étude</b>	<b>3</b>
1.1 Allianz France . . . . .	4
1.2 Bilan prudentiel . . . . .	6
1.3 Meilleure estimation des engagements futurs . . . . .	7
1.4 Principes d'un modèle ALM . . . . .	7
1.4.1 Modélisation Stochastique . . . . .	9
1.4.2 Modélisation Déterministe . . . . .	10
1.5 Approche Flexing . . . . .	11
1.6 Approche Full ALM . . . . .	12
<b>2 Description du portefeuille assurantiel</b>	<b>14</b>
2.1 L'activité de l'assurance vie . . . . .	15
2.1.1 Définition de l'assurance vie . . . . .	15
2.1.2 Répartition et rémunération du capital . . . . .	15
2.1.3 Options et garanties . . . . .	16
2.1.4 Frais liés à l'assurance vie . . . . .	17
2.2 Description du portefeuille . . . . .	17
2.3 Différents niveaux d'agrégation . . . . .	19
2.3.1 Premier niveau d'agrégation . . . . .	19
2.3.2 Second niveau d'agrégation . . . . .	21
<b>3 L'optimisation linéaire</b>	<b>24</b>
3.1 Modèle de projection du passif et calcul des flux . . . . .	25
3.1.1 Primes . . . . .	26
3.1.2 Prestations . . . . .	26
3.1.3 Revalorisation des contrats . . . . .	29
3.1.4 Chargements et pénalités . . . . .	30
3.1.5 Coûts et frais de gestion des placements . . . . .	31
3.1.6 Commissions . . . . .	31
3.1.7 Taxes et impôts . . . . .	31
3.1.8 Flux considérés pour la meilleure estimation . . . . .	32

3.2	Principes de l'algorithme . . . . .	32
3.3	Résultats . . . . .	34
3.4	Seuil de tolérance . . . . .	35
3.5	Critiques de l'approche linéaire . . . . .	37
<b>4</b>	<b>Approche par Clustering</b>	<b>38</b>
4.1	Probabilités de sortie . . . . .	39
4.1.1	Motivation et formulation mathématique . . . . .	39
4.1.2	Calcul des probabilités de sortie . . . . .	40
4.2	Phase de classification du portefeuille . . . . .	42
4.2.1	Description du problème de classification . . . . .	43
4.2.2	Algorithme du kmeans . . . . .	44
4.2.3	Métriques de base des séries temporelles . . . . .	45
4.2.4	Initialisation de l'algorithme . . . . .	47
4.2.5	Evaluation de la classification . . . . .	49
4.3	Phase d'agrégation du portefeuille . . . . .	51
4.3.1	Moyenne pondérée par la provision mathématique de l'âge et de l'ancienneté . . . . .	51
4.3.2	Méthode par développement limité en série de Taylor . . . . .	53
<b>5</b>	<b>Validation de la compression par kmeans</b>	<b>60</b>
5.1	Procédure de la validation . . . . .	61
5.2	Avant compression par kmeans . . . . .	62
5.3	Après compression par kmeans . . . . .	62
5.4	Impact de la compression par kmeans . . . . .	62
5.5	Critiques de l'approche par kmeans . . . . .	64
5.5.1	Critiques générales . . . . .	64
5.5.2	Améliorations possibles de notre étude . . . . .	64
	<b>Conclusion</b>	<b>65</b>
	<b>Bibliographie</b>	<b>i</b>
	<b>Annexes</b>	<b>ii</b>

# Table des figures

1.1	Chiffres clés d'Allianz France . . . . .	4
1.2	Organisation de la direction de l'Actuariat au sein d'Allianz France . . . . .	5
1.3	Structure du bilan prudentiel . . . . .	6
1.4	Schéma synthétique d'un modèle ALM . . . . .	8
1.5	Schéma synthétique de la modélisation en Flexing . . . . .	12
1.6	Schéma synthétique de la modélisation en Full ALM . . . . .	13
2.1	Présentation du portefeuille de contrats . . . . .	18
2.2	Répartition du périmètre MSRE d'AZ Vie par type de contrat . . . . .	19
2.3	Illustration Précision et Compression . . . . .	21
2.4	Représentation de la structure d'un sous-portefeuille . . . . .	22
2.5	Variable d'agrégation . . . . .	23
3.1	Somme des écarts absolus en fonction du Taux de compression . . . . .	36
4.1	Probabilités de sortie associées à un sous-portefeuille . . . . .	42
4.2	Comparaison entre la mesure $p$ -normée et le $DTW$ . . . . .	46
4.3	Exemple d'application du kmeans sur les prob. de sortie pour $k = 5$ . . . . .	49
4.4	Répartition des indices de Davies-Bouldin par mesure de distance . . . . .	51
4.5	Exemple de la méthode d'agrégation par moyenne pondérée de l'âge et de l'ancienneté . . . . .	52
4.6	Comparaison Centroides VS Model Points . . . . .	53
4.7	Age et ancienneté selon les clusters . . . . .	55
4.8	Somme des écarts absolus en fonction du facteur $a$ . . . . .	59
5.1	Erreur de BEL sur les 1000 trajectoires stochastiques . . . . .	63
5.2	Représentation des chroniques de cash-flows et de la valeur cible . . . . .	iii
5.3	La cible et de la combinaison des cashflows à l'initialisation . . . . .	iv
5.4	La cible et de la combinaison des cashflows à première itération . . . . .	v
5.5	La cible et de la combinaison des cashflows à deuxième itération . . . . .	vi
5.6	La cible et de la combinaison des cashflows à la troisième itération . . . . .	vii
5.7	La cible et de la combinaison des cashflows à la quatrième itération . . . . .	viii
5.8	La cible et de la combinaison des cashflows à la cinquième itération . . . . .	ix

# Introduction

La directive Solvabilité 2 a généralisé l'utilisation des modèles ALM (*Asset Liability Management*) pour l'évaluation des provisions des produits d'assurance vie. Cette évolution réglementaire s'est imposée à l'ensemble du marché : mutuelles, institutions de prévoyance et compagnies d'assurance. La nouvelle norme comptable internationale (IFRS 17) de valorisation des contrats d'assurance vient imposer une utilisation encore plus poussée de ces modèles ALM.

Le modèle ALM historiquement utilisé par Allianz France est un modèle en Flexing. Cette méthode permet d'avoir rapidement des résultats, mais à une large maille de calcul. Cependant, avec l'entrée en vigueur de la norme IFRS 17 et sa nouvelle définition de segmentation de portefeuille, le modèle ALM n'est plus en mesure de répondre aux besoins de la compagnie. La solution trouvée par Allianz pour s'adapter à cet environnement changeant a été de développer un nouveau modèle ALM dit en Full ALM. Celui-ci ayant l'avantage de répondre aux nouvelles contraintes imposées par la norme IFRS 17, puisqu'il permet d'effectuer des projections à une maille plus fine.

Toutefois, en raison d'une demande de reporting exigeante afin de prouver la conformité aux différentes normes comptables (IFRS 17 et Solvabilité 2), l'un des défis majeurs des modèles Actif-Passif en Full ALM est le temps nécessaire pour obtenir les résultats. En effet, pour des portefeuilles avec un grand nombre de contrats, le nombre de calculs peut vite devenir considérable : le temps et le volume disque alloués deviennent alors des problématiques opérationnelles non négligeables.

Conformément au Règlement Délégué publié par la commission européenne en janvier 2015, il est possible de faire des regroupements de polices dès lors que les polices considérées et les risques sous-jacents sont similaires et que l'utilisation de ces regroupements de polices (model points) ne fausse pas les résultats. L'objet de cette étude réside dans la construction de ces Model Points. En d'autres termes, il s'agit de transformer le portefeuille assurantiel des périmètres gérés par Allianz France en une nouvelle table de données contenant beaucoup moins de lignes tout en représentant le plus fidèlement possible le portefeuille initial. Cette technique de transformation est appelée compression ou agrégation. Une première méthode d'agrégation a été déjà proposée. Celle-ci est basée sur une approche d'optimisation linéaire. Mais sa configuration actuelle ne per-

met pas d'atteindre les résultats souhaités par Allianz. Il sera donc étudié dans quelle mesure cette approche peut être améliorée. Dans l'éventualité d'un échec, nous proposerons une méthodologie fondée sur des algorithmes de classification statistique.

Nous présenterons dans un premier chapitre, le contexte de changement de modèle dans lequel s'inscrit cette étude. Le deuxième chapitre proposera une description du portefeuille assurantiel soumis à notre analyse. La méthode d'optimisation linéaire comme technique de compression du portefeuille sera décrite dans le troisième chapitre, puis une toute autre approche de compression basée sur les techniques de clustering que nous avons proposée sera présentées et discutées en quatrième chapitre. Pour finir, le cinquième chapitre exposera les résultats du test de robustesse de l'approche par clustering et identifiera à la fois les limites de l'étude ainsi que des pistes de recherches complémentaires.

# Chapitre 1

## Cadre de l'étude

### Sommaire

---

<b>1.1</b>	<b>Allianz France</b> . . . . .	<b>4</b>
<b>1.2</b>	<b>Bilan prudentiel</b> . . . . .	<b>6</b>
<b>1.3</b>	<b>Meilleure estimation des engagements futurs</b> . . . . .	<b>7</b>
<b>1.4</b>	<b>Principes d'un modèle ALM</b> . . . . .	<b>7</b>
1.4.1	Modélisation Stochastique . . . . .	9
1.4.2	Modélisation Déterministe . . . . .	10
<b>1.5</b>	<b>Approche Flexing</b> . . . . .	<b>11</b>
<b>1.6</b>	<b>Approche Full ALM</b> . . . . .	<b>12</b>

---



Ce chapitre présente le contexte de changement de modèle dans lequel s'inscrit cette étude. Il justifie la nécessité de la compression et introduit les notions essentielles de l'étude comme les types de modélisation ALM et les différents modes de projection.

## 1.1 Allianz France

Le Groupe Allianz est un des leaders mondiaux de l'assurance et des services financiers. Il est implémenté dans plus de 70 pays et compte plus de 88 millions de clients. Le siège social est situé à Munich en Allemagne et le groupe opère dans la quasi-totalité des pays européens (dont la France), aux Etats-Unis et dans la région Asie-Pacifique.

En France, le Groupe Allianz est composé d'Allianz France, dont le directeur général est Fabien WATHLE, et d'autres entités telles que Allianz Partners, Euler Hermes ou encore Allianz Global Investors France. Assureur généraliste, Allianz France a pour activité l'assurance dommages, l'assurance de personnes (vie, retraite, santé, prévoyance) et la banque.

Allianz France c'est aujourd'hui :



FIGURE 1.1 – Chiffres clés d'Allianz France

La branche assurance d'Allianz France est organisée en plusieurs unités et directions dont l'**unité Finance**, dirigée par Remi SAUCIE. Celle-ci est décomposée en 6 directions :

- Direction Contrôle & Pilotage Financier
- Direction Process et Transformation Finance
- **Direction de l'Actuariat**
- Direction de la Comptabilité
- Direction de la Fiscalité
- Direction des Risques

Ce stage de fin d'études s'est déroulé au sein de l'équipe **Modélisation Vie** de la Direction Actuariat dont l'organisation est présentée ci-dessous.

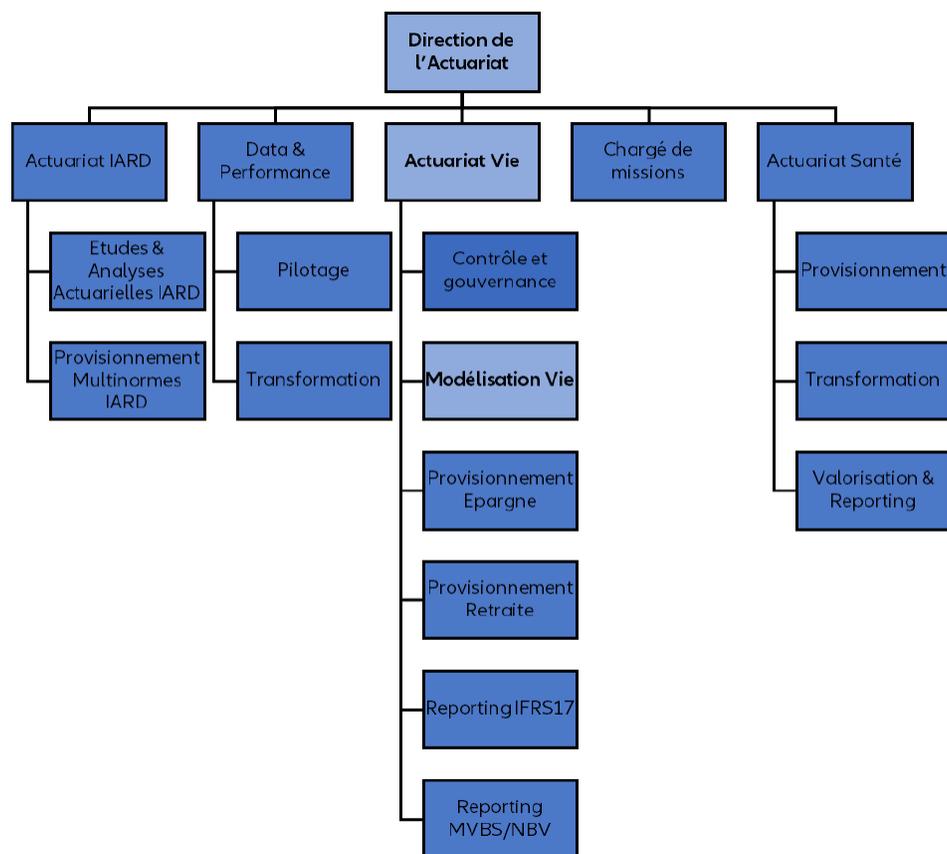


FIGURE 1.2 – Organisation de la direction de l'Actuariat au sein d'Allianz France

L'équipe Modélisation Vie a la charge de veiller sur le modèle interne d'Allianz. Elle effectue au quotidien des missions telles que :

- Assurer la veille méthodologique et accompagner l'évolution du modèle interne pour les aspects passifs d'assurance, quelle que soit l'origine des demandes (projet interne, évolution réglementaire, demande ACPR, etc.) notamment en menant les études d'impact et en participant au processus de validation.
- Assurer la maintenance et l'évolution du modèle interne d'assurance et assurer la documentation du modèle.
- Assurer la conduite régulière des revues actuarielles relatives à la souscription et la tarification (nouveaux produits, appels d'offres, suivi post-lancement) : revue des business plan / profit testing, sign-off des analyses de rentabilité.

Et depuis peu, elle est chargée d'assurer la migration du modèle en Flexing vers le modèle en Full ALM. C'est dans le cadre de cette migration que s'inscrit cette étude.

Dans ce chapitre, nous allons d'abord définir la structure du bilan prudentiel d'une compagnie d'assurance, puis nous formaliserons une des notions centrales de ce bilan à savoir le **Best Estimate of Liabilities**. Ensuite seront abordés les principes généraux d'un modèle ALM. Ces principes sont communs aux deux grandes approches existant pour la mise en œuvre d'un modèle ALM qui elles, seront abordés en fin de chapitre.

## 1.2 Bilan prudentiel

La notion de bilan prudentiel a été instaurée par la Directive 2009/138/CE du Parlement européen et du Conseil du 25 novembre 2009 ou plus connue sous le nom de Solvabilité II. Cette nouvelle définition permet d'avoir une vision économique du bilan. L'actif et le passif sont évalués à leurs justes valeurs.

Au passif, on retrouve les provisions techniques qui représentent l'anticipation des sinistres que l'assureur devra prendre en charge. Elles sont au passif car il s'agit d'une estimation de ce que l'assureur devra payer aux assurés. Il s'agit d'une dette envers les assurés.

La méthode de calcul de ces provisions a été homogénéisée au niveau européen par la réglementation Solvabilité II. Elle implique un calcul des « meilleures estimations » ou « best estimate » des flux probables dans le futur, actualisés à la date d'aujourd'hui, et intègre aussi une marge de risque.

À l'actif, ce sont les placements qui représentent le plus grand poste car l'assureur place les primes qu'il reçoit en instruments financiers, dans l'attente du paiement des sinistres.

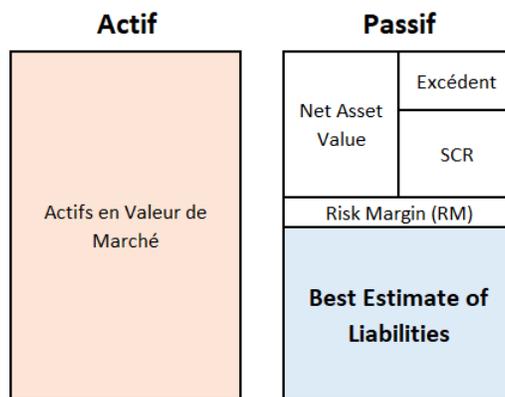


FIGURE 1.3 – Structure du bilan prudentiel

Le capital de solvabilité requis (SCR) correspond au capital économique pour limiter la probabilité de ruine à 1 an à 0,5%

### 1.3 Meilleure estimation des engagements futurs

La meilleure estimation en vision économique des engagements futurs, ou encore les provisions Best Estimate ou *Best Estimate of Liabilities* - BEL est une notion centrale du bilan économique défini par Solvabilité 2. Il représente l'estimation sans biais de la valeur actuelle des flux de trésorerie (cashflow) futurs attendus. Le Best Estimate se calcule comme l'espérance des flux de trésorerie futurs actualisés avec la courbe des taux sans risque. En d'autres termes, il s'agit de la moyenne des flux financiers sortants en valeur actuelle pondérés par leur probabilité d'occurrence sur une période définie par un horizon de projection.

$$BEL = \mathbb{E}^{\mathbb{P} \otimes \mathbb{Q}} \left[ \sum_{t \geq 1} Flux_t \times \delta_t \right] \quad (1.1)$$

avec :

- $\mathbb{P}$  la probabilité sous laquelle est modélisée la durée de vie des contrats des assurés ;
- $\mathbb{Q}$  la probabilité risque neutre ;
- $\delta_t$  le facteur d'actualisation s'exprimant en fonction du taux sans risque ;
- $Flux_t$  les flux futurs de trésorerie ou cashflows.

### 1.4 Principes d'un modèle ALM

A partir d'un portefeuille initial et de scénarios économiques, le modèle ALM permet de construire des bilans économiques sur un horizon de projection donné, sur la base de flux de trésoreries projetés. Compte tenu des résultats fournis par le modèle, il est possible d'évaluer les engagements Best Estimate d'une compagnie d'assurance.

En général, il est nécessaire de choisir un horizon de projection suffisamment grand pour que les engagements de l'assureur soient négligeables en fin de projection. Le choix de cet horizon pourrait dépendre de plusieurs facteurs comme le type de produit modélisé, l'âge moyen des assurés ou aussi le niveau de la courbe des taux. Par exemple, un portefeuille constitué principalement de jeunes assurés nécessite un plus grand horizon de projection, puisque l'assureur est susceptible d'être engagé sur une plus longue période du fait de l'effet relativement faible de la mortalité et donc d'une durée de vie résiduelle plus longue. Dans le cas d'Allianz France, cette décision est prise par le régulateur qui impose à la compagnie, un horizon de projection de 60 ans.

De façon schématique, un modèle ALM est composé de trois différentes composantes : (i) des Model Points de l'actif et du passif, (ii) des hypothèses économiques et

techniques et (iii) un logiciel de projection.

Le schéma ci-dessous présente une vision synthétique d'un modèle ALM.

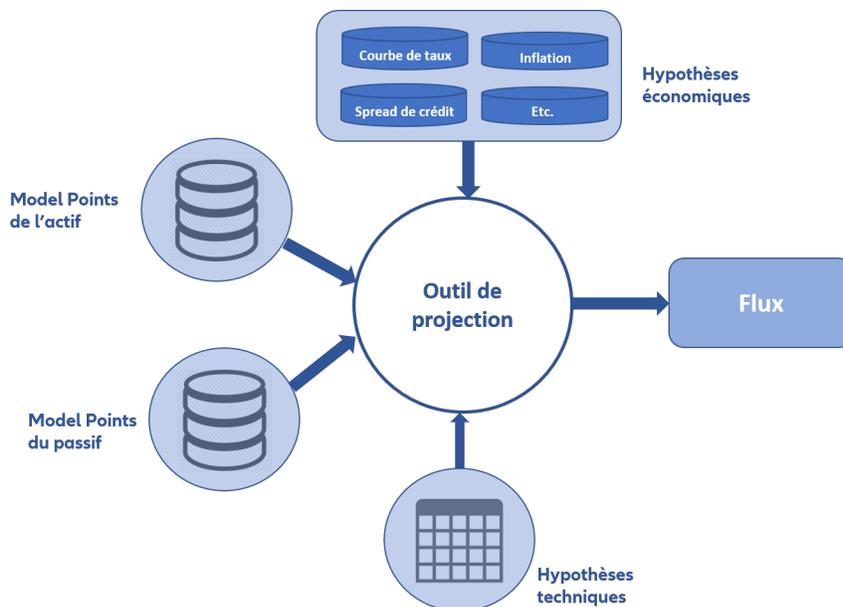


FIGURE 1.4 – Schéma synthétique d'un modèle ALM

Les Model Points, que ce soit pour l'actif ou le passif, sont des tables de données du portefeuille de la compagnie, agrégées en groupes homogènes tout en étant représentatives du portefeuille initial.

Les Model Points de l'actif représentent les actifs détenus par la compagnie, regroupés selon leurs caractéristiques telles que : la nature (obligation d'état, actions, immobilier, etc.), le type de taux (fixe, variable, indexé), la maturité, etc.

Les Model Points du passif, quant à eux représentent originellement les contrats dans le portefeuille de l'assureur. Cependant, du fait de la diversité des actifs en portefeuille et la multitude de contrats commercialisés ; projeter ligne par ligne serait très consommateur en ressources voire impossible avec la technologie actuelle. C'est là que se situe tout l'intérêt de la compression du portefeuille de contrats de sorte à réduire au maximum le nombre de lignes tout en conservant une certaine représentativité.

Afin de limiter la complexité des calculs dans le modèle, on y intègre généralement des hypothèses, établies en amont de la projection. On peut distinguer deux types d'hypothèses : les hypothèses économiques et les hypothèses techniques.

Les hypothèses économiques ont un rôle majeur dans les calculs de l'assureur puisqu'elles permettent de prévoir les différentes situations économiques dans lesquelles

l'assureur pourra potentiellement se retrouver. Ces hypothèses sont généralement regroupées au sein d'un Générateur de Scénarios Economiques (GSE). Il s'agit d'un outil de simulation stochastique qui permet de diffuser pour chaque simulation et sur un horizon de projection, les évolutions des facteurs de risques financiers auxquels l'assureur est exposé : les courbes de taux, les facteurs d'actualisation, les indices actions, l'indice d'inflation, le spread de crédit, etc.

Les hypothèses techniques sont l'ensemble des hypothèses entrant en compte dans la projection du résultat technique de l'assureur vie. Ces hypothèses sont établies grâce à des études statistiques sur les données du portefeuille de l'assureur, parmi elles, on peut retrouver : les tables de mortalité, les lois de rachats, les hypothèses de frais, etc.

Enfin, pour travailler avec des modèles ALM, les assureurs utilisent des progiciels de modélisation, le plus souvent sous le langage C++ (Moses, Prophet, RAFM, ...), qui permettent de développer et d'exécuter des projections financières et actuarielles. A ce titre, plusieurs modes de projection existent. Ils sont présentés dans les parties suivantes.

### 1.4.1 Modélisation Stochastique

Un scénario économique correspond à la projection de grandeurs financières et/ou économiques sur un horizon donné. Le générateur de scénarios économiques fournit en entrée un millier de scénarios avec le même horizon temporel. S'en suivent alors deux boucles imbriquées : une première sur le pas temporel et l'autre sur les scénarios économiques.

Dans le cadre d'une modélisation stochastique, pour l'ensemble des scénarios financiers ( $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_N$ ), le Best Estimate défini dans l'équation 1.1 peut s'estimer par la méthode de Monté-Carlo de la façon suivante :

$$BEL = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^T Flux_j^{\mathbf{F}_i} \times \delta_j^{\mathbf{F}_i} \quad (1.2)$$

avec :

- $N$  le nombre de scénarios économiques ;
- $T$  l'horizon de projection ;
- $Flux_j^{\mathbf{F}_i}$  les flux de trésorerie ou cashflow de l'année  $j$  pour le scénario  $\mathbf{F}_i$  ;
- $\delta_j^{\mathbf{F}_i}$  le déflateur relatif à l'année  $j$  et au scénario  $\mathbf{F}_i$ .

Dans une modélisation stochastique, le passif est projeté en même temps que les actifs puis les contrats sont revalorisés à chaque période via la stratégie de participation aux bénéfices modélisée via un algorithme de partage des richesses entre assureur

et assuré. Les scénarios considérés dans une modélisation stochastique permettent de déclencher des options et garanties financières qui ne le seraient pas sur un scénario moyen. Ces options et garanties financières sont de plusieurs natures et sont associées à des risques pour l'assureur d'où l'importance de les prendre en compte. Le tableau suivant en donne quelques exemples. Ces options et garanties seront vues plus en détail dans la suite.

TABLE 1.1 – Exemples d'options et garanties avec les risques associés

Option/Garantie	Définition	Risque pour l'assureur
Taux minimum garanti	garantit un taux de revalorisation minimal de l'épargne	rendement de l'actif inférieur à ce taux
Option de rachat	donne le droit de récupérer à tout moment, totalement ou partiellement son épargne	risque d'illiquidité pour faire face aux vagues de rachats

### 1.4.2 Modélisation Déterministe

Contrairement à la modélisation stochastique, la modélisation déterministe n'utilise qu'un seul scénario dit scénario moyen  $\mathbf{F}$ . Ce scénario ne déclenche aucune option et garantie. Le coût de ces options et garanties est déterminé par la différence entre le BEL stochastique et le BEL déterministe.

Pour le scénario  $\mathbf{F}$ , le Best Estimate peut s'estimer de la façon suivante :

$$BEL^{\mathbf{F}} = \sum_{j=1}^T Flux_j^{\mathbf{F}} \times \delta_j^{\mathbf{F}} \quad (1.3)$$

avec :

- $T$  l'horizon de projection ;
- $Flux_j^{\mathbf{F}}$  les flux de trésorerie ou cashflow de l'année  $j$  pour le scénario  $\mathbf{F}$  ;
- $\delta_j^{\mathbf{F}}$  le déflateur relatif à l'année  $j$  et au scénario  $\mathbf{F}$ .

Le modèle déterministe, à l'instar du modèle stochastique, projette à la fois le passif et l'actif, ce qui rend les projections assez chronophages. Dans toute la suite de ce mémoire, sauf mention explicite du contraire, tous les engagements de BEL seront calculés sous un mode de projection particulier appelé le « déterministe du Passif seul

» ou plus simplement « Passif seul ». La particularité de ce mode de projection est de ne projeter que le passif sans interaction actif-passif. Les hypothèses de rendements financiers utilisées correspondent alors à celles des taux de rendements prévisionnels des actifs financiers tels que déterminés dans le scénario central (risque neutre). Même si elle n'est pas réaliste, l'intérêt d'une modélisation en passif seul réside dans le gain de temps qu'elle procure puisque les calculs sont réalisés uniquement sur le passif qui concentre toute notre attention.

Combinés aux modes de projections vus précédemment, deux principales approches existent pour la mise en œuvre d'un modèle ALM : **l'approche Flexing** et **l'approche Full ALM**. La première approche est celle historiquement utilisée par Allianz France, tandis que la seconde est celle vers laquelle souhaite migrer l'entité.

## 1.5 Approche Flexing

ALIM (*Asset and Liabilities Interaction Management*) est le modèle ALM utilisé par le groupe Allianz. Il a par la suite été adapté afin de couvrir l'ensemble des activités des différentes sociétés de l'entité Allianz France. Le modèle est à pas annuel et a été produit pour permettre à la compagnie de répondre aux contraintes du superviseur et également pour effectuer des études transverses. Il est développé sous le logiciel RAFM (RiskAgility Financial Modeler) en langage C++. ALIM est basé sur l'approche Flexing.

La méthode Flexing est une méthode dans laquelle les données du passif sont d'abord projetées ligne par ligne de façon déterministe via des modèles déterministes du passif seul appelés les MoDet. Chez Allianz, ces projections sont réalisées à pas mensuel jusqu'à extinction des engagements et ce pour obtenir les flux de trésorerie ou cashflow futurs. Ensuite, un outil (RAI TP) regroupe les chroniques de cashflow projetées en groupes de produits selon différents critères (type de contrat, TMG, etc.), puis annualise les cashflows mensuels regroupés afin de créer les « Liabs » qui serviront à alimenter le modèle stochastique ALIM. Ces « Liabs » ou chroniques de cashflows annuels regroupés seront par la suite ajustés ou tordues (d'où le Flexing) par le modèle ALIM via une prise en compte des scénarios économiques, donnant ainsi le taux de rendement du portefeuille d'actifs induisant par la suite les options de rachat et les garanties de taux.

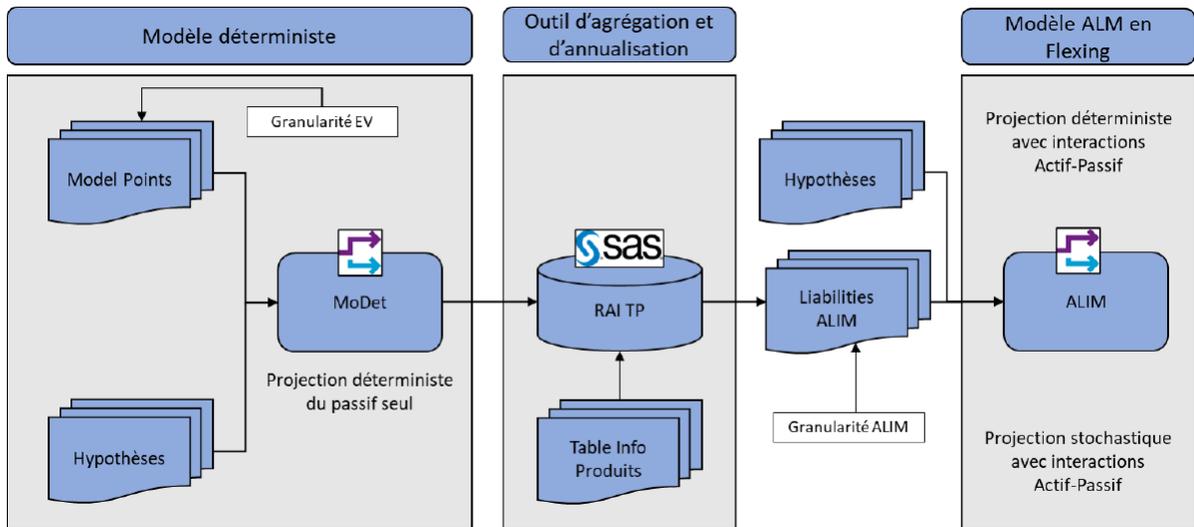


FIGURE 1.5 – Schéma synthétique de la modélisation en Flexing

## 1.6 Approche Full ALM

LSIM (*Liability Stochastic Interaction Model*) est le futur modèle en Full ALM d'Allianz, c'est-à-dire un modèle à pas annuel tout intégré ne nécessitant ni de projection déterministe au préalable, ni de torsion de flux.

La méthode Full ALM est une méthode dans laquelle les données du passif sont utilisées directement dans les modèles stochastiques, sans projection préalable en passif seul ni regroupement à une maille plus grossière.

Par rapport au Flexing, une approche Full ALM permet :

- Une modélisation plus fine et réaliste du passif
- Une évaluation plus juste du coût des options et garanties
- Une simplification du processus de calcul

Pour produire un tel modèle, les modifications à effectuer se situent principalement au niveau du passif. Pour cela, Allianz France s'est servie de la modélisation ALM actuelle. En effet, la construction du modèle LSIM s'est faite selon le cheminement suivant :

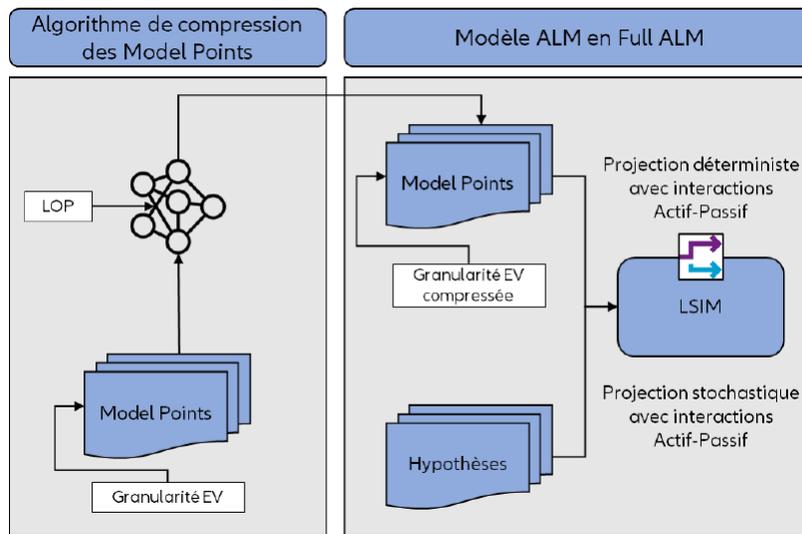


FIGURE 1.6 – Schéma synthétique de la modélisation en Full ALM

- Import des modèles de passifs (MoDet) dans le modèle ALM
- Adaptation des dimensions des modèles de passif à celles du modèle Full ALM (passage de la maille ALIM à la maille EV)
- Revue de la dimension de la *crediting strategy* (dans le modèle en Flexing, la *crediting strategy* est effectuée à la maille ALIM)
- Evolution des outputs de la *crediting strategy* et adaptation des formules de passif (passage de montants crédités à taux crédités)

Les inputs du nouveau modèle LSIM demeurent, à quelques modifications près, les mêmes, à savoir :

- Les hypothèses économiques et techniques
- Les model points de l'actif

On observe cependant, une différence au niveau des Model Points du passif : dans le modèle ALIM en Flexing, les inputs de passif correspondent à des cashflows agrégés (à la maille ALIM) et annualisés obtenus à l'aide d'une unique projection déterministe. Tandis que dans le modèle LSIM en Full ALM, les inputs du passif restent des contrats qui seront projetés à chaque nouveau scénario. D'où la nécessité de compresser ces contrats par un algorithme de compression pour limiter le temps de calcul.

# Chapitre 2

## Description du portefeuille assurantiel

### Sommaire

---

<b>2.1</b>	<b>L'activité de l'assurance vie</b>	<b>15</b>
2.1.1	Définition de l'assurance vie	15
2.1.2	Répartition et rémunération du capital	15
2.1.3	Options et garanties	16
2.1.4	Frais liés à l'assurance vie	17
<b>2.2</b>	<b>Description du portefeuille</b>	<b>17</b>
<b>2.3</b>	<b>Différents niveaux d'agrégation</b>	<b>19</b>
2.3.1	Premier niveau d'agrégation	19
2.3.2	Second niveau d'agrégation	21

---



Après une présentation détaillée l'activité de l'assurance vie, ce chapitre propose une description du portefeuille assurantiel soumis à notre étude. Il décrit la nature et le mécanisme de fonctionnement des contrats constituant le portefeuille.

## 2.1 L'activité de l'assurance vie

Couramment dénommée "le placement préféré des français", l'assurance vie est actuellement en France le produit d'épargne le plus important en termes d'encours. Selon la Fédération Française des Assurances (FFA), l'encours total des contrats d'assurance vie s'élevait à 1 818 milliards d'euros à la fin avril 2021.

### 2.1.1 Définition de l'assurance vie

L'assurance-vie peut être définie comme un contrat d'épargne sur lequel l'assuré verse une somme d'argent de manière ponctuelle ou régulière et, en contrepartie, l'assureur s'engage à verser un capital ou une rente à une ou plusieurs personnes déterminées (l'assuré ou ses bénéficiaires en cas de décès). Il est possible de distinguer deux événements générateurs de la prestation.

- L'assurance en cas de décès : la compagnie s'engage à verser un capital ou une rente à un ou plusieurs bénéficiaires lorsque le décès de l'assuré survient avant la fin du contrat.
- L'assurance en cas de vie : à la différence de la précédente, un capital ou une rente est versé à l'assuré si ce dernier est toujours en vie.

### 2.1.2 Répartition et rémunération du capital

Durant la vie du contrat, le capital est réparti sur un ou plusieurs supports : d'une part le fonds en euros et d'autre part un ou plusieurs supports en unités de compte (UC) qui visent à rechercher un meilleur rendement, mais qui nécessitent d'assumer un risque potentiel de perte en capital.

La rémunération du capital en assurance vie dépend principalement du support sur lequel est placée l'épargne. Sur les fonds en euros, le capital de l'assurance vie est garanti. L'assuré n'a donc à supporter aucun risque de perte en capital. Quant aux unités de compte, il s'agit de supports d'investissement, telles que des obligations, des actions ou encore des parts d'OPCVM (Organisme de placement collectif en valeurs mobilières). Avec ces supports, il n'y a aucune garantie sur le capital de l'assurance vie, car leur valeur et leur taux de rendement varient en fonction de l'évolution des marchés financiers. Cependant ils présentent un fort intérêt car ils peuvent octroyer un meilleur rendement que le fonds en euros, à condition d'accepter un risque de perte en capital.

Le plus souvent, l'assurance vie prendra la forme d'un contrat multisupport. Comme son nom l'indique, ce contrat prévoit que l'épargne soit à la fois placée sur un fonds en euros et sur des unités de compte. Il permet ainsi de protéger une partie du capital sur le fonds en euros, tout en visant un rendement plus attractif sur les supports en unités

de compte. À titre d'exemple, il est possible de placer toute son épargne sur un fonds en euros et de n'investir que les intérêts dégagés sur un ou plusieurs supports en unités de compte.

Pour récupérer son épargne assurance-vie et les éventuels intérêts et/ou plus-values générés, l'assuré dispose de plusieurs options selon les contrats :

- effectuer un rachat intégral pour percevoir la totalité du capital et mettre ainsi fin au contrat d'assurance-vie ;
- réaliser un rachat partiel pour toucher une partie du capital, tandis que le reliquat continue de fructifier ;
- réaliser une sortie en rente viagère, permettant de percevoir régulièrement une fraction du capital, selon une périodicité définie dans le contrat (mensuelle, trimestrielle ou encore annuelle), et ce, jusqu'à son décès ;
- opter pour une sortie mixte, combinant à la fois un rachat du capital et le versement d'une rente viagère.

### 2.1.3 Options et garanties

Les options et garanties sont des prérogatives accordées par l'assureur à l'assuré dans le cadre du contrat liant les deux parties. Ces prérogatives sont associées à des risques que l'assureur se doit de couvrir. Il en existe plusieurs, mais nous aborderons les plus importantes.

#### **Les garanties : Le Taux Minimum Garanti et la Participation aux Bénéfices**

Le taux minimum garanti (TMG) est un taux contractuel qui doit être versé à l'assuré quels que soient les rendements financiers de la compagnie d'assurance ;

La participation aux bénéfices (PB) représente une partie des bénéfices financiers et techniques réalisés par la compagnie d'assurance avec l'épargne des assurés. Selon l'Article A331-4 du Code des Assurances, une compagnie d'assurance doit reverser à ses assurés à minima 85% du résultat financier et 90% du résultat technique de l'année si celui-ci est positif et 100% s'il est négatif. Il faut cependant noter que la compagnie peut décider d'en conserver une partie en provision pour participation aux bénéfices afin de lisser les rendements sur plusieurs années, plutôt que de les reverser directement aux assurés. La durée maximale de détention avant distribution totale est de 8 ans.

## Option de rachat

Dans le cas des contrats d'épargne, les compagnies d'assurance proposent à leurs clients une option de rachat. Cette option permet aux assurés de retirer leur épargne à tout moment et, en général, sans pénalité.

Les assurés exercent le droit de racheter leur contrat dans différents contextes. Le comportement de rachat est modélisé à travers deux mécanismes : les rachats structurels et les rachats conjoncturels.

### 2.1.4 Frais liés à l'assurance vie

Dans de nombreux cas, l'assureur est susceptible d'appliquer à l'assuré des frais sur l'assurance vie. On distingue principalement :

- les frais de dossier facturés lors de la souscription du contrat ;
- les frais d'entrée sur l'assurance vie, généralement payés à chaque versement que réalise l'assuré ;
- les frais de gestion sur l'assurance vie, facturés tout au long du contrat ;
- les frais d'arbitrage sur l'assurance vie qui sont prélevés lors des transferts réalisés entre deux supports.

## 2.2 Description du portefeuille

Les activités d'Allianz France sont modélisées au sein de trois compagnies différentes : Allianz Vie (AZ Vie), Allianz Retraite (AZ Ret) et Génération Vie (GV). Chacune de ces compagnies contient un ou plusieurs périmètres.

Seuls deux périmètres dans la figure 2.1 sont concernés pour l'instant par la migration vers la modélisation en Full ALM. Il s'agit des périmètres MSRE des compagnies AZ Vie et AZ Ret. Cependant, le périmètre retenu pour cette étude est le périmètre MSRE contenant tous les produits **MonoSupport**, **MultiSupports** (Euro/UC) et **REntes** (phase de constitution et phase de restitution) de la compagnie AZ Vie. En outre, la portée de cette étude se limitera uniquement aux produits Mono et MultiSupports (Euro/UC) de ce périmètre et aux produits de rente en phase de constitution ou capitalisation. En effet, les contrats de rente en phase de capitalisation sont assimilables aux contrats d'épargne "pure" car ceux-ci sont placés sur les mêmes supports pour la constitution du capital avant restitution de celui-ci à l'assuré sous forme d'arrérages. De ce fait, dans toute la suite de ce document, le terme "contrat d'épargne retraite individuel" fera référence à ces deux types de produit qui représentent plus de 70% du périmètre MSRE (voir fig. 2.2).

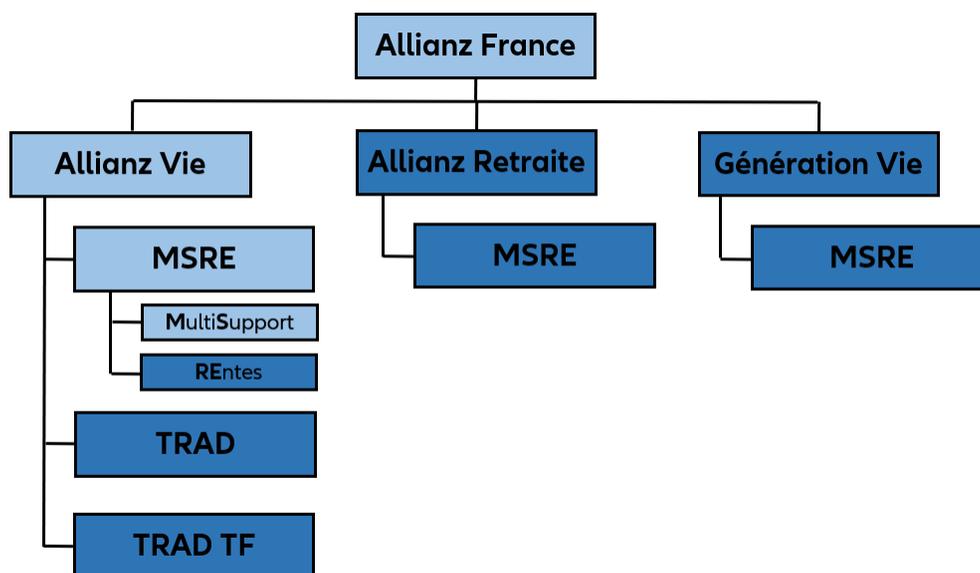


FIGURE 2.1 – Présentation du portefeuille de contrats

Dans le cadre d'un contrat d'assurance vie de type épargne retraite individuel, l'assuré effectue à la souscription un premier versement. Ce capital initial est investi sur différents supports d'investissement dont le rendement est fonction des marchés financiers. L'épargne est capitalisée en fonction de ces rendements selon le support d'investissement. L'assureur doit être en mesure, à tout instant, de verser une somme égale à la valeur de l'épargne à l'assuré dans le cas où celui-ci décide de racheter avant le terme du contrat. Ou encore dans l'objectif de percevoir le capital de son assurance vie, l'assuré peut opter pour une sortie en rentes viagères. Ce mode de sortie consiste à transformer le capital en un revenu garanti et régulier jusqu'à son décès.

Les autres périmètres de AZ Vie sont conservés en Flexing car les enjeux ALM y sont moins importants au sens de la volumétrie des réserves mathématiques et donc présentent peu d'intérêt de passer à une modélisation en Full ALM. Le périmètre MSRE de la compagnie GV est en cours de développement.

Notre portefeuille d'assurés est donc initialement constitué de plusieurs contrats d'épargne retraite individuels. Chaque contrat est défini selon des éléments qui lui sont propres (propriétaire du contrat, âge, frais...). On dit que les contrats sont à la maille<sup>1</sup> individuelle. Le nombre de contrats à la maille individuelle que comprend notre portefeuille est d'environ 1 millions de contrats.

1. Par définition, une maille ou une granularité est une combinaison élémentaire de différentes variables caractéristiques d'un contrat (code produit, tarif, réseau, etc.)

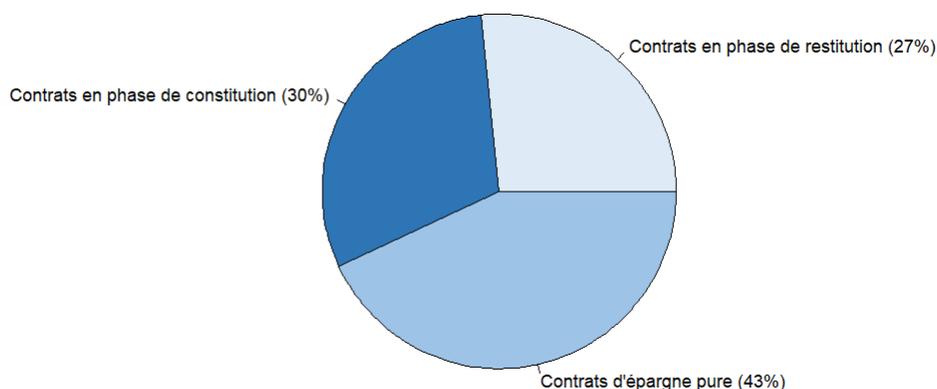


FIGURE 2.2 – Répartition du périmètre MSRE d'AZ Vie par type de contrat

### Création des classes d'agrégation

La création des classes d'agrégation est une étape qui consiste à partitionner le portefeuille en plusieurs sous-portefeuilles. Une classe d'agrégation est une variable créée à partir d'autres variables suivant des règles spécifiques et qui va servir comme critère d'agrégation. Elle permet d'agréger des données ayant le même critère d'agrégation, c'est à dire que seuls les contrats appartenant à une même classe d'agrégation pourront être agrégés dans tout le processus.

La création des classes d'agrégation est une étape de réorganisation du portefeuille qui ne modifie pas la maille des contrats et qui n'engendre aucune perte d'information.

## 2.3 Différents niveaux d'agrégation

### 2.3.1 Premier niveau d'agrégation

Le premier niveau d'agrégation est une étape qui consiste à regrouper au sein d'un même sous-portefeuille, les contrats ayant des caractéristiques identiques ou proches sous contrainte d'une perte d'information minimale dans les résultats et d'un volume de données post agrégation limité. Les données sont regroupées suivant différents critères d'agrégation dont les principaux sont :

- Groupe de produits aux caractéristiques similaires,
- Réseau de distribution,
- Famille de taux minimum garanti,
- Tranche d'âge à la souscription,
- Année de souscription,

- Famille de taux de chargement de gestion,
- Durée de contrat pour les contrats non prorogables ou avec sortie en rente obligatoire

Pour chaque critère d'agrégation, les données quantitatives (primes, provisions mathématiques, nombre de contrats) sont additionnées et les durées et âges sont calculés par moyenne pondérée par les encours.

Cette étape fait passer les contrats de la maille individuelle à la maille EV pour *Embedded Value*, et réduit la taille des données d'un facteur 30 pour atteindre un peu moins de 35 000 contrats.

Le tableau ci-dessous présente les informations que nous retrouvons pour chaque contrat à la maille EV à l'issue du premier niveau d'agrégation :

TABLE 2.1 – Liste des caractéristiques des contrats

Libellé	Description des champs	Exemple
TYPE_PROD	Type de produit	MS
TOP_RE	Indice de rente	N
BU	Business Unit	VI
CODEPROD_M	Code produit pour la phase de capitalisation	Ideavie
CODEPROD_R	Code produit pour la phase de rente	Ideavie
CODE_RES	Code réseau	AF
DATE_SOUS	Date de souscription	01/04/2009
DATPRIME_M	Date de la prime	01/07/2014
DATE_NAI_X	Date de naissance de l'assuré	01/04/1961
AGE_SOUS_M	Age à la souscription	48
AGE_ACT_X	Age actuariel de l'assuré	53
SEXE_X	Sexe de l'assuré (1 = homme, 2 = femme)	2
DURECNT_M	Durée du contrat (phase de capitalisation)	1188
PP_DUREE_M	Durée de la prime périodique	1188
FREQ_PP_MS	Fréquence de la prime périodique (1 = annuel)	1
NB_CNT_INI	Nombre de contrats	0
PRIME_PU	Montant en euro de la prime unique	699436,39
NB_PR_PU_I	Nombre de prime unique en versement libre	54
PC_PU_EURO	Pourcentage de prime unique sur le support en Euro	51,02079
PRIME_PP	Montant en euro de la prime périodique	0
PC_PP_EURO	Pourcentage de prime périodique sur le support en Euro	0

Même si le premier niveau d'agrégation permet de réduire considérablement la taille du portefeuille, le nombre de lignes reste encore trop élevé pour réaliser des projections déterministes ou stochastiques sans risquer des délais déraisonnables dans le temps d'exécution de ces modèles.

### 2.3.2 Second niveau d'agrégation

Le second niveau d'agrégation est le sujet de mémoire. Son but est de compresser davantage le portefeuille de contrats en regroupant cette fois les lignes ayant des caractéristiques différentes mais des évolutions similaires dans le temps.

#### Principe général

Il existe plusieurs manières de regrouper les contrats d'un portefeuille. Cependant, toutes les méthodes partagent l'idée commune de pouvoir remplacer l'ensemble des polices d'assurance, par un ensemble plus réduit de model points de sorte que les provisions techniques (BEL) calculées contrat par contrat d'une part, puis de façon agrégée par des model points d'autre part, soient sensiblement identiques. Ainsi, plus les provisions sont proches, plus l'erreur induite par la création des model points est petite et plus on dira que notre jeu de model points est une représentation fidèle de l'ensemble des polices.

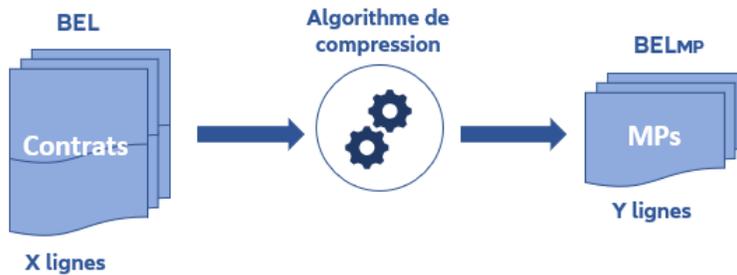


FIGURE 2.3 – Illustration Précision et Compression

La capacité d'une approche à restituer un ensemble de model points fidèle à l'ensemble des polices s'appelle la **précision**. Si l'on désigne par  $BEL$  et  $BEL_{MP}$  les Best Estimate calculés respectivement sur les données avant compression et les données après compression, la précision ou de façon équivalente l'erreur absolue induite par le procédé d'agrégation se calcule de la façon suivante :

$$EcartAbsolu = |BEL - BEL_{MP}|$$

où encore de façon relative :

$$EcartAbsoluRelatif = \frac{|BEL - BEL_{MP}|}{BEL}$$

L'écart absolu relatif sera exprimé en pour 10 000 dans ce mémoire.

Cela étant, il est aussi nécessaire de réduire autant que possible le nombre de lignes de l'ensemble de model points par rapport à l'ensemble des polices. Ce second critère est connu sous le nom de **taux de compression**. Il est exprimé en général en pourcentage.

$$Compression = \frac{X - Y}{X}$$

où  $X$  désigne le nombre de lignes contenu dans l'ensemble des polices (avant compression) et  $Y$  désigne le nombre de lignes contenu dans l'ensemble des model points (après compression).

Il apparait clairement que ces deux notions paraissent antagonistes. En effet, une précision maximale implique de conserver un grand nombre de model points donc un taux de compression proche de zéro. A l'inverse, plus la compression est forte, moins l'ensemble de model points est susceptible de répliquer fidèlement les caractéristiques de l'ensemble de polices initial, et donc plus la précision est faible. Il existe donc un compromis **précision vs compression** qu'il est nécessaire de prendre en compte dans tout approche de regroupement de polices. Dans la majorité des cas, ce problème est résolu en fixant en amont un taux de compression, par exemple 95%, puis choisir l'approche avec la précision la plus élevée. Dans notre étude, nous ciblons un nombre de Model Points de 4500, soit une compression du portefeuille de 87%. Cette valeur est choisie en fonction de la capacité de la plateforme de projection dont dispose Allianz. C'est le nombre maximal de lignes à considérer pour conserver un temps de calcul raisonnable sur ladite plateforme pour les modèles Déterministes et Stochastiques.

### Notion de variable d'agrégation

Comme vu précédemment, un contrat est défini par les éléments comme la Provision Mathématique (PM), le sexe, l'âge, l'ancienneté, etc.

Contrats	Age	Ancienneté	...	PM	Nb contrats
contrat 1	52	237		1000	5
contrat 2	66	236		2500	4
contrat 3	66	204		3000	10
contrat 4	63	226		1200	7
contrat 5	68	199		4000	8
contrat 6	27	209		3700	6
contrat 7	85	217		2000	4
contrat 8	44	199		5000	3
contrat 9	85	210		1600	11
contrat 10	87	238		2700	13

FIGURE 2.4 – Représentation de la structure d'un sous-portefeuille

Une approche naïve pour regrouper ces contrats serait de regrouper directement les valeurs les plus proches pour l'âge et l'ancienneté et de sommer les valeurs des autres

variables pour chaque groupe constitué. Il est aisé de constater qu'il n'est pas pertinent de regrouper d'une telle façon. D'abord pour la pénibilité que cela représente et procéder ainsi ne donne aucune assurance sur la représentativité du model point constitué, du point de vue du BEL, des contrats qui ont permis sa constitution. Aussi le choix des seuils est subjectif. Enfin un tel regroupement impliquerait de négliger les interactions entre les variables. L'idée est de regrouper de façon automatisée et plus intelligente. Une meilleure approche serait d'extraire à partir de ces variables, une dynamique probable de chaque contrat, puis regrouper les contrats sur la base de cette dynamique. C'est la quantité qui reflète cette évolution probable de chaque contrat et sur la base de laquelle seront regroupées les contrats que nous appelons variables d'agrégation.

Contrats	Age	Ancienneté	...	PM	Nb contrats	Variables d'agrégation							
						Var 1	Var 2	Var 3	Var 4	Var 5	...	Var n	
contrat 1	52	237		1000	5								
contrat 2	66	236		2500	4								
contrat 3	66	204		3000	10								
contrat 4	63	226		1200	7								
contrat 5	68	199		4000	8								
contrat 6	27	209		3700	6								
contrat 7	85	217		2000	4								
contrat 8	44	199		5000	3								
contrat 9	85	210		1600	11								
contrat 10	87	238		2700	13								

FIGURE 2.5 – Variable d'agrégation

Chez Allianz France, cette variable d'agrégation est calculée par le biais des modèles de projection implémentés avec un horizon de projection de 60 ans.

# Chapitre 3

## L'optimisation linéaire

### Sommaire

---

<b>3.1</b>	<b>Modèle de projection du passif et calcul des flux</b>	<b>25</b>
3.1.1	Primes	26
3.1.2	Prestations	26
3.1.3	Revalorisation des contrats	29
3.1.4	Chargements et pénalités	30
3.1.5	Coûts et frais de gestion des placements	31
3.1.6	Commissions	31
3.1.7	Taxes et impôts	31
3.1.8	Flux considérés pour la meilleure estimation	32
<b>3.2</b>	<b>Principes de l'algorithme</b>	<b>32</b>
<b>3.3</b>	<b>Résultats</b>	<b>34</b>
<b>3.4</b>	<b>Seuil de tolérance</b>	<b>35</b>
<b>3.5</b>	<b>Critiques de l'approche linéaire</b>	<b>37</b>

---



Ce chapitre présente en première partie le modèle de projection du passif. Ce modèle permet le calcul des cashflows de BEL utilisés comme variable d'agrégation par l'approche par optimisation linéaire présentée en second partie.

## 3.1 Modèle de projection du passif et calcul des flux

L'objectif de cette partie est de présenter l'approche de modélisation retenue pour la projection des cashflows d'assurance d'épargne individuelle et de rentes individuelles dans le modèle MSRE. Ces flux sont agrégés à chaque période pour constituer les cashflows de BEL. L'approche générale de la modélisation est une modélisation déterministe sans interaction actif passif (Passif seul). Les mécanismes de revalorisation des contrats ne tiennent pas compte de la gestion de la PPE (Provision pour Participation aux Excédents).

Les hypothèses retenues pour la modélisation des cashflows du passif sont principalement de six natures :

- les hypothèses démographiques et biométriques : elles regroupent les tables de mortalité, les lois de rachats et d'arbitrages,
- les hypothèses financières : elles regroupent les chroniques de rendement financier, les taux d'actualisation et les taux d'inflation,
- les hypothèses de revalorisation des contrats : il s'agit des taux minimums garantis, des taux de participation aux bénéfices (PB) et des taux de clauses de PB,
- les hypothèses de coûts : elles se composent des coûts de gestion sur encours, des coûts d'acquisition et des coûts de structure,
- les hypothèses de chargements et de commissions : elles regroupent les commissions et les chargements d'acquisition et sur encours,
- les hypothèses de taxes.

Pour chacune des lignes du portefeuille modélisé, et sur la base des hypothèses énumérées ci-dessus, le modèle permet de projeter dans le temps les éléments constitutifs des engagements.

les cashflows d'assurance

- les primes d'assurance payées par l'assuré,
- les prestations à la charge de l'assureur,
- la revalorisation des contrats,
- les chargements et pénalités
- les coûts
- les commissions
- les taxes et impôts

### 3.1.1 Primes

Le modèle MSRE permet de modéliser des primes uniques, des versements libres et des versements réguliers ou primes périodiques.

#### Primes uniques

Les primes uniques versées en phase de constitution correspondent au versement effectué à la souscription du contrat. Du point de vue de leur modélisation, elles correspondent à un flux égal à ce versement au pas de temps correspondant à la date du versement.

#### Versements réguliers

Pour les contrats d'épargne ou les contrats de retraite en phase de capitalisation, la méthode de projection des versements réguliers (mensuels, trimestriels, annuels) est basée sur un taux d'évolution considérant la probabilité de survie des assurés issue des tables de mortalité selon l'âge du Model Point et d'une loi d'indexation des versements réguliers dépendant de l'année civile renseignées dans les hypothèses du modèle.

#### Versements libres

Les versements libres correspondent à des versements exceptionnels effectués par les assurés durant la vie du contrat. Contrairement aux versements réguliers leur fréquence n'est pas contractuellement définie et ceux-ci sont exprimés en fonction d'un montant de versement livre de première année et une loi d'indexation de ce versement sur les années futures dépendant de l'ancienneté du contrat.

### 3.1.2 Prestations

Les prestations couvrent les éléments suivants : les termes, les rachats totaux, les rachats partiels, les décès, les arbitrages (entrant et sortant) et les sorties en rentes des

contrats multi-supports.

### Termes

Le terme est une prestation de l'assureur vie consistant à rembourser le capital existant et ce en date de maturité du contrat. Les montants de prestations brutes modélisés correspondent à la provision mathématique de fin d'année correspondant à la date de maturité calculée au taux garanti annuel. Cette assiette de calcul ne tient pas compte de la participation aux bénéfices de l'année, cette dernière n'étant acquise au contrat qu'en fin d'année.

### Rachats totaux

Les taux de rachats totaux sont appliqués aux nombres de contrats (nets des termes) en début d'année et à une assiette de provision mathématique.

L'assiette de calcul retenue correspond à la provision mathématique calculée au taux garanti en vigueur en milieu d'année. Cette assiette ne tient pas compte de la participation aux bénéfices de l'année, cette dernière n'étant acquise au contrat qu'en fin d'année.

Les éléments permettant de calculer les prestations de rachats totaux sont :

- les hypothèses de lois de rachat total *Best Estimate*,
- les hypothèses de lois de décès *Best Estimate*,
- le montant de provisions mathématiques au taux annuel garanti en vigueur en milieu d'année.

Le calcul des montants de prestations de rachats totaux s'effectue en plusieurs étapes :

1. calcul du taux de rachat total non combiné à partir du taux de rachats issu de la loi de rachat établie à la maille produit x réseau et par année d'ancienneté,
2. calcul du taux de rachat total combiné
3. calcul du montant de prestations brutes,
4. calcul des nombres de rachats,
5. calcul des montants des déductions devant être appliquées aux prestations brutes :
  - a. Pour le support Euro : calcul des montants de prélèvements sociaux associés aux intérêts techniques des prestations de rachats, du montant des pénalités de rachats, des charges fixes liées aux actes de rachats et celles liées à la gestion de l'encours de rachats.

- b. Pour le support UC : les déductions ne tiennent pas compte des prélèvements sociaux.

6. calcul des prestations nettes de rachats totaux.

### Rachats partiels

La méthodologie de calcul des rachats partiels est identique à celle retenue sur les rachats totaux. Il convient de noter qu'une différence réside dans le calcul des nombres de rachats partiels utilisés pour le calcul des coûts d'acte de gestion. Au contraire des nombres de décès ou de rachats totaux qui contribuent à déterminer le nombre de contrats en fin de période, les nombres de rachats partiels ne sont utilisés que pour la détermination des coûts fixes de gestion liés aux rachats partiels.

### Décès

Les taux de décès sont appliqués aux nombres de contrats (nets des termes) en début d'année et à une assiette de provisions mathématiques.

L'assiette de calcul retenue correspond à la provision mathématique revalorisée au taux garanti en vigueur en milieu d'année. Cette assiette ne tient pas compte de la participation aux bénéficies de l'année, cette dernière n'étant acquise au contrat qu'en fin d'année.

Les éléments permettant de calculer les prestations de décès sont :

- les hypothèses de lois de rachat total *Best Estimate*,
- les hypothèses de lois de décès *Best Estimate*,
- le montant de provisions mathématiques au taux annuel garanti en vigueur en milieu d'année.

### Arbitrages

La modélisation des arbitrages tient compte du sens de l'arbitrage, de la nature des supports et du fait que le support fasse ou non l'objet d'une gestion sécurisée.

La gestion sécurisée se traduit par un arbitrage automatique d'une partie des encours investis en UC vers le fonds euro. Cet arbitrage se déclenche dès que le ratio PM Euro/PM UC est inférieur à un certain seuil prédéfini contractuellement. Ces arbitrages dits automatiques tiennent compte d'une fréquence.

Pour les arbitrages ne relevant pas de la gestion sécurisée, les éléments contributeurs au calcul de la prestation sont :

- une fréquence d'arbitrage (annuelle, semestrielle, trimestrielle ou mensuelle),
- une loi d'arbitrage fonds euros vers fonds UC,
- une loi d'arbitrage fonds UC vers fonds euro,
- une loi d'arbitrage fonds UC vers fonds UC,

Ainsi, deux types de prestations sont calculés :

- Ceux au titre des arbitrages des supports Euro vers les supports UC :
  - o Ils sont assis sur la provision mathématique de fin d'année, après paiement des prestations au titre des décès, et rachats (totaux et partiels), avant attribution de la participation aux bénéfices pour les arbitrages Euro vers UC, nette de chargements sur encours (fixe ou en pourcentage).
- Ceux au titre des arbitrages des supports UC vers les supports Euro (si le contrat fait l'objet d'une sécurisation, le montant arbitré correspond au montant de l'arbitrage automatique) :
  - o Ils sont assis sur le montant de la provision mathématique UC de fin d'année après paiement des prestations au titre des décès et rachats (totaux et partiels).

Selon que le contrat fasse ou non l'objet de prélèvements sur arbitrage, la prestation nette correspond au montant brut diminué des chargements sur arbitrages. Ces derniers sont calculés en pourcentage de la prestation brute, selon le paramétrage applicable au contrat. En effet, le modèle prévoit dans le paramétrage, une table de paramètres permettant d'indiquer si les chargements sur arbitrage sont prélevés sur le mouvement sortant ou sur le support entrant.

### **Sorties en rentes des contrats multi-supports**

Les sorties obligatoires en rentes des contrats d'épargne sont réalisées en début d'année suivant le terme. Les conversions en rente dépendent d'un seuil de capital constitutif et sont effectuées à compter du terme de la capitalisation donnée en input du contrat.

Le capital constitutif correspond à la provision mathématique de fin d'année précédant la date de terme de la phase de capitalisation.

### **3.1.3 Revalorisation des contrats**

La revalorisation au titre de la participation aux bénéfices des contrats d'épargne s'effectue en fin d'année selon les modalités de participation applicable au model-point et en vigueur au cours de l'année de calcul.

Pour les contrats en Euro, les modalités possibles de calcul de participation aux bénéfices sont au nombre de 4 :

- un taux de clause de participation aux bénéfices à appliquer au taux de produits financiers bruts d'intérêts techniques,
- un taux de clause de participation aux bénéfices à appliquer au taux de produits financiers nets d'intérêts techniques,
- un taux de clause de participation aux bénéfices à appliquer à un montant de produits financiers,
- un taux de revalorisation lu dans les fichiers d'hypothèses du modèle.

Pour les contrats en unités de compte, le taux de revalorisation correspond à un taux de produit financier lu en hypothèse diminué d'un taux de coût sur placement lu en hypothèse.

### 3.1.4 Chargements et pénalités

Les primes sont sujettes à deux types de prélèvements :

- Des chargements d'acquisition fixes prélevés à la souscription : ils sont déterminés à l'aide d'un montant unitaire renseigné en paramètre qui est appliqué au nombre de contrats de début de période.
- Des chargements sur prime : calculés comme un pourcentage du montant de la prime, ils font l'objet d'hypothèses différenciées à la maille réseau\*produit selon que la prime soit unique ou périodique.

Les prestations font quant à elles l'objet de trois types de prélèvements :

- les arbitrages : ils font l'objet d'un prélèvement calculé comme un taux appliqué au montant de la prestation brute, dépendant du sens de l'arbitrage (au départ du support Euro ou du support UC)
- les rachats totaux ou partiels : ils font l'objet
  - o de frais de dossier fixes prélevés au moment du paiement de la prestation pour chaque contrat mouvementé,
  - o de pénalités de rachats calculées comme un taux de pénalité (exprimé en pourcentage fixe applicable selon l'ancienneté) appliqué aux montants de prestations brutes diminuées des frais de dossier (pour les seuls rachats totaux),

- o pour les seuls rachats totaux, de chargements fixes complémentaires en plus des pénalités de rachat au titre de la gestion des contrats ayant fait l'objet d'une prestation de rachat. Applicable par contrat, ces chargements sont répartis au prorata de l'encours de début de mois entre support Euro et UC.
- les décès : ils font l'objet de chargements fixes complémentaires au titre de la gestion des contrats ayant fait l'objet d'une prestation de décès. Ils sont calculés de la même façon que ceux prélevés au titre des rachats totaux.

### 3.1.5 Coûts et frais de gestion des placements

La modélisation prévoit deux types de coûts répartis entre acquisition et gestion. Selon le type de contrat, le modèle prévoit une modélisation de ces coûts en pourcentage de PM ou en coût unitaire (un montant par contrat).

La modélisation des frais de gestion des placements est différente que l'on soit en mode ALM (interaction Actif-Passif) ou en mode passif seul. En mode ALM, les frais de gestion des placements sont modélisés au niveau du fonds et alloués aux MPs proportionnellement aux réserves. Dans le mode Passif seul, un taux de frais de gestion des placements est lu comme une hypothèse et appliqué aux réserves.

### 3.1.6 Commissions

Les commissions désignent la rémunération versée aux intermédiaires (commerciaux, apporteurs, délégataires et banques). A chaque pas de projection, elles sont calculées en pourcentage d'une assiette et peuvent tenir compte d'une fréquence de commissionnement.

### 3.1.7 Taxes et impôts

#### Contribution sociale de solidarité de société (C3S)

Conformément à la réglementation en vigueur, la modélisation intègre un calcul de taxe au titre de la C3S. Cette dernière est calculée annuellement et assise sur :

- Pour les supports Euro et les rentes en cours de service : les produits financiers nets de coûts financiers et les primes brutes,
- Pour les supports en unités de compte : les primes brutes.

#### Fiscalité relative aux prestations

La modélisation prend en compte les prélèvements sociaux pour les contrats en euro (mono support, bon de capitalisation) au moment des sorties. Pour les supports Euro,

les calculs de prélèvements sociaux sont effectués sur les rachats partiels, les rachats totaux et les termes.

### Fiscalité relative aux encours des contrats

Seuls les supports en Euro font l'objet d'un calcul de prélèvements sociaux sur encours.

### 3.1.8 Flux considérés pour la meilleure estimation

En conclusion, le cashflow de BEL à la date  $t$  est égal à :

$$\begin{aligned} \text{cashflow de BEL} = & - \text{Primes commerciales (Primes pures + Chargements)} \\ & + \text{Prestations de décès(+PB)} \\ & + \text{Prestations de rachats(+PB)} \\ & + \text{Prestations de rentes(+PB)} \\ & + \text{Prestations de maturité(+PB)} \\ & + \text{Coût} \\ & + \text{Commissions} \end{aligned}$$

## 3.2 Principes de l'algorithme

L'approche par l'optimisation linéaire ou Linear OPTimization (LOP en anglais) est une approche basée sur l'algorithme nls (Non-Negative Least Squared). Cette approche utilise comme variable d'agrégation les cashflows de BEL.

Pour un sous-portefeuille donné, l'algorithme prend en entrée l'ensemble des cashflows de BEL projetés contrat par contrat en mode passif seul. On définit alors la matrice des BEL cashflow de sorte que chaque police soit en colonne, où  $CF_i(j)$  désigne le BEL cashflow de la police  $i$  pour la période  $j$

$$A = \begin{pmatrix} CF_1(1) & CF_2(1) & \cdots & CF_m(1) \\ CF_1(2) & CF_2(2) & \cdots & CF_m(2) \\ \vdots & \vdots & \ddots & \vdots \\ CF_1(T) & CF_2(T) & \cdots & CF_m(T) \end{pmatrix}$$

Ensuite, on calcule le vecteur des valeurs cibles qui doit être le même avant et après la compression.  $VC(j) = \sum_i CF_i(j)$  représente la valeur totale de cashflow de toutes les polices à la période  $j$ .

$$b = \begin{pmatrix} VC(1) \\ VC(2) \\ \vdots \\ VC(T) \end{pmatrix} = A * \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Avec :

- $m$  : le nombre de polices initialement en portefeuille ;
- $T$  : l'horizon de projection, ici 60 ans.

L'approche du LOP consiste donc à trouver le vecteur des poids (positifs)  $x$  qui résolve l'équation suivante :

$$x_1 * \begin{bmatrix} CF_1(1) \\ CF_1(2) \\ \vdots \\ CF_1(T) \end{bmatrix} + x_2 * \begin{bmatrix} CF_2(1) \\ CF_2(2) \\ \vdots \\ CF_2(T) \end{bmatrix} + \dots + x_m * \begin{bmatrix} CF_m(1) \\ CF_m(2) \\ \vdots \\ CF_m(T) \end{bmatrix} = \begin{bmatrix} VC(1) \\ VC(2) \\ \vdots \\ VC(T) \end{bmatrix}$$

ou de façon équivalente le vecteur  $x$  solution du programme d'optimisation suivant :

$$\min_x \frac{1}{2} \|b - Ax\|_2^2 \quad ; \quad x \geq 0 \quad (3.1)$$

Ici,  $x \geq 0$  signifie que chaque composante du vecteur  $x$  devrait être positive, et  $\|\cdot\|_2$  représente la norme Euclidienne.

Le gradient de la fonction objectif du programme 3.1 s'écrit  $w = A^T(b - Ax)$  et représente le vecteur des corrélations entre chaque police et le vecteur  $b - Ax$ . L'algorithme itératif NNLS utilisé pour résoudre ce programme repose sur la méthode "active-set" publiée par Lawson and Hanson dans leur ouvrage *Solving Least Squares Problems* paru en 1974.

Cette méthode consiste à définir dans un premier temps deux ensembles d'indices nommés respectivement active-set et passif-set.

Les inputs de l'algorithme sont :

- La matrice  $A$  des cashflows de dimension  $T \times m$  ;
- Le vecteur  $b$  des valeurs cibles de dimension  $T$  ;

L'algorithme est initialisé de la façon suivante :

L'active-set est composé initialement des indices de toutes les polices, c'est à dire  $R = \{1, 2, \dots, m\}$ , tandis que le passif-set reste vide :  $P = \emptyset$ . Le vecteur des poids  $x$  de dimension  $m$  est initialisé au vecteur nul. Et enfin on note  $w^R$  le sous-vecteur extrait de  $w$  avec les indices de l'active-set  $R$ .

La boucle principale de l'algorithme comprend deux phases : une phase sélective et une phase de mise à jour.

- Etape 1 : Sélectionner parmi les cash-flows de l'active-set celui ou ceux qui possèdent la corrélation positive la plus élevée avec le vecteur cible :  $\max(w^R)$
- Etape 2 : Les cashflows sélectionnés à l'étape 1 sont déplacés de l'actif-set vers le passif-set. Puis les poids de tous les cashflows du passif-set sont mis à jour via une régression linéaire entre le vecteur des valeurs cibles et les cash-flows du passif-set.
- Etape 3 : Calculer le nouveau gradient  $w$  et extraire le nouveau  $w^R$ . Si  $R \neq \emptyset$  et  $\max(w^R) > 0$  aller à l'étape 1.

Un exemple illustratif est proposé en Annexe B pour une meilleure compréhension de l'algorithme.

L'output principal de l'algorithme du LOP est le vecteur des poids  $x$  aussi appelé `rescale_x`. Ainsi, la  $i$ ème composante de  $x$  représente le poids accordé à la  $i$ ème police à l'issue de l'algorithme.

L'algorithme LOP n'agrège pas les polices. La construction des Model Points se fait d'abord en supprimant les polices avec des poids nuls puis en remplaçant chacune des autres polices par une autre police avec les mêmes caractéristiques d'âge, de sexe et d'ancienneté, mais avec des provisions mathématiques, primes, etc. rescalées par son poids dans le vecteur  $x$ . Et ce dans le but de minimiser la différence entre les valeurs cibles calculées police par police et ces mêmes valeurs cibles recalculées après rescale.

### 3.3 Résultats

Avant de présenter les résultats, il est important de noter que nous ne sommes pas responsable de cette approche. L'approche du LOP est une solution proposée par un cabinet externe. Elle est intégrée au logiciel de projection livré à Allianz France et codée intégralement en langage C++. Toutefois, un pan de notre objectif était de rendre cet algorithme accessible au personnel d'Allianz et de proposer une amélioration à cet algorithme afin d'améliorer les résultats. C'est ce que nous avons fait et que nous détaillerons plus loin.

## Analyse de la compression

L'approche par l'optimisation linéaire permet de passer d'un portefeuille de 33 482 lignes à 14 993 lignes soit un taux de compression de 55%. Elle permet donc de réduire de moitié la taille du portefeuille pour une erreur de l'ordre de 36 millions d'euros. Même si la précision est assez bonne, le nombre de lignes reste encore trop élevé pour que l'on puisse observer un changement conséquent dans le temps de calcul lors du passage aux modèles déterministe et stochastique.

Comme on peut le constater dans le tableau 3.1, près de 3 sous-portefeuilles sur 4 (74%) n'ont pas été compressés par l'algorithme. C'est à dire que tous les contrats de ces sous-portefeuilles ont été conservés et devraient être considérés comme MPs pour le modèle.

TABLE 3.1 – Répartition du taux de compression par sous-portefeuille

Compression	0%	0-10%	10-30%	30-50%	50-70%	70-100%
Proportion	73.99	7.08	5.64	5.49	4.34	3.46

Ces résultats sont ceux d'Allianz France avant notre intervention au sein de la compagnie. Le problème auquel nous avons été confronté fut dans un premier temps de trouver un moyen d'améliorer la capacité de compression du LOP. Et dans un second temps de trouver de meilleures alternatives à cet algorithme.

## 3.4 Seuil de tolérance

Notre contribution à l'approche de l'optimisation linéaire a été d'intégrer à l'algorithme nnls une valeur réelle positive  $\epsilon$  qui représente la tolérance pour le critère d'arrêt. De ce fait, le critère d'arrêt de l'algorithme passe de  $\max(w^R) > 0$  à  $\max(w^R) > \epsilon$ . Il est important de remarquer que la précision apportée par le ou les nouveaux cash-flows sélectionnés est décroissante au fil des itérations. En d'autres termes, un cash-flow sélectionné lors d'une itération donnée, apporte une contribution plus élevée dans le processus d'approximation de la valeur cible que tous les autres cash-flows qui seront sélectionnés dans les itérations suivantes. Ainsi, la quantité  $\epsilon$  peut être vue comme la contribution minimale souhaitée pour un cash-flow sélectionné à une itération pour que cette itération soit considérée comme pertinente. Sinon l'itération est considérée comme trop coûteuse et doit être abandonnée.

Le tableau ci-dessous récapitule les résultats obtenus selon différentes valeurs du paramètre  $\epsilon$  en première colonne. La deuxième colonne « Nb lignes » renseigne sur l'évolution du nombre de model points. Par exemple pour  $\epsilon = 1e-06$ , le portefeuille

passé de 33 482 lignes à 9 361 soit un taux de compression de 72%.

La colonne « Ecart en euro » représentent les écarts absolus (différence de BEL avant et après compression) cumulés sur tous les sous-portefeuilles. Par exemple, toujours pour  $\epsilon = 1e-06$ , le montant total des engagements provenant de l'ensemble des placements en fonds Euro comme en UC cumulé sur tous les sous-portefeuilles représente un peu plus de 62 milliards et demi d'euros, et le portefeuille compressé s'écarte de ce montant de 37 millions d'euros. En terme relatif, cela signifie que pour 10 000 euros d'engagements, la compression entraîne une perte de 6 euros.

TABLE 3.2 – Répartition du taux de compression par groupe

epsilon	Nb lignes	Compression en %	Ecart en euro	Ecart relatif <sup>a</sup>
0	14 993	55	35 449 810	6
1e-06	9 361	72	37 002 732	6
1e-05	7 708	77	46 446 097	7
1e-04	6 092	82	53 953 300	9
1e-03	4 563	86	124 609 384	20
0.01	3 363	90	212 440 438	34
0.1	2 482	93	930 829 986	149
<b>Portefeuille</b>	<b>33 482</b>		<b>62 640 384 409</b>	

<sup>a</sup> Les erreurs relatifs sont exprimées en pour 10 000 du BEL total

Le graphique suivant montre comment évolue la somme des écarts absolus en fonction du taux de compression. On constate une progression exponentielle de la somme des écarts absolus au-delà d'une compression de 80%.

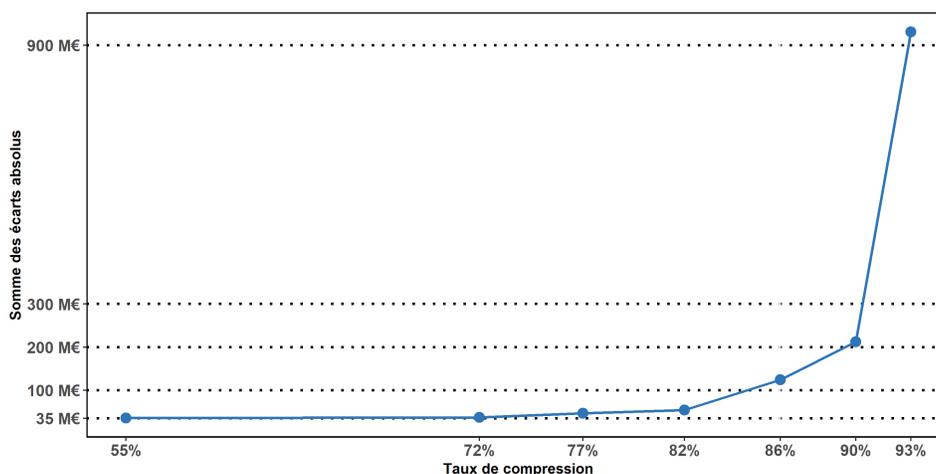


FIGURE 3.1 – Somme des écarts absolus en fonction du Taux de compression

La compression la plus proche de notre compression cible (87%) est obtenue avec  $\epsilon = 1e-03$ . Pour cette valeur du paramètre de tolérance, on observe un nombre de Model Points de 4 563 avec une somme des écarts absolus d'un peu plus de 124 M€.

### 3.5 Critiques de l'approche linéaire

Le LOP possède un avantage majeur qui est de ne pas nécessiter de phase d'agrégation. Les MPs sont directement créés en sélectionnant un sous-ensemble de polices puis en effectuant quelques ajustements (scaling) sur certaines variables, notamment les variables monétaires comme la PM, les primes, etc. En plus c'est un algorithme très rapide à mettre en œuvre et dans son exécution.

Cependant, plusieurs défauts tant théoriques que pratiques viennent limiter l'intérêt de cette approche. Même s'il existe la possibilité d'intégrer un paramètre pour le contrôler, le nombre de MPs final reste inhérent au programme donc implicite : il ne peut être que constaté post-exécution. En outre, comparativement à d'autres approches, le rapport précision/compression est relativement faible (*voir tab. 3.2*). Aussi, des fuites considérables de PM ou de primes peuvent être observées : le montant total de PM ou de prime du portefeuille n'est pas conservé après agrégation. Enfin, le LOP utilise comme variable d'agrégation les cash-flows de BEL qui dépendent du scénario financier. Les MPs ainsi constitués ne sont valables que pour le scénario à partir duquel ils ont été constitués.

# Chapitre 4

## Approche par Clustering

### Sommaire

---

<b>4.1 Probabilités de sortie</b>	<b>39</b>
4.1.1 Motivation et formulation mathématique	39
4.1.2 Calcul des probabilités de sortie	40
<b>4.2 Phase de classification du portefeuille</b>	<b>42</b>
4.2.1 Description du problème de classification	43
4.2.2 Algorithme du kmeans	44
4.2.3 Métriques de base des séries temporelles	45
4.2.4 Initialisation de l'algorithme	47
4.2.5 Evaluation de la classification	49
<b>4.3 Phase d'agrégation du portefeuille</b>	<b>51</b>
4.3.1 Moyenne pondérée par la provision mathématique de l'âge et de l'ancienneté	51
4.3.2 Méthode par développement limité en série de Taylor	53

---



Dans cette partie nous présentons notre méthode d'agrégation du portefeuille de contrats d'assurance vie. Il s'agit d'une procédure en deux étapes. La première étape consiste à utiliser des méthodes statistiques de classification afin de créer des groupes homogènes de contrats. La seconde étape permet la définition d'un contrat représentatif ou "moyen" pour chacune des classes.

## 4.1 Probabilités de sortie

### 4.1.1 Motivation et formulation mathématique

L'approche par optimisation linéaire utilise les chroniques de cash-flow futures comme variable d'agrégation. Cependant, comme cela peut être constaté dans les équations 1.2 et 1.3, ces cash-flows dépendent du scénario financier considéré. Cette dépendance au scénario financier pourrait s'avérer problématique notamment dans un cadre de modélisation stochastique où plusieurs scénarios différents sont considérés. Il est donc indispensable de trouver une nouvelle variable d'agrégation qui décrive à la fois la dynamique future des contrats et qui ne soit pas tributaire du scénario considéré.

Dans le cadre d'un contrat épargne individuelle, trois événements peuvent engendrer l'arrêt du contrat,

- L'arrivée à échéance du contrat,
- Le décès de l'assuré,
- Le rachat du contrat, qui correspond à une décision de la part de l'assuré de solder son compte.

Le rachat et le décès de l'assuré sont des événements probabilisables et sont susceptibles de se réaliser à tout moment pendant la durée de vie du contrat.

Dans les expressions 1.2 et 1.3 de la provision Best Estimate, nous avons fait l'hypothèse implicite que le contrat arrivait nécessairement à terme, c'est à dire jusqu'à l'horizon de projection  $T$ . Dans la pratique, cette hypothèse n'est pas réaliste. Par exemple on ne peut supposer qu'un contrat avec un assuré âgé de 90 ans atteindra l'horizon de projection  $T = 60$  ans. Pour prendre en compte ces événements aléatoire sur la durée de vie des contrats, il est nécessaire de réécrire le BEL.

Pour cela, en se basant sur les notations de 1.3 on note :

$$PM_t^{\mathbf{F}} = \sum_{j=0}^t Flux_j^{\mathbf{F}} \times \delta_j^{\mathbf{F}}$$

la provision mathématique du contrat au cours de l'année de projection  $t$ , c'est à dire la valeur actuelle de l'épargne accumulée jusqu' à l'année  $t$ . Il s'agit du versement initial augmenté des éventuelles primes périodiques, capitalisé au taux de rendement des supports de comptes sur lesquelles l'épargne est investie, actualisé au taux Zéro-Coupon.

Avec cette nouvelle quantité, il est possible de réécrire le BEL d'un contrat comme suit :

$$BEL^{\mathbf{F}} = \sum_{t=0}^{T-1} PM_t^{\mathbf{F}} \times p_t + PM_T^{\mathbf{F}} \times p_T \quad (4.1)$$

où  $p_t$  est la probabilité de sortie associée au contrat entre l'année  $t$  et  $t + 1$ , c'est-à-dire la probabilité jointe de ne pas être sorti au cours des années précédentes et de sortir durant l'année  $t$ . Contrairement aux cash-flows de BEL utilisé par le LOP, il s'agit d'une caractéristique du contrat qui ne dépend pas du scénario financier (hypothèses économiques) mais des hypothèses techniques (*voir 1.4*) donc plus robuste dans une modélisation stochastique. Plus précisément, elle dépend de la probabilité de décès et de la probabilité de rachat qui sont respectivement fonction de l'âge de l'assuré et de l'ancienneté du contrat.

$PM_T$  et  $p_T$  correspondent respectivement à la valeur de l'épargne à l'échéance du contrat et à la probabilité d'arriver effectivement à l'échéance du contrat.

Il est aisé de vérifier que pour les contrats certains d'arriver à terme, c'est à dire  $p_T = 1$  et  $p_t = 0$  pour tout  $t \in \{0, 1, \dots, T - 1\}$ , le BEL devient :

$$BEL^{\mathbf{F}} = PM_T^{\mathbf{F}} = \sum_{j=0}^t Flux_j^{\mathbf{F}} \times \delta_j^{\mathbf{F}}$$

On retrouve bien l'expression de l'équation 1.3.

### 4.1.2 Calcul des probabilités de sortie

Pour calculer la probabilité de sortie d'un contrat au cours de l'année  $t$ , il est nécessaire de définir certaines quantités au préalable.

- **La probabilité de décès  $d_t$**

$d_t$  est la probabilité de décès de l'assuré au cours de l'année de projection  $t$ . Elle est fonction de l'âge de l'assuré au cours de cette même année et est calculée à l'aide des tables de mortalité fournies avec les hypothèses techniques.

- **La probabilité de rachat  $r_t$**

$r_t$  est la probabilité que l'assuré rachète son contrat durant l'année de projection  $t$ . Elle est fonction de l'ancienneté du contrat au cours de l'année de projection  $t$  et est calculée par le biais des lois de rachat structurel mises au point dans les hypothèses techniques.

• **La probabilité de survie cumulée  $q_t$**

$q_t$  est la probabilité de survie cumulée jusqu'à l'année  $t$ , de la police au sein du portefeuille de contrat. Elle s'exprime en fonction des deux probabilités définies précédemment. En effet, soit  $T$  l'échéance du contrat ou l'horizon de projection. Soient également  $\tau_d$  et  $\tau_r$  respectivement les instants de décès de l'assuré et de rachat du contrat. Alors le véritable instant de sortie est  $\tau = \tau_d \wedge \tau_r \wedge T = \min(\tau_d, \tau_r, T)$ .

Pour toute date de sortie anticipée  $t < T$  (avant l'arrivée à échéance du contrat).

$$q_t = \mathbb{P}(\tau > t - 1) = \mathbb{P}(\min(\tau_d, \tau_r, T) > t - 1)$$

$$q_t = \mathbb{P}(\tau_d > t - 1, \tau_r > t - 1, T > t - 1)$$

En supposant une indépendance des instants de décès de l'assuré et de rachat du contrat :

$$q_t = \mathbb{P}(\tau_d > t - 1) \times \mathbb{P}(\tau_r > t - 1) \times \mathbb{P}(T > t - 1)$$

$$q_t = \mathbb{P}(\tau_d > t - 1) \times \mathbb{P}(\tau_r > t - 1) \quad \text{car} \quad \mathbb{P}(T > t - 1) = 1$$

$$q_t = \prod_{j=0}^{t-1} \mathbb{P}(\tau_d > j + 1 | \tau_d > j) \mathbb{P}(\tau_r > j + 1 | \tau_r > j)$$

Ainsi :

$$q_t = \prod_{j=0}^{t-1} (1 - d_j)(1 - r_j)$$

La probabilité de sortie du contrat au cours de l'année  $t$  s'écrit alors,

$$p_t = \mathbb{P}(\tau > t - 1, \tau \leq t) = q_t - q_{t+1}$$

or :

$$q_{t+1} = q_t \times [(1 - d_t)(1 - r_t)]$$

donc :

$$p_t = q_t \times [1 - (1 - d_t)(1 - r_t)]$$

Finalelement :

$$p_t = q_t \times \{d_t + (1 - d_t)r_t\}$$

Enfin, la probabilité que le contrat parvienne à son terme est égale à :

$$p_T = 1 - \sum_{t=0}^{T-1} p_t$$

Dans la pratique, notamment lors de l'écriture du code informatique pour le calcul des probabilités de sortie, les quantités présentes dans l'expression ne sont pas immédiatement accessibles. Celles-ci sont lues directement par le logiciel dans les tables fournies. Les probabilités de sortie sont donc estimées à partir de l'expression suivante :

$$p_t = \frac{N_{t-1}}{N_0} \left( 1 - \frac{N_t}{N_{t-1}} \right) = \frac{N_{t-1} - N_t}{N_0}$$

où :

- $N_t$  désigne le nombre de contrats encore en cours de vie à la date  $t$
- $N_0$  désigne le nombre de contrats initial de chaque contrat issu de la première phase d'agrégation

Dans le cadre des contrats d'assurance vie de type épargne, les contrats ayant les mêmes probabilités de sortie sont simplement associés à des assurés de même caractéristiques biométriques et ayant souscrit lors de la même année. Cette conception peut être étendue aux contrats ayant des probabilités de sortie "proches" en considérant que ces contrats auront aussi des caractéristiques assez similaires. L'idée est d'utiliser les méthodes statistiques de classification afin de regrouper les contrats ayant des probabilités de sortie proches au cours de la projection. Ces probabilités de sortie forment des trajectoires qui permettent une visualisation graphique du portefeuille de contrats.

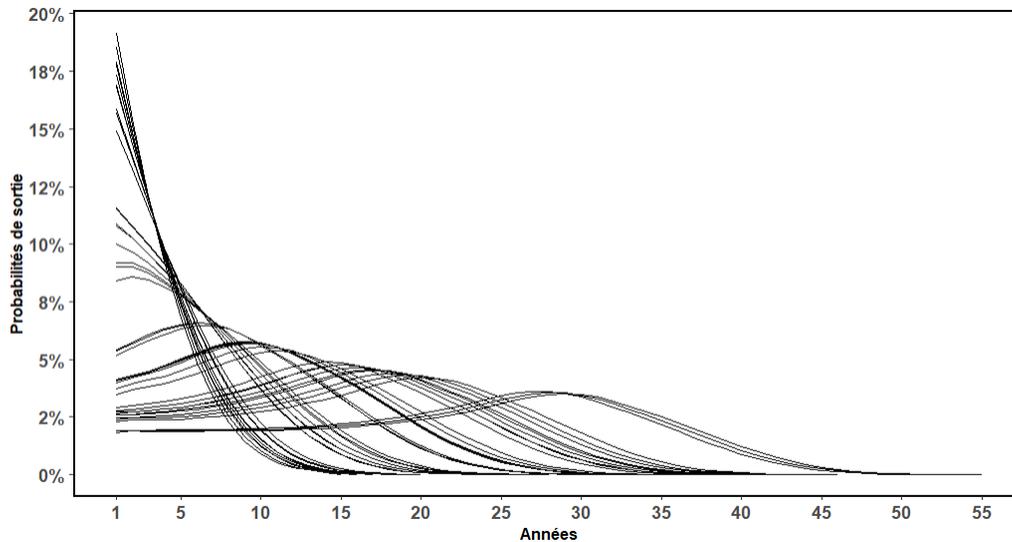


FIGURE 4.1 – Probabilités de sortie associées à un sous-portefeuille

## 4.2 Phase de classification du portefeuille

Le but de la phase de classification est d'affecter à chaque ligne du portefeuille une classe d'appartenance. Tout comme le LOP, il est nécessaire de conserver les classes

d'agrégation. L'idée est de classifier séparément les sous-portefeuilles en regroupant les contrats en fonction de leurs vecteurs de probabilité de sortie.

### 4.2.1 Description du problème de classification

Soit un sous-portefeuille contenant  $n$  contrats. Les variables actives de la méthode de classification seront les composantes du vecteur de probabilités de sortie. La probabilité de sortie à l'échéance du contrat est exclue des variables actives car elle est obtenue par combinaison linéaire des autres variables.

Le problème de classification est analogue à un problème de classification vu dans la littérature : **la classification de séries temporelles**. Une série temporelle, ou série chronologique, est une suite de valeurs numériques représentant l'évolution d'une quantité spécifique au cours du temps.

Différentes méthodes de classification de séries temporelles existent. L'une d'entre elle, dite *Model based* consiste à appliquer un Modèle Semi-Markovien Caché (MSMC) pour extraire de chaque série deux modèles, un premier caractérisant le cas positif et un second modèle de référence caractérisant le cas négatif. Cela permet de calculer les log-vraisemblances associées aux deux modèles puis celles-ci sont comparées via un test pour classifier la série, voir (J. Dumont et al. 2009). Même si cette technique s'est avérée extrêmement efficace<sup>1</sup>, elle n'est applicable que dans les problèmes de classification binaire. D'autres techniques adaptées aux problèmes de classification multi-classes existent cependant. C'est par exemple le cas du clustering par construction de dictionnaire, voir (E. Goffinet et al. 2020). Il s'agit d'une approche à trois étapes : la segmentation par modèle de mélange de chaque série temporelle dans le but de déterminer les points de changement de régimes, la construction d'un dictionnaire de régimes communs et le clustering des séquences de régimes ainsi produites. Si la méthode possède de nombreux avantages comme l'indépendance à la différence de longueur, elle a l'inconvénient de n'être pertinente que pour les séries qui peuvent alterner différentes phases de fonctionnement distinctes que l'on appelle des régimes : par exemple, l'observation de la vitesse d'une voiture qui peut montrer des phases d'accélération, une vitesse de croisière, des phases de freinage, etc. Une approche similaire à celle par dictionnaire, mais ne nécessitant pas d'hypothèse de construction par régime des séries fut explorée par (Till Wohlfarth, 2016) quelques années plutôt. Celle-ci est également basée sur le mélange de modèles en appliquant l'algorithme Espérance-Maximisation (EM). Elle modélise chaque classe par une distribution de probabilité et groupe les séries afin de maximiser la vraisemblance aux centroïdes. Mais face à la complexité de cette approche, Wohlfarth teste un autre algorithme de partitionnement classique et relativement plus simple à savoir le kmeans avec comme mesure de distance entre deux séries la distance euclidienne. C'est une approche que l'on peut aussi voir chez (Pierre-O. Goffard, 2015).

---

1. Les résultats obtenus montrent un très bon de taux de détection (sensibilité de 96% et spécificité de 80%).

Le kmeans est une approche très répandue, assez intuitive, généraliste et plus aisée en terme de programmation. Au regard des outils dont nous disposons et dans un souci d'automatisation du processus de compression, nous optons pour cette approche.

Intéressés par les problèmes que pose l'utilisation d'algorithmes d'apprentissage non supervisé de type kmeans sur un jeu de séries temporelles, (Rémi G. et Nicolas N. 2006) mettent en évidence le problème lié à l'utilisation de la distance euclidienne comme mesure de similarité. En effet, pour ces derniers, parmi les principales mesures de similarité existantes l'une d'entre elles (appelée Dynamic Time Warping) paraît la plus adaptée à l'apprentissage non supervisé. Ils vont encore plus loin en proposant une méthode originale d'agrégation de séries temporelles, compatible avec le Dynamic Time Warping, qui améliore ainsi les résultats obtenus à l'aide de l'algorithme des kmeans.

## 4.2.2 Algorithme du kmeans

Étant donné des points et un entier  $k$ , le problème est de diviser les points en  $k$  groupes, souvent appelés clusters, de façon à minimiser une certaine fonction. Si l'on munit l'espace des points d'une métrique, et on considère la distance d'un point à la moyenne des points de son cluster ; la fonction à minimiser est la somme des carrés de ces distances.

En formulation mathématique, étant donné un ensemble de points  $(x_1, x_2, \dots, x_n)$ , on cherche à partitionner les  $n$  points en  $k$  ensembles  $S = \{S_1, S_2, \dots, S_k\}$  ( $k \leq n$ ) en minimisant la distance entre les points à l'intérieur de chaque partition :

$$\arg \min_S \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (4.2)$$

où  $\mu_i$  est le barycentre des points dans  $S_i$ .

L'algorithme du kmeans a été proposé comme une solution à ce problème (4.2). Il s'agit d'un algorithme classique très utilisé en pratique et considéré comme efficace bien que ne garantissant ni l'optimalité, ni un temps de calcul polynomial. L'algorithme est le suivant :

1. Choisir la valeur de  $k$ .
2. Initialiser les centres des  $k$  clusters (aléatoirement si nécessaire).
3. Affecter chaque série au cluster dont le centre est le plus proche.
4. Ré-estimer les centres des  $k$  clusters, en supposant que toutes les affectations des séries sont correctes.

5. Si aucune des séries n'a changé de cluster, alors fin de l'algorithme. Sinon, retour à l'étape 3.

### 4.2.3 Métriques de base des séries temporelles

L'algorithme de kmeans nécessite la définition d'une distance entre les points. Dans le cas particulier des séries temporelles, il est indispensable de définir une métrique appropriée. Dans le cas de nos travaux, nous avons testé les métriques les plus connues que l'on peut regrouper en deux catégories : les métriques basées sur la valeur (*value-based*) et les métriques basées sur le comportement (*behavior-based*).

- **Métriques basées sur la valeur**

**La distance Euclidienne** - est la fonction de distance la plus utilisée. Elle est généralement considérée comme la mesure de distance la plus simple entre les séries. La distance Euclidienne  $d_E$  (ou la norme  $L_2$ ) entre deux séries temporelles  $\mathbf{x}_i = (x_{i1}, \dots, x_{iT})$  et  $\mathbf{x}_j = (x_{j1}, \dots, x_{jT})$  de longueur  $T$  est définie comme :

$$d_E(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{\sum_{t=1}^T (x_{it} - x_{jt})^2}$$

**La distance de Minkowski** - La généralisation de la distance Euclidienne est la distance de Minkowski, aussi appelée la norme  $L_p$ . Elle est définie comme :

$$d_{L_p}(\mathbf{x}_i, \mathbf{x}_j) = \sqrt[p]{\sum_{t=1}^T |x_{it} - x_{jt}|^p}$$

Nous testerons uniquement le cas  $p = 1$  qui correspond à la distance de Manhattan.

**Dynamic time warping (DTW)** - La particularité du DTW est de savoir gérer les décalages temporels qui peuvent éventuellement exister entre deux séries (Berndt et Clifford 1994). Au lieu de comparer chaque point d'une série avec celui de l'autre série qui intervient au même instant  $t$  comme le fait les mesures telle que la  $p$ -normée (Minkowski), on permet à la mesure de comparer chaque point d'une série avec un ou plusieurs points de l'autre série, ceux-ci pouvant être décalés dans le temps (fig. 4.2)<sup>2</sup>.

2. Cette illustration est extraite de (Rémi G. et Nicolas N. 2006)

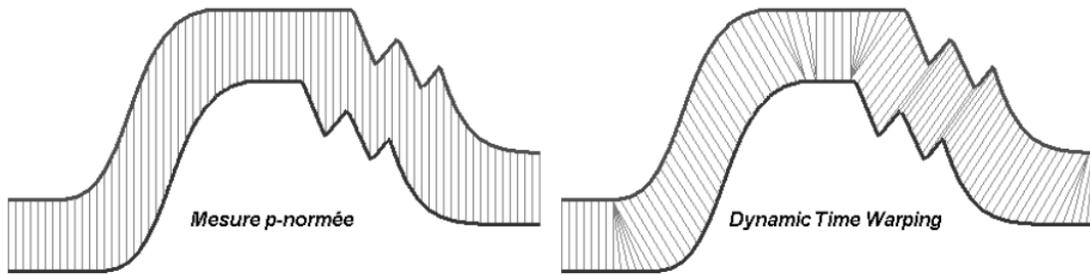


FIGURE 4.2 – Comparaison entre la mesure  $p$ -normée et le DTW

Mathématiquement, la dissimilarité entre les séries  $Q = q_1, q_2, \dots, q_m$  et  $C = c_1, c_2, \dots, c_n$  se calcule de la façon suivante :

Soit  $D(i, j)$  la distance entre les sous-séquences  $q_1, q_2, \dots, q_i$  et  $c_1, c_2, \dots, c_j$  (avec  $1 \leq i \leq m$  et  $1 \leq j \leq n$ ) :

$$D(i, j) = \begin{cases} |q_1 - c_1|, & \text{si } i = j = 1; \\ |q_i - c_j| + \min\{D(i-1, j), D(i-1, j-1), D(i, j-1)\}, & \text{sinon.} \end{cases}$$

La distance finale entre  $Q$  et  $C$  est donnée par  $D(n, m)$

**La distance Dissim** - Elle a été introduite par (Frentzos, Gratsias et Theodoridis 2007) et est spécialement conçue pour les séries collectées à différents taux d'échantillonnage. Cela signifie que chaque série sera définie dans un ensemble fini d'instantanés de temps, mais ceux-ci peuvent être différents pour chaque série. La distance Dissim nécessite une représentation continue des séries et ainsi, les séries comparées sont supposées être linéaires entre les points d'échantillonnage. Une fois cela fait, l'intégrale définie de la distance euclidienne qui les sépare est calculée :

$$Dissim(X, Y) = \sum_{i=0}^{K-1} \int_{t_i}^{t_{i+1}} D_{X,Y}(t) dt \tag{4.3}$$

où,  $T = \{t_0, \dots, t_{K-1}\}$  est un indice temporel global qui fusionne les indices temporels des deux séries en prenant tous les points qui apparaissent dans les deux ensembles.  $D_{X,Y}(t)$  représente la distance euclidienne entre les séries au pas de temps  $t$ . Enfin, afin d'éviter les temps de calcul conséquent de l'intégrale, Frentzos et al. (2007) présentent une approximation de Dissim qui simplifie l'équation 4.3 en utilisant la règle du trapèze. De cette manière, la formule plus simple suivante est obtenue :

$$Dissim\_approx(X, Y) = \sum_{i=0}^{N-1} (D_{X,Y}(t_i) + D_{X,Y}(t_{i+1})) \cdot (t_{i+1} - t_i)$$

C'est cette approximation que nous allons calculer dans ce mémoire.

- **Métriques basées sur le comportement**

Deux séries temporelles  $\mathbf{x}_i$  et  $\mathbf{x}_j$  ont des **comportements similaires**, si pour chaque période  $[t_i, t_{i+1}]$ ,  $i \in \{1, \dots, T\}$ , elles croissent et décroissent simultanément avec le même taux de croissance.

Deux séries temporelles  $\mathbf{x}_i$  et  $\mathbf{x}_j$  ont des **comportements opposés**, si pour chaque période  $[t_i, t_{i+1}]$ ,  $i \in \{1, \dots, T\}$ , quand une série croit, l'autre décroît avec le même taux de croissance en valeur absolue et vice-versa.

Deux séries temporelles  $\mathbf{x}_i$  et  $\mathbf{x}_j$  ont des **comportements différents**, si leurs comportements ne sont ni similaires, ni opposés.

**Le coefficient de corrélation temporelle (CORT)** - est une variante du coefficient de corrélation de Pearson prenant en compte une dépendance temporelle avec une fenêtre  $r \in \{1, \dots, T\}$ . Le coefficient de corrélation temporelle entre deux séries temporelles  $\mathbf{x}_i = (x_{i1}, \dots, x_{iT})$  et  $\mathbf{x}_j = (x_{j1}, \dots, x_{jT})$  est définie comme :

$$CORT(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{t,t'} m_{tt'} (x_{it} - x_{it'}) (x_{jt} - x_{jt'})}{\sqrt{\sum_{t,t'} m_{tt'} (x_{it} - x_{it'})^2} \sqrt{\sum_{t,t'} m_{tt'} (x_{jt} - x_{jt'})^2}}$$

où  $m_{tt'} = 1$  si  $|t - t'| \leq r$ , sinon 0. CORT appartient à l'intervalle  $[-1, +1]$ . La valeur  $CORT(\mathbf{x}_i, \mathbf{x}_j) = 1$  signifie qu'à chaque période observée  $[t, t']$ ,  $\mathbf{x}_i$  et  $\mathbf{x}_j$  ont des comportements similaires. La valeur  $CORT(\mathbf{x}_i, \mathbf{x}_j) = -1$  signifie qu'à chaque période observée  $[t, t']$ ,  $\mathbf{x}_i$  et  $\mathbf{x}_j$  ont des comportements opposés. Enfin, la valeur  $CORT(\mathbf{x}_i, \mathbf{x}_j) = 0$  signifie que les deux séries sont linéairement indépendante. Le CORT n'est pas au sens stricte du terme une distance. Pour les besoins du kmeans, nous avons réalisé un ajustement du coefficient de corrélation temporelle par la transformation suivante :  $x \mapsto \exp(-x)$ . Cette transformation permet d'obtenir des valeurs toujours positives et attribue les valeurs les plus faibles aux séries aux comportements similaires et les valeurs les plus élevées aux séries aux comportements opposés.

#### 4.2.4 Initialisation de l'algorithme

Une étape importante de l'algorithme de kmeans est l'initialisation. Une mauvaise initialisation peut avoir des répercussions importantes sur les résultats notamment si le nombre d'itération n'est pas assez élevé. Plusieurs méthodes existent pour initialiser les centroïdes.

**First pols / Last pols** - Cette méthode consiste à prendre les valeurs des variables actives des  $k$  premières polices différentes pour le first pols ou alternativement pour les

$k$  dernières polices différentes pour le last pols.

**Simple Cluster Seeking (SCS)** - Cette méthode utilise une distance minimale  $dmin = \frac{1}{k}$ . Les  $k$  centroïdes initiaux sont choisis de la manière suivante : (i) Ajouter la première police de la base. (ii) Parcourir toutes les polices en boucle et ajouter chacune à l'ensemble si sa distance par rapport à toutes les polices précédemment ajoutées est supérieure à  $dmin$ . Il faut noter que la distance minimale doit être diminuée s'il n'y a pas  $k$  polices différentes qui sont séparées deux à deux par des distances dépassant  $dmin$ .

**Random pols** - Elle consiste à choisir aléatoirement les centroïdes initiaux parmi les polices.

**Random pols fixed** - Il s'agit d'une méthode similaire à la méthode Random pols et consiste à choisir les centroïdes initiaux de façon pseudo-aléatoire parmi les polices disponibles. Le pseudo-aléatoire signifie que le choix aléatoire peut être reproduit si les exécutions sont effectuées avec exactement les mêmes hypothèses. Les données doivent être placées dans le même ordre.

**Read from file** - Cette méthode consiste à lire les centroïdes initiaux à partir d'un fichier. Par exemple, une fois le processus de clustering calibré, les centroïdes finaux de l'année  $t$  peuvent être utilisés comme centroïdes initiaux pour l'année  $t + 1$ .

**Max distance** - Sélectionner la première police, puis ajouter de manière itérative les polices pour lesquelles la distance par rapport à toutes les polices précédemment sélectionnées est maximisée.

Dans le cadre de notre étude, nous avons d'abord commencé par une initialisation « Random pols ». Le problème lié à cette méthode est l'instabilité des résultats qui peuvent être différents entre deux exécutions. Les méthodes « SCS » et « Max distance » nécessitent des calculs de distances qui auraient pour conséquence d'augmenter le temps d'exécution de notre algorithme. Nous avons finalement opté pour une variante du « first pols » en maximisant l'hétérogénéité de nos centroïdes selon une variable importante de nos données. La variable choisie est l'âge. Cette méthode a consisté à subdiviser l'intervalle des âges en  $k$  parts égales en définissant :

$$b_0 = \min(age) \quad ; \quad b_k = \max(age)$$

puis

$$b_j = b_0 + j * \frac{b_k - b_0}{k} \quad \forall j \in \{1, \dots, k - 1\}$$

et à choisir dans chaque intervalle  $[b_j, b_{j+1}[$ , la première police de la base dont l'âge se situe dans l'intervalle. Il s'agit d'une méthode combinant à la fois les avantages d'un « first pols » pour sa rapidité et d'un « Random pols fixed » pour la stabilité des résultats.

Tous ces éléments évoqués ci-dessus permettent de mettre en œuvre le kmeans. Le graphique 4.3 illustre la classification effectuée avec l'algorithme du kmeans sur les probabilités de sortie du même sous-portefeuille que la figure 4.1. Il s'agit d'une classification avec  $k = 5$  et Manhattan comme mesure de distance utilisée.

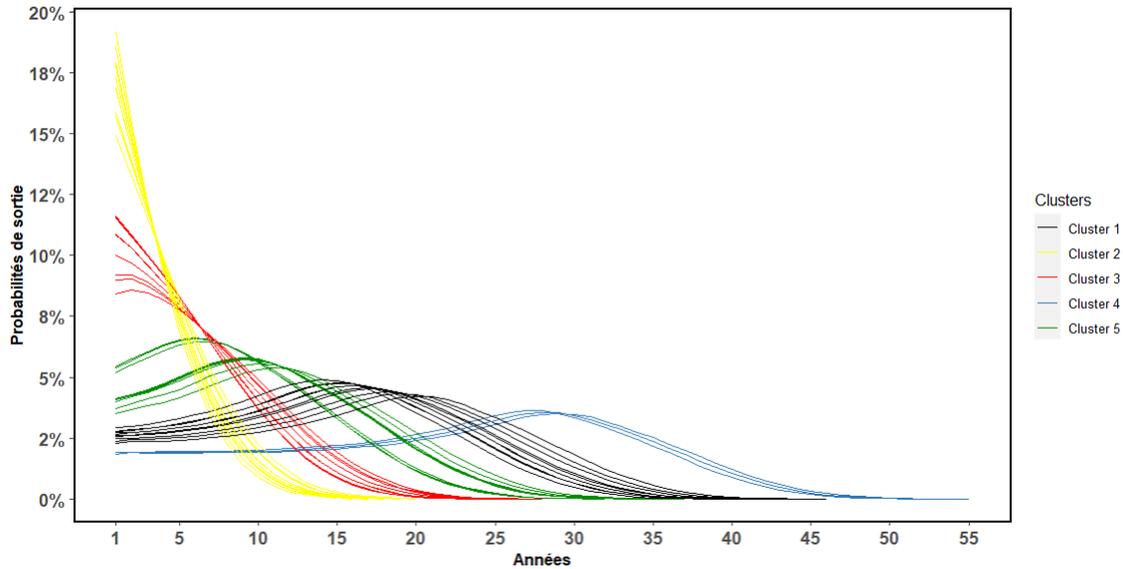


FIGURE 4.3 – Exemple d'application du kmeans sur les prob. de sortie pour  $k = 5$

### 4.2.5 Evaluation de la classification

Pour évaluer la qualité des clusters obtenus par un algorithme de partitionnement tel que le kmeans, il est possible de s'appuyer sur le souhait formulé de regrouper entre elles les observations similaires. Ainsi, les observations appartenant à un même cluster doivent être proches, tandis que les observations dissimilaires doivent appartenir à des clusters différents. Pour quantifier ces notions, nous allons avoir besoin d'un indicateur synthétique. Un nombre incalculable d'indicateurs existent dans la littérature pour évaluer la qualité d'une classification. Notre choix s'est porté sur l'indice de Davies-Bouldin.

**Homogénéité ou tightness (en anglais)** - On appelle homogénéité du cluster  $C_k$ , la moyenne des distances des observations de ce cluster à son centroïde :

$$T_k = \frac{1}{|C_k|} \sum_{x \in C_k} d(x, \mu_k)$$

où  $\mu_k$  est le centroïde de  $C_k$ .

L'homogénéité globale d'un clustering se calcule comme la moyenne des homogénéités des clusters :

$$T = \frac{1}{K} \sum_{k=1}^K T_k$$

**Séparabilité** - On appelle séparabilité des clusters  $C_k$  et  $C_l$  la distance entre leurs centroïdes :

$$S_{kl} = d(\mu_k, \mu_l)$$

La séparabilité globale d'un clustering se calcule comme la moyenne des séparabilités des clusters deux à deux :

$$S = \frac{2}{K(K-1)} \sum_{k=1}^K \sum_{l=k+1}^K S_{kl}$$

Plutôt que de considérer les deux critères de séparabilité (que l'on souhaite élevée) et d'homogénéité (que l'on souhaite faible) séparément, il est possible de les comparer l'un à l'autre grâce à l'indice de Davies-Bouldin.

**Indice de Davies-Bouldin** - On appelle indice de Davies-Bouldin du cluster  $C_k$  la valeur :

$$D_k = \max_{l \neq k} \frac{T_k + T_l}{S_{kl}}$$

L'indice de Davies-Bouldin global d'un clustering se calcule comme la moyenne des indices de Davies-Bouldin des clusters :

$$D = \frac{1}{K} \sum_{k=1}^K D_k$$

Plus cet indice est faible (proche de zéro), meilleure est la classification. Pour chacune des mesures de distance, un indice de Davies-Bouldin est calculé à l'issue de la classification sur chaque sous-portefeuille (692 au total). Puis on étudie la répartition de ces indices par mesure de distance afin d'appréhender la qualité globale de la classification de chaque distance par rapport aux autres.

Les distances Manhattan, Dissim et Euclidienne semblent produire de meilleures classifications. La proximité des résultats de ces trois distances n'est guère étonnante car elles sont la même famille. Au regard de la figure 4.4, il est plus logique de penser qu'il existe une distance adaptée pour chaque sous-portefeuille, mais la distance de Manhattan est adaptée à une plus grande majorité.

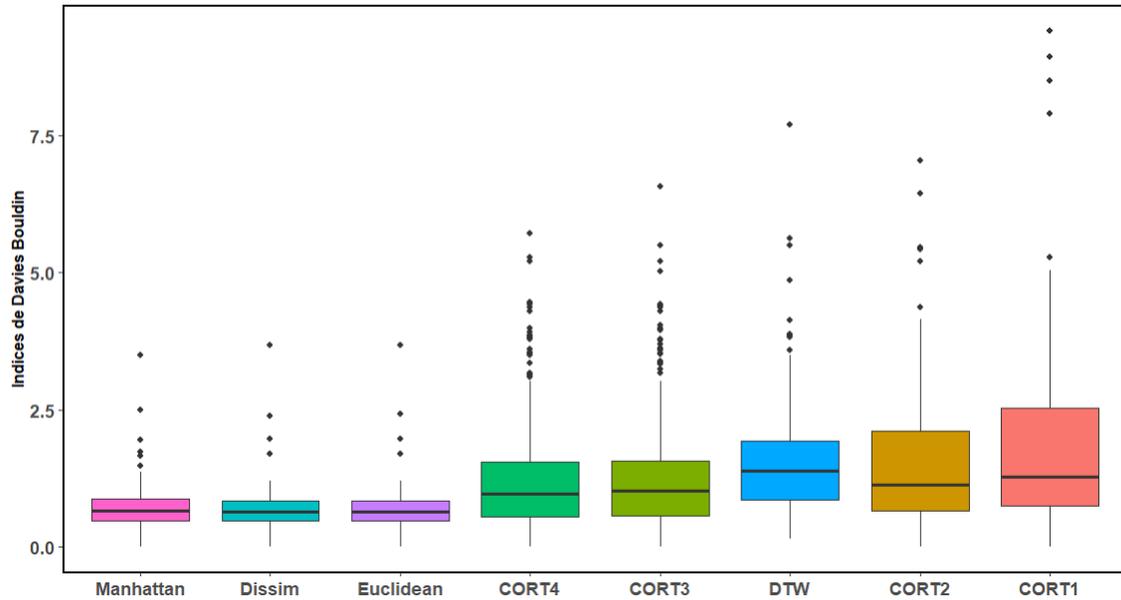


FIGURE 4.4 – Répartition des indices de Davies-Bouldin par mesure de distance

### 4.3 Phase d'agrégation du portefeuille

Cette section décrit la phase d'agrégation. Deux méthodes d'obtention du contrat générique pour chaque cluster seront étudiées. D'une façon générale, toutes les caractéristiques des contrats d'un même cluster représentant des quantités monétaires comme la provision mathématique, les primes, etc., mais aussi le nombre de contrats seront additionner pour obtenir les caractéristiques du Model Point qui est censé les représenter. L'objectif de la phase d'agrégation se résume donc à attribuer au Model Point un âge et une ancienneté.

#### 4.3.1 Moyenne pondérée par la provision mathématique de l'âge et de l'ancienneté

Cette première méthode consiste à calculer naïvement la moyenne pondérée par la PM de l'âge et de l'ancienneté des contrats dans chaque classe pour renseigner l'âge et l'ancienneté des Model Points. Cette méthode est illustrée dans la figure 4.5 sur un sous-portefeuille fictif de 10 contrats et deux classes.

Contrats	Age	Ancienneté	PM	Classe
contrat 1	52	237	1000	1
contrat 2	66	236	2500	1
contrat 3	66	204	3000	2
contrat 4	63	226	1200	1
contrat 5	68	199	4000	2
contrat 6	27	209	3700	2
contrat 7	85	217	2000	2
contrat 8	44	199	5000	1
contrat 9	85	210	1600	1
contrat 10	87	238	2700	2



Model Points	Age	Ancienneté	PM
MP 1	57	215	11300
MP 2	63	212	15400

FIGURE 4.5 – Exemple de la méthode d’agrégation par moyenne pondérée de l’âge et de l’ancienneté

Le tableau 4.1 récapitule les résultats obtenus par cette méthode d’agrégation. Il faut noter qu’un nombre total de 4500 Model Points a été ciblé, soit une compression de 87%. Ce nombre a été réparti entre les sous-portefeuilles au prorata de leur BEL. Les sous-portefeuilles sont rangés par ordre décroissant de leur BEL et les sous-portefeuilles avec les BELs les plus élevés se voient attribuer un grand nombre de MPs. Si le nombre de MPs attribué dépasse le nombre de contrats présents dans le sous-portefeuille, celui-ci n’est pas compressé et le surplus est réparti, encore au prorata du BEL, sur les autres sous-portefeuilles où le nombre de MPs attribué est inférieur au nombre de contrats.

TABLE 4.1 – Résultats du kmeans par fonction de distance

Distance	Nb lignes	Compression en %	Ecart en euro	Ecart relatif <sup>a</sup>
Dissim	4 483	87	55 028 157	8,78
DTW	3 376	90	76 485 778	12,21
Euclidean	4 483	87	55 869 290	8,91
Manhattan	4 477	87	54 095 317	8,63
CORT ( $r = 1$ )	4 460	87	70 511 348	11,25
CORT ( $r = 2$ )	4 465	87	70 637 884	11,27
CORT ( $r = 3$ )	4 456	87	70 157 479	11,20
CORT ( $r = 4$ )	4 457	87	72 114 773	11,51
<b>Portefeuille</b>	<b>33 482</b>		<b>62 640 384 409</b>	

<sup>a</sup> Les erreurs relatives sont exprimées en pour 10 000 du BEL total

Dans toute la suite de ce mémoire, seuls les résultats obtenus avec la distance de Manhattan seront présentés. A taux de compression égal, elle fournit le plus faible écart de BEL entre le portefeuille initial et le portefeuille agrégé.

## Backtesting

Le backtesting consiste à vérifier si les Model Points créés restituent bien le comportement moyen des contrats qu'ils représentent. Ce test permet de mettre à l'épreuve la méthode d'agrégation choisie. L'idée est de calculer les probabilités de sortie de chacun des Model Points et de comparer ces probabilités de sortie aux centroïdes des clusters. Nous avons réalisé ce test sur notre sous-portefeuille illustratif (*voir fig. 4.6*).

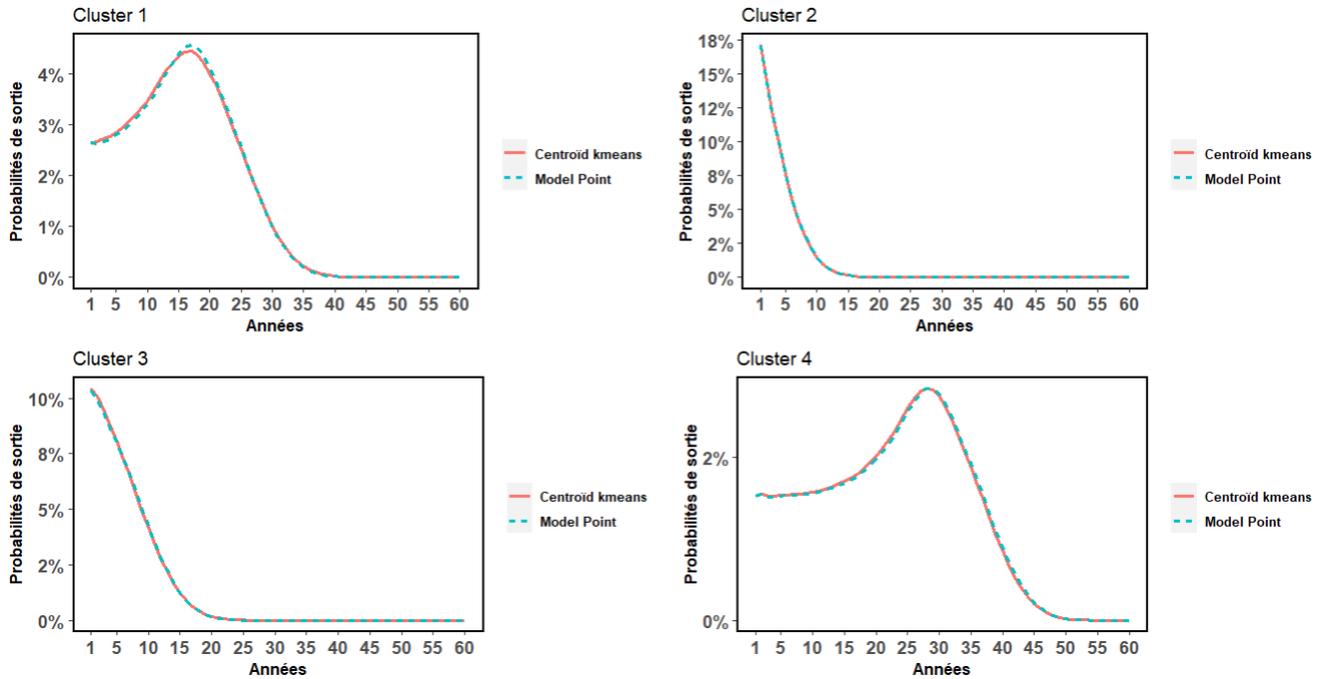


FIGURE 4.6 – Comparaison Centroïdes VS Model Points

Dans chacun des clusters, on constate un alignement quasi parfait du centroïde et du vecteur de probabilités de sortie du Model Point. La moyenne pondérée semble bien marcher pour l'agrégation des contrats. Nous verrons plus tard que sous certaines conditions, il s'agit même de la méthode optimale. Toutefois, il est important de signifier que cet alignement quasi parfait n'est pas observé pour tous les clusters dans d'autres sous-portefeuilles. La prochaine méthode d'agrégation que nous proposons se veut être une tentative pour pallier les insuffisances de la moyenne pondérée par la PM.

### 4.3.2 Méthode par développement limité en série de Taylor

La méthode que nous proposons est décrite par la fonction suivante :

$$C : [0, 1]^{60} \rightarrow \mathbb{R}^2$$

$$p \mapsto C(p)$$

La fonction  $C$  permet de calculer pour chaque vecteur de probabilités de sortie  $p = (p_1, \dots, p_{60})$ , l'âge et l'ancienneté  $(x, y)$  du contrat qui lui correspond tel que  $C(p) = (C_1(p), C_2(p)) = (x, y)$ . L'objectif de notre méthode est de trouver la meilleure forme fonctionnelle qui permettrait de calculer pour un cluster donné, l'âge et l'ancienneté du centroïde de ce cluster.

Soit  $A$  un cluster contenant  $n$  vecteurs de probabilités de sortie notés  $p^1, \dots, p^n$ . Chaque élément de ce cluster est donc un vecteur de taille 60 de sorte que  $p_j^i$  représente la probabilité de sortie du contrat  $i$  à l'année de projection  $j$  avec  $i \in \{1, \dots, n\}$  et  $j \in \{1, \dots, 60\}$ . Le centroïde de ce cluster se calcule de la façon suivante :

$$\bar{p} = \left( \frac{1}{n} \sum_{i=1}^n p_1^i, \dots, \frac{1}{n} \sum_{i=1}^n p_{60}^i \right)$$

Afin d'étudier le comportement local de  $C$  autour du centroïde, on utilise le développement en série de Taylor de  $C(p) = (C_1(p), C_2(p))$  au point  $\bar{p}$ .

Au voisinage de  $\bar{p}$  :

$$C(p) = C(\bar{p}) + \frac{dC}{dp}(\bar{p})(p - \bar{p}) + \sum_{q=2}^{\infty} \frac{1}{q!} \frac{d^q C}{dp^q}(\bar{p})(p - \bar{p})^q \quad (4.4)$$

où  $\frac{d^q C}{dp^q}$  est la dérivée d'ordre  $q$  de  $C$ .

La matrice Jacobienne de  $C$  s'écrit :

$$J_{\bar{p}}(C) = \frac{dC}{dp}(\bar{p}) = \begin{pmatrix} \frac{\partial C_1}{\partial p_1}(\bar{p}) & \frac{\partial C_1}{\partial p_2}(\bar{p}) & \dots & \frac{\partial C_1}{\partial p_{60}}(\bar{p}) \\ \frac{\partial C_2}{\partial p_1}(\bar{p}) & \frac{\partial C_2}{\partial p_2}(\bar{p}) & \dots & \frac{\partial C_2}{\partial p_{60}}(\bar{p}) \end{pmatrix}$$

Avant d'aller plus loin, nous allons introduire une première hypothèse simplificatrice qui consiste à ne considérer que l'âge comme variable cible. Cette hypothèse n'est cependant pas anodine et est basée sur le constat selon lequel, pour la grande majorité des sous-portefeuilles constituant notre portefeuille, l'effet de la mortalité serait beaucoup plus fort sur les probabilités de sortie que l'effet des rachats. Ce constat est illustré avec notre sous-portefeuille d'exemple à la figure 4.7. Suite à la classification à la figure 4.3 et l'obtention des clusters, nous représentons un nuage de points des âges et des anciennetés des contrats correspondants en associant chaque point au cluster auquel il appartient via le même système de couleurs.

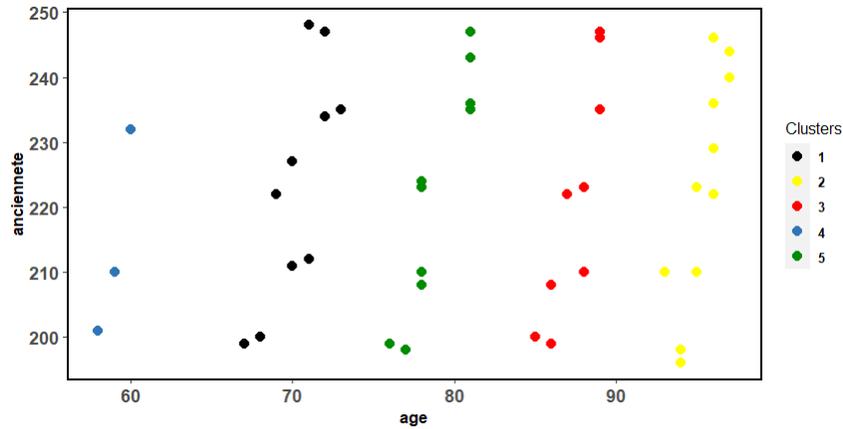


FIGURE 4.7 – Age et ancienneté selon les clusters

Par comparaison des deux figures, on peut constater une nette corrélation entre les âges des contrats et les probabilités de sortie qui en découlent. Les âges les plus élevés sont associés aux probabilités de sortie des débuts de projection les plus élevées. C’est donc naturellement que nous nous focalisons dans la suite sur l’agrégation optimale de l’âge.

On redéfinit la fonction  $C$  comme suit :

$$C : [0, 1]^{60} \rightarrow \mathbb{R}$$

$$p \mapsto C(p) = x$$

Avec cette nouvelle définition, le gradient de  $C$  s’écrit :

$$\nabla_p C = \frac{dC}{dp} = \left( \frac{\partial C}{\partial p_1}, \frac{\partial C}{\partial p_2}, \dots, \frac{\partial C}{\partial p_{60}} \right)$$

La matrice Hessienne de  $C$  s’écrit :

$$H = \frac{d^2 C}{dp^2} = \begin{pmatrix} \frac{\partial^2 C}{\partial p_1^2} & \dots & \frac{\partial^2 C}{\partial p_1 \partial p_{60}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 C}{\partial p_{60} \partial p_1} & \dots & \frac{\partial^2 C}{\partial p_{60}^2} \end{pmatrix}$$

On considère  $\{x_1, \dots, x_n\}$ , les âges des contrats correspondants aux  $n$  vecteurs de probabilités du cluster  $\mathbf{A}$ . Ainsi, on sait que  $C(p^i) = x_i$  pour tout  $i$ . Pour simplifier le raisonnement, nous allons supposer, sans perte de généralité, que tous les contrats possèdent le même montant de PM. Par conséquent, les moyennes pondérées par la PM

deviennent des moyennes arithmétiques simples<sup>3</sup>.

Si on somme l'équation 4.4 sur tous les contrats du cluster  $\mathbf{A}$  on obtient la formule suivante :

$$\begin{aligned} \sum_{i=1}^n C(p^i) &= \sum_{i=1}^n C(\bar{p}) + \frac{dC}{dp}(\bar{p}) \sum_{i=1}^n (p^i - \bar{p}) \\ &\quad + \frac{1}{2} \sum_{i=1}^n (p^i - \bar{p})^T \frac{d^2C}{dp^2}(\bar{p}) (p^i - \bar{p}) \\ &\quad + \sum_{i=1}^n \sum_{q=3}^{\infty} \frac{1}{q!} \frac{d^q C}{dp^q}(\bar{p}) (p^i - \bar{p})^{(q)} \end{aligned}$$

- **Hypothèse de comportement linéaire**

Dans un premier temps, on suppose que  $\frac{d^q C}{dp^q} = 0$  pour  $q \geq 2$ , ce qui se traduit par la linéarité de  $C$ . En reprenant l'équation précédente :

$$\begin{aligned} \sum_{i=1}^n C(p^i) &= n \times C(\bar{p}) + \frac{dC}{dp}(\bar{p}) \sum_{i=1}^n (p^i - \bar{p}) \\ &= n \times C(\bar{p}) + \frac{dC}{dp}(\bar{p}) \left( \sum_{i=1}^n p^i - n \times \frac{\sum_{i=1}^n p^i}{n} \right) \\ &= n \times C(\bar{p}) \end{aligned}$$

Finalement :

$$C(\bar{p}) = \frac{1}{n} \sum_{i=1}^n C(p^i) = \frac{1}{n} \sum_{i=1}^n x_i$$

Si la fonction  $C$  qui à chaque vecteur de probabilités associe l'âge du contrat correspondant est linéaire, alors aucune erreur n'est commise en estimant l'âge du centroïde par la moyenne pondérée par la PM des âges des contrats du cluster. On retrouve ainsi notre première méthode naïve d'agrégation.

---

3. Cette hypothèse est faite uniquement dans un but didactique, sinon nous avons tenu compte de la pondération par la PM dans nos travaux

• **Hypothèse d'une tendance quadratique**

On suppose maintenant que  $\frac{d^q C}{dp^q} = 0$  pour  $q \geq 3$ . Par un raisonnement similaire, on obtient :

$$\sum_{i=1}^n C(p^i) = n \times C(\bar{p}) + \frac{1}{2} \sum_{i=1}^n (p^i - \bar{p})^T \bar{H} (p^i - \bar{p})$$

avec  $\bar{H}$  représentant la hessienne évaluée au centroïde  $\bar{p}$ .

Développons l'expression obtenue :

$$\sum_{i=1}^n C(p^i) = n \times C(\bar{p}) + \frac{1}{2} \sum_{i=1}^n (p^i - \bar{p})^T \begin{pmatrix} \frac{\partial^2 C}{\partial p_1^2}(\bar{p}) & \cdots & \frac{\partial^2 C}{\partial p_1 \partial p_{60}}(\bar{p}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 C}{\partial p_n \partial p_1}(\bar{p}) & \cdots & \frac{\partial^2 C}{\partial p_{60}^2}(\bar{p}) \end{pmatrix} \begin{pmatrix} p_1^i - \bar{p}_1 \\ p_2^i - \bar{p}_2 \\ \vdots \\ p_{60}^i - \bar{p}_{60} \end{pmatrix}$$

$$\sum_{i=1}^n C(p^i) = n \times C(\bar{p}) + \frac{1}{2} \sum_{i=1}^n (p^i - \bar{p})^T \begin{pmatrix} \sum_{j=1}^{60} \frac{\partial^2 C}{\partial p_1 \partial p_j}(\bar{p})(p_j^i - \bar{p}_j) \\ \vdots \\ \sum_{j=1}^{60} \frac{\partial^2 C}{\partial p_{60} \partial p_j}(\bar{p})(p_j^i - \bar{p}_j) \end{pmatrix}$$

Pour  $j_1, j_2 \in \{1, \dots, 60\}$  on a :

$$\sum_{i=1}^n C(p^i) = n \times C(\bar{p}) + \frac{1}{2} \sum_{i=1}^n \sum_{j_1=1}^{60} \sum_{j_2=1}^{60} \frac{\partial^2 C}{\partial p_{j_1} \partial p_{j_2}}(\bar{p})(p_{j_1}^i - \bar{p}_{j_1})(p_{j_2}^i - \bar{p}_{j_2}) \quad (4.5)$$

Nous nous arrêtons dans ce mémoire à l'hypothèse d'un comportement quadratique pour la fonction  $C$ . Cette hypothèse constitue une des limites de notre méthode d'agrégation. Il est quasiment impossible d'aller au-delà comme par exemple une hypothèse de comportement cubique, car les calculs deviennent trop complexes. Déjà que l'expression obtenue à l'équation 4.5 est extrêmement difficile à résoudre. Nous sommes dans l'obligation de faire des hypothèses simplificatrices pour la résolution. Ce problème peut néanmoins être surmonté via l'utilisation de techniques numériques très abouties qui ne seront pas abordées dans cette étude.

On suppose d'abord que les dérivées croisées  $\frac{\partial^2 C}{\partial p_{j_1} \partial p_{j_2}}$  sont toutes nulles pour  $j_1 \neq j_2$ . Par conséquent l'équation 4.5 devient :

$$\sum_{i=1}^n C(p^i) = n \times C(\bar{p}) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{60} \frac{\partial^2 C}{\partial p_j^2}(\bar{p})(p_j^i - \bar{p}_j)^2$$

Enfin la dernière hypothèse concerne les éléments diagonaux de la matrice hessienne de  $C$ . Nous supposons que  $\frac{\partial^2 C}{\partial p_j^2}(\bar{p}) = \theta_j = 2a_j$  pour tout  $j$  avec  $a_j \in \mathbb{R}$ .

Nous convenons que cela représente un postulat théorique assez fort et que nous ne disposons d'aucun moyen de le vérifier. Mais grâce à cette hypothèse, on a :

$$\sum_{i=1}^n C(p^i) = n \times C(\bar{p}) + \sum_{i=1}^n \sum_{j=1}^{60} a_j (p_j^i - \bar{p}_j)^2$$

Si de plus  $a_j = a \in \mathbb{R}$  pour tout  $j$  on a :

$$\sum_{i=1}^n C(p^i) = n \times C(\bar{p}) + a \sum_{i=1}^n \|p^i - \bar{p}\|_2^2$$

Finalement l'âge du centroïde s'obtient comme suit :

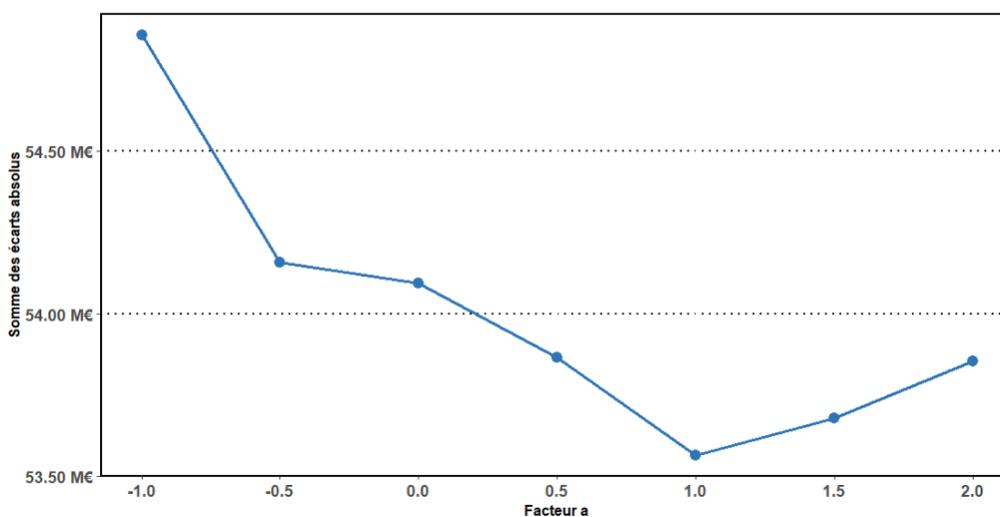
$$C(\bar{p}) = \frac{1}{n} \sum_{i=1}^n C(p^i) - \frac{a}{n} \sum_{i=1}^n \|p^i - \bar{p}\|_2^2$$

$$C(\bar{p}) = \frac{1}{n} \sum_{i=1}^n C(p^i) - a \times T_A$$

avec  $T_A$  qui désigne l'indice d'homogénéité ou tightness (vu en 4.2.5) du cluster  $\mathbf{A}$ .

Il s'agit de la moyenne pondérée par la PM des âges des contrats du cluster, pénalisée par une quantité proportionnelle à la moyenne des distances de tous les vecteurs de probabilités par rapport au centroïde. Ce facteur de pénalisation montre tout l'intérêt d'une bonne classification. Et nous offre un outil d'explication des résultats de la première méthode d'agrégation. Les clusters pour lesquels le backtesting a révélé un mauvais alignement du centroïde et du vecteur des probabilités de sortie du Model Point sont certainement des clusters assez hétérogènes d'où un facteur de pénalisation assez élevé qui dévie considérablement l'âge du centroïde de la moyenne pondérée des âges.

Remarquons qu'il est possible de faire varier le coefficient  $a$  afin d'obtenir le meilleur facteur de pénalisation. Le domaine de variation de  $a$  doit être choisi de telle sorte que le résultat ne soit pas négatif et n'excède pas l'âge maximal dans la table de mortalité.

FIGURE 4.8 – Somme des écarts absolus en fonction du facteur  $a$ 

Dans notre étude, nous avons décidé de retenir la valeur 1 pour le coefficient  $a$  après une recherche dans l'intervalle  $[-1, 2]$ . Il est important de noter qu'il pourrait éventuellement s'agir d'une solution locale. Les résultats sont présentés dans le tableau ci-dessous :

TABLE 4.2 – Récapitulation des résultats

Méthode	Nb lignes	Compression en %	Ecart en euro	Ecart relatif <sup>a</sup>
LOP	4 563	86	124 609 384	20
kmeans M1	4 477	87	54 095 317	8,63
kmeans M2	4 477	87	53 564 162	8,55
<b>Portefeuille</b>	<b>33 482</b>		<b>62 640 384 409</b>	

<sup>a</sup> Les erreurs relatives sont exprimées en pour 10 000 du BEL total

Bien que pas très significatif, la méthode d'agrégation basée sur le développement en série de Taylor permet une baisse de l'écart absolu de BEL de 1% par rapport à la méthode d'agrégation par moyenne pondérée simple.

# Chapitre 5

## Validation de la compression par kmeans

### Sommaire

---

5.1	Procédure de la validation . . . . .	61
5.2	Avant compression par kmeans . . . . .	62
5.3	Après compression par kmeans . . . . .	62
5.4	Impact de la compression par kmeans . . . . .	62
5.5	Critiques de l’approche par kmeans . . . . .	64
5.5.1	Critiques générales . . . . .	64
5.5.2	Améliorations possibles de notre étude . . . . .	64

---



Après construction des Model Points sous un mode de projection passif seul, ce chapitre a pour objectif de tester la robustesse de la compression après passage aux modes de projection déterministe et stochastique. Il permet d’évaluer le véritable impact de la compression dans le processus de production de la compagnie.

## 5.1 Procédure de la validation

La validation de la compression est une étape qui vise à tester la robustesse des model points, construits sous le mode de projection passif seul, dans des conditions beaucoup plus réalistes. C'est à dire dans les modes de projection déterministe et stochastique. Le premier projette les données à partir d'un seul scénario moyen  $\mathbf{F}$ , tandis que le second projette sur un exemple de scénarios  $(\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_N)$ . Ici  $N = 1000$ .

Les étapes de la validation peuvent être déclinées en trois points :

1. Fixer un mode de projection (déterministe ou stochastique)
2. Projeter le portefeuille ligne par ligne selon le mode de projection défini au point 1, puis déterminer le total des engagements de BEL pour ce portefeuille
3. Projeter les model points issus de la compression selon le même mode de projection, puis déterminer le total des engagements de BEL pour l'ensemble de ce portefeuille agrégé
4. Comparer ces deux engagements de BEL et mesurer l'impact de la compression sur l'évaluation des engagements.

Pour ce qui est de la première étape, nous réaliserons la validation sous les deux modes de projection. Cela permettra par ailleurs d'évaluer le coût des options et garanties.

En ce qui concerne le deuxième point, compte tenu de la taille de notre portefeuille de contrats, il ne nous est pas possible de projeter l'ensemble des lignes ni en déterministe ni en stochastique. C'est d'ailleurs ce qui justifie la nécessité de la compression en passif seul. Nous allons donc restreindre la validation à trois sous-portefeuilles volontairement choisis par Allianz pour être plus ou moins représentatifs de l'ensemble des sous-portefeuilles. Ces trois sous-portefeuilles sont présentés dans le tableau suivant :

TABLE 5.1 – Présentation des sous-portefeuilles pour la validation

	Nb lignes avant compression	Nb lignes après compression	taux de de compression	Type
SP1	108	26	76%	<i>MultiSupport</i>
SP2	329	6	98%	<i>MonoSupport EURO</i>
SP3	126	16	87%	<i>MultiSupport</i>
<b>Total</b>	<b>563</b>	<b>48</b>	<b>91%</b>	

## 5.2 Avant compression par kmeans

Les projections avant compression concernent au total 563 contrats. Ces contrats sont répartis parmi trois sous-portefeuilles. Le portefeuille constitué des trois sous-portefeuilles génère au total pour la compagnie des engagements de BEL de 713 M€ et de 725 M€ respectivement sous le mode déterministe et stochastique. Les engagements de BEL déterminés à partir du modèle déterministe sous-estiment le véritable montant des engagements car ceux-ci n'intègrent pas les Options et Garanties (O&G).

TABLE 5.2 – Résultats de la projection avant compression par kmeans

	BEL Déterministe	BEL Stochastique	O&G
SP1	367 883 509	371 267 940	3 384 431
SP2	86 192 623	87 071 097	878 474
SP3	259 362 551	267 074 790	7 712 239
<b>Total</b>	<b>713 438 683</b>	<b>725 413 827</b>	<b>11 975 143</b>

## 5.3 Après compression par kmeans

Les projections après compression concernent au total 48 contrats. Le portefeuille compressé génère également pour la compagnie des engagements de BEL de 713 M€ et de 725 M€ respectivement sous le mode déterministe et stochastique. Les engagements de BEL déterminés à partir du modèle déterministe sont encore inférieurs à ceux du modèle stochastique en raison des Options et Garanties (O&G).

TABLE 5.3 – Résultats de la projection après compression par kmeans

	BEL Déterministe	BEL Stochastique	O&G
SP1	367 816 529	371 194 628	3 378 099
SP2	86 211 237	87 130 212	918 976
SP3	259 501 324	267 261 989	7 760 665
<b>Total</b>	<b>713 529 090</b>	<b>725 586 830</b>	<b>12 057 740</b>

## 5.4 Impact de la compression par kmeans

Pour le modèle déterministe comme stochastique, l'impact de la compression par kmeans sur les engagements de BEL reste limité et s'évalue seulement en millier d'euros.

En terme relatif, la compression par kmeans permet de réduire le nombre de ligne du portefeuille de 91% avec seulement une déviation du montant de BEL de 0,01% pour le modèle déterministe et 0,02% pour le modèle stochastique.

TABLE 5.4 – Impact relatif de la compression par kmeans

	BEL Deterministe	BEL Stochastique	O&G
SP1	-0,02%	-0,02%	-0,19%
SP2	0,02%	0,07%	4,61%
SP3	0,05%	0,07%	0,63%
<b>Total</b>	<b>0,01%</b>	<b>0,02%</b>	<b>0,69%</b>

Le résultat du modèle stochastique étant une moyenne, il est important de mettre en évidence le niveau de volatilité de cette moyenne afin de s'assurer que ce résultat n'est pas le fruit d'une compensation de scénarios favorables et défavorables. Pour cela, nous traçons le nuage de points qui représente l'erreur relative du BEL par rapport au numéro d'itération ou scénario.

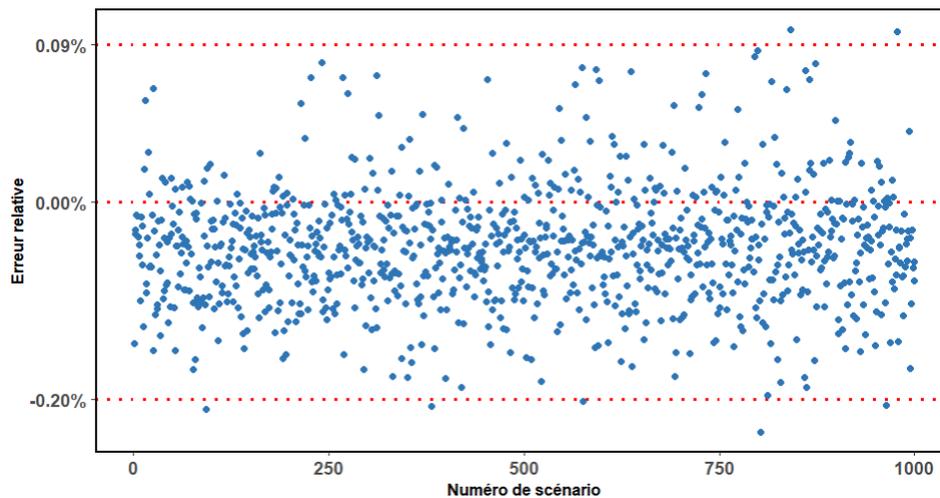


FIGURE 5.1 – Erreur de BEL sur les 1000 trajectoires stochastiques

On constate que l'erreur relative n'est pas volatile et reste majoritairement comprise entre -0,2% et +0,09%. Cela est lié l'approche retenue de regroupement des contrats à partir des probabilités de sortie.

## 5.5 Critiques de l'approche par kmeans

### 5.5.1 Critiques générales

L'algorithme du kmeans est un algorithme assez pratique pour la compression de portefeuille de contrats d'assurance vie. Il possède de nombreux avantages comme la possibilité de contrôler le nombre de MPs cible via son paramètre  $k$ . Il permet de compresser considérablement le portefeuille tout en conservant une précision très satisfaisante. En plus, la phase d'agrégation permet d'assurer la représentativité du portefeuille compressé vis-à-vis des encours : pas de perte de PM, prime, nombre de contrats, etc. avant et après compression. Cependant, cette approche n'est pas sans limite. Il est important de bien choisir la méthode d'initialisation de l'algorithme ainsi que la métrique. De plus, le processus de production de MPs est un peu plus lent avec cette méthode qu'avec le LOP.

Enfin, l'utilisation des probabilités de sortie à la place des cashflows de BEL permet de rendre la compression plus robuste.

### 5.5.2 Améliorations possibles de notre étude

Il est possible de trouver une méthode de répartition plus optimale du nombre de Model Points entre les sous-portefeuilles par exemple en prenant en compte l'hétérogénéité du sous-portefeuille, en plus du BEL dans les critères d'allocation.

Considérer le développement de Taylor de la fonction  $C$  à des ordres supérieurs à 2.

# Conclusion

Au cours de ce mémoire, il a été question de développer un algorithme de compression de portefeuille d'assurance vie. L'étude s'est limitée au cadre restreint de contrats de type épargne retraite individuelle. Le portefeuille agrégé contient 4500 contrats avec en input un portefeuille contenant plus d'un million de contrats avant le premier niveau d'agrégation et 35 000 contrats avant le second niveau d'agrégation.

Une première approche basée sur de l'optimisation linéaire a été proposée au chapitre 3. Mais le niveau de compression de celle-ci a été jugé insuffisant par Allianz. Il a fallu donc trouver un mécanisme qui permettrait d'améliorer cette capacité de compression tout en détériorant le moins possible la précision. C'est ce que nous avons réussi en intégrant à l'algorithme un paramètre de tolérance. Celui-ci nous a permis d'atteindre le nombre de Model Points cible de 4500 avec une erreur relative de BEL se chiffrant à 20 euros pour 10 000. Ce qui représente une performance assez notable. Malgré cette amélioration apportée à l'algorithme, son utilisation des cashflows de BEL comme variable d'agrégation met à mal sa robustesse dans des modèles stochastiques.

L'approche alternative exposée à travers ce mémoire au chapitre 4 a elle aussi permis d'atteindre le nombre de Model Points cible, mais avec une erreur relative de BEL de moins de 9 euros pour 10 000. Les résultats de la validation en chapitre 5 sur un portefeuille plus restreint confirment la robustesse de cette approche au passage à des modélisations déterministe et stochastique avec un impact de la compression sur les options et garanties qui s'évalue à 0,69%. La méthode est déjà passée en production au sein d'Allianz France.

Pour finir, d'autres approches d'agrégation auraient pu être explorées, mais souhaitant privilégier l'aspect pratique, tout le processus de compression a été développé dans l'outil RAFM qui est un environnement de développement assez contraignant mais qui a l'avantage d'être rapide (1 heure pour tout le processus : classification, agrégation et validation) et facilement utilisable par les collaborateurs en interne (contrairement à d'autres outils qui nécessitent une formation et qui peuvent être source d'erreurs opérationnelles).

# Bibliographie

**Goffinet, Etienne & Lebbah, Mustapha & Azzag, Hanane & Giraldi, Loïc. (2020).** Modèles de mélange pour le clustering de séries temporelles basée sur un dictionnaire, Laboratoire Informatique de Paris Nord, 93430 Villetaneuse, France.

**Jérôme Dumont, Alfredo Hernández, Pedro Gomis, Guy Carrault. (2009)** Classification de séries temporelles par modèles markoviens cachés application à l'ischémie myocardique. XXII Colloque GRETSI, Sep 2009, Dijon, France. pp.552. fihal-00913757

**Pierre-Olivier Goffard. (2015)** Optimisation de l'agrégation d'un portefeuille de contrat d'assurance vie. AXA France, Université de Aix-Marseille, Juillet 2015, Marseille, France.

**Rémi Gaudin, Nicolas Nicoloyannis. (2006).** Apprentissage non supervisé de séries temporelles à l'aide des K-Means et d'une nouvelle méthode d'agrégation de séries, LABORATOIRE ERIC 3038 Université Lumière - Lyon2, Dec 2006, 69676 BRON, France. pp. 213-218, doi : 10.1109/ICMLA.2006.12.

**Saeid Soheily-Khah. (2016).** Generalized k-means based clustering for temporal data under time warp. Artificial Intelligence. Universite Grenoble Alpes, Oct 2016. France.

**Till Wohlfarth. (2016).** Machine-learning pour la prédiction des prix dans le secteur du tourisme en ligne, Analyse de données, Statistiques et Probabilités [physics.data-an]. Télécom ParisTech, 2013. Français.

# Annexes

## Annexe A : Outils mathématiques nécessaires pour le LOP

Soit  $V$  un vecteur, on dit que  $V$  est centré si la moyenne des composantes de  $V$  est égale à 0. On dit que  $V$  est réduit si la variance ou l'écart-type de  $V$  est égal à 1.

Soient  $V_1$  et  $V_2$  deux vecteurs centrés réduits, alors

$$\langle V_1, V_2 \rangle = V_1^T V_2 = \text{Corr}(V_1, V_2)$$

Soient  $V_1$  et  $V_2$  deux vecteurs, plus le coefficient de corrélation entre  $V_1$  et  $V_2$  est élevé, plus  $V_1$  est susceptible d'expliquer  $V_2$  et vice-versa.

Soit  $A$  une matrice de taille  $n \times m$  et  $U$  un vecteur de taille  $m$ . On suppose que  $A$  est constituée en colonne de vecteurs  $V_1, V_1, \dots, V_m$ . Alors :

$$A^T U = (V_1^T U, V_2^T U, \dots, V_m^T U) = (\text{Corr}(V_1, U), \text{Corr}(V_2, U), \dots, \text{Corr}(V_m, U))$$

## Annexe B : Exemple illustratif du LOP

Considérons les écoulements de BEL cash-flow de cinq contrats et appliquons à ceux-ci l'algorithme du LOP.

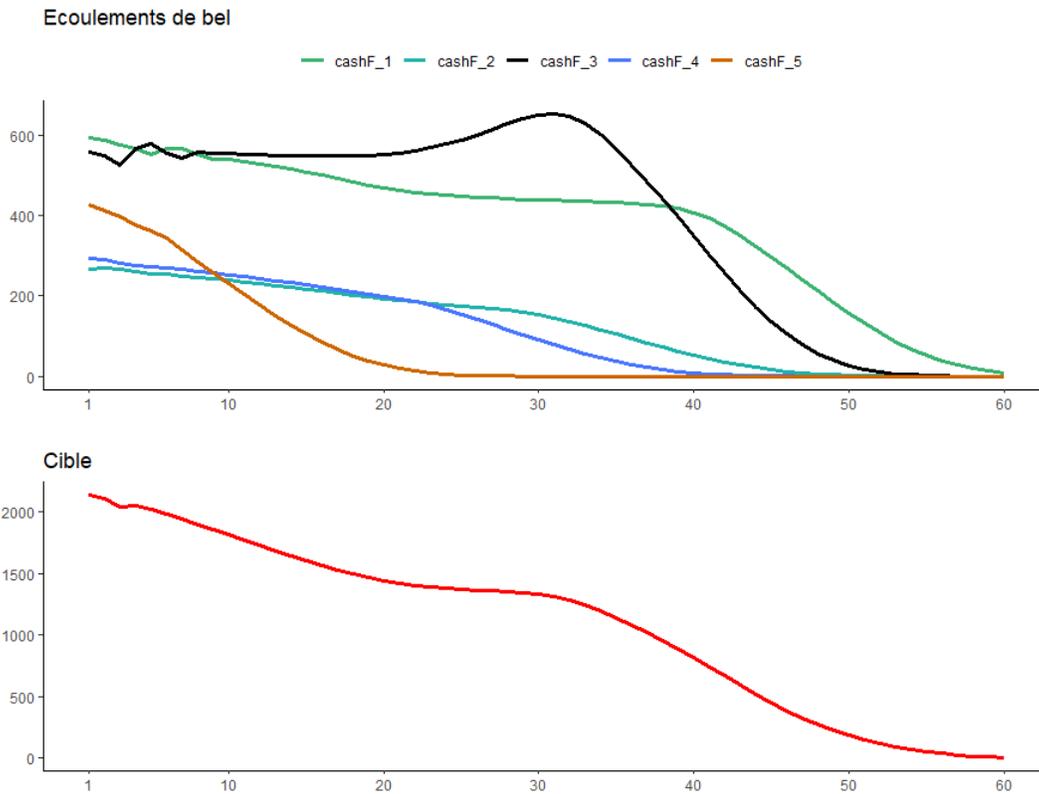


FIGURE 5.2 – Représentation des chroniques de cash-flows et de la valeur cible

## L'initialisation

A l'initialisation, l'algorithme choisit des poids nuls pour chaque cashflow  $\implies x = (0, 0, 0, 0, 0)$ , ce qui correspond à la ligne horizontale bleue.

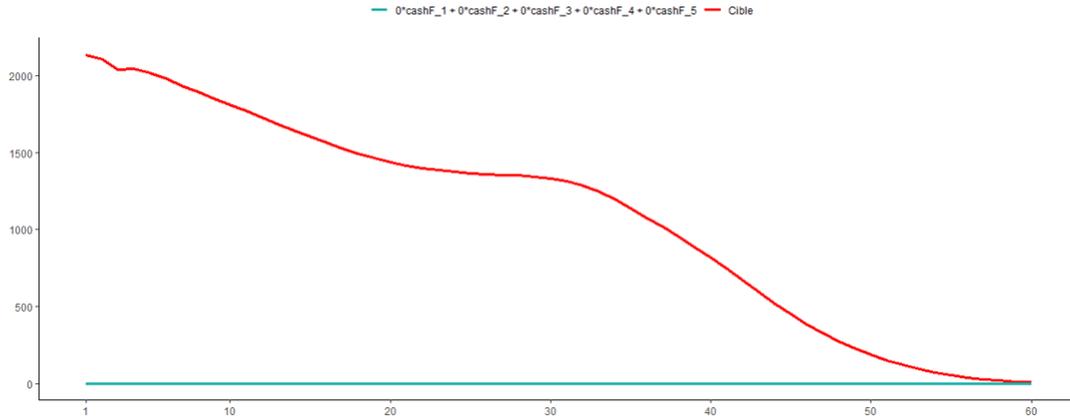


FIGURE 5.3 – La cible et de la combinaison des cashflows à l'initialisation

L'algorithme va chercher à réduire au maximum l'écart entre la ligne bleue :  $Ax$  et la ligne rouge :  $b$

**Remarque :** L'écart initial correspond aussi à la cible.

## La première itération

A la première itération, on calcule les corrélations entre tous les cashflows et l'écart initial qui correspond à la valeur cible pour cette itération

	cashF1	cashF2	cashF3	cashF4	cashF5
Corrélations	0,940	0,965	0,883	0,901	0,699

Ensuite on choisit le cashflow pour lequel la corrélation est positive et maximale  $\implies \text{cashF2}$

**Remarque :** Une corrélation aussi élevée (proche de 1) est synonyme d'une contribution considérable de la part de ce cashflow.

### Mise à jour des poids

Régression linéaire de la cible sur le ou les cashflows sélectionnés :

$$cible = \alpha_0 + \alpha_2 * cashF2$$

Dans notre exemple, les coefficients obtenus sont :  $\alpha_0 = 228,073$  et  $\alpha_2 = 6,818$ .

Le nouveau vecteur de poids à l'issue de cette étape est :  $x = (0, 6.818, 0, 0, 0)$

### Calcul du nouvel écart

Le nouvel écart est la différence entre la cible et la progression :

$$ecart = cible - \alpha_2 * cashF2$$

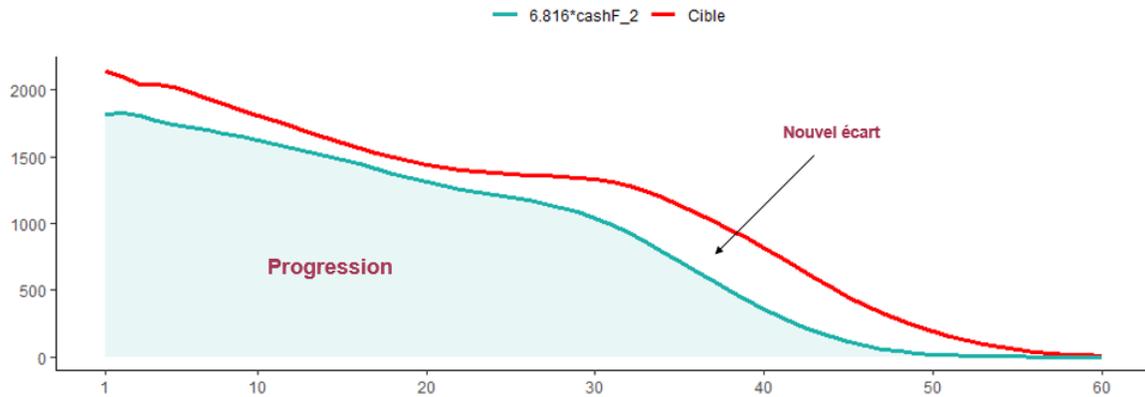


FIGURE 5.4 – La cible et de la combinaison des cashflows à première itération

### La deuxième itération

A la deuxième itération, on calcule les corrélations entre les cashflows non sélectionnés et le nouvel écart

	cashF1	cashF3	cashF4	cashF5
Corrélations	0,077	0,058	-0,032	0,000

Ensuite on choisit le cashflow pour lequel la corrélation est positive et maximale  
 $\Rightarrow cashF1$

**Remarque :** Avec une corrélation de 0,077, la contribution de ce cashflow sera moindre par rapport à celle de cashflow précédent.

### Mise à jour des poids

Régression linéaire de la cible sur tous les cashflows déjà sélectionnés jusque là :

$$cible = \beta_0 + \beta_1 * cashF1 + \beta_2 * cashF2$$

Les coefficients obtenus sont :  $\beta_0 = -41,394$ ,  $\beta_1 = 1,523$  et  $\beta_2 = 4,377$ .

Le nouveau vecteur de poids à l'issue de cette étape est :  $x = (1.523, 4.377, 0, 0, 0)$

### Calcul du nouvel écart

Le nouvel écart est la différence entre la précédente et la progression :

$$ecart = cible - \beta_1 * cashF1 - \beta_2 * cashF2$$

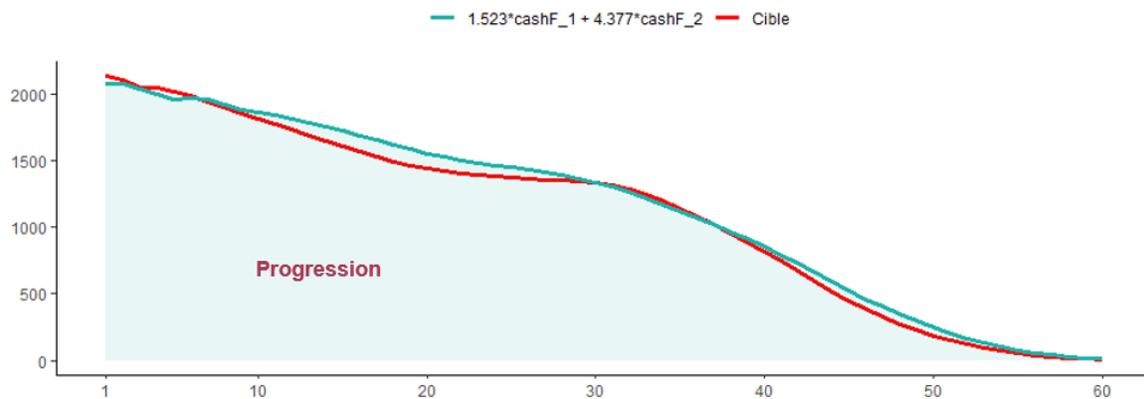


FIGURE 5.5 – La cible et de la combinaison des cashflows à deuxième itération

### La troisième itération

A la troisième itération, on calcule les corrélations entre les cashflows non sélectionnés et le nouvel écart

	cashF3	cashF4	cashF5
Corrélations	0,003	-0,005	0,025

Ensuite on choisit le cashflow pour lequel la corrélation est positive et maximale  $\implies cashF5$

**Remarque :** La contribution sera encore moindre que toutes les précédentes (corr = 0,025).

### Mise à jour des poids

Régression linéaire de la cible sur tous les cashflows déjà sélectionnés jusque là :

$$cible = \gamma_0 + \gamma_1 * cashF1 + \gamma_2 * cashF2 + \gamma_5 * cashF5$$

Les coefficients sont :  $\gamma_0 = -40,073$ ,  $\gamma_1 = 1,596$ ,  $\gamma_2 = 3,696$  et  $\gamma_5 = 0,300$ .

Les poids à l'issue de cette étape sont :  $x = (1.596, 3.696, 0, 0, 0.300)$

### Calcul du nouvel écart

Le nouvel écart est la différence entre la précédente et la progression :

$$ecart = cible - \gamma_1 * cashF1 - \gamma_2 * cashF2 - \gamma_5 * cashF5$$

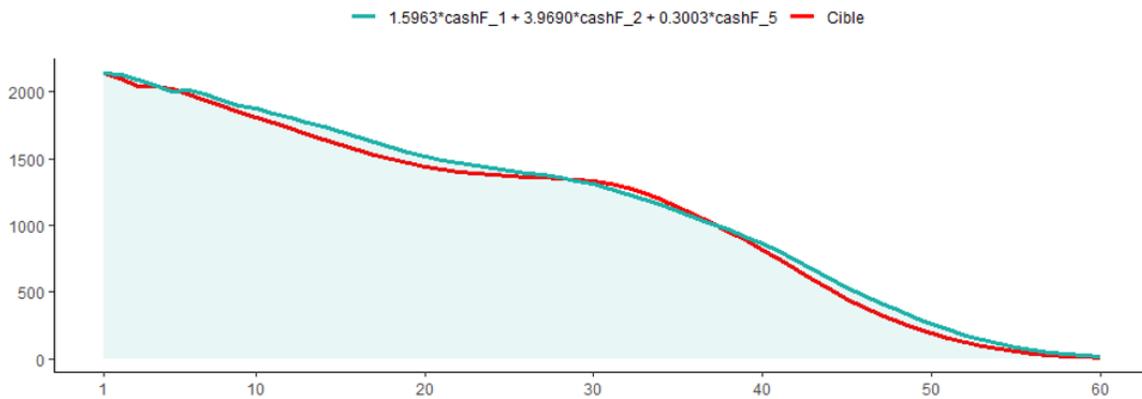


FIGURE 5.6 – La cible et de la combinaison des cashflows à la troisième itération

### La quatrième itération

À la quatrième itération, on calcule les corrélations entre les cashflows non sélectionnés et le nouvel écart

	cashF3	cashF4
Corrélations	0,014	-0,009

Ensuite on choisit le cashflow pour lequel la corrélation est positive et maximale  $\Rightarrow cashF3$

**Remarque :** Une contribution encore plus marginale (corr = 0,014).

### Mise à jour des poids

Régression linéaire de la cible sur tous les cashflows déjà sélectionnés jusque là :

$$cible = \theta_0 + \theta_1 * cashF1 + \theta_2 * cashF2 + \theta_3 * cashF3 + \theta_5 * cashF5$$

Les coefficients :  $\theta_0 = -2,28$ ,  $\theta_1 = 1,05$ ,  $\theta_2 = 2,78$ ,  $\theta_3 = 0,696$  et  $\theta_5 = 0,878$ .

Les poids à l'issue de cette étape sont :  $x = (1.5, 2.78, 0.696, 0, 0.878)$

### Calcul du nouvel écart

Calculer le nouvel écart !

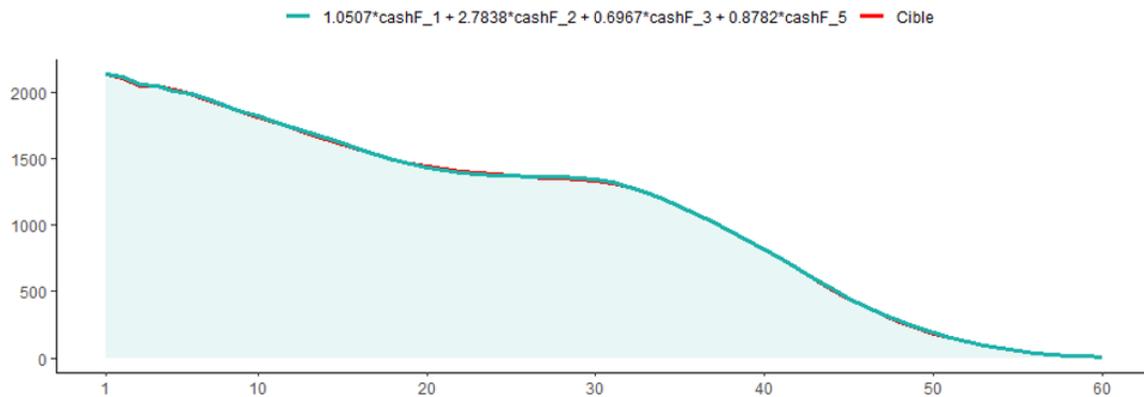


FIGURE 5.7 – La cible et de la combinaison des cashflows à la quatrième itération

### La cinquième itération

A la cinquième itération, on calcule la corrélation entre le dernier cashflow non sélectionné et le nouvel écart.

	cashF4
Corrélations	0,0002

Puisque la corrélation est positive, on la conserve  $\implies cashF4$

**Remarque :** Une contribution est la plus petite (corr = 0,0002).

### Mise à jour des poids

Régression linéaire de la cible sur tous les cashflows déjà sélectionnés jusque là :

$$cible = \lambda_0 + \lambda_1 * cashF1 + \lambda_2 * cashF2 + \lambda_3 * cashF3 + \lambda_4 * cashF4 + \lambda_5 * cashF5$$

Les coefficients :  $\lambda_0 = 0$ ,  $\lambda_1 = 1$ ,  $\lambda_2 = 1$ ,  $\lambda_3 = 1$ ,  $\lambda_4 = 1$  et  $\lambda_5 = 1$ .

Les poids à l'issue de cette étape sont :  $x = (1, 1, 1, 1, 1)$

### Calcul du nouvel écart

Le nouvel écart est nul !

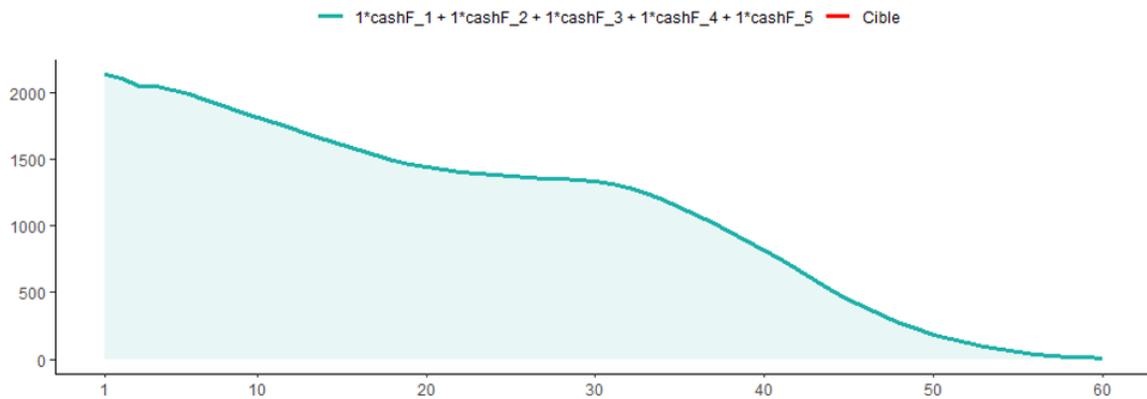


FIGURE 5.8 – La cible et de la combinaison des cashflows à la cinquième itération

L'algorithme possède deux critères d'arrêt naturels :

- Tous les cashflows ont été sélectionnés
- Toutes les corrélations restantes sont négatives

# Note de Synthèse

## Introduction

La directive Solvabilité 2 a généralisé l'utilisation des modèles ALM (*Asset Liability Management*) pour l'évaluation des provisions des produits d'assurance vie. Cette évolution réglementaire s'est imposée à l'ensemble du marché : mutuelles, institutions de prévoyance et compagnies d'assurance. La nouvelle norme comptable internationale (IFRS 17) de valorisation des contrats d'assurance vient imposer une utilisation encore plus poussée de ces modèles ALM.

Le modèle ALM historiquement utilisé par Allianz France est un modèle en Flexing nommé ALIM. Cette méthode permet d'avoir rapidement des résultats, mais à une maille assez grossière. Cependant, avec l'entrée en vigueur de la norme IFRS 17 et sa nouvelle définition de segmentation de portefeuille, le modèle ALIM n'est plus en mesure de répondre aux besoins de la compagnie. La solution trouvée par Allianz pour s'adapter à cet environnement changeant a été de développer un nouveau modèle ALM dit en Full ALM nommé LSIM. Celui-ci ayant l'avantage de répondre aux nouvelles contraintes imposées par la norme IFRS 17, puisqu'il permet d'effectuer des projections à une maille plus fine.

Toutefois, l'un des défis majeurs des modèles Actif-Passif en Full ALM est le temps nécessaire pour obtenir les résultats. En effet, pour des portefeuilles avec un grand nombre de contrats, le nombre de calculs peut vite devenir considérable : le temps et le volume disque alloués deviennent alors des problématiques opérationnelles non négligeables.

Conformément au Règlement Délégué publié par la commission européenne en janvier 2015, il est possible de faire des regroupements de polices dès lors que les polices considérées et les risques sous-jacents sont similaires et que l'utilisation de ces regroupements de polices (model points) ne fausse pas les résultats. L'objet de cette étude réside dans la construction de ces Model Points. En d'autres termes, il s'agit de transformer le portefeuille assurantiel des périmètres gérés par Allianz France en de nouvelles tables de données représentant le plus fidèlement possible ces portefeuilles. Pour de grands portefeuilles, le regroupement « manuel » des contrats n'est pas optimal. Une

première méthode d'agrégation a été déjà proposée. Celle-ci est basée sur une approche d'optimisation linéaire. Mais sa configuration actuelle ne permet pas d'atteindre les résultats souhaités par Allianz. Nous étudierons donc dans quelle mesure cette approche peut être améliorée. Dans l'éventualité d'un échec, nous proposerons une méthodologie fondée sur des algorithmes de classification statistique.

## Description du portefeuille

Notre portefeuille de contrats est initialement constitué de plusieurs contrats d'épargne retraite individuels.

Dans le cadre d'un contrat d'assurance vie de type épargne retraite individuel, l'assuré effectue à la souscription un premier versement. Ce capital initial est investi sur différents supports d'investissement dont le rendement est fonction des marchés financiers. L'épargne est capitalisée en fonction de ces rendements selon le support d'investissement. L'assureur doit être en mesure, à tout instant, de verser une somme égale à la valeur de l'épargne à l'assuré dans le cas où celui-ci décide de racheter avant le terme du contrat. Ou encore dans l'objectif de percevoir le capital de son assurance vie, l'assuré peut opter pour une sortie en rentes viagères. Ce mode de sortie consiste à transformer le capital en un revenu garanti et régulier jusqu'à son décès.

Chaque contrat de notre portefeuille est défini selon des éléments qui lui sont propres (propriétaire du contrat, âge, frais...). On dit que les contrats sont à la maille individuelle. Le nombre de contrats à la maille individuelle que comprend notre portefeuille est d'environ 1 millions de contrats.

Un premier traitement est réalisé sur le portefeuille de contrats individuels. La création des classes d'agrégation est une étape qui consiste à partitionner le portefeuille en plusieurs sous-portefeuilles. Une classe d'agrégation est une variable créée à partir d'autres variables suivant des règles spécifiques et qui va servir comme critère d'agrégation. Elle permet d'agréger des données ayant le même critère d'agrégation, c'est à dire que seuls les contrats appartenant à une même classes d'agrégation pourront être agrégés dans tout le processus.

La création des classes d'agrégation est une étape de réorganisation du portefeuille qui ne modifie pas la maille des contrats et qui n'engendre aucune perte d'information.

Ensuite le portefeuille partitionné subit un premier niveau d'agrégation. Le premier niveau d'agrégation est une étape qui consiste à regrouper au sein d'un même sous-portefeuille, les contrats ayant des caractéristiques identiques ou proche sous contrainte d'une perte d'information minimale dans les résultats et d'un volume de données post agrégation limité. Les données sont regroupées suivant différents critères d'agrégation

comme le réseau de distribution, la famille de taux minimum garanti, etc.

Pour chaque critère d'agrégation, les données quantitatives (primes, provisions mathématiques, nombre de contrats) sont ensuite additionnées et les durées et âges sont calculés par moyenne pondérée par les encours.

Cette étape fait passer les contrats de la maille individuelle à la maille EV pour *Embedded Value*, et réduit le taille des données d'un facteur 30 pour atteindre un peu moins de 35 000 contrats.

Le second niveau d'agrégation est l'étape ultime de la compression, et constitue le cœur de ce mémoire. Son but est de compresser davantage le portefeuille de contrats à la maille EV en regroupant cette fois les lignes ayant des caractéristiques différentes mais des évolutions similaires dans le temps.

L'idée principale de ce second niveau d'agrégation est de remplacer l'ensemble des polices d'assurance à la maille EV, par un ensemble plus réduit de lignes appelées Model Points de sorte que les provisions techniques (BEL) calculées contrat par contrat à la maille EV d'une part, puis de façon agrégée par des Model Points d'autre part, soient sensiblement identiques. Ainsi, plus les provisions sont proches, plus l'erreur induite par la création des Model Points est petite et plus on dira que notre jeu de Model Points est une représentation fidèle de l'ensemble des polices.

Deux méthodes d'agrégation sont présentées dans ce mémoire. La première, appelée Optimisation linéaire, est proposée par un cabinet de conseil et présente des insuffisances auxquelles nous tenterons de remédier. La seconde est une approche basée sur le clustering.

## L'optimisation linéaire

L'approche par l'optimisation linéaire ou Linear Optimization (LOP en anglais) est une approche basée sur l'algorithme nls (Non-Negative Least Squared). Cette approche utilise comme variable d'agrégation les cashflows de BEL.

Pour un sous-portefeuille donné, l'approche du LOP consiste à trouver le vecteur des poids (positifs)  $x$  solution du programme d'optimisation suivant :

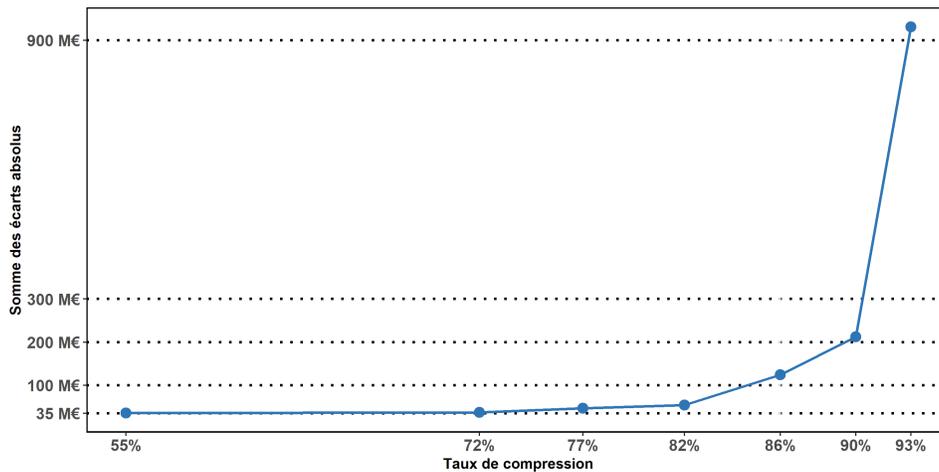
$$\min_x \frac{1}{2} \|b - Ax\|_2^2 \quad ; \quad x \geq 0$$

où  $A$  est la matrice des cashflows de BEL projetés contrat par contrat en mode passif seul. La matrice  $A$  est constituée de sorte que les cashflows de BEL relatifs à chaque contrat sont rangés en colonne. Et  $b$  est le vecteur des valeurs cibles de longueur

égale au nombre de lignes de la matrice  $A$  et dont chaque composante est obtenue par la somme sur toutes les colonnes de la ligne correspondante.

Ici,  $x \geq 0$  signifie que chaque composante du vecteur  $x$  devrait être positive ou nulle, et  $\|\cdot\|_2$  représente la norme Euclidienne.

A cet algorithme, nous avons intégré un paramètre permettant de modifier le critère d'arrêt et de rendre l'algorithme plus flexible. Ainsi, nous avons réussi à atteindre le niveau de compression (87%) exigé par Allianz, mais au prix d'un écart absolu de BEL se chiffrant à 125 M€.



Somme des écarts absolus en fonction du Taux de compression

## Approche par Clustering

L'approche par optimisation linéaire utilise les chroniques de cashflow futures comme variable d'agrégation. Cependant, ces cashflows dépendent du scénario financier considéré. Cette dépendance au scénario financier pourrait s'avérer problématique notamment dans un cadre de modélisation stochastique où plusieurs scénarios différents sont considérés. Il est donc indispensable de trouver une nouvelle variable d'agrégation qui décrive à la fois la dynamique future des contrats et qui ne soit pas tributaire du scénario considéré.

Dans le cadre d'un contrat épargne retraite individuelle, trois événements peuvent engendrer l'arrêt du contrat : (i) L'arrivée à échéance du contrat, (ii) Le décès de l'assuré et enfin (iii) Le rachat du contrat, qui correspond à une décision de la part de l'assuré de solder son compte.

Le rachat et le décès de l'assuré sont des événements probabilisables et sont susceptibles de se réaliser à tout moment pendant la durée de vie du contrat.

Pour prendre en compte ces évènements aléatoire sur la durée de vie des contrats, nous devons exprimer le Best Estimate of Liabilities comme suit :

$$BEL^{\mathbf{F}} = \sum_{t=0}^{T-1} PM_t^{\mathbf{F}} \times p_t + PM_T^{\mathbf{F}} \times p_T$$

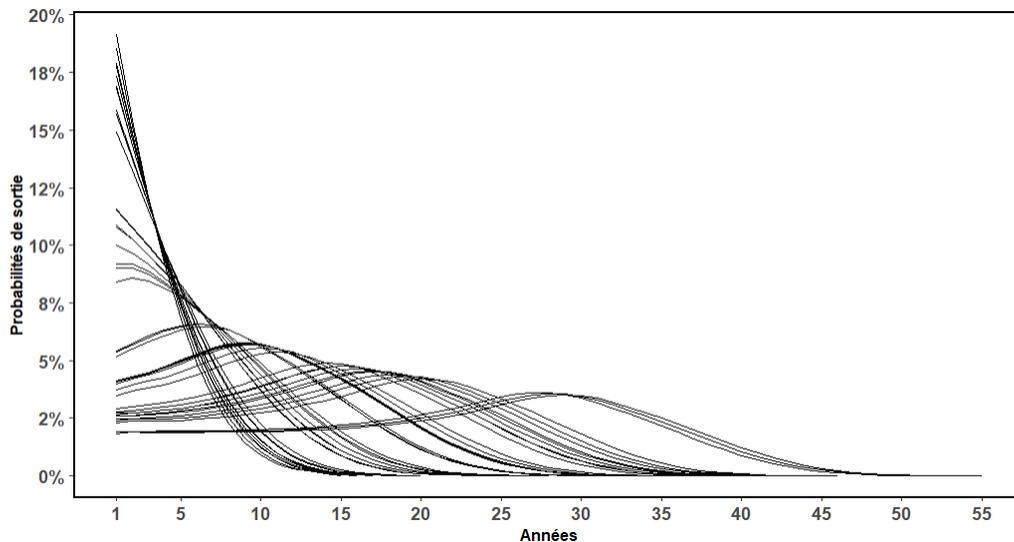
où  $PM_t^{\mathbf{F}}$  est la provision mathématique du contrat au cours de l'année de projection  $t$ , c'est à dire la valeur actuelle de l'épargne accumulée jusqu' à l'année  $t$ . Il s'agit du versement initial augmenté des éventuelles primes périodiques, capitalisé au taux de rendement des supports de comptes sur lesquelles l'épargne est investie, actualisé au taux Zéro-Coupon.

$p_t$  est la probabilité de sortie associée au contrat entre l'année  $t$  et  $t + 1$ , c'est-à-dire la probabilité jointe de ne pas être sorti au cours des années précédentes et de sortir durant l'année  $t$ .

Les probabilités de sortie sont des caractéristiques du contrat qui ne dépendent pas du scénario financier donc plus robuste dans une modélisation stochastique.

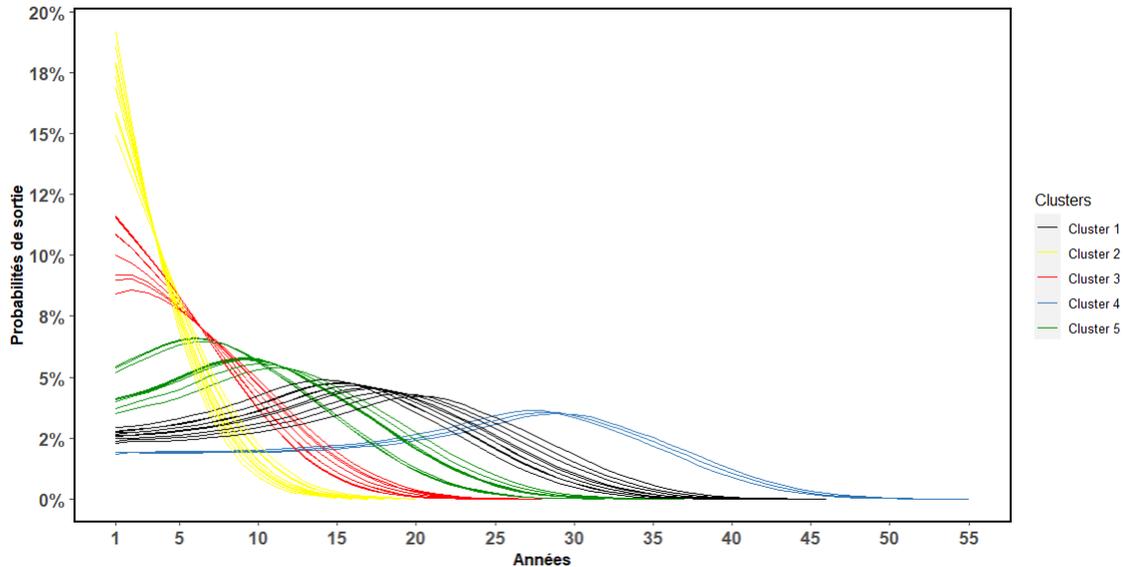
$PM_T$  et  $p_T$  correspondent respectivement à la valeur de l'épargne à l'échéance du contrat et à la probabilité d'arriver effectivement à l'échéance du contrat.

Ces probabilités de sortie forment des trajectoires qui permettent une visualisation graphique du portefeuille de contrats.



Probabilités de sortie associées à un sous-portefeuille

Une fois la variable d'agrégation définie, la première phase de l'approche par Clustering est la phase de classification qui consiste à utiliser des méthodes statistiques de classification afin de créer des groupes homogènes de contrats. La méthode pour laquelle nous avons opté est le kmeans.



Application du kmeans sur les prob. de sortie pour  $k = 5$

Une fois la classification terminée, la seconde étape de l'approche par clustering est la phase d'agrégation qui consiste en la définition d'un contrat représentatif ou "moyen" pour chacune des classes ou clusters. Cette tâche se résume à attribuer au "moyen" un âge et une ancienneté. Pour cela, nous avons proposé une première méthode dite naïve qui consiste à attribuer au contrat "moyen" les moyennes pondérées des âges et anciennetés des différents contrats du cluster. Nous avons ensuite proposé une généralisation de cette méthode avec une approche basée sur le développement en série de Taylor de la fonction qui à chaque vecteur de probabilités de sortie associe le couple d'âge et ancienneté du contrat correspondant.

Grâce à notre approche par clustering, et notre méthode d'agrégation à partir des développements en série de Taylor, nous avons réussi, pour le même taux de compression (87%) à réduire l'écart absolu de BEL à un peu plus de 53 M€ soit une réduction de plus de 57% par rapport au LOP.

# Executive Summary

## Introduction

The Solvency 2 directive has generalized the use of ALM models (Asset and Liability Management) for the evaluation of life insurance product provisions. This regulatory change has been imposed on the entire market : mutuals, insurance companies, etc. The new International Financial Reporting Standard (IFRS 17) that was issued by the International Accounting Standards Board for the valuation of insurance contracts requires even more extensive use of these ALM models.

The ALM model historically used by Allianz France is a Flexing model called ALIM. This method provides quick results, but at a much too aggregated and therefore approximate level. However, with the implementation of IFRS 17 and its new definition of portfolio segmentation, the ALIM model is no longer able to meet the company's needs. The solution found by Allianz to adapt to this changing environment was to develop a new ALM model called LSIM with the Full ALM approach. This approach has the advantage of meeting the new constraints imposed by IFRS 17, since it allows projections to be made at a finer scale.

However, one of the major challenges of Asset-Liability models in Full ALM approach is the time required to obtain the results. Indeed, for portfolios with a large number of contracts, the number of calculations can quickly become considerable : the time and disk volume allocated then become significant operational issues.

In accordance with the Delegated Regulation published by the European Commission in January 2015, it is possible to group policies as long as the considered policies and the underlying risks are similar and the use of these policy groupings (model points) does not distort the results. The purpose of this study is to construct these Model Points. In other words, it is a matter of transforming the insurance portfolio managed by Allianz France into new data tables representing these portfolios as faithfully as possible. For large portfolios, the "manual" aggregation of contracts is not optimal. A first aggregation method has already been proposed. It is based on a linear optimization approach. But its current configuration does not allow to reach the results desired by Allianz. We will therefore study to what extent this approach can be improved. In

the event of failure, we will propose a new methodology based on statistical clustering algorithms.

## Description of our portfolio

Our portfolio of contracts is initially made up of several individual savings contracts.

Within the framework of a life insurance contract of the individual savings type, the insured makes an initial payment at the time of subscription. This initial capital is invested in different investment vehicles whose return depends on the financial markets. The savings are capitalized according to these returns depending on the investment medium. The insurer must be able, at any time, to pay a sum equal to the value of the savings to the insured in the event that the latter decides to surrender before the end of the contract. Or, in order to receive the capital of his life insurance, the insured can opt for an exit in life annuities. This method of withdrawal consists of transforming the capital into a guaranteed and regular income until his death.

Each contract in our portfolio is defined according to its own elements (contract owner, age, fees, etc.). The contracts are said at the individual level. The number of individual contracts in our portfolio is approximately 1 million.

A first processing is performed on the portfolio of individual contracts. The creation of aggregation classes is a step that consists in partitioning the portfolio into several sub-portfolios. An aggregation class is a variable created from other variables according to specific rules and which will serve as an aggregation criterion. It allows to aggregate data having the same aggregation criterion, i.e. only the contracts belonging to the same aggregation class can be aggregated in the whole process.

The creation of aggregation classes is a portfolio reorganization step that does not modify the contract level and does not result in any loss of information.

Then the partitioned portfolio undergoes a first level of aggregation. The first level of aggregation is a step that consists of grouping together within the same sub-portfolio, contracts with identical or similar characteristics under the constraint of minimal loss of information in the results and a limited volume of post-aggregation data. The data is grouped according to different aggregation criteria such as the distribution network, the minimum guaranteed rate family, etc.

For each aggregation criterion, the quantitative data (premiums, mathematical reserves, number of contracts) are then added together and the durations and ages are calculated as an average weighted by the outstanding amounts.

This step moves the contracts from the individual level to the EV level for *Embedded Value*, and reduces the size of the data by a factor of 30 to just under 35,000 contracts.

The second level of aggregation is the final stage of compression, and is the heart of this paper. Its purpose is to further compress the portfolio of contracts at the EV level by grouping together lines with different characteristics but similar dynamics over time.

The main idea of this second level of aggregation is to replace the set of insurance policies at the EV level by a smaller set of lines called Model Points so that the technical reserves (BEL) calculated contract by contract at the EV level on the one hand, and calculated with Model Points on the other hand, are more or less identical. Thus, the closer the reserves are, the smaller the error induced by the creation of the Model Points and the more we can say that our set of Model Points is a faithful representation of all the policies.

Two aggregation methods are presented in this paper. The first one, called Linear Optimization, is proposed by a consulting firm and has some shortcomings that we will try to remedy. The second one is an approach based on clustering.

## The Linear OPTimization approach

The Linear OPTimization (LOP) approach is an approach based on the nls (Non-Negative Least Squared) algorithm. This approach uses the BEL cash flow as an aggregation variable.

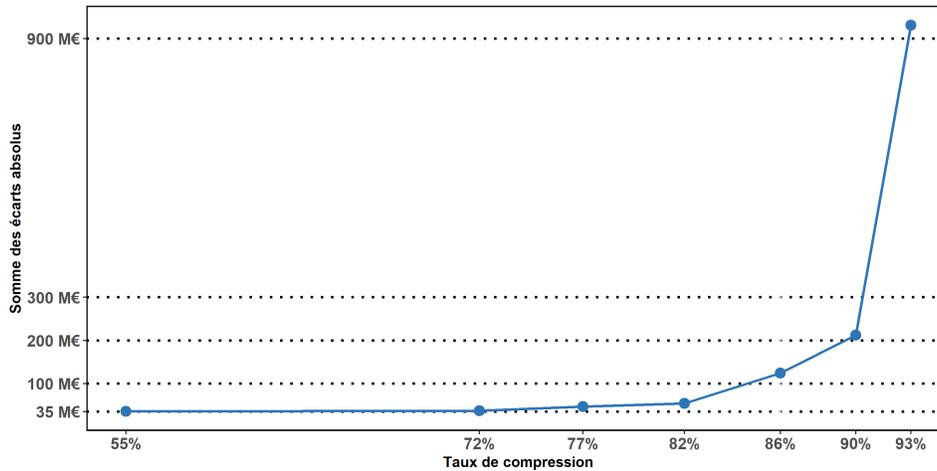
For a given sub-portfolio, the LOP approach consists in finding the vector of (non-negative) weights  $x$  solution of the following optimization program :

$$\min_x \frac{1}{2} \|b - Ax\|_2^2 \quad ; \quad x \geq 0$$

where  $A$  is the matrix of BEL cash flows projected on a contract-by-contract basis in passive-only mode. The matrix  $A$  is constructed so that the BEL cash flows for each contract are arranged in columns. And  $b$  is the vector of target values of length equal to the number of rows of matrix  $A$  and each component of which is obtained by summing over all the columns of the corresponding row.

Here,  $x \geq 0$  means that each component of the vector  $x$  should be non-negative, and  $\|\cdot\|_2$  represents the Euclidean norm.

To this algorithm, we have integrated a parameter that allows us to modify the stopping criterion and to make the algorithm more flexible. In this way, we were able to achieve the level of compression (87%) required by Allianz, but at the cost of an absolute BEL deviation of 125 M€.



Sum of the absolute deviations as a function of the compression ratio

## Clustering approach

The linear optimization approach uses the future cash flows as the aggregation variable. However, these cash flows depend on the considered financial scenario. This dependence on the financial scenario could be problematic especially in a stochastic modeling framework where several different scenarios are considered. It is therefore essential to find a new aggregation variable that describes both the future dynamics of the contracts and that is not dependent on the scenario considered.

In the context of an individual savings contract, three events can lead to the termination of the contract : (i) The maturity of the contract, (ii) The death of the insured and finally (iii) The surrender of the contract, which corresponds to a decision on the part of the insured to settle his account.

The surrender and death of the insured are probable events and are likely to occur at any time during the life of the contract.

To take into account these random events over the life of the contracts, we must express the Best Estimate of Liabilities as follows

$$BEL^{\mathbf{F}} = \sum_{t=0}^{T-1} PM_t^{\mathbf{F}} \times p_t + PM_T^{\mathbf{F}} \times p_T$$

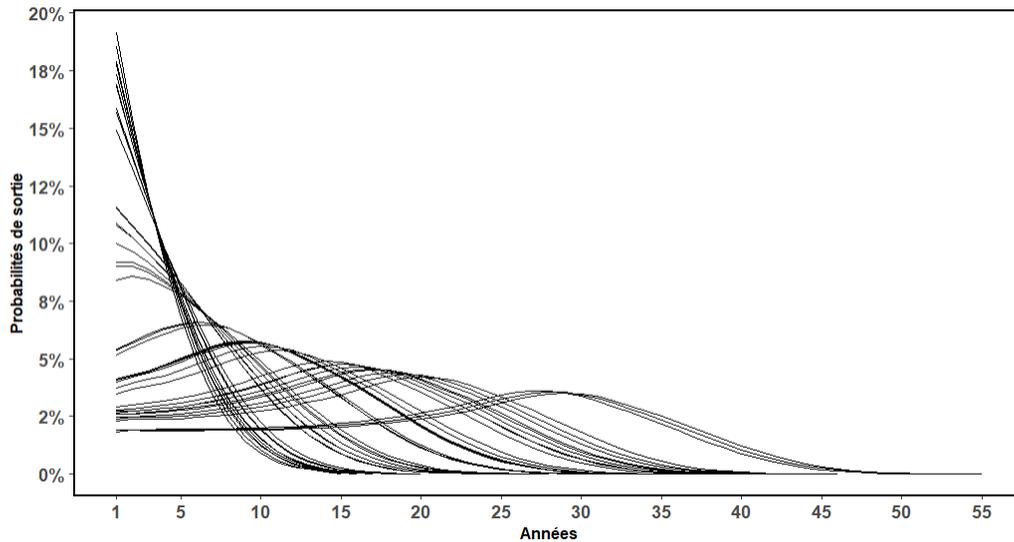
where  $PM_t^{textbf{F}}$  is the mathematical reserve of the contract during the projection year  $t$ , i.e. the present value of the savings accumulated up to year  $t$ . This is the initial payment plus any periodic premiums, capitalized at the rate of return on the account media in which the savings are invested, discounted at the Zero-Coupon rate.

$p_t$  is the probability of exit associated with the contract between year  $t$  and  $t + 1$ , i.e., the joint probability of not having exited in previous years and of exiting in year  $t$ .

Probabilities of exit are characteristics of the contract that do not depend on the financial scenario and are therefore more robust in a stochastic model.

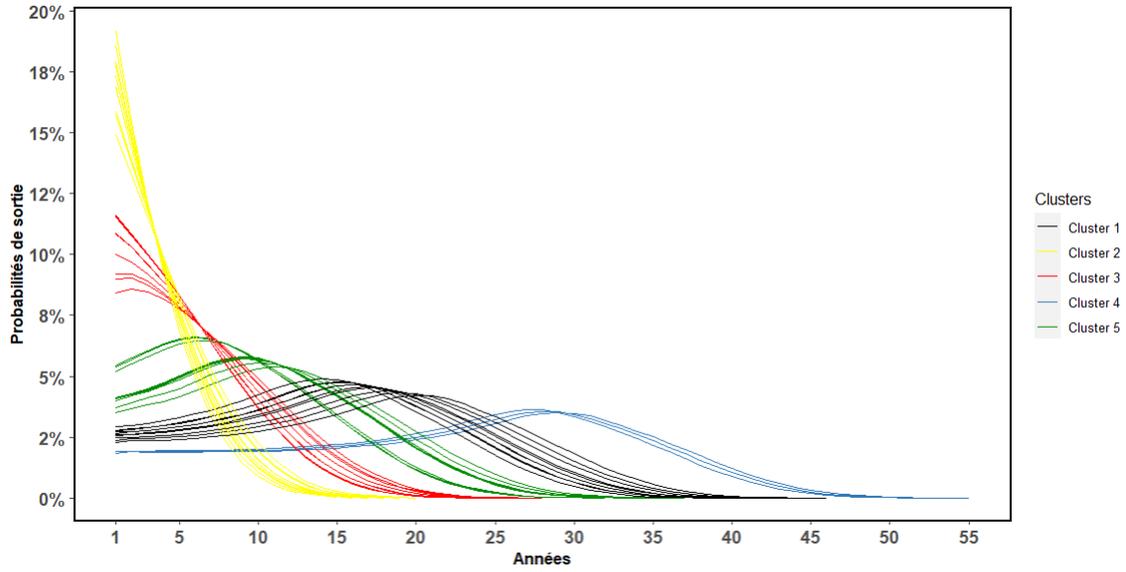
$PM_T$  and  $p_T$  correspond respectively to the value of the savings at the end of the contract and to the probability of actually reaching the end of the contract.

These probabilities of exit form trajectories that allow a graphical visualization of the contract portfolio.



Probabilities of exit associated with a sub-portfolio

Once the aggregation variable is defined, the first phase of the clustering approach is the classification phase which consists in using statistical classification methods to create homogeneous groups of contracts. The algorithm we have chosen is the kmeans.



Application of kmeans on the prob. of exit for  $k = 5$

Once the classification is completed, the second step of the clustering approach is the aggregation phase, which consists in defining a representative or "average" contract for each of the clusters. This task consists in assigning to this "average" contract an age and a seniority. To do this, we proposed a first so-called naive method which consists in attributing to the "average" contract the weighted averages of the ages and seniority of the different contracts of the cluster. We then proposed a generalization of this method with an approach based on the development in Taylor series of the function which associates to each vector of probabilities of exit the age and seniority pair of the corresponding contract.

Thanks to our clustering approach, and our aggregation method based on Taylor series developments, we succeeded, for the same compression rate (87%), in reducing the absolute BEL gap to just over 53 M€, i.e. a reduction of more than 57% compared to the LOP.