

Mémoire présenté devant l'Université de Paris-Dauphine  
pour l'obtention du Certificat d'Actuaire de Paris-Dauphine  
et l'admission à l'Institut des Actuares

le 08/04/2021

Par : Valentin CERISIER

Titre : ORSA : Application de méthodes de machine learning dans le calcul de la solvabilité infra-annuelle

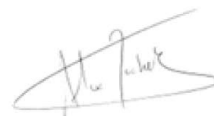
Confidentialité : ☒ Non ☐ Oui (Durée : ☐ 1 an ☐ 2 ans)

---

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

*Membres présents du jury de l'Institut  
des Actuares :*

*Entreprise :*  
Nom : Mazars  
Signature :



*Membres présents du Jury du Certificat  
d'Actuaire de Paris-Dauphine :*

*Directeur de Mémoire en entreprise :*  
Nom : Abdelwahab MAHJOUB  
Signature :

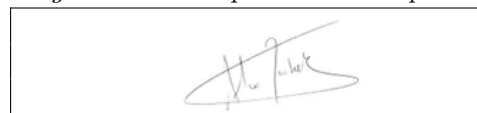


---

*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*



*Signature du candidat*

*Bibliothèque :*





## Résumé

---

L'ORSA (*Own Risk and Solvency Assessment*) est une étude qualitative et prospective que les assureurs européens, soumis à la directive Solvabilité II, sont supposés effectuer à chaque modification dans leur stratégie, et donc plusieurs fois par an. Elle a pour objectif d'inciter les assureurs à mieux connaître et appréhender leurs risques et à s'assurer que leurs exigences de solvabilité sont respectées à tout moment. Notre mémoire porte sur l'application d'une approche *machine learning* pour pallier les contraintes de temps et de puissance de calcul nécessaires à la projection d'indicateurs de solvabilité dans le cadre d'une étude de type ORSA. Notre étude se concentre sur le calcul d'un indicateur bilantiel, le *Best Estimate* ou BE, d'une compagnie d'assurance-vie commercialisant des produits d'épargne en euros.

Le *Best Estimate* est calculé grâce à un modèle de gestion Actif/Passif développé par Mazars. Celui-ci prend en entrées 1000 scénarios économiques projetés sur 30 ans, des variables relatives à la situation et la stratégie de l'assureur au moment de l'évaluation et produit en sortie le BE. Cet indicateur est approché par méthode de Monte Carlo, en faisant la moyenne sur les 1000 scénarios des provisions que l'assureur doit posséder afin d'honorer ses engagements. Nous appellerons par la suite ces provisions des BE par scénario. Les 1000 scénarios sont obtenus grâce à un GSE (Générateur de Scénarios Economiques) prenant en entrée la courbe de taux sans risque fournie par l'EIOPA (*European Insurance and Occupational Pensions Authority*).

Ainsi, notre étude commence par l'analyse méticuleuse du modèle ALM afin de sélectionner les variables ayant, selon nous, le plus grand pouvoir explicatif sur le BE. Les variables sont sélectionnées avec attention car les résultats du mémoire en dépendent. De plus, afin de réduire le nombre de variables explicatives de nos bases d'apprentissage et éviter l'information économique redondante, nous effectuons une analyse des corrélations et des Analyses en Composante Principales (ACP). Nous utilisons ensuite ces variables pour générer nos bases de données. Pour ce faire, nous choisissons des plages sur lesquelles faire varier nos entrées de l'outil ALM, puis nous récupérons les 1000 BE par scénario et le BE en sortie. Nous obtenons ainsi une base d'apprentissage pour nos BE par scénario et une pour nos BE.

Avant de commencer l'étude *machine learning*, nous nous familiarisons avec les données via une exploration descriptive et statistique. Celle-ci nous permet d'identifier des *outliers* dans nos données et mettre en avant les variables pouvant les expliquer. De plus, nous pouvons intuitivement les variables qui semblent expliquer le plus la variabilité du BE.

S'ensuit l'étude *machine learning*. Celle-ci n'est pas concluante sur les BE par scénario avec des modèles qui peinent à capter la complexité des mécanismes du modèle ALM. En revanche, l'étude sur les BE est très satisfaisante et nous sélectionnons finalement le modèle XGBoost pour la suite de notre étude.

Le mémoire se conclut par des applications à des études de type ORSA. Une première application consiste à utiliser la méthode d'interprétabilité *a posteriori* de modèles "boîtes noires" appelée LIME. Elle permet de montrer que les prédictions du modèle XGBoost sont fiables et met en exergue le potentiel d'une telle méthode comme instrument de pilotage de la stratégie de l'assureur. Notre seconde application consiste à projeter le BE de notre compagnie fictive sur 3 années. Nous définissons pour cela des scénarios "Monde Réel" permettant de faire évoluer dans le temps notre compagnie et nous comparons chaque année notre BE prédit au BE calculé par l'outil ALM. Nous considérons un scénario central et 4 scénarios stressés simulant la crise du Covid-19. Nous montrons ainsi que notre approche permet un gain de temps très important, pour une erreur minimale sur le scénario central. Concernant les scénarios COVID-19 plus adverses, l'erreur est plus grande et nous pensons que deux approches sont à envisager afin d'améliorer les prédictions sur la projection. La première consisterait à recalibrer le modèle XGBoost chaque année de projection et la seconde à compléter la base de données pour apporter plus de variabilité. Finalement, nous ne projetons que le BE et non le SCR, ce qui limite l'exhaustivité de notre étude dans le cadre de l'ORSA. Des travaux équivalents pourraient être envisagés afin de compléter notre approche.

## Abstract

---

ORSA (*Own Risk and Solvency Assessment*) is a qualitative and prospective study that European insurers, subject to the Solvency II directive, are supposed to conduct each time there is a change in their strategy, and therefore several times a year. Its aim is to encourage insurers to better know and understand their risks and to ensure that their solvency requirements are met at all times. Our thesis focuses on the application of a machine learning approach to overcome the time and computing power constraints required to project solvency indicators in the context of an ORSA-type study. Our study focuses on the calculation of a balance sheet indicator, the Best Estimate or BE, of a life insurance company marketing savings products in euros.

The Best Estimate is calculated using an Asset and Liability Management model developed by Mazars. This model takes as inputs 1000 economic scenarios projected over 30 years, variables relating to the situation and strategy of the insurer at the time of the valuation and produces the BE as an output. This indicator is approached using the Monte Carlo method, by averaging over the 1000 scenarios the provisions that the insurer must have in order to honour its commitments. We will subsequently refer to these reserves as BE by scenario. The 1000 scenarios are obtained thanks to a ESG (Economic Scenario Generator) using the risk-free yield curve provided by EIOPA (*European Insurance and Occupational Pensions Authority*) as an input.

Thus, our study begins with a meticulous analysis of the ALM model in order to select the variables that we believe have the greatest explanatory power on the BE. The variables are carefully selected because the results of the paper depend on them. Moreover, in order to reduce the number of explanatory variables in our databases and to avoid redundant economic information, we perform a correlation analysis and Principal Component Analysis (PCA). We then use these variables to generate our databases. To do so, we choose ranges on which we vary our variables as inputs to the ALM model and retrieve the 1000 BE per scenario and the BE as output. We thus obtain two databases, one for our BE by scenario and one for our BE.

Before starting the machine learning study, we familiarise ourselves with the data through a descriptive and statistical exploration. This allows us to identify outliers in our data and highlight the variables that could explain them. In addition, we can intuit the variables that seem to explain the most the variability of the BE.

This is followed by the machine learning study. This study is not conclusive on the BE per scenario with models which have difficulties to capture the complexity of the mechanisms of the ALM model. On the other hand, the study on the BE is very satisfactory and we finally selected the XGBoost model for the continuation of our study.

The dissertation concludes with applications to ORSA-type studies. A first application consists in using a model-agnostic method called LIME, which allows us to show that the predictions of the XGBoost model are reliable and highlights the potential of such a method as a tool for steering the insurer's strategy. Our second application consists in projecting the BE of our fictitious company over 3 years. We define "Real World" scenarios allowing us to make our company evolve over time and we compare each year our predicted BE to the BE calculated by the ALM tool. We consider a central scenario and 4 stressed scenarios simulating the Covid-19 crisis. This application shows that our approach allows us a very important time saving, for a small error. Concerning the more adverse COVID-19 scenarios, the error is larger and we believe that two approaches should be considered in order to improve the predictions on the projection. The first would be to recalibrate the XGBoost model each projection year and the second would be to complete the database to provide more variability. Finally, we only project the BE and not the SCR, which limits the completeness of our study in the ORSA framework. Equivalent work could be envisaged to complete our approach.

*Keywords : ORSA; Machine Learning, Best Estimate, Euro savings, LIME.*

# Note de Synthèse

La projection du *Best Estimate* (ou BE) sur un horizon temporel allant de 3 à 5 ans fait partie des exigences de l'ORSA et plus largement du pilier 2 de la directive Solvabilité II. Cependant, cette projection par le biais d'un modèle de gestion Actif/Passif est coûteuse en puissance et en temps de calcul. Ces études ne sont donc pas menées aussi fréquemment que la directive le préconise. Ce mémoire vient apporter une réponse à cette problématique en étudiant une méthode alternative, ayant déjà fait ses preuves dans de nombreux domaines et sur des sujets du même acabit : le *machine learning*.

Ce mémoire se concentre sur le calcul du BE d'un assureur-vie proposant des contrats d'épargne en euros. Cet assureur fictif cherche à répondre aux exigences d'une étude de type ORSA, c'est-à-dire à mieux comprendre ses risques et anticiper voire piloter les évolutions de son BE sur plusieurs années, en se détachant des contraintes temporelles et de puissance de calcul.

## Modélisation ALM et sélection des variables explicatives

Afin de mener à bien cette étude, nous utilisons le modèle ALM développé par Mazars. Il prend en entrée 1000 scénarios économiques projetés sur 30 ans, des variables relatives à la situation et la stratégie de l'assureur au moment de l'évaluation. Il produit en sortie le BE. Cet indicateur est approché par méthode de Monte Carlo, en faisant la moyenne sur les 1000 scénarios des provisions que l'assureur doit posséder afin d'honorer ses engagements. Nous appellerons par la suite ces provisions des BE par scénario. Les 1000 scénarios sont obtenus grâce à un GSE (Générateur de Scénarios Economiques) prenant en entrée la courbe de taux sans risque fournie par l'EIOPA (*European Insurance and Occupational Pensions Authority*).

Notre travail débute par l'analyse des mécanismes du modèle ALM dans le but de sélectionner les variables explicatives qui, selon nous, permettent au mieux de prédire le BE. Ces variables sont à la fois relatives à l'environnement économique dans lequel évolue notre compagnie d'assurance, mais aussi relatives à sa stratégie et sa situation au moment du calcul. De plus, notre étude *machine learning* s'est tournée vers le BE ainsi que le BE par scénario, nécessitant deux bases d'apprentissage et une adaptation dans le choix de nos variables économiques.

Cette étape préliminaire de sélection de variables a été effectuée précautionneusement, nous permettant de générer les deux bases de données d'apprentissage. Nous présentons dans les tableaux 1 et 2 les variables explicatives sélectionnées.

TABLE 1: Les variables économiques sélectionnées

Nom de la variable	Code de la variable
Les rendements actions	Taux_Action_Y_i (i allant de 1 à 30)
Les rendements immobiliers	Taux_Immo_Y_i (i allant de 1 à 30)
Les taux zéro-coupon	Taux_ZC_Y_i (i allant de 1 à 30)
Les taux des concurrents sur le marché	Taux_Concu_Y_i (i allant de 1 à 30)
La courbe de taux EIOPA	Taux_Y_i (i allant de 1 à 30)

TABLE 2: Les variables bilantielles et stratégiques sélectionnées

Nom de la variable	Code de la variable
L'âge moyen du MP	Age
L'ancienneté moyenne du MP	Anciennete
Le Taux Minimum Garanti	TMG
Pourcentage de plus ou moins-values latentes	Plus_Moins_VL
Le pourcentage de PPE en fonction des PM	Taux_PPE
Le pourcentage d'obligation dans l'allocation cible	Oblig
Le pourcentage d'action dans l'allocation cible	Action
Taux de frais de gestion	Frais_Gestion

De plus, afin de réduire le nombre de variables explicatives de nos bases d'apprentissage et éviter l'information économique redondante, nous effectuons une analyse des corrélations et des Analyses en Composante Principales (ACP). A l'issue de ces études, nous avons sélectionné les variables qui nous semblent les plus à même d'expliquer les variations du BE.

### Construction des bases d'apprentissage et exploration descriptive

Il s'ensuit la création des bases de données. Pour ce faire, nous faisons varier nos variables sélectionnées en entrée du modèle ALM, sur des plages préalablement définies, en récupérant en sortie 1000 BE par scénario et un BE à chaque variation. Des contraintes de temps et de puissance de calcul ont limité le pas de variation sur ces plages, mais nous obtenons finalement une base de 12 375 000 BE par scénario, soit 12 375 BE.

Notre étude se poursuit par une exploration descriptive et statistique de nos données afin de mieux les appréhender et déceler de potentielles anomalies. Nous n'avons cependant pas de valeurs manquantes, ni de valeurs de variables extrêmes car nous avons généré nous même les bases de données. Notre étude se concentre sur le BE et les figures ci-dessous nous permettent de le visualiser. Sur le *boxplot* des BE (figure 1), nous remarquons la présence d'*outliers*, c'est-à-dire des BE dont les valeurs sont plus grandes que les autres et pouvant atteindre l'ordre de 1851 M€. A titre comparatif, nos BE ont une valeur moyenne de 661 M€ et les BE par scénario 663 M€. L'histogramme de nos BE (figure 2) est aussi représenté ci-dessous et nous permet de juger de la distribution nos BE.

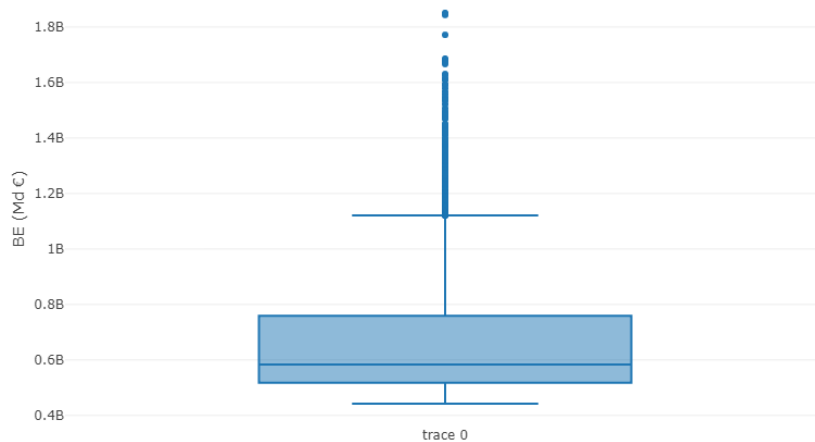


FIGURE 1: Boxplot des BE

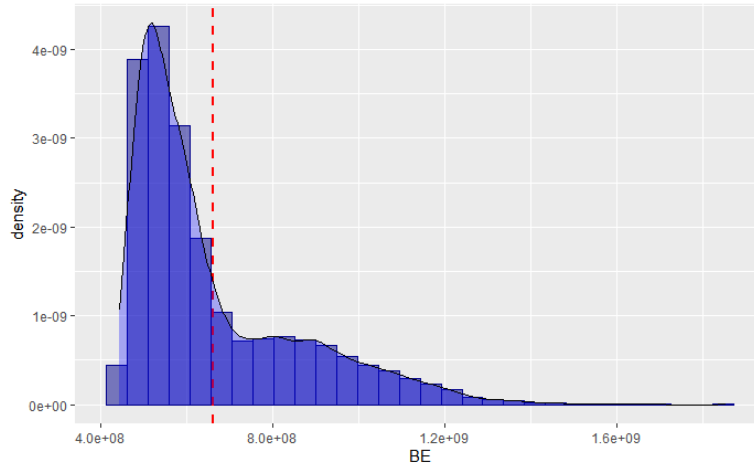


FIGURE 2: Histogramme et densité des BE

Pour expliquer ces BE extrêmes, nous représentons les distributions des BE en fonction des modalités des variables. Cela nous permet d'intuiter les variables pouvant expliquer les comportements extrêmes du BE et plus généralement présentant le plus gros potentiel explicatif. Nos études montrent que les valeurs extrêmes de BE sont toutes pour un TMG de 3,5%. Cependant, la distribution de nos BE est très étendue pour cette valeur de TMG et nous n'observons pas une tendance à la hausse du BE en fonction du TMG. Nous en concluons ainsi que le TMG n'est pas l'unique variable pouvant expliquer la présence d'*outliers* dans notre base de données. Encouragé par ces résultats, nous poursuivons en représentant le BE en fonction de deux variables sur une carte de fréquentation (ou *heatmap*), la valeur moyenne de tous les BE ayant comme paramètre notre couple de variables. Cette étude nous permet d'étudier les effets croisés. Les résultats les plus probants sont obtenus pour le couple de variable "TMG"/"Taux\_Y\_1", c'est-à-dire une variable symbolisant la courbe de taux. Nous constatons qu'un environnement de taux bas, combiné à un TMG élevé du MP, favorise les BE les plus élevés. Des études équivalentes sont menées sur toutes nos variables et nous observons une variabilité significative de la distribution de nos BE en fonction des modalités des variables suivantes : "TMG", "Age", "Oblig" ainsi que la Courbe de taux en général.

### Etude *machine learning* et approximation du BE

Nous sélectionnons 4 modèles pour l'étude *machine learning* : le modèle de régression linéaire, Random Forest, XGBoost et les réseaux de neurones. Nous adoptons une démarche d'optimisation des hyperparamètres de chaque modèle par la méthode des *k-fold cross validation*. Ainsi, nous obtenons les modèles les plus performants dans le sens où nous minimisons leurs biais et variances.

Après optimisation des hyperparamètres et entraînement des modèles sur les deux bases, nous avons pu analyser leurs performances sur les bases tests et les bases complètes. Les résultats de nos modèles sur la base des BE par scénario n'ont pas été probants. Nos métriques de performances montrent que les erreurs de prédictions sont élevées sur la base de test et ainsi que sur la base complète. Nous obtenons notamment un RMSE de 566 M€ pour notre modèle le plus performant, Random Forest, ce qui est du même ordre de grandeur que la valeur moyenne des BE par scénario. C'est un indicateur d'erreurs de prédictions importantes. Ce constat est vérifié en confrontant nos prédictions avec les observations. La représentation des prédictions en fonction de nos BE par scénario devrait idéalement être linéaire, ce qui n'est pas le cas, avec des écarts entre observations et prédictions non négligeables.

Aux vues des performances non satisfaisantes de l'étude sur les BE par scénario, nous avons décidé de nous concentrer sur l'apprentissage des BE. Les résultats de cette étude sont eux très satisfaisants, comme le montre le tableau 3 ci-dessous :

TABLE 3: Les résultats de nos modèles sur la base complète

Modèle	RMSE (en M €)	MAE (en M €)	MAPE (en %)
<b>Modèle Linéaire</b>	93,98	65,33	0,1
<b>Random Forest</b>	31,70	16,03	0,02
<b>XGBoost</b>	24,38	10,71	0,01
<b>Réseau de Neurones</b>	25,51	10,81	0.012

Nous remarquons que XGBoost et le réseau de neurones sont les plus performants avec une MAPE, c'est-à-dire la moyenne des erreurs absolue en pourcentage de 0.01 %, ce qui est remarquable. Afin de mieux juger de la performance de nos modèles, nous traçons les observations en fonction du BE. Pour les deux modèles, les résultats sont très satisfaisants, comme on peut l'observer sur la figure 3 dans le cas de XGBoost, avec des prédictions très proches des observations. Nous remarquons cependant que la qualité des prédictions se détériore pour les plus grands BE. Nous choisissons finalement XGBoost comme modèle pour notre application à l'ORSA.

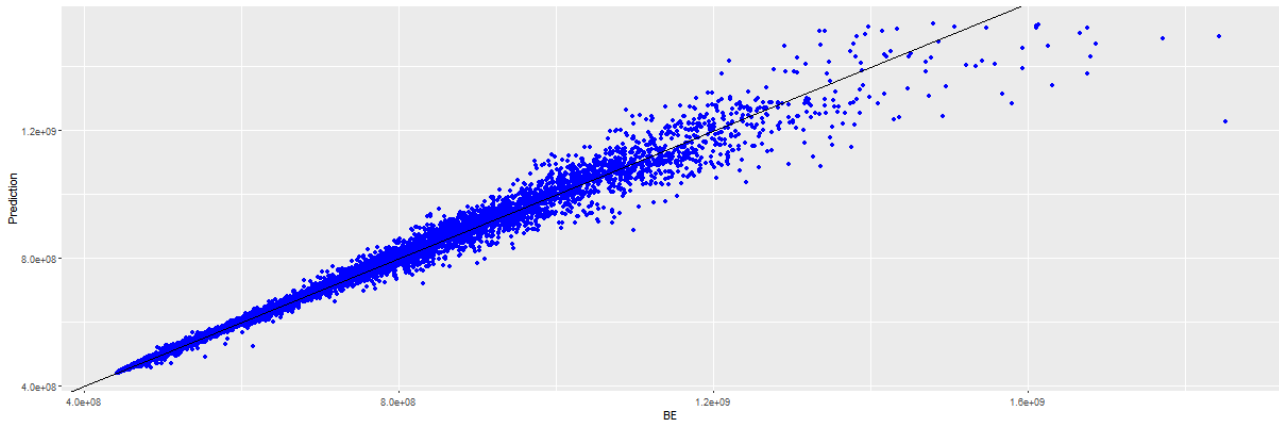


FIGURE 3: BE et prédictions du XGBoost

A ce stade, nous pouvons d'ores et déjà percevoir l'intérêt de l'approximation par le modèle XGBoost, le calcul du BE étant instantané, alors que 10 minutes sont nécessaires avec le modèle ALM.

## Applications à l'ORSA

Ce mémoire se clôture par la mise en application de notre étude *machine learning* à une étude de type ORSA. Notre première application vise à expliquer les sensibilités d'observations du BE à nos paramètres par des méthodes d'interprétabilité a posteriori des modèles dits "boîtes noires". Les objectifs sont doubles : montrer que nous pouvons avoir confiance dans les prédictions du modèle, indispensable dans un cadre réglementaire comme Solvabilité 2, puis exposer le potentiel d'un tel outil pour piloter la stratégie de l'assureur concernant le BE. Nous avons implémenté plus spécifiquement la méthode LIME, qui consiste à approximer localement une prédiction de notre modèle "boîte noire" par un modèle interprétable, un modèle linéaire généralisé dans notre étude, permettant ainsi d'expliquer nos prédictions. Nous obtenons des résultats comme représentés ci-dessous :



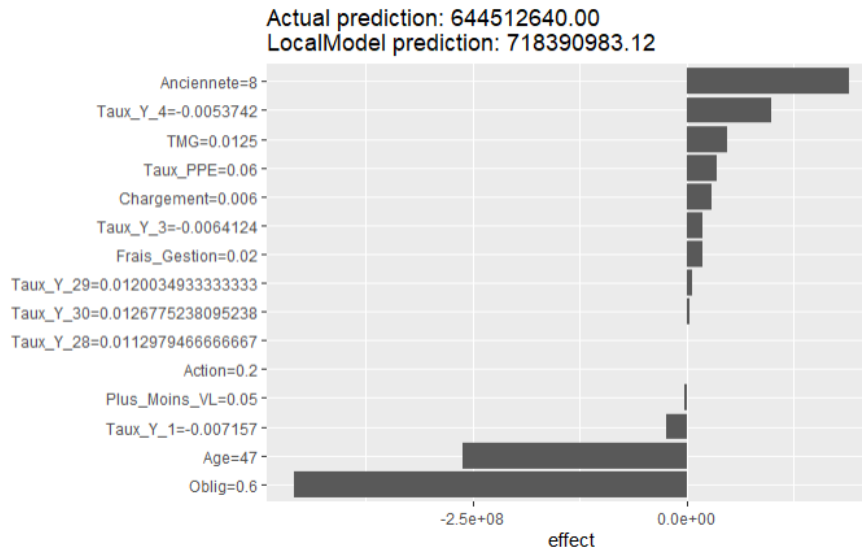


FIGURE 4: Explications d'un BE de valeur centrale par la méthode LIME

Sur l'exemple de la figure 4, nous expliquons la prédiction d'un BE central. Pour ce faire, nous présentons les variables pour lesquelles les coefficients du modèle linéaire approximant notre prédiction sont les plus importants. Ainsi, nous remarquons qu'une augmentation de la variable "Oblig", c'est-à-dire le pourcentage d'obligations ciblé dans notre portefeuille d'actif, a un impact négatif sur le BE. Ainsi, l'assureur peut choisir de modifier son allocation d'actif en augmentant la part des obligations, recherchant ainsi à faire baisser son BE.

Notre seconde application vise à projeter le BE de notre compagnie fictive sur 3 années, que nous considérons être le plan stratégique de notre assureur. Notre démarche consiste, dans un premier temps, à définir un scénario "Monde Réel" permettant de faire évoluer dans le temps notre compagnie et, dans un second temps, à comparer chaque année notre BE prédit au BE calculé par l'outil ALM.

Notre première projection selon cette démarche nous donne ainsi une erreur moyenne de 0,56%, l'erreur ayant une tendance croissante avec le temps de projection. Ceci s'explique par la déformation de la courbe de taux qui entraîne des prédictions de BE sur des combinaisons de paramètres sur lesquelles XGBoost n'a pas effectué l'apprentissage. Cela induit une erreur due à la variance de notre modèle. Cette application nous permet aussi de juger du temps de calcul économisé par cette approche *machine learning* : l'outil ALM met près d'une heure à projeter le BE sur les 3 années alors qu'en comparaison XGBoost ne met que 5 minutes.

Ces résultats étant encourageants, nous poursuivons cette étude en projetant notre compagnie dans l'environnement de crise sanitaire que nous connaissons actuellement. Pour ce faire, nous définissons 4 scénarios se distinguant par l'évolution des rendements actions et de la crise économique. Nous nous basons sur les évolutions de l'indice du CAC 40 pour définir nos scénarios qui se distinguent par leur forme d'évolution et leur point de chute à la date de fin de projection. Ainsi, nous modélisons deux scénarios d'évolution en forme U, avec une baisse des rendements actions qui restent sur un plateau bas 2 ans puis remontent et deux scénarios d'évolution en V avec une baisse de rendements actions mais qui remontent directement l'année suivante. Nous tâchons d'expliquer les variations de BE d'un scénarios à l'autre grâce aux projections d'indicateurs comme le niveau de PPE ou de plus-values. Nos résultats sur cette étude montrent des erreurs de prédiction à compter de 2020 plus grandes que lors de notre première application, de l'ordre de 5% d'erreur moyenne. Ces erreurs sont cette fois-ci conséquentes et symptomatiques d'un manque de variabilité dans notre base de données, d'autant plus sur un scénario choqué entraînant des grandes variations de nos variables.

## Conclusion

Nous jugeons cette étude très encourageante dans le sens où l'objectif est atteint : nous obtenons, en un temps largement inférieur qu'avec un outil ALM, une projection de nos BE sur 3 ans et avec une erreur minime sur le scénario central. Concernant les scénarios COVID-19, plus adverses, nous pensons que deux approches sont à envisager afin d'améliorer les prédictions sur la projection. La première consisterait à recalibrer le modèle XGBoost chaque année de projection et la seconde à compléter la base de données pour apporter plus de variabilité. Nous en concluons que notre modèle peut être utilisé pour de "petites" variations, ce qui reste prometteur. De plus, les modèles d'interprétabilité comme LIME sont des outils de pilotage pour l'assureur qui peut ainsi adapter sa stratégie à tout moment.

Notre approche présente des limites. Pour mener à bien cette étude, nous avons dû faire des hypothèses simplificatrices qui ne sont pas envisageables pour un assureur et qui font que notre base de données manque de variabilité, spécialement concernant l'environnement économique. En effet, dans le cadre de l'assurance vie, une des sources principales de variations du BE vient de l'environnement économique et de la capacité de l'assureur à tirer profit de son actif. Compléter la base de données en jouant sur le paramétrage du GSE et en rendant notre allocation d'actif plus diversifiée et représentative du marché serait donc une bonne piste d'amélioration. Cependant, ces hypothèses ont été motivées par des contraintes logistiques de temps et de puissance de calcul qu'un assureur a les moyens financiers et matériels de surmonter s'il veut développer une approche comme celle que nous proposons. Finalement, nous ne projetons que le BE et non le SCR ce qui limite l'exhaustivité de notre étude dans le cadre de l'ORSA. Des travaux équivalents sur cet indicateur pourraient être envisagés afin de compléter l'approche.

# Synthesis note

The projection of the Best Estimate (or BE) over a time horizon of 3 to 5 years is part of the ORSA (*Own Risk and Solvency Assessment*) requirements and more broadly of Pillar 2 of the Solvency II Directive. However, this projection by means of an Asset and Liability Management model is costly in terms of power and calculation time. These studies are therefore not carried out as frequently as the directive recommends. This thesis provides an answer to this problem by studying an alternative method, which has already proven itself in many fields and on subjects of the same nature: the machine learning.

This paper focuses on the calculation of the BE of a life insurer offering savings contracts in euros. This fictitious insurer seeks to meet the requirements of an ORSA-type study, i.e. to better understand its risks and anticipate or even pilot the evolution of its BE over several years, while detaching itself from the constraints of time and computing power.

## The ALM model and the selection of explanatory variables

In order to conduct this study, we use the ALM model developed by Mazars. It takes as inputs 1000 economic scenarios projected over 30 years, variables relating to the situation and strategy of the insurer at the time of the valuation. It produces as output the BE. This indicator is approximated by Monte Carlo method, by averaging over the 1000 scenarios the provisions that the insurer must have in order to meet its commitments. We will subsequently call these reserves "BE per scenario". The 1000 scenarios are obtained using a ESG (Economic Scenario Generator) using the risk-free yield curve provided by EIOPA (*European Insurance and Occupational Pensions Authority*) as an input.

Our work begins by analysing the mechanisms of the ALM model in order to select the explanatory variables that we believe allow us to better predict the BE. These variables are related to the economic environment in which our insurance company operates, but also to its strategy and its situation at the time of the calculation. In addition, our machine learning study is oriented towards the BE as well as BE per scenario, requiring two learning bases and an adaptation in the choice of our economic variables. This preliminary step of variable selection has been carefully carried out, allowing us to generate the two learning databases. We present in tables 4 and 5 the selected explanatory variables.

Table 4: The economical variables selected

Name of the variables	Code of the variables
Equity returns	Taux_Action_Y_i (i from 1 to 30)
Real estate returns	Taux_Immo_Y_i (i from 1 to 30)
The zero-coupon interest rates	Taux_ZC_Y_i (i from 1 to 30)
Competitors' market rates	Taux_Concu_Y_i (i from 1 to 30)
The EIOPA risk-free rates	Taux_Y_i (ii from 1 to 30)

Table 5: The balance sheet and strategic variables selected

Name of the variables	Code of the variables
The average age of the MP	Age
The mean seniority of the MP	Anciennete
The Minimum Rate Guaranteed	TMG
Percentage of unrealised gains and losses	Plus_Moins_VL
The percentage of PPE with regard to the PM	Taux_PPE
The percentage of bonds in the targeted allocation	Oblig
The percentage of equities in the targeted allocation	Action
Management fee rate	Frais_Gestion

In addition, in order to reduce the number of explanatory variables in our databases and to avoid redundant economic information, we perform correlation analysis and Principal Component Analysis (PCA). At the end of these studies, we have selected the variables that seem to us most likely to explain the variations in the BE.

### Construction of the databases and data mining

This is followed by the creation of databases. To do this, we vary our selected variables at the input of the ALM model, over previously defined ranges, by recovering at the output 1000 BE per scenario and one BE at each variation. Constraints of time and computing power limited the variation on these ranges, but we finally obtain a database of 12 375 000 BE per scenario, i.e. 12 375 BE.

Our study continues with a descriptive and statistical exploration of our data in order to better apprehend them and detect potential anomalies. However, we do not have missing values or extreme variable values because we generated the databases ourselves. Our study focuses on the BE and the figures below allow us to visualize it. On the boxplot of the BE (figure 5), we notice the presence of outliers, i.e. BE whose values are larger than the others and which can reach the order of €1851 million. As a comparison, our BE have an average value of €661 million and the BE per scenario €663 million. The histogram of our BE (figure 6) is also represented below and allows us to judge the distribution of our BE.

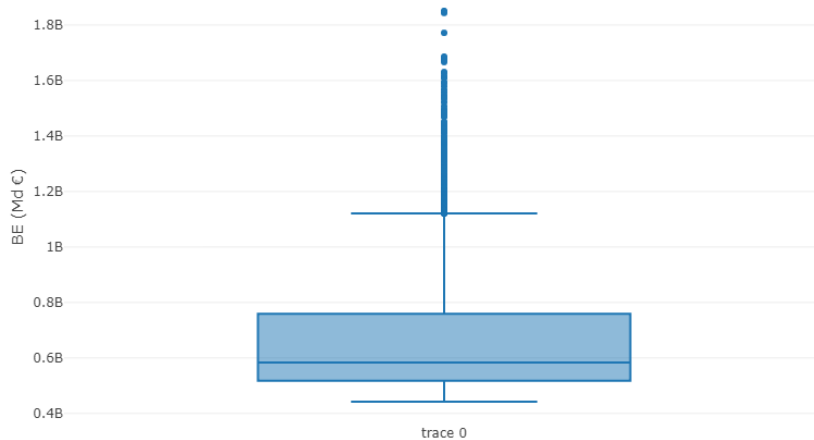


Figure 5: Boxplot of the BE

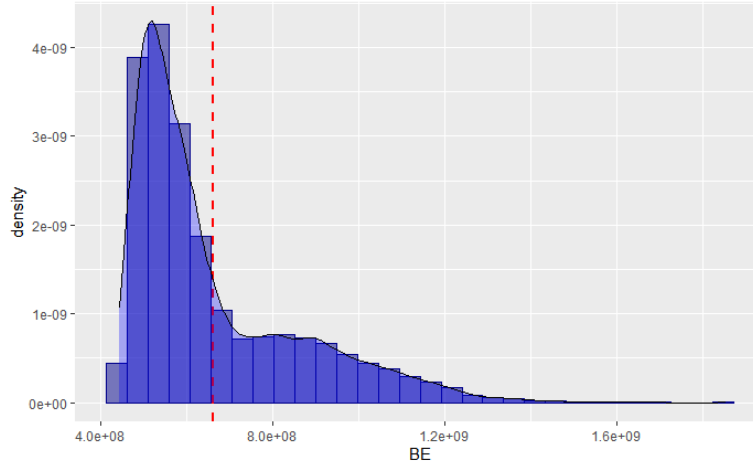


Figure 6: Histogram and density of the BE

To explain these extreme BE, we represent the distributions of the BE according to the modalities of the variables. This allows us to intuit the variables that can explain the extreme behaviours of the BE and more generally have the greatest explanatory potential. Our studies show that the extreme values of BE are all for a TMG of 3.5%. However, the distribution of our BE is very wide for this TMG value and we do not observe an upward trend in BE as a function of TMG. We thus conclude that the TMG is not the only variable that can explain the presence of *outliers* in our database. Encouraged by these results, we continue by representing the BE as a function of two variables on a heatmap, the mean value of all the BE having as a parameter our pair of variables. This study allows us to study cross effects. The most convincing results are obtained for the "TMG"/"Taux\_Y\_1" variable pair, i.e. a variable symbolising the yield curve.

We observe that a negative interest rate environment, combined with a high TMG, leads to the highest BE. Equivalent studies are conducted on all our variables and we observe a significant variability in the distribution of our BE according to the modalities of the following variables: "TMG", "Age", "Oblig" as well as the rate curve in general.

### Machine learning study and approximation of the Best Estimate

We select 4 models for the Machine Learning study: the linear regression model, Random Forest, XGBoost and neural networks. We adopt an approach of optimisation of the hyperparameters of each model by the k-fold cross validation method. Thus, we obtain the best performing models in the sense that we minimise their biases and variances.

After optimising the hyperparameters and learning the models on the two learning bases, we were able to analyse their performance on the test and complete databases. The results of our models on the databases of BE per scenario were not conclusive. Our performance metrics show that the prediction errors are high on the test base and therefore on the complete base. In particular, we obtain an RMSE of €566 million for our best-performing model, Random Forest, which is of the same order of magnitude that the mean value of the BE per scenario and is an indicator of large prediction errors. This is verified by comparing our predictions with observations. The representation of predictions according to our BE per scenario should ideally be linear, which is not the case, with significant discrepancies between observations and predictions.

Has a result of this unsatisfactory performance on the BE per scenario, we decided to focus on the BE. On the contrary, the results of this study are very satisfactory, as shown in Table 6 below :

Table 6: Results on the complete database

Model	RMSE (in Mn €)	MAE (in Mn €)	MAPE (in %)
<b>Linear regression</b>	93, 98	65, 33	0, 1
<b>Random Forest</b>	31, 70	16, 03	0, 02
<b>XGBoost</b>	24, 38	10, 71	0, 01
<b>Neural Network</b>	25, 51	10, 81	0.012

We notice that XGBoost and the neural network are the best performing with a MAPE, that is to say the average absolute error in percentage of 0.01%, which is remarkable. In order to better judge the performance of our models, we plot the observations according to the BE. For both models, the results are very satisfactory, as can be seen on figure 7 in the case of XGBoost, with predictions very close to the observations. However, we notice that the quality of the predictions deteriorates for the largest BE. We finally choose XGBoost as a model for our application in ORSA.

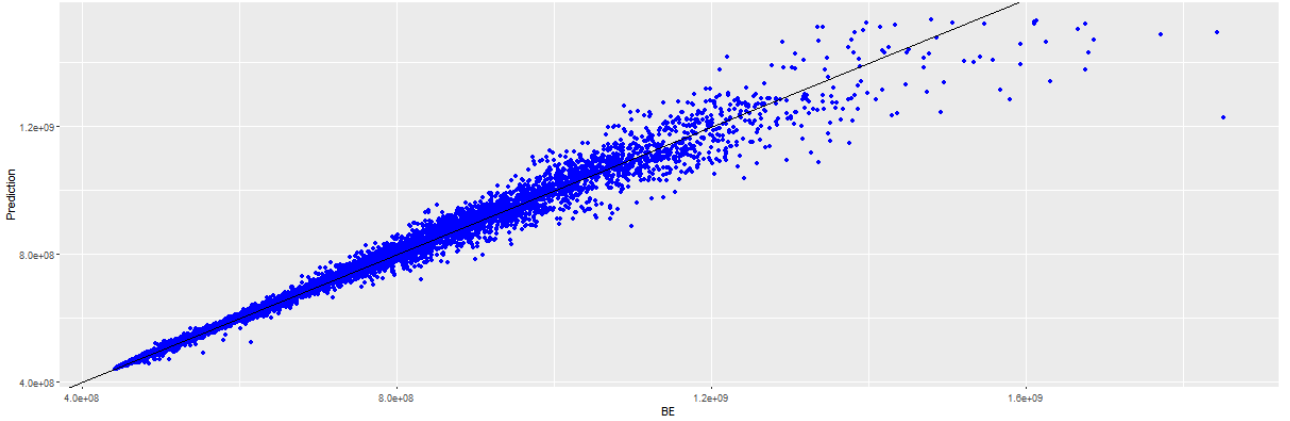


Figure 7: BE and predictions of XGBoost

At this stage of the study, we can already perceive the interest of the approximation by the XGBoost model, the calculation of the BE is instantaneous, whereas 10 minutes are necessary with the ALM model.

## Applications to ORSA

This dissertation concludes with the application of our machine learning study to ORSA.

Our first application aims at explaining the sensitivities of BE observations to our parameters by methods of a posteriori interpretability of the so-called "black box" models. The objectives are double : to show that we can have confidence in the model predictions, which is essential in a regulatory framework such as Solvency II, and then to expose the potential of such a tool to steer the insurer's strategy concerning the BE. More specifically, we have implemented the LIME method, which consists in locally approximating a prediction of our "black box" model by an interpretable model, a linear model generalised in our study, thus allowing us to explain our predictions. We obtain results as shown below:

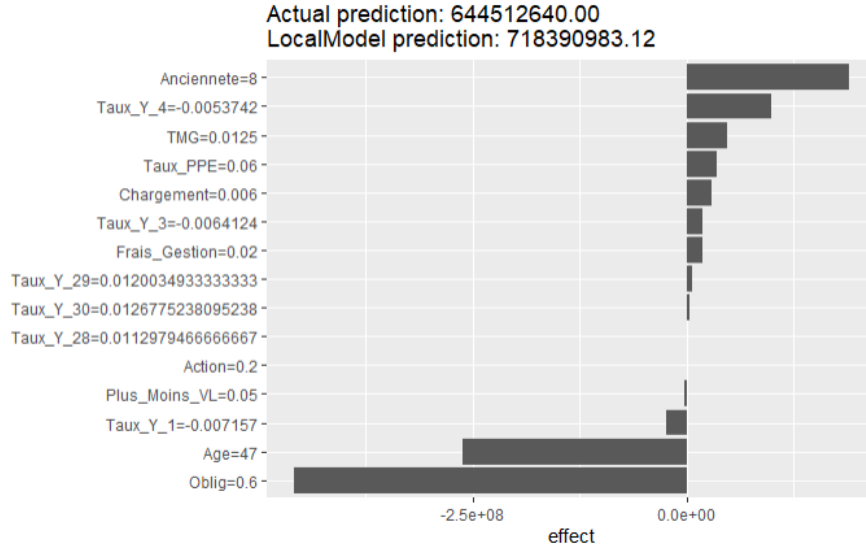


Figure 8: Explanations of a central BE with LIME

Using the example of figure 8, we explain the prediction of a central BE. To do so, we represent the variables for which the coefficients of the linear model approximating our prediction are the most important. In this case, we note that an increase in the "oblig" variable, i.e. the targeted percentage of bonds in our asset portfolio, has a negative impact on the BE. Thus, the insurer may choose to change its asset allocation by increasing the proportion of bonds, thus seeking to lower its BE

Our second application aims at projecting the BE of our fictitious company over 3 years, which we consider to be the strategic plan of our insurer. Our approach consists, firstly, in defining a "Real World" scenario allowing our company to evolve over time and, secondly, in comparing each year our predicted BE to the BE calculated by the ALM tool.

Our first projection using this approach gives us an average error of 0.56%, with the error increasing with the projection time. This can be explained by the distortion of the rate curve which leads to predict the BE on combinations of parameters on which XGBoost has not performed the training. This induces an error due to the variance of our model. This application also allows us to judge the computing time saved by this machine learning approach: the ALM tool takes almost an hour to project the BE over the 3 years whereas in comparison XGBoost takes only 5 minutes.

As these results are satisfactory, we are continuing this study by projecting our company into the health crisis environment we are currently experiencing. To do so, we define 4 scenarios different in terms of the evolution of equity returns and the economic crisis. We base ourselves on the evolution of the CAC 40 index to define our scenarios, which are distinguished by their form of evolution and their falling point at the end of the projection. Thus, we model two U-shaped scenarios, with a drop in equity returns that remain at a low level for two years and then rise again, and two V-shaped scenarios with a drop in equity returns but which rise again the following year. We try to explain the variations in BE from one scenario to the other by projecting indicators such as the level of PPE or capital gains. Our results on this study show that the prediction errors from 2020 onwards are larger than in our first application, in the order of 5% average error. This time, these errors are significant and symptomatic of a lack of variability in our database, all the more so in a shocked scenario leading to large variations in our variables.

## Conclusion

We find this study very encouraging in the sense that the objective is achieved: we obtain, in a much shorter time than with an ALM tool, a projection of our BE over 3 years and with a minimal error on the central scenario. Concerning the more adverse COVID-19 scenarios, we believe that two approaches should be considered in order to improve the predictions on the projection. The first one would consist in recalibrating the XGBoost model each projection year and the second one would consist in completing the database to bring more variability. In addition, interpretability models such as LIME are steering tools for the insurer who can thus adapt his strategy at any time.

There are limits to our approach. In order to carry out this study, we had to make simplifying assumptions that are not feasible for an insurer and that make our database lack variability, especially with regard to the economic environment. Indeed, in the life insurance context, one of the main sources of variation in BE comes from the economic environment and the insurer's ability to leverage its assets. Completing the database by playing on the ESG parameterisation and making our asset allocation more diversified and representative of the market would therefore be a good way of improvement. However, these assumptions were motivated by logistical constraints of time and computing power that an insurer has the financial and material means to overcome if it wants to develop an approach such as the one we are proposing. Finally, we are only projecting the BE and not the SCR, which limits the completeness of our study in the ORSA. Equivalent work could be considered to complement our approach.



# Remerciements

Je souhaiterais remercier les associés de Mazars Actuariat Alice THOU, Alexandre GUCHET et Grégory BOUTIER pour m'avoir permis de réaliser mon mémoire au sein du cabinet ainsi que toute l'équipe de Mazars Actuariat pour leur accueil et leur soutien tout au long de ce mémoire, malgré des conditions de travail inédites.

En particulier, je remercie Abdelwahab EL MAHJOUB, Julien MATTEI et Thomas YAGUES pour leur accompagnement, leur encadrement et leurs conseils précieux qui m'ont permis de mener cette étude passionnante et enrichissante.

Je remercie également Quentin GUIBERT, encadrant ce mémoire pour le compte du Master Actuariat de l'Université Paris Dauphine, qui s'est toujours rendu disponible pour répondre à mes questions et a su me guider dans mes travaux.

Merci à Anthony REVEILLAC, enseignant à l'INSA Toulouse de la matière d'introduction à l'actuariat et Christophe DUTANG responsable du Master Actuariat de l'Université Paris Dauphine, pour la qualité de leur enseignement et sans qui je ne me serais pas tourné vers la voie de l'actuariat.

Je remercie finalement ma famille pour son soutien au quotidien.



# Table des matières

Résumé	3
Abstract	4
Note de Synthèse	5
Synthesis note	11
Remerciements	17
Table des matières	19
Introduction	21
<b>1 Cadre de l'étude</b>	<b>23</b>
1.1 Le cadre réglementaire de Solvabilité II . . . . .	23
1.2 Les spécificités de l'assurance vie . . . . .	30
<b>2 L'outil ALM</b>	<b>39</b>
2.1 Le calcul du <i>Best Estimate</i> par méthode de Monte Carlo . . . . .	39
2.2 Génération des scénarios économiques . . . . .	41
2.3 Les mécanismes de l'outil ALM . . . . .	44
2.4 La modélisation du passif . . . . .	44
2.5 La modélisation de l'actif . . . . .	51
2.6 Le périmètre de l'étude : modélisation d'une assurance vie fictive . . . . .	57
<b>3 Base de données et <i>machine learning</i></b>	<b>61</b>
3.1 La stratégie mise en place . . . . .	61

3.2	Les variables explicatives et leurs plages de variation . . . . .	62
3.3	Générer les bases de données à partir de l'outil ALM . . . . .	75
3.4	Visualisation des données et études préliminaires . . . . .	77
3.5	Les mesures de l'erreur de prédiction . . . . .	81
3.6	Les modèles utilisés . . . . .	82
3.7	Optimisation des hyperparamètres des algorithmes . . . . .	90
3.8	Les résultats obtenus et la sélection du modèle . . . . .	93
<b>4</b>	<b>Etudes de type ORSA</b>	<b>101</b>
4.1	Interprétation de notre modèle de <i>machine learning</i> . . . . .	101
4.2	Application à l'ORSA : la projection de BE . . . . .	106
	<b>Conclusion</b>	<b>123</b>
	<b>Bibliographie</b>	<b>125</b>
<b>A</b>	<b>Annexes</b>	<b>127</b>
A.1	Chocs sur les taux S2 . . . . .	128
A.2	Matrices de corrélations . . . . .	129

# Introduction

Depuis son entrée en application le 1<sup>er</sup> janvier 2016, Solvabilité 2 impose des règles prudentielles aux assureurs de l'Union Européenne. L'ORSA, acronyme d'*Own Risk and Solvency Assessment*, s'inscrit dans le cadre de cette directive et vise à répondre aux exigences en matière de supervision et d'évaluation interne des risques et de la solvabilité.

L'ORSA impose ainsi à l'assureur d'identifier et cartographier ses risques tout en jugeant de la tolérance qu'il leurs accorde puis de vérifier que les indicateurs de solvabilité soient calculés avec des hypothèses adéquates et en accord avec les exigences de la directive. En somme, la compagnie d'assurance doit être en mesure de s'assurer **à tout moment** de la couverture de son capital réglementaire et de la compréhension de ses risques. Pour ce faire, l'assureur projette, sur un horizon temporel allant de 3 à 5 ans, différents indicateurs de solvabilité afin de quantifier son risque et assurer sa solvabilité. Néanmoins, dans le cas des compagnies d'assurance vie, ce calcul nécessite la mise en place de modèles de gestion Actif/Passif simulant des interactions complexes et nécessitant une puissance et un temps de calcul significatifs.

Il n'est donc pas envisageable pour les compagnies d'assurance de répondre aux exigences de l'ORSA à tout moment comme ils le devraient. Elles se tournent déjà vers des méthodes alternatives (LSMC, proxy...) pour estimer ces indicateurs avec une précision qui respecte les exigences de l'ORSA et qui leur permettent de réduire significativement leurs coûts et temps de calcul.

En réponse à cette problématique, nous proposons une approche *Machine Learning* qui a déjà montré l'étendue de son potentiel dans de multiples domaines dont la tarification en assurance non-vie. Les méthodes de *Machine Learning* présenteraient l'intérêt de prédire presque instantanément les indicateurs sans avoir à relancer les modèles Actif-Passif. De plus, elles mettent à disposition des méthodes d'analyse permettant de surpasser leur mauvaise réputation de "boîte noire" et ainsi piloter au mieux les contrats en épargne, enjeu dans un contexte de crise sanitaire et de taux bas sans précédent.

Pour mener à bien cette étude, nous commencerons par rappeler le cadre réglementaire et les spécificités de l'épargne en assurance vie dans le premier chapitre. Dans le second, nous étudierons en détail le modèle de gestion Actif-Passif développé par Mazars afin de sélectionner les variables pertinentes pour expliquer le *Best Estimate*, indicateur sur lequel nous concentrerons notre étude. Nous développerons notre approche *Machine Learning* dans le troisième chapitre et clôturerons ce mémoire par un chapitre dans lequel nous utiliserons le modèle sélectionné pour des études et applications rentrant dans le processus ORSA d'une compagnie d'assurance vie fictive.



# Chapitre 1

## Le cadre de l'étude : la réglementation Solvabilité II et les spécificités de l'assurance vie.

### 1.1 Le cadre réglementaire de Solvabilité II

La directive européenne Solvabilité II est en vigueur depuis le 1<sup>er</sup> janvier 2016 et s'inscrit dans la lignée de la réforme Bâle 2 qui est venue réguler le monde financier. Elle a entraîné l'instauration de nouvelles méthodes d'évaluation du bilan économique des assureurs et réassureurs et la mise en place de nouvelles exigences de capital de solvabilité. Solvabilité II incite les organismes d'assurance à mieux connaître et gérer leurs risques et à adapter leurs stratégies et leurs exigences réglementaires à leur profil de risque. Comme son prédécesseur Solvabilité I, cette directive s'articule autour de trois piliers.

#### 1.1.1 Pilier I : les exigences quantitatives

Le pilier I détaille les exigences en termes de calcul du capital économique et de valorisation du bilan économique. Contrairement à Solvabilité I, nous ne parlons plus de bilan comptable mais de bilan économique. Les actifs sont en valeur de marché et les passifs sont évalués sous une vision *Best Estimate* que nous présenterons par la suite.

#### Les provisions techniques

Sous Solvabilité I, les provisions techniques sont évaluées selon le principe de prudence en accord avec les normes de comptabilité françaises. A l'inverse, Solvabilité II apporte une modification majeure dans le sens où les provisions ne sont plus estimées de manière prudente mais sont une estimation moyenne des engagements sans marge de prudence. Nous distinguons deux éléments à additionner dans les provisions techniques : la marge pour risque et le *Best Estimate*.

Le *Best Estimate* (ou BE) correspond à la valeur économique des passifs sans marge de prudence. Il est égal à la valeur actuelle probable des flux de trésorerie futurs sortant nets des flux entrants, estimés sur la base de la courbe des taux sans risque. En assurance vie, son calcul est stochastique et nécessite la modélisation des interactions actif-passif ou ALM (*Asset and Liabilities Management*)

permettant de valoriser les options et garanties que l'assureur s'engage contractuellement à servir. Le développement d'un outil ALM par Mazars permet ce calcul du *Best Estimate* et nous en détaillerons les mécanismes dans le chapitre 2.

La marge pour risque (ou RM pour *Risk Margin*) représente le coût lié aux risques non couvrables entraînant un aléa sur le *Best Estimate*. Elle est souvent définie comme la marge telle que les provisions techniques soient équivalentes au montant qu'une entité tierce demanderait pour reprendre et honorer les engagements d'assurance et de réassurance existants. Elle est évaluée suivant la méthode de coût du capital pour toutes les activités.

### Les fonds propres économiques (FP)

Sous Solvabilité II, les fonds propres économiques ou *Net Asset Value* représentent la richesse réelle de la compagnie et sont égaux à la différence entre l'actif et les provisions techniques. De plus, cette directive prévoit deux exigences de capital : le MCR et le SCR. Ce sont deux niveaux de fonds propres de référence définis ainsi :

- Le Capital Minimum Requis (*Minimum Capital Requirement* : MCR) représente le niveau minimal de fonds propres qu'une compagnie d'assurance doit détenir pour pouvoir exercer son activité. En dessous du MCR, l'assureur ne serait plus en mesure de respecter ses engagements vis à vis des assurés.
- Le Capital de Solvabilité Requis (*Solvency Capital Requirement* : SCR) est le montant de fonds propres dont doit disposer l'assureur pour faire face à une ruine à horizon un an avec une probabilité de 99,5%. Les compagnies d'assurance disposent de deux moyens pour calculer le SCR : la formule standard ou le modèle interne complet ou partiel. La formule standard est une formule générique établie par le régulateur sur la base d'hypothèses standards supposées applicables à tout assureur. Le modèle interne est développé au sein même de l'entreprise qui doit recevoir une autorisation préalable délivrée par l'ACPR \* avant de pouvoir être utilisé lors du calcul du SCR.

Le ratio de couverture  $RC$  est obtenu en faisant le ratio entre les fonds propres économiques  $FP$  et le SCR, ce qui se réécrit ainsi :

$$RC_t = \frac{FP_t}{SCR_t}$$

Nous appellerons surplus la différence entre la valeur totale des fonds propres et le SCR.

### Le bilan économique Solvabilité II

Un bilan type Solvabilité II est schématisé ci-dessous :

---

\* Autorité de Contrôle Prudentiel et de Résolution



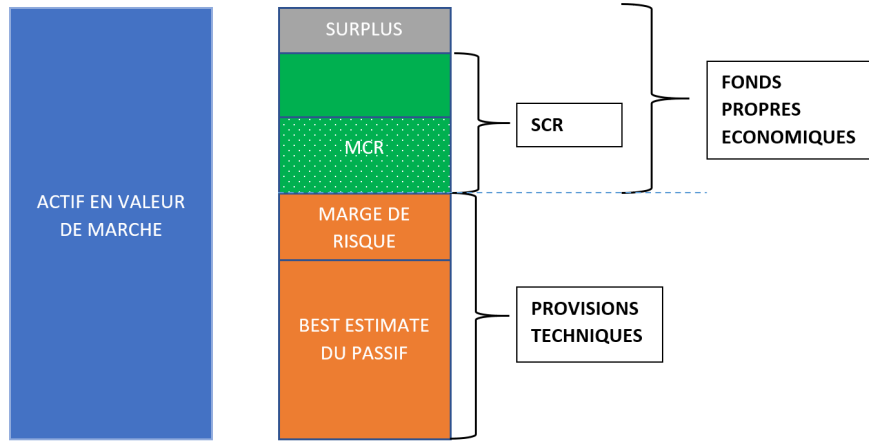


FIGURE 1.1: Bilan économique sous Solvabilité II

### La mise en place du bilan économique

Notre étude se concentre sur le *Best Estimate* défini précédemment et se calculant ainsi :

$$BE_t = \mathbb{E}^Q \left[ \sum_{u \geq t+1}^{t+H} \left( \frac{\delta_u}{\delta_t} \times P_u \right) | \mathcal{F}_t^{MR} \right]$$

Avec  $P_u$  les flux de trésorerie de l'instant  $u$ .

La marge pour risque est le coût de portage du capital réglementaire au titre des risques non couvrables, et cela jusqu'à extinction du passif. Elle est calculée comme le "coût d'opportunité" annuel généré par la détention du SCR estimé sur un an (le coût du capital étant fixé à 6%) mais aussi du SCR futur dans 1 an, 2 ans... jusqu'à la liquidation complète du passif. La marge pour risque (RM) se calcule à un niveau global de la manière suivante :

$$RM_t = 6\% \times \sum_{u \geq t} \frac{SCR(u)}{(1 + r_u)^u}$$

Ce calcul fastidieux nécessite de projeter le SCR sur plusieurs années, les assureurs peuvent recourir à des proxys pour calculer les SCR globaux ou les SCR modulaires et sont même amenés à chercher directement une formule fermée de la marge pour risque, c'est ce que nous verrons par la suite.

Nous déduisons de ces variables la valeur des Fonds Propres en soustrayant à la valeur de marché de l'actif le *Best Estimate* et la *Risk Margin*, ce qui se réécrit :

$$FP_t = Actif_t - BE_t - RM_t$$

avec  $Actif_t$  l'actif en valeur de marché.

Les assureurs sont amenés à approcher ces variables bilantielles par des méthodes qui nécessitent des projections sur plusieurs années et pour un grand nombre de scénarios. A cette fin, des générateurs de scénarios économiques et des modèles ALM permettant de modéliser chaque année les interactions entre l'actif et le passif et ainsi simuler les flux de trésorerie et résultats financiers nécessaires au calcul de ces variables. Ces calculs nécessitent une puissance de calcul considérable et sont très chronophages. Nous aborderons plus en détail le calcul du BE dans le second chapitre.

## Le calcul du capital économique Solvabilité II

Le SCR représente le capital permettant d'absorber un éventuel choc dû à la survenance dans l'année d'événements extrêmes ne se produisant qu'une fois tous les 200 ans. La ruine économique est la situation où la valeur de marché de l'actif est inférieure au BE. Ainsi, on peut définir mathématiquement ce capital économique comme étant :

$$SCR = FP_0 - q_{0,5\%}(D(1).FP_1)$$

Où  $D(1)$  est le facteur d'actualisation et  $q_{0,5\%}$  est le quantile de niveau de confiance 0,5% de la distribution des FP à horizon un an.

Deux approches sont à la disposition des assureurs pour mener à bien ce calcul. La première, l'approche par "formule standard" est proposée par le régulateur. La seconde consiste à approcher la distribution des fonds propres à horizon un an afin d'en déduire le quantile de niveau 0,5%. C'est l'approche par "modèle interne".

**Approche de type Formule Standard** : c'est une approche modulaire dans laquelle l'assureur calcule un SCR par variation entre des fonds propres centraux et des fonds propres choqués au titre d'un seul facteur de risque (action, taux, mortalité,...). Ces SCR par choc sont ensuite agrégés à l'aide de matrices de corrélations. Les chocs ainsi que les coefficients de corrélations sont fournis par le régulateur. Les chocs sont appliqués en  $t=0$ . La figure 1.2 détaille les différents modules et leurs sous-modules de risques. L'agrégation s'effectue du bas vers le haut. La première étape consiste à calculer les SCR pour chaque sous-module, qui sont ensuite agrégés par module grâce aux matrices de corrélation. La seconde étape est ensuite d'agréger les SCR modulaires pour obtenir le BSCR pour *Basic SCR*. On en déduit finalement le SCR global de l'assureur.

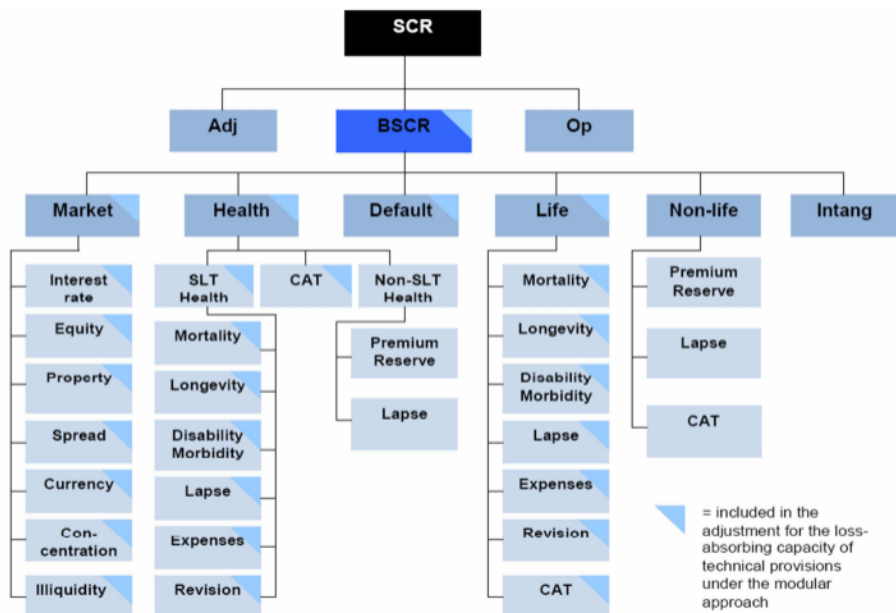


FIGURE 1.2: Agrégation et calcul du SCR sous Solvabilité II

**Approche de type Modèle Interne** : l'objectif est d'obtenir la distribution des fonds propres à horizon un an, afin d'en déduire le quantile à 0,5%. Pour ce faire, deux types de projections sont

nécessaires et il en découle une forte complexité. On nomme aussi cette approche Simulations dans les Simulations.

Les premières projections, les **projections primaires**, s'effectuent en **univers historique**. Le bilan de la compagnie est ainsi projeté de manière à prendre en compte différents facteurs de risques. Ces projections doivent respecter les évolutions historiques observées ainsi que les corrélations entre les risques.

Les secondes projections, les **projections secondaires**, sont effectuées en **univers risque-neutre**. Partant de la situation de l'entreprise obtenue par la projection monde réelle, on effectue un certain nombre de projections secondaires risque-neutres, permettant ainsi de valoriser les fonds propres économiques. La valeur de FP économique est ainsi obtenue pour chaque projection primaire et par conséquent une distribution empirique des FP en est déduite.

La seconde approche présente l'intérêt d'être plus adaptée au profil de risque de l'assureur, mais nécessite une puissance et un temps de calcul non négligeable, une problématique d'autant plus marquée dans le cadre du pilier 2, où des projections du bilan économique sur plusieurs années sont à effectuer. Pour pallier à cela, des méthodes alternatives ont été étudiées comme :

- **Les Simulations dans les Simulations accélérées.** Cette méthode a été introduite par DEVINEAU et LOISEL, 2009 et l'objectif est de détecter les scénarios primaires les plus adverses en matière de solvabilité et de n'effectuer les simulations secondaires que sur ceux-ci.  
Pour mettre en place cette méthode, il s'agit dans un premier temps d'introduire des facteurs de risque véhiculant l'intensité du risque de chaque simulation primaire. Ceux-ci permettent de définir un nuage de point ainsi qu'une région de confiance en-dehors de laquelle se trouvent les scénarios primaires les plus adverses. Les simulations primaires et secondaires sont effectuées au-delà de cette frontière de confiance, jusqu'à ce que les "pires valeurs" des FP soient stabilisées.
- **Les Replicating portfolio ou portefeuilles répliquant.** Un portefeuille répliquant est un portefeuille d'actifs qui reproduit les cash flows de passifs pour chaque simulation et chaque date future. Ainsi, à l'horizon de ruine, la valeur du portefeuille répliquant équivaut à celle du *Best Estimate*. On utilise ainsi le portefeuille répliquant comme un proxy des passifs, qui peut être valorisé rapidement par des formules fermées ou des techniques de calcul rapides.
- **Les formes paramétriques.** L'objectif est ici d'obtenir la distribution des fonds propres ou du BE à l'aide d'une formule fermée (polynôme), dont les paramètres sont les facteurs de risques. Deux approches sont alors possibles. La première consiste à calibrer la forme paramétrique sur les fonds propres économiques, c'est la méthode de *Curve Fitting*. La seconde approche consiste à la calibrer sur les VAN de marges, c'est la méthode des *Least Square Monte Carlo* (LSMC). Pour plus de détails à ce sujet, une implémentation de la méthode LSMC a été effectuée par HAGUET, 2013.

Notre travail se place dans la continuité de ces approches et vise à implémenter une approche alternative de type *machine learning* afin d'accélérer le temps de calcul du BE.

### 1.1.2 Pilier 2 : les exigences qualitatives

Le pilier 2 définit les exigences en matière de supervision. Les objectifs sont de renforcer la gouvernance et la gestion du risque des assureurs.

Solvabilité II introduit la démarche ERM (*Enterprise Risk Management*) qui définit la gestion du risque comme étant l'identification, l'évaluation et la priorisation des risques suivies par la coordination et l'application de ressources pour minimiser, piloter et contrôler la probabilité et/ou l'impact d'événements non voulus et maximiser la réalisation d'opportunités. Cela passe notamment par l'instauration de quatre fonctions clés définies dans les articles 44,46 à 48 de la directive : l'audit interne, la fonction de conformité, la fonction de gestion des risques et la fonction actuarielle.

De plus, Solvabilité II exige un dispositif d'évaluation interne des risques et de la solvabilité, plus communément appelé ORSA pour *Own Risk and Solvency Assessment* et défini par l'article 45 de la directive 2009/138/CE. Dans le cadre de l'ORSA, l'assureur doit procéder annuellement ou à chaque modification de son profil de risque aux trois études suivantes :

**Le besoin global de solvabilité.** Cela doit permettre d'identifier les spécificités du profil de risque de la société, plus particulièrement les risques non quantifiés tels que le risque de réputation. L'assureur cartographie ainsi les risques auxquels il est assujetti et en définit le niveau de tolérance selon une métrique qu'il prend soin de définir. Ainsi, chaque compagnie va définir des procédures qui sont proportionnées à la nature et la complexité des risques inhérents à son activité.

**Le respect permanent des exigences de capital.** Permet d'évaluer la capacité de la société à couvrir ses exigences de capital dans le temps et s'effectue généralement via des stress tests dans la durée.

**La déviation du profil de risque** avec la formule standard ou le modèle de calcul du SCR. L'objectif est de vérifier l'adéquation des hypothèses effectuées au moment du calcul du SCR avec le profil de risque spécifique de l'assureur.

Pour mener à bien ces études, l'assureur met en place un ensemble de processus et procédures pour identifier, évaluer, surveiller et gérer les risques à court et long terme et déterminer le niveau des fonds propres nécessaires pour assurer que les besoins en solvabilité sont satisfaits continuellement. L'ORSA s'inscrit donc dans une vision prospective et pluriannuelle, l'horizon de temps devant être adapté à l'activité et à la stratégie globale de la compagnie. Cet horizon temporel dépend du profil de risque de la compagnie, mais nous retrouvons souvent des horizons stratégiques de l'ordre de 3 à 5 ans.

L'ORSA est intégré au processus de décision d'une compagnie et nécessite par conséquent la mise en place d'indicateurs multiples permettant un pilotage adapté. Voici une liste non exhaustive d'indicateurs fréquemment trouvés :

- Les indicateurs bilantiels comme le SCR, le *Best Estimate* ou encore le ratio de solvabilité permettant de mesurer la santé financière et l'évolution des engagements d'une compagnie d'assurance.
- Le RORAC pour *Return On Risk Adjusted Capital* qui permet de mesurer la rentabilité. Pour plus de détail sur cet indicateur, nous renvoyons le lecteur au mémoire de Sophie Décupère (DECUPÈRE, 2011) qui détaille sa définition ainsi que la propriété de RORAC compatibilité qui doit être vérifiée pour qu'une méthode d'allocation de capital soit cohérente.
- Le capital alloué comme mentionné précédemment est aussi un bon indicateur de risque et est utile pour piloter l'activité. Comme spécifié dans le mémoire de Sophie Décupère (DECUPÈRE, 2011), cela consiste à "déterminer le volume de capital consommé par un segment de risque après avoir alloué les bénéfices de diversification".
- Le ROE (*Return On Equity*), le chiffre d'affaires, le taux d'affaires nouvelles... et bien d'autres.

Ces indicateurs permettent de juger de la pertinence d'une nouvelle orientation stratégique et d'en assurer le pilotage. L'ORSA donne lieu à la rédaction d'un rapport et doit permettre la définition des risques et de la stratégie les concernant, et ainsi l'évolution des processus stratégiques.

Dans ce cadre, les compagnies d'assurances devraient être capables à tout moment d'évaluer leurs *Best Estimate* en fonction de l'évolution de sa stratégie et/ou de son profil de risque afin de pouvoir piloter au mieux son activité. Cependant cela n'est pas réalisable pour les problèmes de temps et de puissance de calcul mentionné précédemment. Les compagnies d'assurances se limitent ainsi à une revue de type ORSA trimestrielle. C'est dans ce cadre que notre étude prend place. **Notre objectif est d'obtenir un BE directement à partir de données bilantielles et économiques, notre modèle de *machine learning* ayant appris les mécanismes de l'outil ALM et prédisant un BE sans avoir à effectuer les nombreuses simulations.**

### 1.1.3 Pilier 3 : le reporting

Le pilier 3 introduit les reporting publics et réglementaires. Il a un rôle clé à jouer dans Solvabilité II car il assure une plus grande transparence des assureurs.

Ce reporting implique donc deux types de rapports qualitatifs :

- Le SFCR pour *Solvency and Financial Condition Report* qui est un rapport annuel public et donc accessible à tous.
- Le RSR pour *Regular Supervisory Report* qui est un rapport privé entre la société et le régulateur. Celui-ci est à fournir au moins tous les 3 ans. Il contient toutes les informations nécessaires à la supervision.

La différence entre ces deux rapports est le niveau de détail demandé et l'audience. Ainsi, les SFCR se doivent d'être moins détaillés de manières à être clair et compréhensibles pour un public qui n'est pas familier avec les aspects techniques de solvabilité II. Afin de pouvoir comparer les rapports entre les différents assurés, ils doivent avoir la même structure orientée autour de 5 axes :

1. *Business, External Environment and Performance* qui traite de l'activité et des résultats.
2. *Governance and Remuneration policy* qui expose le système de gouvernance.
3. *Risk Profile* où l'assureur dévoile son profil de risque.
4. *Valuation for Solvency Purposes* qui regroupe les méthodes de valorisations utilisées.
5. *Capital Management* dans lequel les exigences de capital et fonds propres sont abordés.

Ces rapports permettent ainsi aux assurés, aux régulateurs et autres acteurs des marchés financiers un accès à une information permettant l'évaluation de la gestion des risques des compagnies d'assurance.

Finalement, les rapports quantitatifs sont effectués dans les QRT (*Quantitative Reporting Templates*). Chaque QRT contient une information spécifique liée aux exigences réglementaires Solvabilité II. Nous trouvons notamment le QRT bilan présentant l'ensemble des actifs et passifs hors fonds propres, les QRT SCR présentant les résultats des calculs des SCR, les QRT MCR idem que les QRT SCR mais pour les MCR et la liste est longue, une compagnie pouvant avoir à remettre plus de 60 QRT annuellement.

Nous avons présenté le cadre réglementaire de notre étude et introduit des notions clés pour ce mémoire : l'ORSA et le *Best Estimate*. Nous avons de plus mis en avant les problématiques de temps de calcul et de complexité qu'implique le calcul du BE, d'autant plus marquées dans le cadre d'une étude ORSA où il doit être projeté sur plusieurs années. Bien que le BE soit calculé par tout type d'assureur, son calcul est plus complexes en assurance vie. Cela s'explique par les garanties et options des contrats d'assurance vie, que nous allons à présent aborder.

## 1.2 Les spécificités de l'assurance vie

Ce mémoire traitant de l'activité d'épargne en assurance vie, nous allons en détailler le fonctionnement. L'assurance vie est une branche de l'assurance qui fait naître des engagements dont l'exécution dépend de la durée de vie humaine. Ainsi, l'assurance vie est incluse dans l'assurance de personnes. De manière simplifiée, elle est souvent décrite comme étant l'assurance dont le calcul des prestations dépend des tables de mortalité.

Ceci est le principe de base, cependant, actuellement, une part importante de l'assurance vie correspond en réalité à de l'activité d'épargne. Ainsi, l'aléa repose plus sur le moment de remboursement des sommes (rachats).

L'assuré verse une prime unique ou périodique qui va alimenter son compte et donc sa provision mathématique qui sera par la suite revalorisée à un certain taux. Le capital de sortie est constitué des différents versements et des intérêts qui sont liés aux taux de revalorisation.

### 1.2.1 Historique de l'épargne préférée des français

La première apparition de l'activité d'assurance vie remonte à 1653, date à laquelle le banquier italien Lorenzo Tonti créa la tontine. Cependant, l'activité d'assurance vie a longtemps posé un problème du fait de son objet (la mort d'un être humain donnant lieu à une compensation financière) et de ses méthodes (calculs statistiques, utilisation de probabilités...). La mauvaise réputation de cette activité entraîna son interdiction par Colbert en 1681 et ce jusqu'à 1818. L'assurance vie s'est ensuite développée, avec notamment la création des contrats d'épargnes dans le début des années 1980. Depuis, l'assurance vie croît exponentiellement grâce à la maîtrise de l'inflation.

Comme nous avons pu l'observer durant cette période de confinement due à la Covid 19, l'activité d'épargne a encore de beaux jours devant elle. Les français avaient déjà épargné 55 milliards d'euro le 20 avril 2020 selon une étude de l'OFCE\* sur l'évolution de l'impact économique de la pandémie de la Covid 19 et des mesures de confinement en France. La banque de France prévoit que l'épargne pourrait atteindre les 100 milliards d'ici la fin de l'année. Bien qu'elle ait été en quelque sorte contrainte, ces chiffres montrent malgré tout l'attrait des français pour l'épargne.

Concernant l'assurance vie, elle restait le placement préféré des français en 2019, comme en atteste le fait que la moitié des 3 300 milliards d'euros épargnés en 2019 a été investie dans l'assurance vie selon la caisse des dépôts. Plusieurs facteurs expliquent le succès de l'assurance vie, d'épargne et de retraite :

- La dimension de protection sociale et de mutualisation des risques
- La fiscalité attractive

---

\*Observatoire Français des Conjonctures Economiques

- La possibilité de sortir en rente ou en capital.

Cependant, les conditions actuelles de taux bas impliquant des rendements sur les contrats d'épargne euro bien moins généreux que par le passé (se tassant autour de 1%) ce à quoi s'ajoute la crise sanitaire, nous avons vu ces derniers mois un regain de confiance des Français pour le livret A et le LDDS \*, au détriment de l'assurance vie. Par conséquent, en août 2020, la décollecte nette des assurance vie a atteint 6,5 milliards d'euros depuis le début de cette même année, ce qui reflète un retrait massif des épargnants.

### 1.2.2 Les différents contrats d'épargne

Nous présentons ici les contrats d'épargne les plus fréquents chez les assureurs :

#### **Contrat Euro :**

Le souscripteur dépose des fonds qui sont placés par l'assureur. En contrepartie des frais de gestion prélevé par l'assureur, l'assuré a droit à une part des revenus financiers. Ces versements d'une partie des résultats financiers entraînent ainsi une interaction entre l'actif et le passif. Le risque est donc porté par l'assureur qui garantit le nombre d'euro.

#### **Contrat en unités de compte (UC) :**

Le souscripteur investit dans une part d'un actif risqué, par exemple une part dans un OPCVM (Organisme de Placement Collectif en Valeurs Mobilières) ou de SCPI (Société Civile de Placement Immobilier). C'est le souscripteur qui porte le risque car l'assureur garantit le nombre d'UC et non leur valeur. L'assureur se rémunère en prélevant des frais de gestions. Afin de rendre plus attractif ces contrats, l'assureur peut proposer une garantie plancher en cas de décès, qui peut prendre la forme d'un capital garanti par exemple.

#### **Contrat multi-supports :**

Ces contrats mélangent une partie UC et une partie Euro. Le montant des garanties fait référence à un ou plusieurs supports en UC et un support en euro. Le risque est donc partagé entre l'assureur (pour la partie euro) et l'assuré (pour la partie UC).

#### **Nouveaux contrats :**

D'autres contrats ont vu le jour, tels que les contrats Eurocroissance ou les contrats Vie génération. Les contrats Eurocroissances sont nés en 2014 mais ont été amendés par la loi Pacte afin d'en simplifier le fonctionnement. Ceux-ci sont à mi-chemin entre un contrat Euro et un contrat UC. Ils permettent une garantie du capital variant de 80% à 100% non plus en permanence comme pour les contrats Euro, mais à une échéance donnée de 8 ans minimum. Les contrats vie-génération sont des contrats en UC destinés à encourager l'investissement dans les PME ou le logement social et ainsi favoriser

---

\*Livret de Développement Durable et Solidaire

l'investissement "utile". Cela se traduit par des investissements en partie dans des sociétés contribuant au financement du logement social ou intermédiaire, dans des entreprises non-cotées et dans l'économie sociale et solidaire.

Nous allons maintenant nous concentrer sur les garanties des contrats Euro, qui seront les seuls contrats que nous considérerons dans notre étude.

### 1.2.3 Les spécificités des contrats Euro

Nous considérerons dans ce mémoire que notre compagnie d'assurance ne possède et ne commercialise que des contrats Euros. Nous allons donc ici énoncer les caractéristiques de ce type de contrats très répandu.

#### Le paiement des primes

L'assuré peut être amené à payer une prime unique ou une prime périodique. Cependant, l'article L 132-20 du Code des assurances prévoit qu'en assurance vie, l'assureur n'a pas le droit d'exiger le paiement des primes. Si l'assuré cesse de payer ses primes, la garantie est alors réduite. L'article L 132-21 du Code des assurances stipule que les modalités de calcul de la valeur de rachat, de transfert ou de réduction doivent être inscrites dans le contrat.

#### Le TMG (Taux minimum garanti)

Le TMG correspond au rendement minimum servi par le contrat. L'assureur s'engage à verser le TMG pendant une durée déterminée dans le contrat ou de façon viagère. L'assureur a l'obligation de revaloriser l'encours de l'assuré au minimum au TMG. Bien évidemment, si l'épargne réalise de meilleurs résultats, le taux de rendement sera supérieur au TMG.

Afin d'éviter que des assureurs proposent des TMG très attractifs dans le but d'attirer de nouveaux clients mais non réalisables, il est décrit dans l'article A132-3 du code des assurances un ensemble de règles encadrant l'utilisation du TMG :

1. Le taux proposé ne peut être supérieur à 85% de la performance des actifs de la société sur deux ans consécutifs.
2. La promesse de revalorisation des contrats bénéficiant d'un TMG ne doit pas se faire au détriment de la rémunération d'autres contrats.
3. Le TMG est fixé sur une durée minimale de six mois.
4. Le TMG ne peut dépasser un plafond dépendant de la stabilité des marchés obligataires dont les exigences sont les suivantes : « Les taux garantis mentionnés au II ne peuvent excéder le minimum entre 150% du taux d'intérêt technique maximal défini aux articles A. 132-1 et A. 132-1-1 par référence à 75% du taux moyen des emprunts d'Etat à la date d'effet de la garantie et le plus élevé des deux taux suivants : 120 % de ce même taux d'intérêt technique maximal et 110 % de la moyenne des taux moyens servis aux assurés lors des deux derniers exercices précédant immédiatement la date d'effet de la garantie. ». Cela peut être réécrit de la manière suivante :

$$TMG_{max} = \min(150\% \times 75\% \times TME; \max(120\% \times 75\% \times TME; 110\% \times TMS))$$



avec TME le taux moyen des emprunts d'Etat et TMS le taux moyen servi aux assurés lors des deux derniers exercices.

Une telle garantie est risquée pour l'assureur qui peine à avoir du rendement dans le contexte de taux bas, c'est pour cela qu'elle tend à disparaître des contrats.

### La participation aux bénéfices (PB)

Les réglementations poussent les assureurs à être prudents au moment du calcul de la prime et du provisionnement (table de mortalité, taux technique plafonné. . .). Cette prudence implique des bénéfices techniques et financiers probables que l'assureur doit partager avec les assurés. Ainsi, depuis 1967 la réglementation définit un niveau minimum de participation aux résultats.

Il est important de distinguer la PB contractuelle, la PB réglementaire et la PB discrétionnaire.

**La PB réglementaire :** Son calcul se fait au niveau global de l'entreprise et non contrat par contrat.

Le principe est le suivant : la rémunération globale de la compagnie est au plus égal à 10% du résultat technique (i.e. le résultat lié à la mortalité et à la gestion), et 15% du résultat financier (hors produits financiers sur fonds propres). Ainsi, soit  $PR$  la participation aux résultats et  $PR_{min}$  le minimum de participation des assurés aux résultats. Nous avons alors les relations suivantes :

$$PR_{min} = \begin{cases} 90\% \text{ du bénéfice technique} \\ \text{ou} \\ 100\% \text{ de la perte technique} \end{cases} + \begin{cases} 85\% \text{ du bénéfice financier} \\ \text{ou} \\ 0\% \text{ de la perte financière} \end{cases}$$

Mais aussi :

$$PR = PB + IT$$

Avec IT les intérêts techniques. On en déduit que la PB minimale s'écrit :

$$PB_{min} = PR_{min} - IT$$

**La PB contractuelle :** Des clauses contractuelles peuvent bonifier le mécanisme de PB. La PB contractuelle peut être définie au niveau du contrat ou d'un ensemble de contrats. Dans la plupart des compagnies, la somme des PB contractuelles est supérieure à la PB réglementaire.

**La PB discrétionnaire :** L'assureur peut décider de distribuer une part des produits financiers réalisés durant l'année à **tous ou une partie** des bénéficiaires des contrats d'assurance vie. C'est comparable à la rémunération variable discrétionnaire qui consiste pour un employeur à attribuer un bonus ou une prime, exceptionnelle ou non, à certains de ses salariés. Ici, l'assureur peut décider de servir en plus de la PB réglementaire ou contractuelle une PB discrétionnaire à certains assurés.

L'assureur dispose tout de même d'une certaine flexibilité vis-à-vis de la distribution de la participation aux bénéfices. Tout d'abord, il dispose de 8 ans à compter de leur constitution pour les distribuer et peut choisir de provisionner la PB dans la PPE (cf partie suivante). Il peut aussi choisir à quels assurés elle sera redistribuée, sans obligations de distribuer de la PB aux assurés qui en ont permis la constitution.

## Les options de rachat

Contrairement aux contrats de retraite ou temporaires décès qui ne sont pas rachetables pour éviter l'antisélection, les contrats d'épargne le sont. Deux types de rachats sont possibles : les rachats totaux par lesquels l'assuré rachète l'intégralité de son contrat et les rachats partiels, où seule une partie de l'épargne est rachetée. Par leur essence même d'option, les assurés peuvent choisir de l'exercer ou non.

Les rachats sont motivés par différentes raisons, dont en voici les deux principales :

- Le cadre fiscal avantageux dont bénéficie l'assurance vie. Ce cadre fiscal avantageux s'explique tout d'abord parce qu'un contrat d'assurance vie n'est pas assimilable à la détention d'un actif car c'est une créance sur un assureur, mais aussi parce qu'historiquement, l'assurance vie représentait un moyen de protection qu'il n'était pas légitime de taxer. Bien que les avantages fiscaux se soient réduits depuis les années 80, ils restent une des raisons principales du succès de l'assurance vie et visent à favoriser l'épargne longue durée.

Ce dispositif concerne essentiellement l'impôt sur le revenu. Pour l'épargne, les plus-values réalisées bénéficient d'un dispositif spécifique qui dépend de l'ancienneté du contrat. Nous considérons dans le tableau suivant que les pourcentages comprennent à la fois les prélèvements sociaux et les prélèvements forfaitaires dus à l'impôt sur le revenu. Nous distinguons aussi les cas avant et après 2017, date de l'introduction de la "flat tax" ou "prélèvement forfaitaire unique".

TABLE 1.1: Tableau simplifié du cadre fiscal sur le revenu

	Pour les gains issus de versements effectués avant le 27/09/2017	Pour les gains issus de versements effectués après le 27/09/2017
<b>Entre 0 et 4 ans</b>	52,20%	30%
<b>Entre 4 et 8 ans</b>	32,20%	30%
<b>Après 8 ans</b>	Abattement annuel sur la part d'intérêts rachetée de € 4600 pour une personne seule, € 9200 pour un couple	
	24,7%	24,7% pour les gains provenant des € 150000 versés et 30% au-delà de € 150000

Cette fiscalité explique que l'on trouve chez les assureurs une augmentation des rachats après 4 ans d'ancienneté et une vague encore plus significative après 8 ans.

En cas de contre-assurance en cas de décès, les contrats d'épargne disposent d'avantages sur l'impôt sur la succession que nous ne développerons pas ici.

- La conjoncture économique et la concurrence. En effet prenons l'exemple d'une remontée quasi immédiate et conséquente des taux. Dans cette situation, un nouvel assureur entrant sur le marché aura un rendement plus compétitif et pourra alors servir un taux plus avantageux à ses assurés comparé à celui servi par un assureur déjà présent sur le marché et ayant dans son portefeuille d'actif une majorité d'anciennes obligations peu fructueuses. L'assuré aura alors intérêt à racheter son contrat pour aller chez le nouveau concurrent. Ces rachats varient essentiellement en fonction du taux servi par l'assureur et le taux attendu par l'assuré.

Nous verrons par la suite que l'outil ALM modélise différemment les rachats dits "structurels" dus en partie au cadre fiscal et les rachats dits "conjoncturels" liés, comme leur nom l'indique, à la conjoncture économique.

### 1.2.4 Le provisionnement en assurance vie

Nous allons maintenant étudier le provisionnement en assurance vie. Dans une démarche de calcul de *Best Estimate* où nous cherchons à évaluer la juste valeur des options contractuelles comme celles que nous venons de mentionner, l'étude des provisions est nécessaire. Ce sont des mécanismes de pilotage permettant aux assureurs de répondre à leurs engagements et des leviers de management de leur stratégie. Nous étudions ici les provisions les plus répandues en assurance vie.

#### La Provision Mathématique (PM)

Elle représente la dette probable de l'assureur vis-à-vis des assurés.

Article R. 331-3 du code des assurances stipule que la Provision Mathématique est la différence entre les valeurs actuelles probables des engagements respectivement pris par l'assureur et les assurés.

$$PM = VAP_{assureur} - VAP_{assurés}$$

#### La Provision pour Participation aux Excédents (PPE)

La PPE représente le montant de participations aux bénéfices attribué aux assurés qui n'est pas versé immédiatement après la liquidation de l'exercice qui les a produits. Toutes les sommes affectées à la PPE doivent être restituées aux assurés sous un délai de 8 ans. L'utilisation de la PPE est laissée à l'appréciation de l'assureur qui peut s'en servir par exemple pour lisser le taux servi aux assurés ou proposer un taux très concurrentiel pour attirer de nouveaux clients. La PPE ne peut pas être utilisée pour servir le TMG.

Depuis l'arrêté du 24 décembre 2019 relatif aux fonds excédentaires en assurance vie et publié dans le journal officiel du Ministère de l'Economie et des Finances, la PPE peut être reprise, après acceptation de l'ACPR, dans le cas bien précis où les deux situations suivantes sont réunies :

1. Le résultat technique de l'assureur vie du dernier exercice comptable est négatif.
2. Le SCR n'est plus couvert.

Ce nouvel arrêté est entré en vigueur à la suite d'une demande des assureurs et doit permettre d'atténuer l'augmentation du SCR et par opposition la diminution du ratio de couverture des assureurs vie dans un environnement de taux bas particulièrement contraignant pour leur activité. Cependant, la PPE étant une provision faisant écho à un engagement envers les assurés, l'arrêté stipule que les montants de PPE prélevés doivent être restitués dans un délai de 8 ans et qu'aucun dividende ne peut être servi tant que cette restitution n'a été effectuée.

#### La Réserve de Capitalisation (RKPI)

La réserve de capitalisation est une provision qui vise à neutraliser les effets de plus et moins-values obligataires. La valeur de la réserve de capitalisation est modifiée en cas de vente d'obligation. Le versement ou prélèvement sur la réserve de capitalisation est fait pour que le taux actuariel du titre après versement ou prélèvement soit égal au taux actuariel à l'achat. En d'autres termes, la réserve de

capitalisation est alimentée par les plus-values réalisées sur les cessions d'obligations et reprise symétriquement en cas de moins-values. Son rôle est double : lisser le résultat et dissuader les compagnies d'assurance de vendre leurs obligations en cas de baisse des taux et de dégager des bénéfices ponctuels avant d'acquérir, plus tard, d'autres obligations moins performantes du fait de cette baisse des taux.

### **La Provision Globale de Gestion (PGG)**

Cette provision vise à couvrir les charges de gestion futures des contrats non couvertes par les chargements sur primes ou sur encours. Elle est calculée par groupes homogènes de contrats. Ainsi, un contrat ou groupe de contrat dont les chargements de gestion sont insuffisants doit faire l'objet d'une PGG. Elle est égale à la différence entre les valeurs actuelles probables des ressources et des charges de gestions.

### **La Provision pour Aléas Financiers (PAF)**

Elle est destinée à compenser la baisse de rendement de l'actif qui risquerait d'entraîner un décalage entre le taux de rendement des actifs et le taux contractuel garanti par l'assureur. Elle est calculée lorsque 80% du rendement financier d'une compagnie est inférieur au taux technique moyen garanti par l'assureur sur les contrats. Elle est égale à la différence entre les provisions mathématiques recalculées sur la base d'un taux technique égal à un taux prudent, par exemple 60% du TME, et les provisions mathématiques comptables. L'inconvénient de cette provision est l'absence de dimension prospective du risque de manque de revenu financier. Elle est familièrement qualifiée de « Provision pour Accélération à la Faillite ».

### **La Provision pour Dépréciation Durable (PDD)**

Le code des assurances détaille les mécanismes de dépréciation des actifs en fonction de leur qualification en titre amortissable ou non amortissable. Nous détaillons les situations dans lesquelles une dépréciation de l'actif est envisageable dans le tableau ci-dessous.

TABLE 1.2: Tableau des classes d'actifs et leur mécanismes respectifs de dépréciation

	Actifs R. 343-9 (titres amortissables)	Actifs R. 343-10 (titres non amortissables)
<b>Exemples d'actifs</b>	Obligations et obligations à coupons indexés	Actions, immobilier et SCI...
<b>Actif déprécié lorsque...</b>	le risque de défaut de l'émetteur est avéré	<ul style="list-style-type: none"> <li>• nous sommes en moins-value latente de 20% à 30% pendant au moins 6 mois consécutifs</li> <li>• nous avons des indices objectifs permettant de prévoir que l'émetteur ne pourra pas recouvrer tout ou partie de sa valeur</li> <li>• il existe déjà une dépréciation durable</li> </ul>

### La Provision pour risque d'exigibilité (PRE)

La PRE concerne seulement les articles relevant de l'article R 343-10 du code des assurances détaillé précédemment. La PRE est calculée au global de tous ces actifs lorsqu'ils se trouvent en situation de moins-values latente nette globale.

La dotation de PRE peut être lissée dans le temps : la dotation annuelle à la PRE au titre de l'exercice est égale à la moins-value latente nette globale divisée par la partie entière de la duration des passifs.

### Autres provisions

D'autres provisions existent comme la provision pour égalisation ou la provision pour frais d'acquisition reportés mais nous n'en détaillerons pas le fonctionnement ni l'utilisation dans ce mémoire.

### Conclusion du chapitre

Dans ce chapitre introductif, nous avons présenté des notions clés à la suite de notre étude. Un assureur vie proposant des produits d'épargne en euro s'engage contractuellement envers ces assurés à respecter des garanties et des options qui dépendent de manière directe au rendement des actifs investis. Dans le cadre de la réglementation Solvabilité 2, l'assureur doit calculer la juste valeur de ses engagements au travers du BE et par conséquent valoriser les options et garanties en question. C'est à cette fin que les assureurs ont développé les outils ALM. Cependant, la complexité de ces outils et le temps de calcul nécessaire limite encore les études ORSA qui devraient être menées à chaque modification du risque, ce qui n'est pas envisageable. Ainsi, nous verrons au travers de ce mémoire si une approche *machine learning* peut permettre de surpasser ces contraintes opérationnelles tout en restant exhaustive.

Nous allons ainsi dans le chapitre suivant développer plus en détails le calcul du BE et présenter l'outil ALM développé par Mazars.



## Chapitre 2

# L'outil ALM et le périmètre de l'étude

Une compagnie d'assurance vie commercialisant des produits d'épargne propose des options et garanties contractuelles qui nécessitent de modéliser les interactions entre l'actif et le passif. En l'occurrence, les contrats que nous considérons dans notre étude ont les caractéristiques suivantes :

- Ce sont des contrats euros et aucune part d'UC n'est considéré ;
- Une garantie de TMG peut être renseignée. Cette garantie assure au détenteur du contrat de voir son encours annuellement revalorisé à minima à ce taux. Nous supposons notre garantie de TMG valable toute la durée du contrat, c'est à dire jusqu'au rachat ou au décès de l'assuré ;
- Les options de rachat partiel ou complet sont possibles et cela à tout moment, mais une pénalité de rachats prenant la forme d'un taux est implémentée. Cette pénalité de rachat vise à les limiter ;
- Une garantie de PB contractuelle peut aussi être proposée. L'objectif est ainsi de proposer aux assurés une participation aux bénéfices supérieure à la PB réglementaire et donc plus avantageuse ;
- L'assuré peut alimenter son contrat d'épargne au travers de primes libres. Un chargement d'acquisition est néanmoins considéré.
- Un chargement sur encours est prélevé annuellement.

Ainsi, afin de valoriser les options et garanties mentionnées ci-dessus, les assureurs ont développé les outils ALM utilisant des nombreux scénarios économiques et modélisant au mieux les interactions entre passif et actif. Comprendre ces interactions entre l'actif et le passif est primordial pour la suite du mémoire car cela permet d'orienter nos choix de variables explicatives. De plus, cela permettra d'affiner notre compréhension des modèles issus de notre étude *machine learning*, ce qui est en accord avec notre démarche ORSA.

### 2.1 Le calcul du *Best Estimate* par méthode de Monte Carlo

Nous rappelons que le BE est la valeur actuelle probable des flux de trésorerie entrant nets de primes. Notre modèle considère trois types de cash flows sortants :

**Les prestations :** notre étude est axée sur l'activité d'épargne d'une assurance vie, nous considérons donc les rachats des contrats et les versements de capitaux en cas de décès comme prestations.

Ces prestations dépendent des garanties proposées dans les contrats. Ainsi, elles vont être impactées par la revalorisation annuelle de l'épargne des assurés, qui découle directement de la PB contractuelle et du TMG et de la capacité de l'assureur à obtenir du rendement sur ses actifs. De plus, des primes peuvent venir augmenter l'assiette de prestation et inversement, la pénalité de rachat va venir la diminuer. Finalement, l'âge des détenteurs des contrats joue un rôle sur les prestations de décès.

**Les taxes :** les prélèvements sociaux essentiellement.

**Les frais :** nous considérons dans notre étude les frais de placements financiers, de gestion de sinistres, d'administration et d'acquisition des primes.

Les cash flows entrant se limiteront aux primes annuelles.

Ainsi, afin de mener à bien ce calcul, Mazars a mis au point un outil prenant en entrée 1000 scénarios économiques et projetant l'actif et le passif de la société sur 30 ans tout en modélisant les interactions entre ces deux. Il permet donc de projeter les flux de trésorerie et calculer le BE grâce à la formule ci-dessous :

$$BE = \frac{1}{1000} \sum_{i=1}^{1000} \sum_{n=1}^{30} \delta_n^i \times P_n^i = \frac{1}{1000} \sum_{i=1}^{1000} BE^i \quad (2.1)$$

Où  $\delta_n^i$  est le facteur d'actualisation en  $t=0$  intervenant l'année  $n$  dans le scénario  $i$  et  $P_n^i$  est la somme des cash flows intervenant l'année  $n$  dans le scénario  $i$ . Les  $BE^i$  sont les BE évalués pour le scénario économique  $i$ . C'est un abus de langage car les  $BE^i$  ne sont pas des BE, leurs moyennes empiriques le sont. Afin de s'assurer une distinction claire entre ces deux variables, nous appellerons dans la suite les  $BE^i$  **BE par scénario** et leur moyenne le **BE**. Le choix de 1000 scénarios fait suite à une étude de convergence.

Voici ci-dessous une schématisation de la démarche menant au calcul du BE :

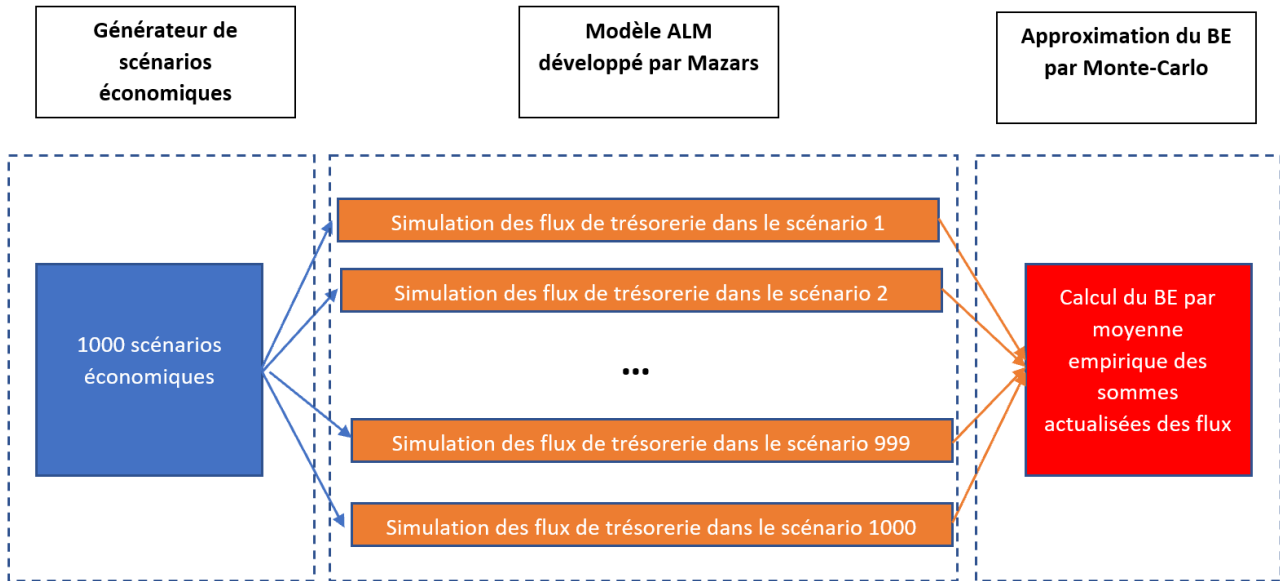


FIGURE 2.1: Architecture de calcul du BE



Le générateur de scénarios projette sur 30 ans les scénarios de taux et de rendement de l'actif. Pour chacun de ces scénarios, l'outil ALM projette année par année les flux de trésoreries  $P_n^i$  permettant le calcul des BE par scénario, puis du BE.

Nous allons maintenant détailler le fonctionnement du générateur de scénario économique, puis nous aborderons les mécanismes du modèle ALM.

## 2.2 Génération des scénarios économiques

Comme vu dans la partie précédente, le BE est approché par la méthode de Monte-Carlo, nous approximations donc l'espérance par une moyenne empirique dont la somme nécessite de pouvoir projeter les flux de prestations futurs de l'assureur selon un nombre important de scénario économiques.

Pour ce faire, nous avons besoin d'un générateur de scénarios économiques. Celui-ci a pour but de projeter les différents paramètres financiers tels que des taux, des volatilités, des rendements... et cela pour chaque simulation et sur un horizon de projection défini. Ces scénarios économiques doivent respecter deux conditions afin d'être correctement calibrés :

- Les scénarios générés doivent être *market consistent*, c'est-à-dire que la trajectoire de l'actif généré doit être en accord avec les conditions actuelles du marché.
- Les scénarios sont générés dans un univers risque-neutre, c'est-à-dire que les rendements moyens des actifs générés doivent être égaux au taux sans risque.

### 2.2.1 La structure à terme des taux d'intérêts

#### Modélisation de la courbe des taux par un modèle de Hull and White

L'outil ALM utilisé a recourt au modèle de Hull and White à un facteur pour diffuser la courbe des taux. Ce modèle prend en input la courbe des taux Spot à l'instant initial fourni par l'EIOPA\*. Nous justifions l'utilisation de ce modèle par sa simplicité et sa capacité à générer des taux négatifs, ce qui nous permet de ne pas allonger le temps de calcul des scénarios et de retranscrire l'univers de taux bas dans lequel nous vivons actuellement. Dans ce modèle, le taux court instantané  $r$  est solution de l'Equation Différentielle Stochastique (EDS) suivante

$$dr_t = (\theta_t - ar_t)dt + \sigma dW_t$$

Avec :

- $a$  une constante positive
- $W_t$  un mouvement brownien sous la probabilité risque neutre
- $\theta_t$  une fonction déterministe définie de manière que la structure des taux à l'instant  $t=0$  corresponde à celle des taux observés sur les marchés, dans notre cas le prix des zéros-coupons fourni

---

\*Autorité Européenne des Assurances et des Pensions Professionnelles : assure la stabilité et l'efficacité du système financier dans l'Union européenne

par l'EIOPA. En effet,  $\theta_t$  doit vérifier l'équation suivante :

$$\theta_t = \frac{df(0,t)}{dt} + af(0,t) + \frac{\sigma^2}{2a}(1 - e^{-2at})$$

avec  $f(0,t)$  le taux forward instantané à l'instant initial donné en input.

Le calibrage de  $a$  et  $\sigma$  a été effectué par le pôle ingénierie financière de Mazars Actuariat et a défini  $a$  comme étant égal à 0.10% et  $\sigma$  à 0.80%. Nous avons seulement eu à renseigner les taux sans risque fourni par l'EIOPA au 31/12/2019. Une fois ce calibrage effectué, il est possible de simuler les taux courts via l'approximation suivante que nous ne démontrerons pas :

$$\forall s < t, r_t \approx r_s e^{-a(t-s)} + f(0,t) - e^{-a(t-s)} f(0,s) + \sigma \sqrt{\frac{(1 - e^{-2a(t-s)})}{2a}} \epsilon_t$$

Avec  $\epsilon_t \sim \mathcal{N}(0,1)$

### Détermination du prix des zéros-coupons

Soit  $P(t,T)$  le prix du ZC à l'instant  $t$  de maturité  $T$ .  $P(t,T)$  est une fonction du temps  $t$  et du taux court  $r_t$  qui se définit ainsi :

$$P(t,T) = A(t,T)e^{-B(t,T)r_t}$$

Avec :

$$A(t,T) = \frac{P(0,T)}{P(0,t)} \times \exp \left[ B(t,T)f(0,t) - \frac{\sigma^2}{4a}(1 - e^{-2at})B(t,T)^2 \right]$$

$$B(t,T) = \frac{1 - e^{-a(T-t)}}{a}$$

Remarque :  $P(0,t) = \frac{1}{(1+f(0,t))^t}$

### Détermination de la structure par termes des taux

Ayant préalablement calculé les taux courts puis les prix des zéros-coupons, nous disposons de la structure par termes de taux  $R(t,T)$  via la relation suivante :

$$R(t,T) = -\frac{\ln(P(t,T))}{T-t}$$

A noter que l'implémentation ne permet pas au taux de dépasser une valeur minimale de  $-2\%$ .

Ces taux nous permettent finalement de calculer la valeur de marché des obligations de nominal  $N$  et de taux de coupon  $T_c$  à tout temps  $t$  et pour toutes les maturités  $T$  en appliquant la formule suivante :

$$V_m(t,T) = \sum_{i=1}^{T-t} \frac{T_c \times N}{(1 + R(t,t+i))^i} + \frac{N}{(1 + R(t,T))^{T-t}}$$

### 2.2.2 Modèle des rendements actions

Nous nous plaçons en univers risque neutre. Nous utilisons le modèle de Black&Scholes (B&S) pour modéliser le cours de l'action, qui vérifie donc l'équation différentielle :

$$dS_t = S_t(bdt + \sigma dW_t)$$

avec  $b$  le drift et  $\sigma$  la volatilité et  $W_t$  le mouvement Brownien.

Par la formule d'Ito, on obtient :

$$S_t = S_0 e^{(b - \frac{\sigma^2}{2})t + \sigma W_t}$$

Il peut en être déduit le rendement de l'actif entre les dates  $t$  et  $t+1$  :

$$\log(S_{t+1}) - \log(S_t) = (b - \frac{\sigma^2}{2})(t+1 - t) + \sigma(W_{t+1} - W_t) \Leftrightarrow \log(S_{t+1}) - \log(S_t) = b - \frac{\sigma^2}{2} + \sigma(W_{t+1} - W_t)$$

Or, par définition du mouvement brownien,  $X_t = W_{t+1} - W_t$  suit une  $\mathcal{N}(0, 1)$ . Nous obtenons finalement :

$$\log(S_{t+1}) - \log(S_t) = b - \frac{\sigma^2}{2} + \sigma X_t$$

Il est important de préciser que sous la probabilité risque neutre, l'espérance du rendement est égale au taux sans risque. De plus, nous avons recourt à la volatilité implicite et à l'inversion du prix d'un call pour obtenir la volatilité. En effet, nous rappelons que dans le modèle de B&S le prix d'un call s'obtient par formule fermée :

$$Black_{call}(S_0, k, r, T, \sigma) = S_0 \times \Phi(d_1) - K e^{-rt} \times \Phi(d_2)$$

Avec :

- $\Phi$  la fonction de répartition d'une  $\mathcal{N}(0, 1)$
- $d_1 = \frac{\ln(\frac{S_0}{K} + (r + \frac{\sigma^2}{2})T)}{\sigma\sqrt{T}}$
- $d_2 = d_1 - \sigma\sqrt{T}$
- $K$  le strike du call
- $r$  le taux sans risque

Nous obtenons finalement une volatilité implicite de 18%

### 2.2.3 Modèle des rendements immobiliers

La modélisation du cours de l'immobilier est identique et nous déduisons la volatilité associée au cours de l'immobilier à partir de la volatilité du cours de l'action. Pour cela nous faisons l'hypothèse que le rapport entre les volatilités actions et immobilier est environ égal au rapport des chocs du QIS5 associés. La volatilité implicite retenue finalement est 10%.

## 2.3 Les mécanismes de l'outil ALM

Nous avons vu dans la partie précédente la manière dont sont générés les scénarios économiques, nous allons maintenant tâcher de détailler la manière dont sont modélisés l'actif et le passif. Nous détaillerons aussi les principales étapes nécessaires à la projection de notre assureur sur 30 ans et à la valorisation du BE.

Nous nous aiderons du schéma ci-dessous, qui décrit la chronologie de l'outil ALM et nous permet de mieux appréhender la projection temporelle des flux de trésorerie.

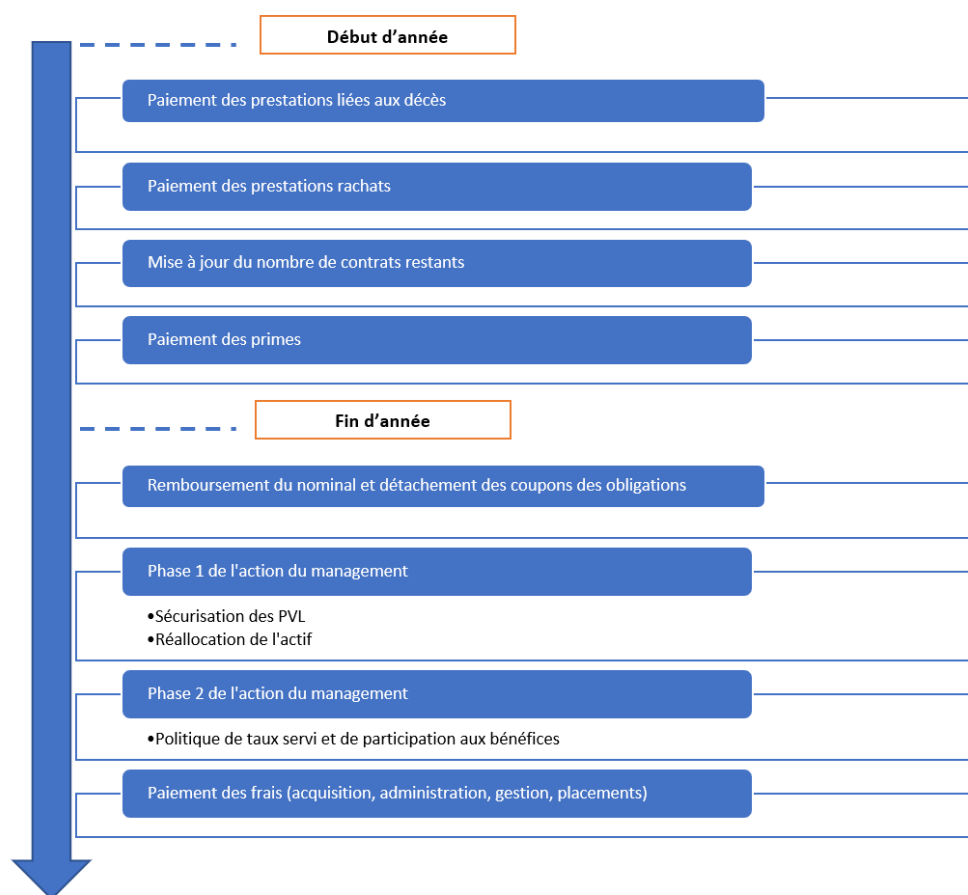


FIGURE 2.2: Schéma simplifié des étapes annuelles de l'outil ALM

Nous allons tâcher à présent d'entrer plus précisément dans ces rouages.

## 2.4 La modélisation du passif

Dans cette partie, nous présentons la modélisation du passif, plus spécialement les hypothèses simplificatrices effectuées dans l'outil ALM. De plus, nous détaillerons l'évolution des différents postes du passif en lien avec le calcul de *Best Estimate*.

### 2.4.1 Les hypothèses de modélisation

Des hypothèses ont été effectuées au moment de la réalisation de l'outil ALM. Ces hypothèses simplificatrices doivent respecter le cadre de Solvabilité 2 et plus spécialement permettre un calcul du *Best Estimate* reflétant au mieux les flux de trésorerie futurs de l'assureur.

Voici la liste des hypothèses simplificatrices :

- La mortalité est supposée déterministe et fonction de l'âge des assurés ;
- Les décès et rachats ont lieu en fin d'année ;
- Les revalorisations ont lieu en fin d'année, avant les rachats et décès ;
- Les contrats sont regroupés homogènement en *model points* (MP) selon des critères discriminants ;
- Tous les éléments du bilan ne sont pas modélisés, les seules provisions modélisées sont la PM, la PPE et la réserve de capitalisation ;
- La société fonctionne en run-off, c'est-à-dire qu'aucune police n'est souscrite lors de la projection et seules les primes futures des contrats en stock sont comptabilisées ;
- L'entreprise prend fin au bout de 30 ans et les assurés restant en fin de projection rachètent leurs contrats.

### 2.4.2 Modélisation et évolutions des éléments du passif apparaissant au bilan

La provision mathématique est calculée à la clôture de la manière suivante :

$$PM_{Cl\acute{o}ture}(t) = PM_{Ouverture} - D\acute{e}c\grave{e}s - Rachats + Primes - Chargement acquisition \\ + Revalorisation - Commission sur encours - Pr\acute{e}l\grave{e}vements sociaux \quad (2.2)$$

Les fonds propres évoluent aussi durant l'année car, selon une approche comptable, les fonds propres sont composés du capital social (part des actionnaires), de la réserve (bénéfices des années précédentes non redistribués aux actionnaires) et du résultat de l'année. Nous supposons dans notre outil que l'assureur ne redistribue jamais les bénéfices aux actionnaires, ce qui nous amène à recalculer les fonds propres chaque année en ajoutant le résultat de l'année aux fonds propres de l'année précédente.

$$FP_N = FP_{N-1} + R_N$$

avec  $FP_N$  les fonds propres de l'assureur lors de l'année N et  $R_N$  le résultat statutaire en normes comptables française de la compagnie cette même année. Nous ne considérons pas d'imposition sur le résultat.

La PPE varie à plusieurs étapes de notre outil en fonction de la conjoncture économique et des actions du management. Celle-ci est un élément indispensable de l'outil ALM, notamment par son rôle dans la revalorisation des engagements. L'évolution de la PPE sera détaillée par la suite, dans la section 2.5.2 dédiée au *Management Actions*.

La réserve de capitalisation est, pour sa part, mise à jour à chaque cession d'obligation.

### 2.4.3 Retour au calcul du BE

En nous référant à l'équation (2.1), le calcul du BE par scénario et par extension du BE passe par le calcul des flux de trésorerie  $P_n^i$  nécessitant l'outil ALM. Nous détaillons maintenant la composition de ces flux de trésorerie :

$$P_n^i = \text{Décès} + \text{Rachats} - \text{Primes} + \text{Prélèvements sociaux} + \text{Frais}$$

#### Païement des prestations liées au décès

Comme nous le remarquons sur la figure 2.2, nous supposons que les prestations de décès sont servies en début d'année. Plusieurs hypothèses sont effectuées : nous supposons premièrement que les décès sont déterministes en fonction de l'âge de l'assuré et deuxièmement, nous considérons que la prestation décès d'un assuré est immédiatement payée par l'assureur. La table de mortalité utilisée est la table "TF 00-02".

Cela étant dit, nous définissons pour chaque *model point* la PM qui doit être libérée pour les bénéficiaires des assurés décédés de la manière suivante :

$$\text{Prestation}_{DC}(t) = \sum_{i=1}^{Nb_{MP}} tx_{DC}(i) PM_i(t-1)$$

Avec  $i$  le numéro du *model point* et  $PM_i(t-1)$  le montant de provision mathématique du model point  $i$  en fin d'année  $t-1$ .

#### Païement des prestations liées aux rachats

Comme mentionné dans le chapitre précédent, nous distinguons les rachats structurels des rachats dynamiques. Cela entraîne donc deux modélisations différentes.

##### Les rachats structurels

Nous rappelons que les rachats structurels sont liés à la fiscalité avantageuse après les paliers de 4 et 8 ans d'ancienneté du contrat d'épargne. Ainsi, les rachats structurels sont fonctions de l'ancienneté et exprimés en pourcentage de la provision mathématique du model point.

Nous gardons la même fonction pour les contrats souscrits après 2017 malgré la mise en place de la "flat tax" qui vient supprimer le palier des 4 ans. Nous supposons également que cette fonction est le fruit d'une étude statistique menée par notre compagnie fictive.

##### Les rachats conjoncturels

Les rachats conjoncturels ou dynamiques modélisent les rachats expliqués par les conditions actuelles des marchés financiers. Cela implique que l'assuré est attentif à la conjoncture économique et aux opportunités offertes par l'ensemble des concurrents sur le marché.

Mazars utilise dans son outil la loi de rachat proposée par l'ACPR dans les orientations nationales complémentaires QIS 5\* dont le résultat dépend de l'écart entre le taux servi par l'assureur  $TS$  et le

---

\*Quantitative Impact Study

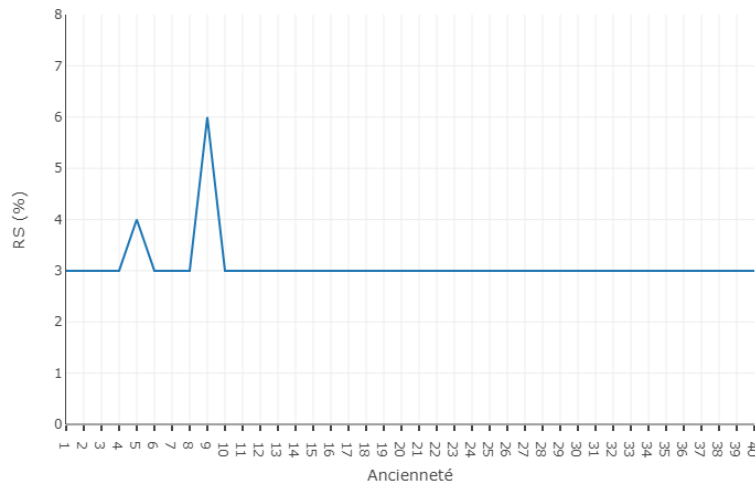


FIGURE 2.3: Courbe des rachats structurels

taux attendu par l'assuré  $TA$  aux vues des conditions économiques. Celle-ci se détaille ainsi :

$$RC = \begin{cases} RC_{max} & \text{si } TS - TA < \alpha \\ RC_{max} \times \frac{TS - TA - \beta}{\alpha - \beta} & \text{si } \alpha \leq TS - TA < \beta \\ 0 & \text{si } \beta \leq TS - TA < \gamma \\ RC_{min} \times \frac{TS - TA - \gamma}{\delta - \gamma} & \text{si } \gamma \leq TS - TA < \delta \\ RC_{min} & \text{si } TS - TA \geq \delta \end{cases}$$

L'ACPR propose des plages de variations pour les 6 paramètres et il a été retenu dans l'outil la valeur moyenne pour chacun d'entre eux. Nous résumons cela dans le tableau suivant :

TABLE 2.1: Paramètres de la loi de rachats conjoncturels proposée par l'ACPR

	Plafond min	Plafond max	Outil Mazars
$\alpha$	-6%	-4%	-5%
$\beta$	-2%	0%	-1%
$\gamma$	1%	1%	1%
$\delta$	2%	4%	3%
$RC_{min}$	-6%	-4%	-5%
$RC_{max}$	20%	40%	30%

et voici l'allure de cette loi en fonction du spread de taux  $TS - TA$ .

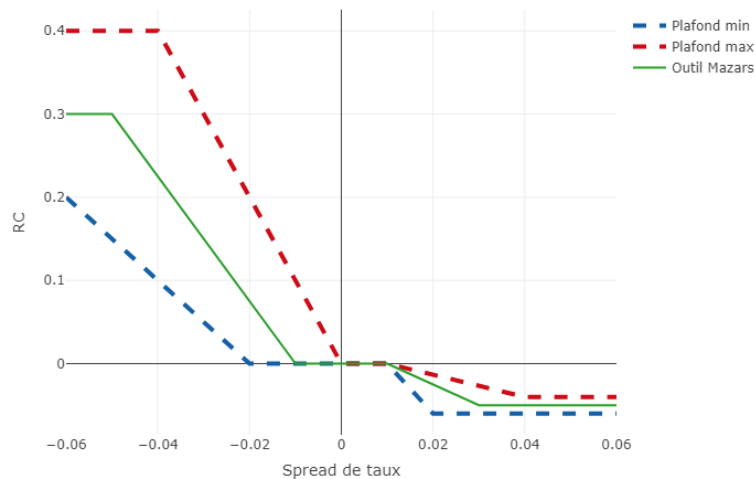


FIGURE 2.4: Rachats conjoncturels en fonction du spread de taux

Un lecteur intéressé pourra se référer au mémoire de Sylvain Detroulleau et Sandrine Mouret (SYLVAIN DETROUILLEAU, 2013) qui ont étudié les apports de la logique floue dans la modélisation des rachats conjoncturels, ce qui pourrait représenter une approche alternative.

Les rachats conjoncturels concernent les assurés sensibles aux taux pratiqués sur le marché, il est ainsi considéré dans cet outil que le taux attendu par les assurés est le taux concurrent que nous allons définir ci-dessous.

### Le taux concurrent

Le taux concurrent modélise le taux pratiqué par la concurrence sur le marché et qui pourrait inciter l'assuré à racheter son contrat d'épargne afin de choisir un nouveau placement plus attractif et fructueux. Un assuré dispose d'une multitude d'informations venant d'acteurs divers, ce qui rend la modélisation de la concurrence très complexe. Chaque assureur propose un produit différent, avec des caractéristiques contractuelles qui lui sont propres et à cela s'ajoute d'autres options d'épargne comme le livret A. C'est donc une simplification forte que de résumer l'information de la concurrence en un unique taux.

Deux approches sont possibles afin de modéliser la concurrence. La première approche est proposée par l'ACPR et suggère l'utilisation du taux moyen des emprunts d'Etat comme estimateur du taux concurrent. La seconde approche, implémentée dans l'outil Mazars, est d'utiliser un taux composite qui se construit à partir de plusieurs références, cela permettant de se rapprocher de la réalité du marché de l'épargne.

L'outil ALM vise à résumer dans ce taux composite l'information de trois profils de concurrences différents :

- Un assureur plus "agressif" dans son allocation cible d'actif ;
- Un nouvel assureur entrant sur le marché ;
- La banque et le livret A.

Un assureur plus agressif va chercher un meilleur rendement sur ses investissements et va tourner son allocation d'actif vers les actions. Ainsi, nous définissons son allocation comme étant composée de



30% d'action et 70% d'obligations, ce qui nous permet de définir le rendement du concurrent agressif à la date  $t$  de la manière suivante :

$$Taux_{Aggressif}(t) = \frac{1}{5} \sum_{i=0}^4 0,3 \times rendement_{actions}(t-i) + 0,7 \times taux_{ZC}(t-i, m)$$

Le rendement de l'actif obligataire est estimé par le taux zéros-coupons de maturité  $m$  égale à 8 ans et le rendement actions est directement fourni par le générateur de scénarios économiques comme vu précédemment. On effectue une moyenne sur 5 ans afin de lisser ce taux et de modéliser une prise d'information sur plusieurs années de la part de l'assuré.

Le nouvel assureur est un nouvel entrant sur le marché et dispose d'un rendement d'actif égal au rendement obligataire d'une OAT 10 ans à la date considérée, contrairement aux assureurs déjà établis sur le marché dont le rendement est lissé par la présence d'obligations anciennes dont les taux de coupons reflètent l'environnement obligataire historique. En cas de remontée brutale des taux, le nouvel entrant bénéficiera pleinement de ce nouvel environnement et proposera des revalorisations supérieures à celles des assureurs établis, qui souffrent de l'inertie obligataire de leur portefeuille.

$$Taux_{Nouveau}(t) = \frac{1}{5} \sum_{i=0}^4 taux_{ZC}(t-i, 10)$$

Finalement, nous modélisons le livret A vers qui, comme nous l'avons vu, se sont tournés les français pour épargner pendant le confinement et s'affirme comme le grand concurrent des assurances vie. Le taux du livret A est ainsi défini dans notre modèle comme le taux d'emprunt court, le taux zéro coupon de maturité 1 an.

$$Taux_{LivretA}(t) = taux_{ZC}(t, 1)$$

Finalement, le taux concurrent est le taux maximal distribué sur le marché, ce qui se traduit dans notre modélisation par le maximum de nos trois taux de référence introduits ci-dessus.

$$Taux_{Concurrent}(t) = \max(Taux_{Aggressif}(t), Taux_{Nouveau}(t), Taux_{LivretA}(t))$$

Le taux concurrent influence directement les rachats conjoncturels car nous modélisons le taux attendu  $TA$  par le taux concurrent  $TC$ . Ainsi, si nous prenons l'exemple où  $TS$  est supérieur à  $TC$ , une augmentation du taux concurrent va venir diminuer le spread de taux  $TS - TA = TS - TC$  et entraîner une augmentation des rachats conjoncturels  $RC$ . Bien évidemment, nous verrons par la suite qu'une action du management vise à augmenter le taux servi  $TS$  et ainsi amortir les rachats conjoncturels. De plus, le taux concurrent synthétise une partie de l'information économique, celui-ci moyennant les taux ZC et les rendements actions sur plusieurs années.

### Le taux de rachat total

Après avoir modélisé les rachats structurels et les rachats conjoncturels, nous définissons le taux de rachat total  $RT$  qui vient agréger ces deux informations de la manière suivante :

$$RT = \min(1, Max(0, RS + RC))$$

Dans le cas où le taux concurrent est largement inférieur au taux servi par notre assureur, le taux de rachat conjoncturel peut se retrouver négatif et entraîner un taux de rachat total inférieur au taux de rachat structurel. Inversement, en cas de taux concurrent supérieur au taux servi par l'assureur, le taux de rachat total est supérieur au taux structurel. Nous calculons un taux de rachat total par *Model Point*, celui-ci dépendant directement des paramètres du *Model Point* (ancienneté, âge...).

Finalement, ce taux de rachat nous permet de calculer les prestations payées par l'assureur au titre des rachats. Nous supposons que les contrats d'épargne modélisés imposent une pénalité en cas de rachat qui se traduit par un taux de pénalité  $tx_{pénalité\text{rachat}}$  qui restera fixe et identique pour tous les contrats.

$$Prestation_{Rachat}(t) = (1 - tx_{pénalité\text{rachat}}) \sum_{i=1}^{Nb_{MP}} RT(i) \times PM_i(t-1)$$

$$Pénalité_{Rachat}(t) = tx_{pénalité\text{rachat}} \times \sum_{i=1}^{Nb_{MP}} RT(i) \times PM_i(t-1)$$

### **Paiement des frais et des taxes**

La dernière étape annuelle de l'outil ALM est le paiement des frais. Ceux-ci sont ventilés de la manière suivante :

- Les frais de placement : calculés en pourcentage de la valeur comptable de l'actif.
- Les frais d'acquisition : calculés en pourcentage des primes versées brutes de chargement.
- Les frais de gestion : calculés en pourcentage des prestations rachats et décès.
- Les frais d'administration : calculés en pourcentage des provisions mathématiques.

Les prélèvements sociaux représentent l'intégralité des taxes payées. Les prélèvements sociaux sont modélisés à travers un taux de prélèvements sociaux fixé et appliqué à la provision mathématique.

### **Paiement des primes**

Nous faisons l'hypothèse que seuls les contrats en stock au moment du calcul du BE émettent des primes annuellement. Lorsque nous considérerons l'entrée en portefeuille de nouveaux contrats dans notre étude, le *New Business* (NB), nous ne projeterons pas de primes les concernant. Les primes émises en année  $t$  pour chaque MP  $i$  s'expriment comme suit :

$$Primes_{émise} = \begin{cases} Primes(1, i) & \text{si ancienneté} = 1 \\ Primes(t-1, i) \times (1 - tx_{chuteprimes}) \times \frac{Nb_{contrats}(t, i)}{Nb_{contrats}(t-1, i)} & \text{sinon} \end{cases}$$

Avec  $Primes(1, i)$  les primes versées par le MP  $i$  à son entrée dans le portefeuille de l'assureur. Nous introduisons ici un nouveau paramètre qui est le taux de chute des primes  $tx_{chuteprimes}$ , qui s'élèvera à 2%. Comme nous l'avons mentionné dans le chapitre dédié à l'assurance vie, un assuré titulaire

d'un contrat d'épargne n'a aucune obligation de versement de primes et le taux de chute des primes représente cette décroissance du volume des primes versées à l'assureur. De plus, afin de payer les frais d'acquisition, l'assureur va prélever des chargements d'acquisition sur les primes brutes. Les primes réellement acquises par l'assuré sont les primes nettes de chargement d'acquisition.

Nous allons maintenant traiter des actions du mangement. Cette partie est importante car elle définit la stratégie de l'assureur tout au long de la projection. Ainsi, ceux-ci sont des actions que l'assureur prend en fonction de la situation dans laquelle il se trouve, afin d'à minima pouvoir répondre à tout moment à ses engagements et aux garanties contractuelles de chaque MP. Au travers de ces actions, l'assureur cherche à rester compétitif et éviter les rachats.

## 2.5 La modélisation de l'actif

L'interaction entre l'actif et le passif va définir et influencer les engagements de l'assureur envers ses assurés. Par conséquent, il choisit précautionneusement les produits qu'il intègre à son portefeuille d'actif, en veillant à piloter les risques qui en découlent. Ainsi, l'outil utilisé projette l'actif de manière à refléter et respecter la conjoncture économique dans laquelle évolue notre assureur. De plus, bien que très simplifié, notre portefeuille d'actif se compose de produit que l'on retrouve très fréquemment chez les assureurs. Plus précisément, nous retrouvons les classes d'actifs suivantes :

- Les actions ;
- Les obligations ;
- Le monétaire ;
- L'immobilier.

### Le portefeuille obligataire

Seules des obligations simples à taux fixes en fin d'année sont implémentées dans l'outil ALM. De plus, celles-ci sont supposées notées AAA par simplification, afin de se détacher du risque de spread. Cela n'est pas très représentatif du marché, les portefeuilles obligataire des assureur ayant généralement un rating moyen proche de A.

L'allocation d'actif d'un assureur est composée en majorité d'obligations, qui présentent l'intérêt d'être des actifs peu risqués et facilitant l'adossement actif/passif, celle-ci distribuant à des échéances connues des coupons de valeurs connus (dans les cas des obligation à taux fixes ou OAT). Nous rappelons qu'une obligation est un titre de dette émis par une société ou un état. La personne titulaire d'une obligation reçoit des intérêts appelés coupons et spécifiés par le taux facial  $T_c$ . La maturité est la date à laquelle la somme est remboursée en totalité. A cette maturité  $T$ , le détenteur de l'obligation reçoit une somme  $N$  appelée nominal. Ainsi, chaque fin d'année, les obligations en portefeuille libèrent des coupons si elles ne sont pas à maturité et le nominal si c'est le cas. De plus, sa valeur de marché se calcule comme suit :

$$VM(t) = \sum_{i=1}^{T-t} \frac{T_c \times N}{(1 + R(t, t+i))^i} + \frac{N}{(1 + R(t, T))^{T-t}}$$

avec  $R(t, t+i)$  le taux zéro-coupon à la date  $t$  et de maturité  $t+i$  (cf 2.2.1). Pour actualiser l'actif obligataire en valeur comptable, le taux zero coupons est remplacé par le taux actuariel à l'achat dans

la formule ci-dessus. Cependant par simplification, nous supposons nos obligations achetées au pair, c'est à dire que le prix mais aussi la valeur nette comptable sont égales au nominal.

### Le portefeuille d'actions

Afin de simplifier l'outil, nous ne considérons qu'un unique model point action. Dans ce mémoire, nous supposons que nos actions évoluent selon l'indice du CAC 40.

### Le portefeuille monétaire

Nous considérons que notre portefeuille monétaire nous rapporte le taux sans risque.

#### 2.5.1 Phase 1 de l'action du management

##### Modélisation des produits financiers

Une fois les coupons sur obligations générés, nous calculons les produits financiers comme suit :

$$PF = Coupons_{Oblig} + Intérêts_{Monet} + 10\%PVL_{Actions} - Frais_{Placements}$$

Avec  $Intérêts_{Monet}$  représentant les intérêts sur l'actif monétaire, le rendement monétaire correspondant au taux ZC à un an projeté. Le terme  $10\%PVL_{Actions}$  suppose que l'assureur effectue systématiquement 10% de plus-values sur les actions, quand la situation le permet. Le terme  $Coupons_{Oblig}$  prend en compte les coupons mais aussi le nominal des obligations arrivées à terme.

De plus, le modèle ne sépare pas les actifs en représentation des capitaux propres, c'est à dire à destination de l'assureur, et ceux en représentation de la provision mathématique à destination des assurés. Or, cette distinction est importante comme en atteste la section 1.2.3, où nous avons vu que les assureurs utilisent, pour calculer la PB réglementaire, la participation des assurés aux résultats qui se calcule grâce aux résultats financiers en représentation de la provision mathématiques et non des fonds propres. Ainsi, l'outil répartit les produits financiers au prorata de la part actionnaires et assurés dans le passif. A cette fin, la PM et la PPE sont jugées à destination des assurés et les FP auxquels sont ajoutés la réserve de capitalisation appartiennent aux actionnaires et donc l'assureur. Finalement, cela se retranscrit ainsi :

$$Part_{assurés}(t) = \frac{PM(t-1) + PPE(t-1)}{Passif(t-1)}$$

$$Part_{assureur}(t) = \frac{FP(t-1) + RKPI(t-1)}{Passif(t-1)}$$

Ce qui nous permet de répartir les produits financier ainsi :

$$PF_{assurés}(t) = Part_{assurés}(t) \times PF(t) \tag{2.3}$$

$$PF_{assureur}(t) = Part_{assureur}(t) \times PF(t)$$

## Réallocation de l'actif

La revalorisation de la provision mathématique ainsi que le paiement des prestations du début d'année modifient inévitablement la structure du bilan et par conséquent de l'actif. Comme vu précédemment, les valeurs de marché et comptables des obligations sont recalculées après récupérations des coupons, à cela s'ajoute les réalisations de possibles plus-values latentes sur les actions au moment de la revalorisation de la provision mathématique. Enfin et surtout l'actif monétaire est impacté par les entrées et sorties de cash subies par l'assureur en fin d'année précédente ainsi que pour payer les prestations de début d'année.

Ainsi, l'ensemble de l'actif est impacté, mais non proportionnellement : l'allocation en est donc modifiée. Nous faisons dans cet outil l'hypothèse que l'assureur cherche à retrouver une allocation cible en valeur de marché constante paramétrée dans l'outil. Cela n'est cependant pas très représentatif du fonctionnement d'un assureur qui va s'adapter à la situation du marché.

Cette méthode de réallocations est basée sur des simplifications fortes. Nous supposons tout d'abord nos actifs parfaitement liquides. De plus, l'assureur peut se retrouver dans une situation où, pour atteindre l'allocation cible, il doit acheter la moitié d'une obligation. Cela est supposé possible dans notre modèle, les actifs étant infiniment divisibles. Finalement, notre assureur peut acheter et vendre des actifs à sa guise et dans les quantités souhaitées.

Dans le cas des obligations, l'outil est implémenté de manière à vendre prioritairement les obligations de maturité les plus proches et à acheter uniquement des obligations d'Etat AAA de maturité 10 ans cotant au pair. Notre modèle attribue un taux de coupon aux nouvelles obligations afin que la valeur comptable soit égale la valeur de marché.

Finalement, les ventes d'actifs entraînent la réalisations de plus ou moins-values qui sont traitées différemment en fonction de l'actif en question : les plus ou moins-values réalisées sur les cessions d'obligations sont affectées à la réserve de capitalisation et celles réalisées sur les cessions d'actions ou d'immobilier s'ajoutent aux produits financiers.

### 2.5.2 Phase 2 de l'action du management

Cette seconde phase du management détermine le taux auquel l'assureur revalorise les PM des assurés. Contrairement à la réallocation de l'actif où l'allocation cible est fixe, le taux de revalorisation ne peut être fixe car il se doit de retranscrire la réalité économique de la simulation. Ainsi, le management se fixe un objectif en matière de revalorisation qui n'est autre que le taux concurrent (cf 2.4.3). Comme vu précédemment, le taux concurrent vise à modéliser le taux observé sur le marché. Ainsi, l'objectif de notre assureur est de servir un taux qui n'est pas inférieur au taux dont l'assuré pourrait disposer sur le marché. Si cet objectif est rempli, l'assuré est satisfait et les rachats conjoncturels sont nuls. L'objectif du management n'est pas de dépasser le taux concurrent mais bien de l'égaliser.

Si le taux contractuel est supérieur au taux concurrent, le management revalorisera la provision mathématique à un taux supérieur au taux concurrent. De plus, l'assureur n'est pas certain de pouvoir atteindre cet objectif, cela dépendant de ses ressources.

Nous étudierons par la suite l'algorithme spécifiant les actions du management pour atteindre cet objectif.

### Le taux contractuel

La revalorisation contractuelle représente la revalorisation due aux engagements de l'assureur envers l'assuré. C'est un calcul qui s'effectue donc par *Model Point*, chaque model point ayant des caractéristiques contractuelles différentes, et qui dépend du TMG et de la PB contractuelle. Pour le *Model Point*  $i$  on a pour revalorisation contractuelle  $RC_i$  :

$$RContr_i = \max(PM_i \times TMG_i; PB_{contractuelle_i} \times PF_{assuré} \times \frac{PM_i}{PM})$$

Avec :

- $TMG_i$  le taux minimum garanti associé au model point  $i$ ;
- $PF_{assuré}$  défini comme dans la formule (2.3) ;
- $\frac{PM_i}{PM}$  correspond à la part de la provision mathématique du model point  $i$  dans la provision mathématique globale.

Finalement, nous introduisons le taux de revalorisation garantie du model point  $i$   $TContrat_i$  comme le rapport entre sa revalorisation contractuelle et sa provision mathématiques :

$$TContr_i = \frac{RContr_i}{PM_i}$$

### Lien entre revalorisation garantie et revalorisation cible

Nous avons introduit l'objectif du management et le taux de revalorisation garantie ; nous pouvons maintenant définir la revalorisation cible  $RC_i$  visée par l'assureur pour le *Model Point*  $i$  :

$$RC_i = PM_i \times \max(TC, TContr_i)$$

en rappelant que  $TC$  est le taux concurrent.

Cette formule est en accord avec nos remarques précédentes : dans le cas où le taux concurrent est supérieur au taux de revalorisation garantie, le management va chercher à revaloriser à hauteur du taux concurrent pour limiter les rachats conjoncturels. A l'inverse, si le taux de revalorisation garantie est supérieur au taux concurrent, le management se doit contractuellement de servir le taux de revalorisation garantie.

Globalement, le niveau de revalorisation cible de l'ensemble du portefeuille se définit comme la somme des revalorisation cible de chaque model point.

$$RC = \sum_i^{Nb_{MP}} RC_i$$

Nous allons maintenant étudier l'algorithme de politique de taux servi, c'est à dire la stratégie d'allocation des ressources financières disponibles pour l'assureur afin d'essayer d'atteindre ce montant de revalorisation.

### L'algorithme de politique de taux servi

La première étape de cet algorithme est de servir la revalorisation contractuelle. L'algorithme va premièrement vérifier si les produits financiers générés par la richesse des assurés ( $PF_{assurés}$ ) suffisent à servir cette revalorisation. S'ils sont suffisants, le nouvel objectif va être d'atteindre la revalorisation cible, dans le cas où le taux concurrent est supérieur au taux contractuel. En cas d'insuffisance, l'assureur va commencer, dans les limites du possible, par effectuer des plus-values latentes sur les actions et si cela n'est toujours pas suffisant, il abandonne les produits financiers en représentation des fonds propres et de la réserve de capitalisation, c'est à dire  $PF_{assureur}$ . Si ce sacrifice n'est toujours pas suffisant l'assureur n'a d'autre choix que d'enregistrer une perte pour servir la revalorisation garantie. Ce schéma résume ces mécanismes :

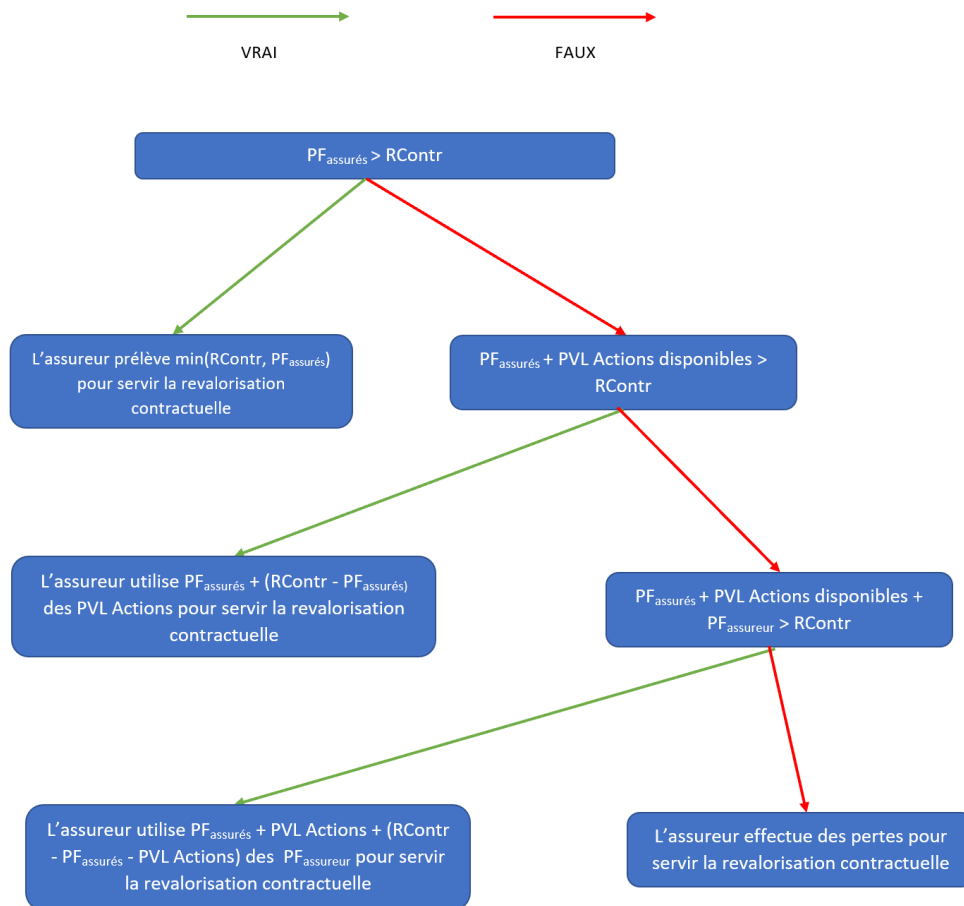


FIGURE 2.5: Etape 1 de l'algorithme, servir la revalorisation contractuelle

A noter qu'un assureur ayant subi des pertes pour servir la revalorisation garantie peut malgré tout atteindre la revalorisation cible, et cela parce qu'il peut utiliser la PPE pour servir le TMG et la revalorisation garantie.

S'ensuit alors la seconde étape de l'algorithme qui va viser à atteindre la revalorisation cible. A ce stade, la revalorisation contractuelle a été servie conformément à l'étape 1 et l'assureur va utiliser les ressources qu'il lui reste. Comme nous l'avons rappelé dans la section 1.2.4 concernant la PPE, celle-ci doit être redistribuée sous 8 ans. Pour satisfaire cela, nous redistribuons chaque année 15% de la PPE

aux assurés. Ainsi, cette seconde étape commence par vérifier si 15% de la PPE est suffisant pour atteindre la revalorisation cible. Si c'est le cas, les objectifs du management sont atteints et l'assureur évite les rachats conjoncturels. A l'inverse, si cela est insuffisant, l'assureur va s'appuyer sur ses autres ressources pour servir la revalorisation cible. Premièrement, si tous les produits financiers à destination des assurés n'ont pas été utilisés dans l'étape 1, l'assureur va utiliser les  $PF_{assurés}$  restants, en ne dépassant pas 85% de les  $PF_{assurés}$  initiaux qui correspond au seuil de PB minimum (nous appellerons ce prélèvement  $PrelevementFi_1$ ). Si cela ne suffit toujours pas, la seconde option est de reprendre de la PPE sans dépasser 85% de la PPE initiale ( $PrelevementPPE_2$ ). Dans un troisième temps, l'assureur va utiliser les 15% restant des  $PF_{assurés}$  pour essayer d'atteindre la revalorisation cible (que nous appellerons  $PrelevementFi_2$ ) et finalement il réalisera des plus-values latentes supplémentaires sur action.

Encore une fois, nous schématisons l'étape 2 afin d'en clarifier les mécanismes :

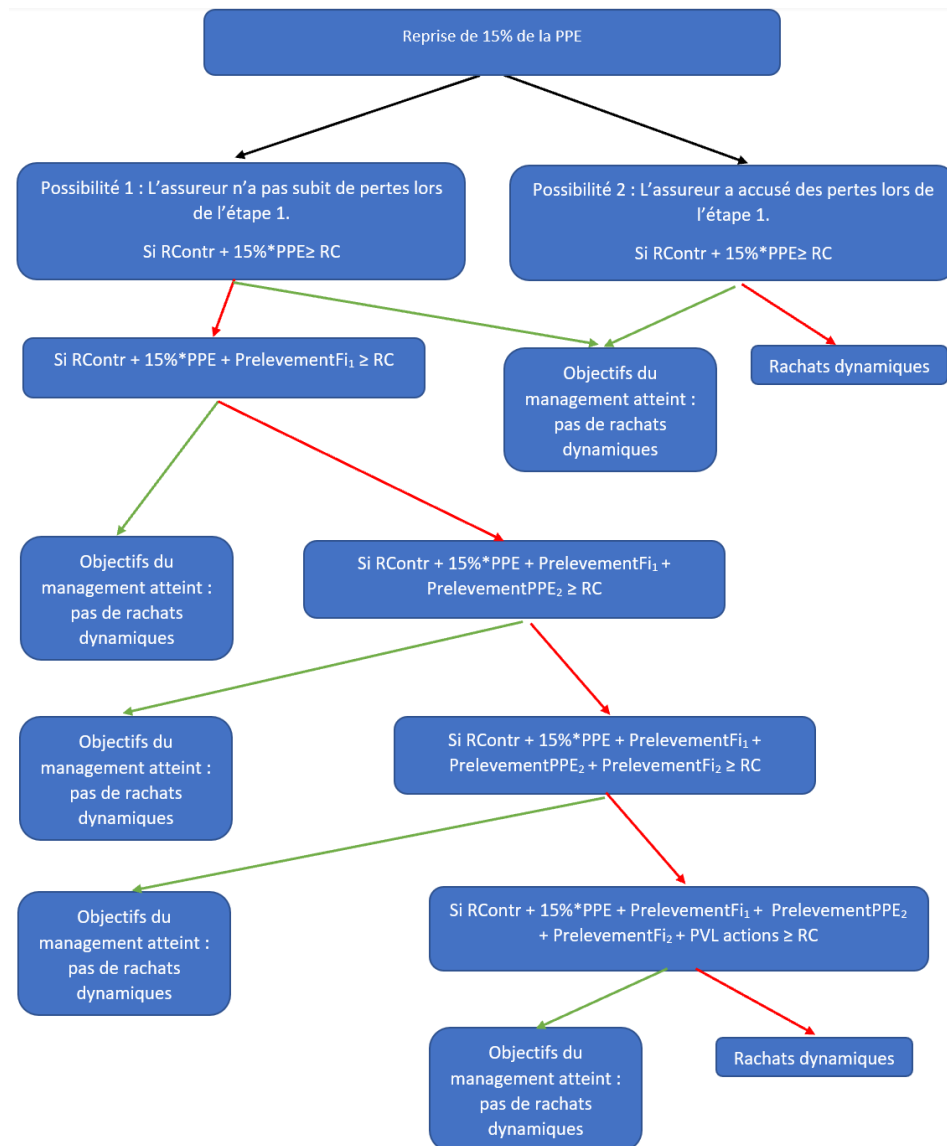


FIGURE 2.6: Etape 2 de l'algorithme, atteindre la revalorisation cible (RC)



Si ces actions du management ne permettent pas d'atteindre la revalorisation cible, Il y a un écart entre le taux de revalorisation servi  $TS$  et le taux concurrent  $TC$ , ce qui entrainera des rachats conjoncturels dont l'intensité dépendra de l'écart  $TS - TC$  (cf 2.4.3).

### La participation aux bénéfices minimum

Finalement, nous devons nous assurer que 85% des produits financiers nets de frais de gestion sont bien redistribués aux assurés, comme le prévoit le code des assurances dans les articles A-132-1 et A-132-2 sur la participation aux bénéfices. Pour ce faire, nous comparons le montant de produits financiers réellement attribué dans le cadre de la revalorisation et nous les comparons à la valeur règlementaire. Si les produits financiers attribués sont inférieurs aux 85% des produits financiers nets de frais de gestion, nous dotons alors la PPE.

## 2.6 Le périmètre de l'étude : modélisation d'une assurance vie fictive

Nous avons tâché de mener notre étude sur une société d'assurance représentative du marché actuel afin d'obtenir des ordres de grandeur de BE cohérents. Nous avons cependant fait des hypothèses simplificatrices afin de ne pas complexifier notre étude, la première étant de ne modéliser dans notre portefeuille que des contrats Euro.

Nous présentons ici la situation centrale de la compagnie d'assurance à la date d'arrêté du 31/12/2019 et nous ferons varier un certain nombre de paramètres pour générer la base de données. Cela sera détaillé dans le chapitre suivant.

### 2.6.1 Projection de scénarios économiques en accord avec la date d'arrêté

Nous avons récupéré les courbes des taux sans risques publiées à la date du 31/12/2019 fournie par l'EIOPA. A partir des courbes de taux sans risque, nous avons put projeter nos 1000 scénarios économiques sur 30 ans. Notre étude ORSA se basera sur cette projection de scénarios. Nous avons utilisé les paramètres définis en interne pour cet arrêté dans les modèles de Hull and White et Black and Scholes. (cf 2.2.1 et 2.2.2)

### 2.6.2 Bilan initial de notre compagnie

Le ratio de PPE/PM a été fixé à 3,5%, dans un ordre de grandeur proche de ce qui peut être trouvé chez les assureurs. Il en est de même pour la réserve de capitalisation, avec un ratio de RKPI/PM de l'ordre de 2%. Les valeurs nettes comptables de nos actifs sont directement déduites de notre allocation initiale, c'est à dire 70% d'obligations, 10% d'actions, 15% d'immobilier et finalement 5% de monétaire. Nous faisons ici l'hypothèse que notre allocation initiale et notre allocation cible sont égales.

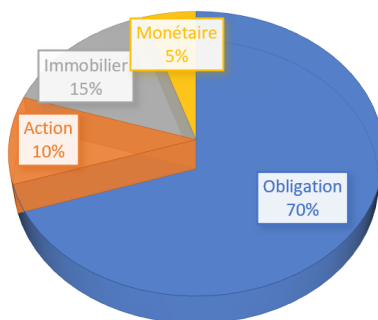


FIGURE 2.7: Allocation de l'actif

Finalement, voici notre bilan initial :

Actif	Passif
Obligations 389 480 000 €	FP 50 000 000 €
Actions 55 640 000 €	PM 480 000 000 €
Immobilier 83 460 000 €	PPE 16 800 000 €
Monétaire 27 820 000 €	RKPI 9 600 000 €

FIGURE 2.8: Bilan initial en norme comptable française implémenté dans l'outil ALM

Finalement, nous supposons initialement un stock de 10% de plus values.

### 2.6.3 Model Points de l'actif

Comme énoncé dans la partie 2.5, nous faisons comme hypothèses que notre société d'assurance n'achète que des obligations de maturité 10 ans et que les obligations présentes dans le portefeuille à la date initiale ont été achetées au pair, ce qui implique que la valeur comptable est égale à la valeur nominale. Nous modélisons des obligations en stock et futures (celles qui seront émises sur le marché et potentiellement achetées sur l'horizon de projection). Nous choisissons donc 10 obligations en stock dans le portefeuille initialement et 30 (pour les 30 années de projections) qui pourront potentiellement servir pour les réinvestissements futurs. La maturité de nos plus vieilles obligations en portefeuille, c'est à dire achetées en 2010, est 2020, celles achetées en 2011 ont pour maturité 2021 et ainsi de suite. Les taux de coupons des obligations en stock sont fixés comme étant la valeur de l'OAT 10 au moment de l'achat de l'obligation. Nous avons récupéré ses valeurs sur la banque de France. Nous fixons la Valeur Nette Comptable (VNC) de chaque obligation en portefeuille initialement comme étant la VNC des obligations divisée par le nombre d'obligation en stock. Par simplification, nous n'avons qu'un MP d'action, d'immobilier et de monétaire.

Ainsi, voici les MP d'obligations en stock à la date d'arrêt :

TABLE 2.2: MP des obligations en stock

	Année de terme	Rating	Devise	Taux de coupon	Nominal (= VNC)
<b>1</b>	2020	AAA	EUR	2,66%	38 948 000
<b>2</b>	2021	AAA	EUR	2,65%	38 948 000
<b>3</b>	2022	AAA	EUR	2,23%	38 948 000
<b>4</b>	2023	AAA	EUR	2,4%	38 948 000
<b>5</b>	2024	AAA	EUR	1,28%	38 948 000
<b>6</b>	2025	AAA	EUR	0,89%	38 948 000
<b>7</b>	2026	AAA	EUR	0,13%	38 948 000
<b>8</b>	2027	AAA	EUR	0,76%	38 948 000
<b>9</b>	2028	AAA	EUR	0,81%	38 948 000
<b>10</b>	2029	AAA	EUR	-0,27%	38 948 000

Concernant les obligations qui ne sont pas en stock nous ne renseignons initialement que leur année de terme, rating et devise. Le rating et la devise sont inchangés et restent respectivement AAA et EUR. Les années de termes sont les valeurs suivantes des années de termes des obligations déjà en stock, la dernière année de terme étant donc 2059. Les taux de coupons de ces obligations est défini comme le taux ZC de maturité 10 ans de l'année d'entrée sur le marché de l'obligation.

#### 2.6.4 *Model Point* du passif

Nous faisons une simplification en ne considérant qu'un unique *Model Point* en stock pour le passif. Ce choix a été effectué afin d'éviter de complexifier les modèles, faciliter l'étude de l'impact des paramètres intrinsèques au *Model Point* (comme l'âge, le TMG...). Cela permet aussi de limiter l'ajout de variables explicatives, la taille de la base de données et le temps de calcul qui, comme nous le verrons dans la partie suivante fait partie des défis du mémoire. Nous avons donc décidé d'effectuer cette hypothèse simplificatrice, que nous retrouvons dans plusieurs mémoires portant sur des études de même nature.

Nous avons initialement 45 MP implémentés dans l'outil ALM, dont 25 en portefeuille. Nous avons défini le TMG de chaque MP pour qu'il soit cohérent avec sa date de souscription. A cette fin, nous avons choisi d'attribuer aux MP le taux technique Vie maximum de son année de souscription fournie par SPAC actuaires. A partir de 2015, aux vues de l'environnement de taux bas et par conséquent la baisse de rendement des assureurs, nous avons implémenté un TMG à 0%. Nous faisons aussi l'hypothèse que le TMG est viager dans cette étude, bien que généralement les contrats prévoient un TMG sur 8 ans.

Nous faisons le choix de transformer ces 25 MP initiaux en un seul MP moyen, c'est à dire de caractéristiques contractuelles la moyenne des caractéristiques des MP. De plus, nous attribuons à ce MP moyen une PM correspondant à la PM de notre bilan, un nombre d'assurés correspondant au nombre d'assurés dans le portefeuille de l'assureur. Nous modélisons aussi des primes pour notre MP moyen qui s'élèveront à 4 000 000 €/par an, avec un taux de chute des primes à 2%.

En voici finalement les caractéristiques :

TABLE 2.3: MP du passif

Stock/NB	Ancienneté	TMG	Chargement (encours)	Nombre de polices	PM	Age moyen	Primes (EUR)
Stock	12	1,85%	0,6%	24000	480 000 000	52	4000000

### 2.6.5 Les frais et taxes appliqués

Nous supposons des frais fixes pour toute la durée de projection et dont les valeurs sont :

- 0.15% pour le taux de frais de placement
- 1% pour le taux frais d'acquisition
- 1% pour les frais de gestion
- 0.2% les frais d'administration

Le taux de prélèvement sociaux est lui fixé à 15,5%.

#### Conclusion du chapitre

Nous avons dans ce chapitre mis en avant les mécanismes de l'outil ALM, qui font intervenir à chaque étape un grand nombre de paramètres. Une attention particulière a été portée aux variables qui influencent de manière directe le calcul du BE.

Dans un second temps, nous avons présenté l'assureur vie sur lequel se base notre étude. Des hypothèses simplificatrices ont été effectuées pour mener à bien cette étude et surpasser les contraintes opérationnelles.

Ce nouveau chapitre s'attache à exposer clairement les variables sélectionnées et à présenter l'étude *machine learning* menée et ses résultats.

## Chapitre 3

# Création de la base de données et approximation du *Best Estimate* par *machine learning*

### 3.1 La stratégie mise en place

Notre stratégie d'apprentissage par méthode de *machine learning* se fractionne en plusieurs étapes allant de la prise en main du modèle ALM à l'apprentissage d'un modèle de *machine learning* en substitution de ce modèle ALM. Nous retrouvons dans ce schéma les différentes étapes, que nous détaillerons dans les parties suivantes.

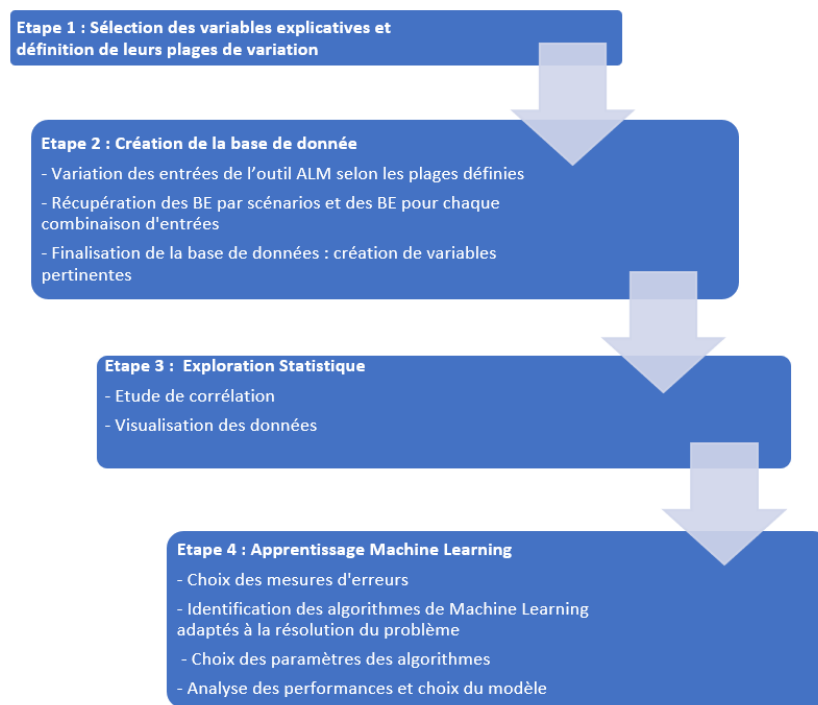


FIGURE 3.1: Stratégie de mise en place de l'étude *machine learning*

Nos études préliminaires sur l'outil ALM nous ont permis de sélectionner les variables qui, selon nous, auraient un bon pouvoir prédictif sur le BE. Nous avons donc défini des plages de variations pour celles-ci en gardant en tête la problématique du temps de calcul et de la taille de la base de données.

S'en est suivi la création de la base de données : faire varier les inputs sur les plages de variations définies en faisant tourner le modèle ALM à chaque pas de variation afin de récupérer les BE par scénario et les BE moyens.

Une fois la base de données simulée, nous avons alors effectué une exploration statistique sur nos données, avec notamment une analyse de corrélation sur nos variables décrivant les scénarios économiques. Nous ajoutons aussi des variables que nous pensons pertinentes à cette étude.

Nous nous servons finalement de cette base de données pour effectuer l'apprentissage *machine learning*.

## 3.2 Les variables explicatives et leurs plages de variation

L'étude de l'outil ALM a mis en exergue les variables qui ont un impact direct et significatif sur notre BE. Nous allons maintenant tâcher de lister les variables retenues et expliquer succinctement les raisons de leur sélection. Rappelons que ce mémoire a pour difficultés de simplifier une information économique projetée temporellement sur 30 ans, mais aussi une information liée directement à la situation et la stratégie de l'entreprise en  $t = 0$ , afin de pouvoir déduire par un modèle de *machine learning* le *Best Estimate*. Le choix des variables est donc une étape clé dont dépend les bons résultats de notre étude.

### 3.2.1 Gérer la taille de la base de données et le temps de calcul

Un des défis de ce mémoire est de jongler avec deux contraintes : garder une taille de la base de données suffisamment restreinte pour pouvoir effectuer une étude *machine learning* exhaustive et respecter les échéances imposées, mais aussi suffisamment complète pour expliquer le BE.

Prenons un exemple simplifié pour comprendre cet enjeu. Supposons que nous ayons sélectionné 10 variables bilantielles explicatives avec une plage de variation de 5 modalités pour chacune de ces variables. Nous nous plaçons dans une approche d'apprentissage sur les BE par scénario et supposons que nous modélisons 5000 scénarios différents. Notre taille de base de données serait donc dans ce cas :

$$n_{bdd} = 10^5 \times 5000 = 5 \times 10^8 BE$$

soit 500 millions de BE. De plus le calcul d'un BE moyen prend 10 minutes, ce qui en d'autres termes signifie que le calcul de 1000 BE par scénario prend 10 minutes. Partant de ce constat, on en déduit le temps de calcul nécessaire :

$$t_{calcul} = \frac{5 \times 10^8}{100} = 5\,000\,000 \text{ minutes} = 83\,333 \text{ heures} = 3\,472 \text{ jours}$$

Nous aurions donc une base de données de 500 millions de BE, pour 10 variables bilantielles, 30 taux projetés, 30 rendements action, 30 rendements immobiliers, 30 taux concurrents soit 130 variables explicatives. Dans cet exemple, la taille de la base de données est trop volumineuse et le temps de calcul trop long et inconcevable dans le temps imparti de notre mémoire.

Notre première démarche a été de réfléchir à un moyen d'accélérer la fréquence de calcul des BE. A cette fin, nous avons modifié l'outil ALM sous R afin de paralléliser les calculs et par conséquent

agir sur le temps de calcul. L'amélioration due à la parallélisation est d'autant plus significative que le nombre de processeurs logiques disponibles sur l'ordinateur est grand. Ainsi, en moyenne nous avons réduit de 60% le temps de calcul sur nos différents appareils, ce qui est conséquent. Cela n'étant pas suffisant, nous avons dû faire des hypothèses simplificatrices et choisir judicieusement nos variables explicatives et leurs plages de variation, c'est ce que nous allons expliciter par la suite.

### 3.2.2 L'information retenue du *Model Point* de passif

Lors de la prise en main de l'outil ALM, la base de MP telle qu'implémentée dans l'outil contenait 45 MP et en retenant seulement 3 variables par MP nous aurions déjà 135 variables. Nous réduisons ainsi le nombre de variable en nous limitant à un seul MP, pour lequel nous avons sélectionné les variables suivantes :

**L'âge moyen du model point** qui aura un impact direct sur le BE au travers des prestations décès. Nous l'appellerons dans nos modèles "Age".

**Le TMG** qui définit les engagements envers les assurés et mécaniquement impact le BE. Généralement, l'assureur ne s'engage à servir le TMG qu'un certain nombre d'années défini contractuellement. Afin de mieux étudier l'impact des variations du TMG, nous considérerons ici que celui-ci est valable pour toute la durée du contrat, soit jusqu'au rachat ou au décès. C'est ce que nous appelons un TMG viager. Nous nous référerons à cette variable par le nom "TMG".

**L'ancienneté** dont les prestations de rachats dépendent au travers de notre modélisation des rachats structurels. Celle-ci sera appelée "Anciennete".

**Les chargements sur encours** qui viennent impacter les revenus de l'assureur sur l'encours et la partie monétaire de l'actif. Cette variable porte le nom de "Chargement".

Le choix d'un unique MP n'est pas sans conséquences et implique de négliger les interactions et les effets de mutualisation des MP qui ont une incidence sur le BE. Prenons comme exemple la situation d'un portefeuille composé de deux MP : un MP dont l'entrée dans le portefeuille est récente et de TMG nul et un MP avec une plus grande ancienneté et au TMG avantageux. Nous supposons une légère remontée des taux permettant aux assureurs d'avoir un rendement obligataire plus favorable. Dans cette situation, on aura une tendance à la hausse des rachats pour le MP récent, mais qui sera diluée par une stagnation des rachats du MP plus ancien, fort de son TMG plus favorable.

Pour choisir nos plages de variations nous avons gardé en tête la contrainte liée au temps de calcul, à laquelle s'ajoute le fait que nos variables ne varient pas indépendamment les unes des autres. Nous voulons donc que les variations effectuées dans notre MP soient cohérentes : il est peu probable de trouver chez un assureur un MP d'âge moyen 20 ans et d'ancienneté 15 ans par exemple. Ces trois variables sont ainsi corrélées par construction, l'âge moyen de souscription à une assurance vie étant d'une quarantaine d'année, une augmentation de l'âge moyen du MP implique une ancienneté plus grande. Dans la même logique, un TMG élevé correspond à des souscriptions plus anciennes quand la situation économique permettait des rendements plus élevés pour les assureurs et donc un âge moyen du MP plus élevé. En revanche, nous ne cherchons pas à avoir des variables corrélées à 100% car dans ce cas nous n'aurions qu'à en garder une sur les 4 et cela n'est pas recherché. Ainsi, nous avons tâché de respecter cela au moment de faire varier les paramètres de notre MP. Notre démarche a consisté à repartir des 45 MP implémentés dans l'outil et d'en extraire 8 de manière à avoir le plus de variations dans nos paramètres. Nous étions alors assurés d'avoir des combinaisons de paramètres plausibles, mais en plus de cela, nous en avons ajouté 3 qui n'étaient pas implémentées initialement. Cela permet de

diversifier notre base et d'avoir un modèle capable de mieux s'adapter à des changements dans les portefeuilles d'assurés. Par exemple, les MP 10 et 11 nous permettent d'envisager un MP plus jeune à la souscription et inversement le MP 9 un MP plus âgé à la souscription.

Mais la principale contrainte nous ayant poussé à raisonner ainsi est le temps de calcul et la taille de base de données. Notre démarche, nous permet d'obtenir 11 combinaisons différentes, là où une démarche de variation sur des plages de taille 5 des paramètres indépendamment les uns des autres nous en aurait donné 625 par exemple.

Voici le détail de nos 11 combinaisons envisagées, les 8 premières étant des combinaisons extraites du portefeuille initial et les 3 dernières des ajouts :

TABLE 3.1: Les 11 combinaisons de paramètres des MP

Combinaison	Age	Anciennete	TMG	Chargement
1	62	23	3,5%	0,3%
2	60	21	2,5%	0,3%
3	58	19	3%	0,6%
4	50	11	2,25%	0,6%
5	48	9	1,75%	0,6%
6	47	8	1,25%	0,6%
7	45	6	1,25%	0,9%
8	43	4	0%	0,9%
9	74	15	2%	0,6%
10	25	4	0%	0,9%
11	47	23	3.5%	0,3%

Nous faisons ici des simplifications fortes qui trouvent une explication autre que le gain de temps. Dans une démarche ORSA qui est propre à la compagnie, nous ne voulons pas un modèle *machine learning* universel, mais un modèle adapté à la conjoncture économique et permettant à notre compagnie d'étudier adéquatement ses risques. Libre à elle d'adapter sa démarche. Nous proposons ici une méthode où elle pourra régénérer une base d'apprentissage en cas de perte de performance du modèle dû à des changements de l'environnement dans lequel elle évolue par exemple, mais aussi en cas de modèle plus performant mis au point. Nous recherchons donc de la flexibilité et de l'adaptabilité.

Au passif, certaines variables du bilan interviennent dans le calcul du BE et c'est ce que nous exposons dans la partie suivante.

### 3.2.3 L'information retenue au passif du bilan

D'un point de vue bilantiel, nous ferons varier le **ratio PPE/PM**. Comme nous l'avons vu précédemment, la PPE est une ressource utilisée dans la phase 2 de stratégie de participation aux bénéfices (2.5.2) pour servir la revalorisation cible et limiter les rachats dynamiques. Ainsi, la PPE représente une "richesse" initiale dont la variation influence directement le BE. Nous ferons aussi varier le **pourcentage de plus ou moins-values initial** (PMVL), l'assureur étant amené à faire des plus-values sur les actions pour servir les revalorisations contractuelles et cibles. Pour des raisons de temps de calcul et de taille de base de données, nous ne faisons pas varier la réserve de capitalisation qui n'impacte qu'indirectement le BE au travers de la variable  $PF_{assureur}$ . Nous utilisons comme nom de variable "Taux\_PPE" pour le ratio PPE/PM et "Plus\_Moins\_VL" pour le pourcentage de plus ou



moins-values. Concernant les plages de variations, nous faisons fluctuer nos variables autour des valeurs centrales telles que définies dans la partie précédente. Nous nous limitons à 5 variations par plage et en voici le détail :

TABLE 3.2: Plage de variation des variables bilantielles

Variable	Min		Central		Max
Plus_Moins_VL	5%	7.5%	10%	15%	20%
Taux_PPE	3%	1.5%	3%	4.5%	6%

### 3.2.4 L'information de l'actif retenue

Nous allons faire varier **l'allocation cible**. Comme nous l'avons mentionné précédemment, celle-ci est fixée tout au long de la projection et est impliquée directement dans la phase 1 de l'action du management. Son impact est multiple et touche les produits financiers, le rendement de l'assureur, donc la revalorisation des provisions lors de la seconde phase de la stratégie de participation aux bénéfices, et par conséquent le BE. Nous avons décidé de ne faire varier que les parts actions et obligations qui sont les plus impliquées dans le mécanisme ALM, ce qui nous amène à ne garder que deux variables "Oblig" et "Action" dans nos modèles.

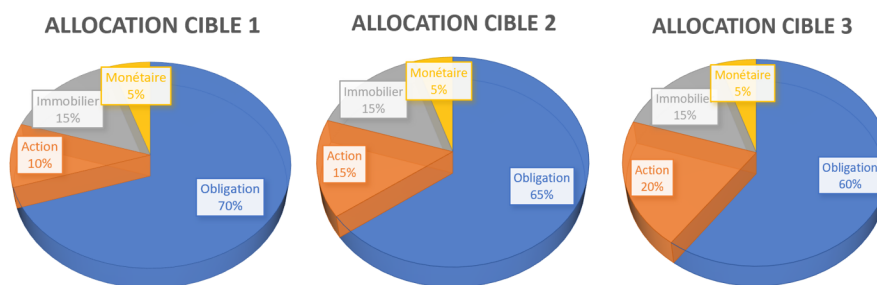


FIGURE 3.2: Les 3 variations de l'allocation cible

### 3.2.5 Les frais et les taxes

Au vu des contraintes opérationnelles, nous n'avons fait varier que les **taux de frais de gestion** sur la plage suivante :

TABLE 3.3: Plage de variation du taux de frais de gestion

Variable	Min	Central	Max
Frais_Gestion	0%	1%	2%

Nous faisons aussi le choix de garder le taux de prélèvement sociaux fixés à 15,5%.

### 3.2.6 Résumer l'information économique

L'environnement économique est déterminant pour une compagnie d'assurance vie par son influence sur son passif au travers de la politique de taux servi et de participation aux bénéfices. De plus, dans le cadre d'un calcul stochastique, les caractéristiques du GSE ont un impact direct sur la valeur temps des options et garanties du passif. L'information économique retenue diffère en fonction de l'approche que nous avons effectuée.

#### Approche d'apprentissage sur les BE par scénario

Premièrement, dans le cadre de l'étude des BE par scénario, nous projetons nos taux zéros-coupons, nos rendements actions et immobiliers sur 30 ans ce qui représente 90 variables : un rendement immobiliers, un rendement actions et un taux ZC par années et cela sur 30 ans. En plus de cela, en nous basant sur les études menées par Julien Mattei dans son mémoire (MATTEI, 2019), nous utilisons aussi le taux concurrent projeté encore une fois sur les 30 années. Cela nous amène à 120 variables pour spécifier l'évolution temporelle de l'environnement économique dans laquelle évolue notre compagnie d'assurance fictive. Nous appellerons ces variables de la manières suivantes dans nos modèles :

**Les taux zéros-coupons :** "Taux\_ZC\_Y 1", "Taux\_ZC\_Y 2", ... , "Taux\_ZC\_Y 30" .

**Les rendements immobiliers :** "Immo\_Y 1", "Immo\_Y 2", ... , "Immo\_Y 30".

**Les rendements actions :** "Action\_Y 1", "Action\_Y 2", ... , "Action\_Y 30".

**Les taux concurrents :** "Taux\_Concu\_Y 1", "Taux\_Concu\_Y 2", ... , "Taux\_Concu\_Y 30".

Nous complétons l'information économique avec des variables caractéristiques aux scénarios, c'est-à-dire les écarts-type, moyennes et optimums par décennie. Nous ajoutons ainsi :

- Les moyennes et écarts-type des 10 premières années, des années 10 à 20, des 10 dernières années et des 30 années et cela pour les taux ZC, les rendements actions et immobiliers.
- Le maximum et le minimum des taux ZC, des rendements actions et immobiliers sur les 30 années.

C'est un ajout de 30 variables aux 120 variables de projection. Le but de cette complétion est de venir substituer dans des cas précis dépendants de l'analyse des corrélations qui suit, nos projections sur 30 ans et ainsi réduire notre nombre de variables. Notre base de données pour l'étude des BE par scénario est finalement composée de 150 variables économiques, dimensions que nous avons tâché de réduire grâce notamment aux études détaillées ci-dessous.

#### Etudes de corrélation pour sélectionner les variables explicatives

L'objectif de cette étude de corrélation est de permettre une réduction du nombre de variables explicatives liées à la projection économique sur 30 années. Nos études de corrélations commencent par l'utilisation de matrices de *scatterplots* (ou nuage de points) permettant de juger visuellement d'une potentielle corrélation entre nos variables. Cette étude préliminaire des scatterplot permet de visualiser les interactions entre les variables, qui ne sont pas nécessairement linéaires et choisir les coefficients de corrélation adéquates. Nous prenons ici l'exemple des rendements immobiliers de la 2<sup>eme</sup> à la 10<sup>eme</sup>

année et des taux ZC sur la même période. Cet exemple a pour but d'expliquer notre démarche qui a été effectuée sur l'ensemble des variables économiques.

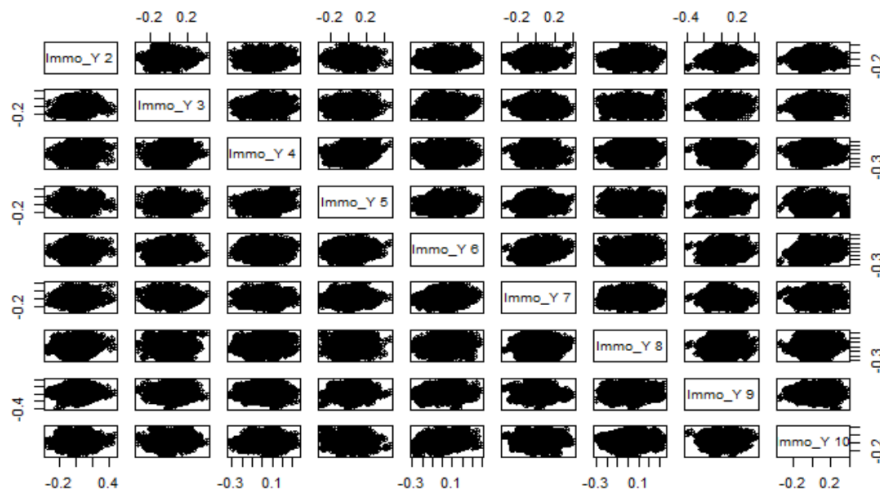


FIGURE 3.3: Scatterplot des rendements immobiliers

Nous voyons ici qu'aucune dynamique ne semble se dégager entre les rendements immobiliers d'une année à l'autre. Cela laisse donc présager des corrélations proches de 0.

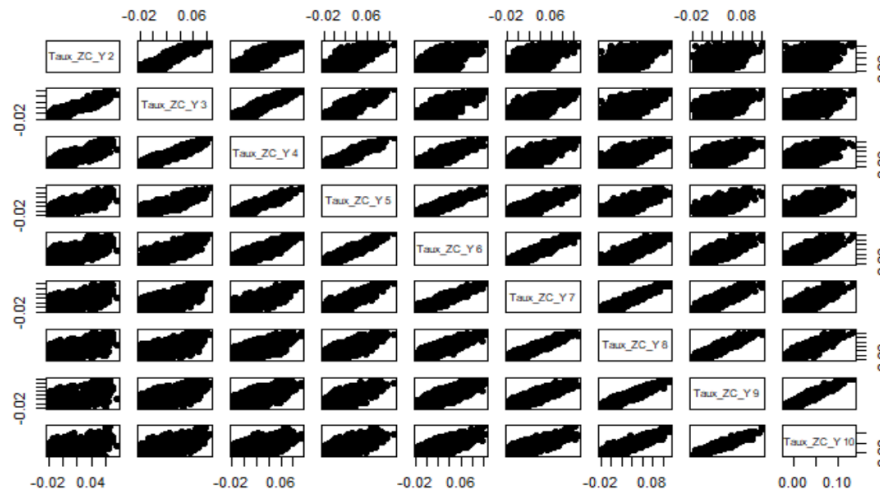


FIGURE 3.4: Scatterplot des taux ZC

Inversement aux rendements immobiliers, nous décelons une relation quasiment affine entre les Taux ZC d'une année à l'autre, voire d'une année à des années antérieures. De plus, cette tendance s'accroît plus l'année de projection augmente. Cette étude laisse ainsi présager une corrélation forte entre les taux ZC. Une fois ce travail de visualisation effectué, nous étudions les corrélations de ces mêmes variables afin de confirmer les observations sur les *scatterplots*.

Nos observations ne semblant pas montrer d'autres liens entre les variables qu'affine, nous faisons

donc le choix d'une mesure des corrélations par le coefficient de Pearson dont nous rappelons la formule :

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

avec

- $\text{cov}$  la covariance
- $\sigma_X$  et  $\sigma_Y$  les écarts types de X et Y
- $\mu_X$  et  $\mu_Y$  les moyennes de X et Y

Nous obtenons ainsi une matrice des corrélations des rendements immobiliers comme suit :

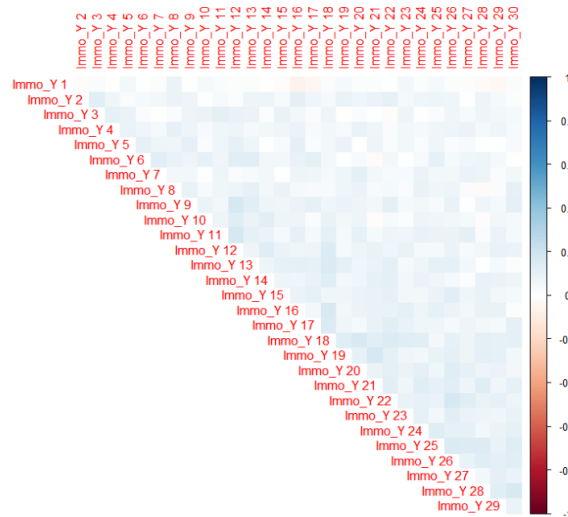


FIGURE 3.5: Matrice des corrélations des rendements immobiliers

En accord avec nos analyses précédentes, nos corrélations entre les rendements immobiliers sont faibles voire nulles. Concernant les taux ZC en voici les résultats :

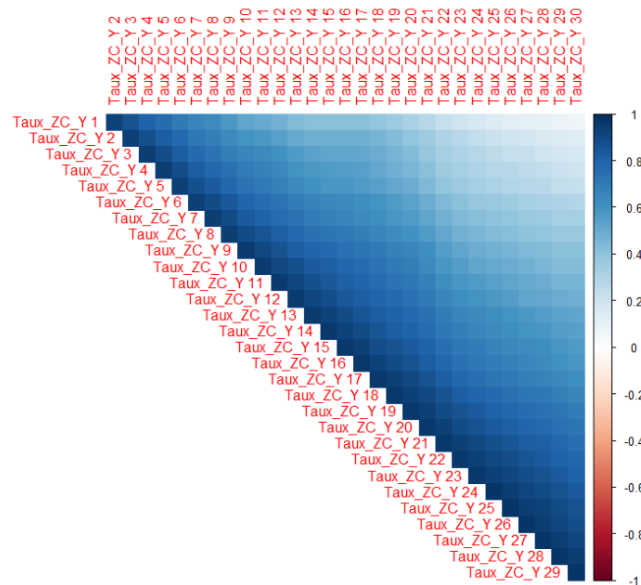


FIGURE 3.6: Matrice des corrélations des taux ZC

Ici nous remarquons que les taux ZC sont très corrélés, et ce sur plusieurs années comme le montre notamment la corrélation de 0,81 entre les taux ZC "Taux\_ZC\_Y 10" et "Taux\_ZC\_Y 7". Ce type d'étude a été mené au **sein de chaque variable** entre les années de projection, et en voici les conclusions :

- Comme nous l'avons vu précédemment, les rendements immobiliers ne sont pas corrélés d'une année à l'autre.
- Les rendements actions ne sont pas corrélés non plus d'une année à l'autre.
- Les taux ZC sont très corrélés et ce sur plusieurs années. De plus cette corrélation s'intensifie avec les années. Prenons par exemple la variable "Taux\_ZC\_Y 25", celle-ci est corrélée à plus de 80% avec les variables allant de "Taux\_ZC\_Y 20" et "Taux\_ZC\_Y 30". Nous pourrions par exemple envisager de garder seulement "Taux\_ZC\_Y 25" pour résumer l'information des 10 dernières années.
- Nous faisons les mêmes remarques que pour les taux ZC concernant les taux concurrents, eux aussi très corrélés.

Concernant les corrélations **entre les variables** :

- Les rendements actions et immobiliers sont corrélés à hauteur de 75% pour la même année de projection, et non corrélés sinon.
- Les corrélations entre les taux ZC et les rendements actions sont moindres, de même que pour les rendements immobiliers. Cependant, les rendements immobiliers restent plus corrélés avec les taux ZC que les rendements actions.
- Concernant finalement les taux concurrents, leur formulation dépend directement du taux ZC ce qui explique une corrélation forte entre ces variables. Bien que les rendements actions interviennent aussi dans sa formulation, les corrélations sont moins grandes et le constat est le même pour les rendements immobiliers.

Finalement, nous étudions la corrélation des variables de complétion :

- Seules les moyennes sont fortement corrélées pour les rendements immobiliers et actions.
- Les moyennes des taux ZC et des rendements immobiliers sont corrélées. L'écart-type des ZC est corrélé à la moyenne et la valeur maximale des rendements immobiliers.
- Les variables ajoutées des rendements actions et des taux ZC ne montrent pas de corrélations remarquables.

L'ensemble des matrices de corrélations sont détaillées en Annexe afin qu'une personne intéressée puisse s'y référer.

### Méthode de l'analyse des composantes principales (ACP)

L'analyse en composantes principales vise à réduire les dimensions d'un problème, ce qui répond bien à notre problématique. Pour ce faire, cette méthode applique une transformation orthogonale sur les variables qui sont supposées colinéaires, ce qui permet de créer de nouvelles variables linéairement indépendantes, résumant l'information de plusieurs variables de bases.

Nous avons appliqué cette méthode à deux situations. Premièrement, nous l'avons appliquée aux 150 variables relatives aux scénarios économiques. Dans un second temps, nous avons remplacé l'information économique par les variables de complétion, passant ainsi d'un problème à 150 dimensions à 30, tout en cherchant à réduire encore cette dimension grâce à l'ACP.

Concernant la première situation, nous obtenons les résultats suivants :

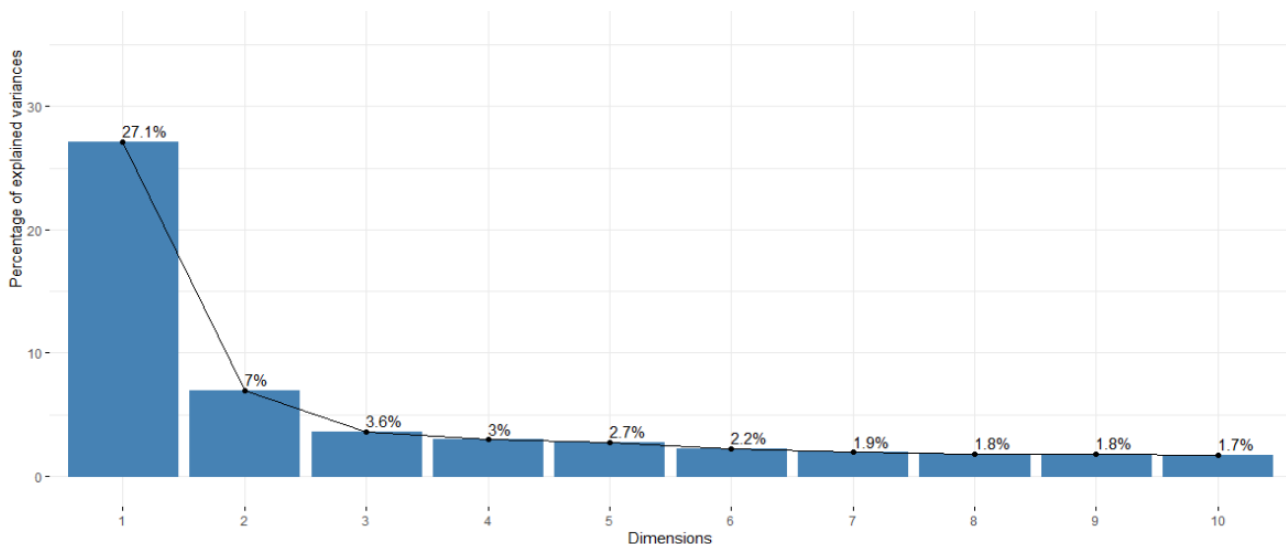


FIGURE 3.7: Evolution de la variance expliquée en fonction de la dimension, base complète

Comme nous pouvons l'observer sur la figure 3.7, le premier axe principal n'explique que 27,1% de la variance et le second 7% seulement. Ainsi, le cumul des 10 premiers axes n'explique que 52,8% de la variance. Ces résultats n'étant pas satisfaisants, nous nous sommes tournés vers la seconde approche en ne considérant que les variables de complétion.

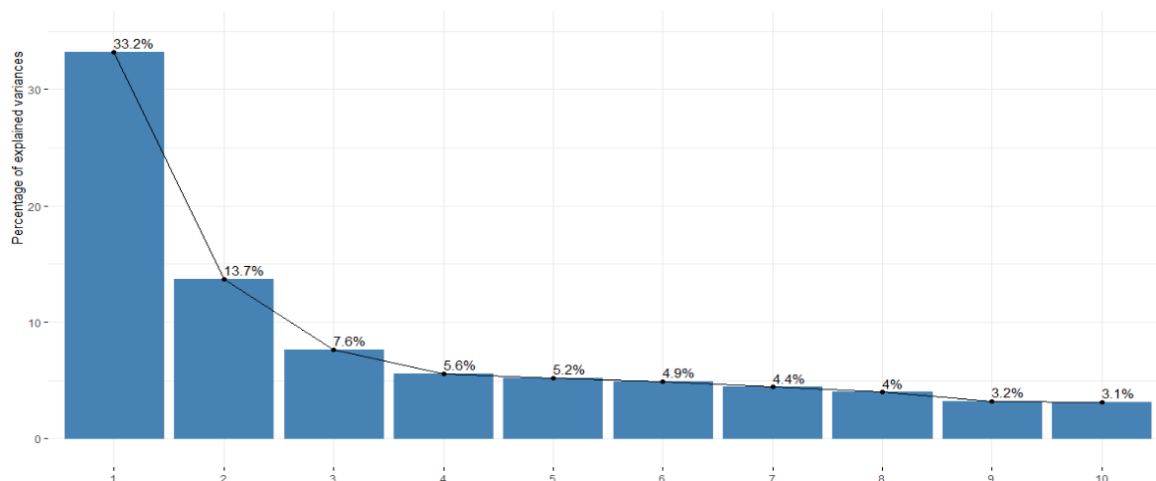


FIGURE 3.8: Evolution de la variance expliquée en fonction de la dimension, base de complétion.

Sur la figure 3.8, nous observons une amélioration de la variance expliquée sur les premiers axes principaux comparé à la figure 3.7. C'est maintenant 65,3% de variance qui est expliquée sur le cumul des 5 premiers axes principaux. Cependant, nous n'observons pas de décroissance exponentielle voire simplement forte de la variance expliquée, avec une part de variance expliqué encore significative sur le 10ème axe principale.

Nous représentons ci-dessous la matrice des cosinus carrés de l'angle entre la représentation de la variable et sa projection sur les 5 premiers axes. Ce graphique nous permet de juger de la qualité de la représentation de chacune des variables sur les 5 premiers axes principaux.

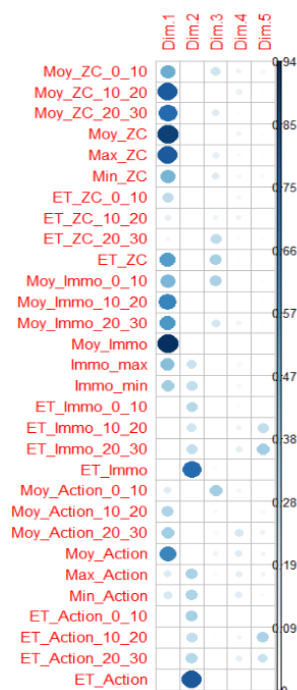


FIGURE 3.9: Matrice des cosinus carrés de l'angle de chaque variable et sa projection

Nous remarquons sur la figure 3.10 que les variables représentant les moyennes des différents scénarios sont les mieux représentées par le premier axe et que les variables représentant les écarts-type ont une meilleure qualité de représentation sur le second axe. Nous représentons finalement nos variables selon les 2 premiers axes principaux. Nous obtenons le résultat ci-dessous.

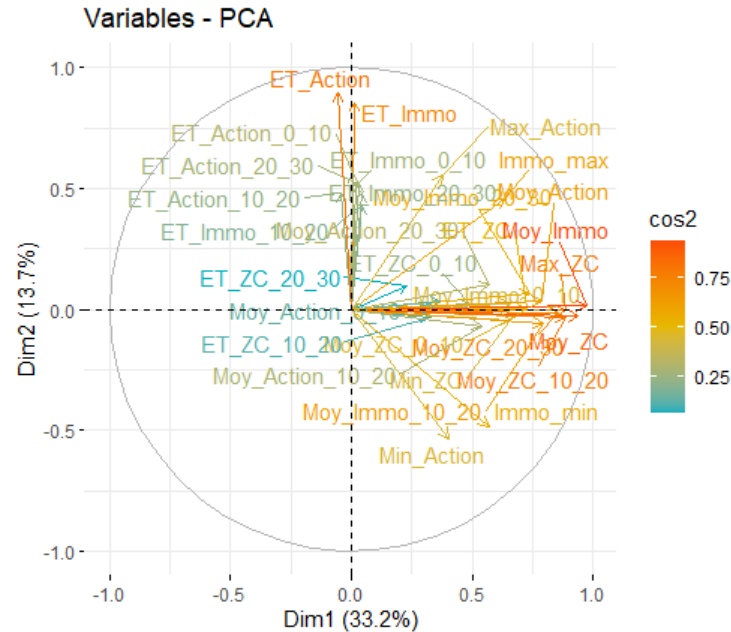


FIGURE 3.10: Représentation des variables sur les 2 premiers axes principaux

Bien que le nombre important de variables rende l'analyse de la figure 3.10 ardue, nous notons, en lien avec les remarques précédentes, que plus nous nous déplaçons vers la droite du premier axe principale, plus les variables moyennes de nos scénarios auront tendance à augmenter. De même pour le second axe, plus nous nous déplaçons vers le haut de celui-ci, plus les écarts-types des rendements actions et immobiliers augmentent, ce qui est symptomatique de scénarios avec des rendements plus volatiles. Une étude plus poussée des scénarios économiques a été effectuée par MATTEI, 2019, mais nos travaux sur l'ACP ont des conclusions proches malgré que notre nombre de scénarios économiques soit plus élevé.

### Conclusion de nos études

Nos résultats d'ACP sont peu satisfaisants, la réduction de la base de données sur les premiers axes principaux ne pouvant se faire qu'au prix d'une perte de variance expliquée non négligeable. De plus, sélectionner un grand nombre d'axes principaux va complexifier l'analyse des effets des variables sur chacun de ces axes, risquant ainsi d'alourdir nos futurs travaux.

Nous avons donc décidé de nous baser sur l'étude des corrélations pour réduire le nombre de variables explicatives. Les rendements immobiliers et actions sont très corrélés, nous faisons le choix de ne garder que les rendements actions. Ce choix est motivé par le fait que les rendements actions sont moins corrélés avec les taux ZC, nous perdons ainsi moins d'information en gardant les rendements actions. De plus, afin de pallier cette perte d'information, nous utilisons les variables de complétions des rendements immobiliers. Nous gardons les rendements actions sur l'intégralité de la projection,



donc les variables de complétions des rendements actions ne sont pas nécessaires.

Les moyennes des taux ZC sont corrélées aux moyennes des rendements immobiliers, nous ne les ajoutons pas à notre base. En revanche, nous gardons les écarts-type des taux ZC sur les 3 premières décennies mais pas l'écart-type sur les 30 années car corrélé aux variables de complétions des rendements immobiliers.

Nous simplifions l'information des taux ZC et des taux concurrents grâce à leurs corrélations à la fois inter-annuelles et inter-variables. Premièrement, les taux ZC et les taux concurrents sont très corrélés, nous choisissons de garder, à partir de la 10<sup>ème</sup> année de projection, seulement les taux ZC. De plus, les taux concurrents sont fortement corrélés d'une année à l'autre, ce qui nous a mené à faire le choix de ne garder que les taux concurrents des années 1, 2, 3, 5, 7, 9. La même observation sur les corrélations entre les taux ZC des différentes années a été effectuée et nous a amené à choisir les taux des années 1, 2, 4, 6, 8, 10, 12, 16, 20, 24, 30.

Des méthodes alternatives tel qu'un autoencoder de variables par un réseau de neurones ont été envisagées. Cependant, bien que ces méthodes aient montré des pouvoirs de compression de base de données très encourageants, il se pose le problème de lisibilité et d'interprétabilité des variables compressées.

### Approche d'apprentissage sur les BE

Concernant la seconde approche pour laquelle nous étudions le BE moyen, nous ne gardons que la **courbe de taux EIOPA** en entrée du générateur de scénarios. Ces variables sont appelées "Taux\_Y 1", "Taux\_Y 2", ... , "Taux\_Y 30". Nous avons sélectionné 5 courbes de taux EIOPA pour notre étude, les courbes de taux au 31/12/2019 et 31/12/2017 centrales, la courbe de taux au 31/12/2015 choquée à la hausse et les courbes de taux au 31/12/ 2019 déformées d'une manière que nous allons expliquer. Dans un premier temps, rappelons que le taux choqué à la hausse au temps  $t$  selon Solvabilité 2 (que nous appellerons  $up_t$ ) se définit en fonction du taux central ( $centr_t$ ) de la manière suivante :

$$up_t = centr_t + \max(0.01, ChocUp_t \times |centr_t|)$$

et le taux choqué à la baisse ( $down_t$ ) :

$$down_t = \begin{cases} centr_t & \text{si } centr_t < 0 \\ centr_t - ChocDown_t \times |centr_t| & \text{sinon} \end{cases}$$

avec les  $ChocDown_t$  et  $ChocUp_t$  tels que spécifiés dans Solvabilité 2 et consultables en annexe. Nous utilisons donc le choc S2 à la hausse décrit ci-dessus sur la courbe de taux au 31/12/2015, ce qui nous permet d'avoir un environnement avec des taux positifs pris en compte dans notre modèle. En plus de cela, nous effectuons deux déformations différentes sur la courbe de taux dans le sens où nous appliquons les chocs S2 directement sur la courbe de taux centrale et non par le biais des formules détaillées ci-dessus, ce qui nous donne les déformations suivantes :

$$Def1_t = centr_t * (1 + ChocUp_t)$$

et

$$Def2_t = centr_t * (1 - ChocDown_t)$$

Nous effectuons ainsi deux déformations différentes de celles S2 sur nos taux afin d'enrichir notre modèle. Nous obtenons ainsi les courbes suivantes :

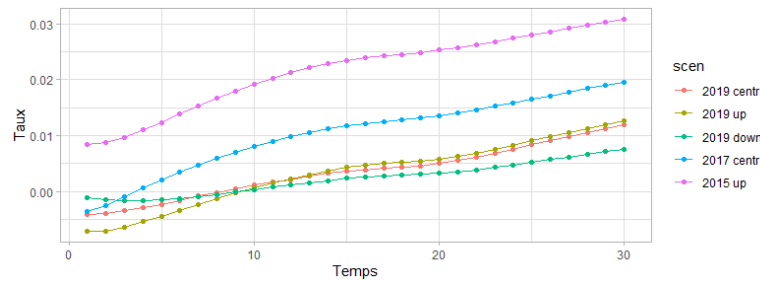


FIGURE 3.11: Courbes de taux EIOPA en entrée du modèle

A partir de ces courbes de taux EIOPA nous obtenons au total 5000 scénarios économiques. Nous devons aussi prendre en compte pour la suite que ces variables sont des séries temporelles, qui doivent donc être manipulées comme telles. Afin de simplifier l'information économique, nous avons utilisé une approche par étude de corrélation.

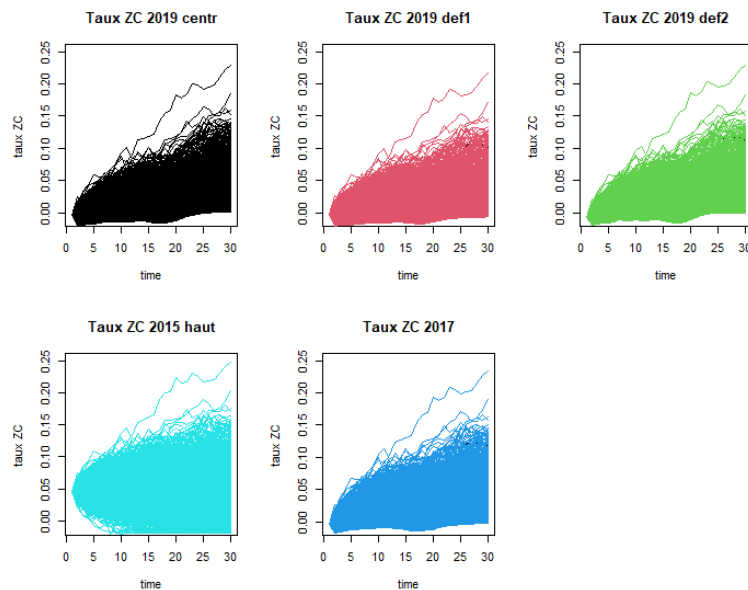


FIGURE 3.12: Courbes de taux ZC en sortie du GSE

Nous remarquons alors que nos courbes de taux sont corrélées entre-elles, malgré les déformations des courbes de taux EIOPA et cela sera quantifié lors de l'analyse des corrélations. Ce phénomène s'observe directement au travers du scénario de croissance des taux plus rapide qui se distingue des autres et que l'on retrouve sur les 5 graphiques. Cela est dû à l'utilisation d'un GSE ayant des paramètres inchangés. C'est une limite de l'approche : une amélioration de la base de données serait d'apporter des variations du GSE qui produirait ainsi des scénarios différents de par leur construction.

### 3.2.7 Résumé des variables de nos bases de données

Les variables relatives à la situation de notre assureur sont les mêmes que l'on considère l'approche sur les BE par scénario ou sur les BE :

- Les variables relatives au MP :
  - "Age" ;
  - "Anciennete" ;
  - "TMG" ;
  - "Chargement" .
- Les variables bilantielles :
  - "Plus\_Moins\_VL" ;
  - "Taux\_PPE" ;
  - "Oblig" ;
  - "Action"
- Les frais avec la variable "Frais\_Gestion"

Les variables relatives à l'environnement économiques diffèrent elles en fonction de l'approche :

- Approche sur les BE par scénario
  - Les rendements actions : "Action\_Y 1", "Action\_Y 2", ... , "Action\_Y 30" ;
  - Les rendements immobiliers : "Moy\_Immo\_0\_10", "Moy\_Immo\_10\_20", "Moy\_Immo\_20\_30", "Moy\_Immo", "ET\_Immo\_0\_10", "ET\_Immo\_10\_20", "ET\_Immo\_20\_30", "ET\_Immo", "Immo\_max" et "Immo\_min" ;
  - Les taux zéro-coupon : "Taux\_ZC\_Y 1", "Taux\_ZC\_Y 2", "Taux\_ZC\_Y 4", "Taux\_ZC\_Y 6", "Taux\_ZC\_Y 8", "Taux\_ZC\_Y 10", "Taux\_ZC\_Y 12", "Taux\_ZC\_Y 16", "Taux\_ZC\_Y 20", "Taux\_ZC\_Y 24", "ET\_ZC\_0\_10", "ET\_ZC\_10\_20", "ET\_ZC\_20\_30" ;
  - Les taux concurrents : "Taux\_Concu\_Y 1", "Taux\_Concu\_Y 2", "Taux\_Concu\_Y 3", "Taux\_Concu\_Y 5", "Taux\_Concu\_Y 7", "Taux\_Concu\_Y 9" .
- Approche sur les BE :
  - La courbe de taux EIOPA : "Taux\_Y 1", "Taux\_Y 2", ... , "Taux\_Y 30" .

## 3.3 Générer les bases de données à partir de l'outil ALM

A ce stade, nous avons choisi nos variables explicatives ainsi que nos plages de variations. Nous devons alors calculer les BE et pour cela nous utilisons l'outil ALM.

Afin de calculer un BE, l'outil ALM a besoin qu'on lui fournisse les paramètres liés à l'environnement économique et à la situation de notre assureur. Concernant l'environnement économique, nous sélectionnons une courbe de taux EIOPA. Celle-ci nous permet, grâce au GSE, de produire les 1000 scénarios économiques projetés en entrée de l'outil ALM. En plus des scénarios économiques, nous

spécifions la situation et les caractéristiques de l'assureur. Nous fournissons ainsi en entrée de l'outil ALM ce que nous appelons des "entrées fixées" qui sont les paramètres de notre assurance vie fictive que nous ne faisons pas varier, à l'inverse des "entrées variables" qui sont les variables explicatives que nous avons sélectionnées. Grâce à ces paramètres, l'outil calcule les 1000 BE par scénario et par conséquent un BE moyen.

A la lumière du processus de calcul de BE que nous venons de détailler, la création de la base de données doit s'effectuer en faisant varier les paramètres de l'outil ALM selon deux blocs distincts : un bloc de variations des paramètres économiques (bloc 2) et un bloc de variation des paramètres de l'assureur (bloc 1).

Le schéma 3.13 détaille le processus de création de la base de données et se lit de la gauche vers la droite. Nous sélectionnons donc une courbe de taux parmi les 5 sélectionnées pour notre étude et générons grâce au GSE 1000 scénarios économiques. A ce stade, nous entrons dans le bloc 1, dans lequel nous parcourons toutes les combinaisons possibles de nos "entrées variables" et nous calculons pour chaque combinaison les 1000 BE par scénario et le BE. Une fois que toutes les combinaisons ont été explorées, nous sortons du bloc 1 pour entrer dans le bloc 2 et sélectionnons une nouvelle courbe de taux, et nous reproduisons la même démarche en relançant le bloc 1.

Voici ci-dessous une schématisation de ce que nous venons d'évoquer :

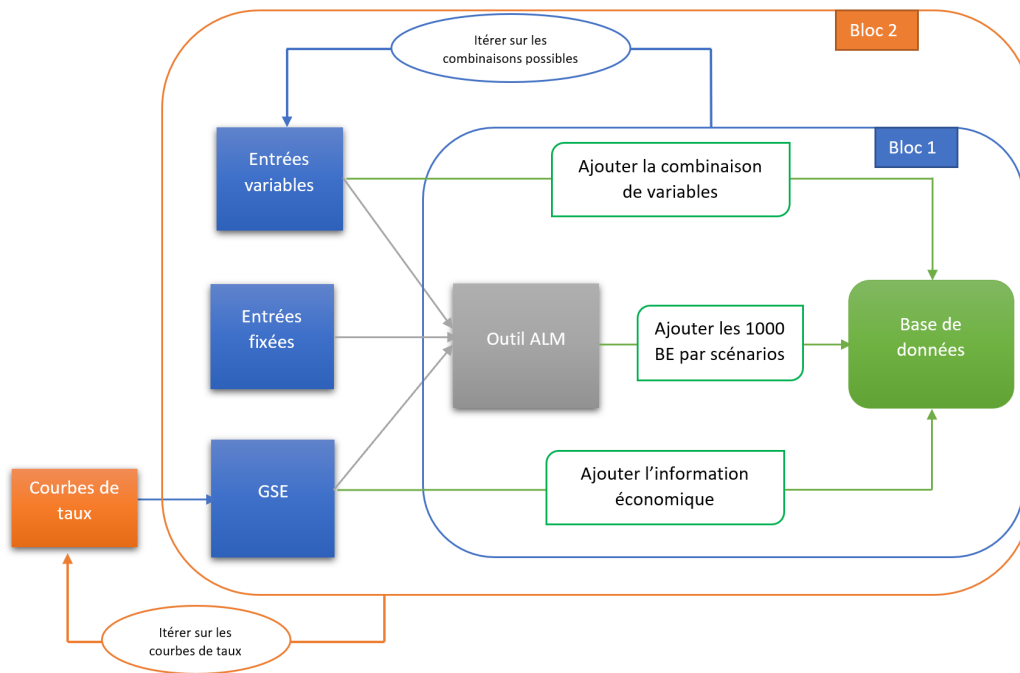


FIGURE 3.13: Schéma du processus de création des bases de données

Concernant le nombre de combinaisons possibles dans le bloc 1, nous multiplions le nombre de modalités par variables explicatives. En voici le détail :

$$nb_{combinaisons} = nb_{MP} \times Variation_{Frais\_Gestion} \times Variation_{Plus\_Moins\_VL} \times Variation_{Taux\_PPE} \times nb_{Alloc\_cibles}$$

avec 11 MP, 3 variations de taux de frais de gestion, 5 variations de taux de PMVL, 5 variations de taux de PPE et 3 allocations cibles. Nous obtenons ainsi 2475 combinaisons soit 2 475 000 BE par scénario et 2475 BE. Nous effectuons cela pour 5 courbes de taux différentes ce qui nous amène à une

base de données de 12 375 000 BE par scénario et 12 375 BE. La base d'apprentissage des BE se déduit directement de la base de données des BE par scénario en remplaçant l'information économique issue du GSE par la courbe de taux et en moyennant nos BE par scénario.

### 3.4 Visualisation des données et études préliminaires

Nous allons dans cette partie effectuer une analyse descriptive de nos données. Nous avons généré nous-même notre base de données, nous n'avons donc pas de valeurs manquantes. De plus, nos variables sont toutes quantitatives. Dans un premier temps, nous nous focalisons sur nos variables à expliquer : les BE et les BE par scénario. Voici ci-dessous les premières informations concernant nos données produites :

TABLE 3.4: Statistiques descriptives des BE et BE par scénario

	BE (M€)	BE par scénario (M €)
<b>Min</b>	442, 1	405, 6
<b>1<sup>er</sup> quartile</b>	517, 5	502, 4
<b>Médiane</b>	583, 7	567, 5
<b>Moyenne</b>	661, 1	662, 8
<b>3<sup>ème</sup> quartile</b>	759, 1	673, 8
<b>Max</b>	1851	175700

Nous remarquons que nous avons des valeurs plus extrêmes pour les BE par scénario, mais dont l'effet est moindre sur nos BE bien qu'ils puissent prendre des valeurs de l'ordre du milliard d'euros. Nous résumons ces informations dans un diagramme moustache, et nous représentons par un histogramme la distribution de nos BE en représentant l'espérance par la ligne pointillée rouge et la densité par la ligne complète bleue.

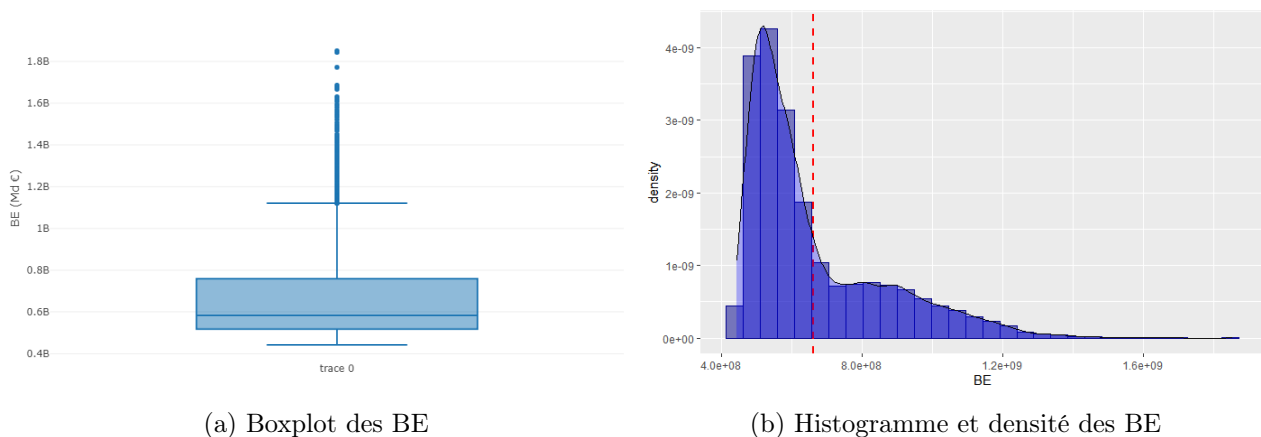


FIGURE 3.14: Statistiques descriptives des BE

Afin d'expliquer au mieux les valeurs extrêmes de BE, nous avons mené une étude permettant de lier ses valeurs extrêmes à nos variables explicatives. Graphiquement, nous traçons les diagrammes moustache de nos BE en fonction des modalités de nos variables et étudions les variations de nos BE. Nous fixons la modalité de notre variable étudiée, laissant les autres libres ce qui nous permet

d'obtenir une distribution de BE par modalité. Nous pouvons ainsi observer des distributions variables en fonction des modalités pour certaines variables, c'est le cas spécialement du TMG :

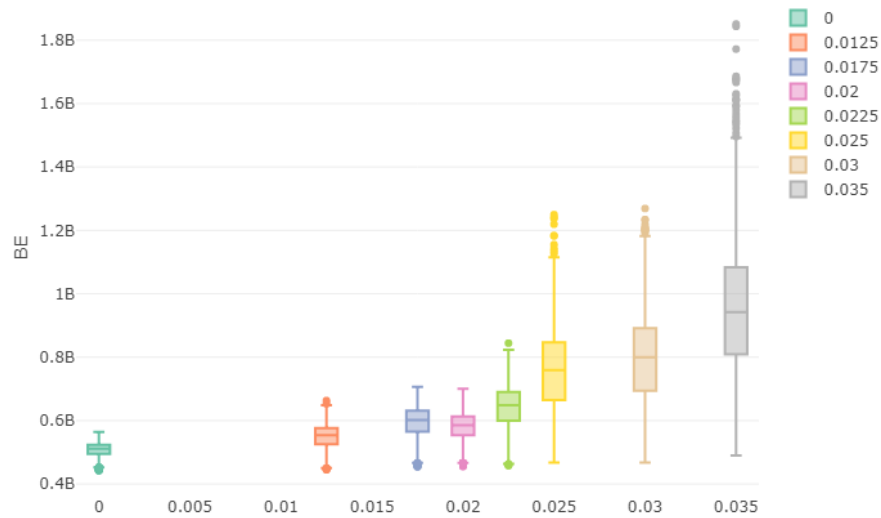


FIGURE 3.15: Boxplot des BE en fonction du TMG

Nous remarquons alors que nos valeurs extrêmes de BE ont toutes le TMG maximal de 3,5%. Une explication plausible est qu'un TMG élevé suppose des engagements grands pour l'assureur qui doit revaloriser chaque année l'encours des assurés au TMG. De plus, cet effet est accru par notre hypothèse de TMG viager. Cependant, nous ne faisons pas ici une relation de cause conséquence entre TMG élevé et BE élevé : le BE augmente lorsque la valorisation du passif excède le rendement de l'actif, ce qui dépend directement de l'environnement économique et de la courbe de taux. Pour vérifier cela, nous traçons le BE en fonction du TMG et de la variable "Taux\_Y\_1", qui nous permet de distinguer les 5 scénarios considérés. Nous obtenons le schéma suivant :

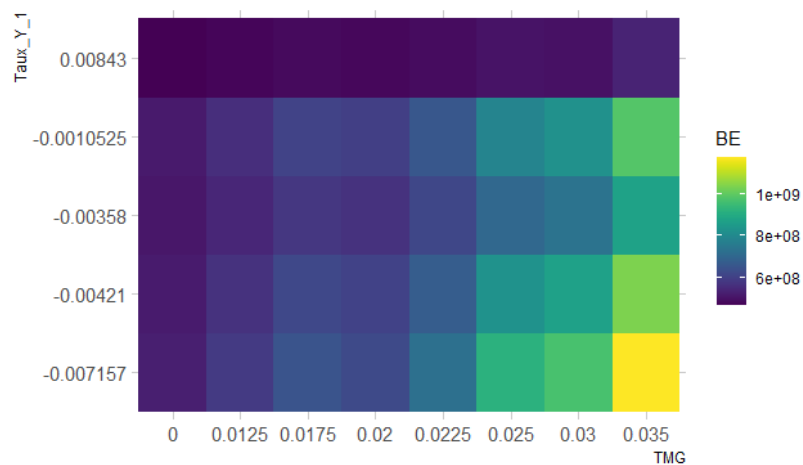


FIGURE 3.16: Heatmap des BE en fonction du TMG et du taux ZC de la première année de projection

Nous traçons plus exactement la moyenne des BE de paramètres "Taux\_Y\_1" et "TMG" dans chacune des "cases". Ce graphique valide nos dires et montre qu'un environnement de taux négatifs combiné à un TMG important entraîne, en moyenne, des BE plus élevés. En revanche, nous notons qu'un TMG élevé dans un environnement de taux positifs donne un BE moyen bien moins élevé. Ceci montre encore une fois que la valeur des garanties et des options est directement impactée par l'environnement économique.

A l'inverse, pour certaines variables, la distribution du BE ne semble pas être impactée par un changement de modalité. C'est le cas du taux de plus ou moins-values latentes comme le montre le graphique qui suit :

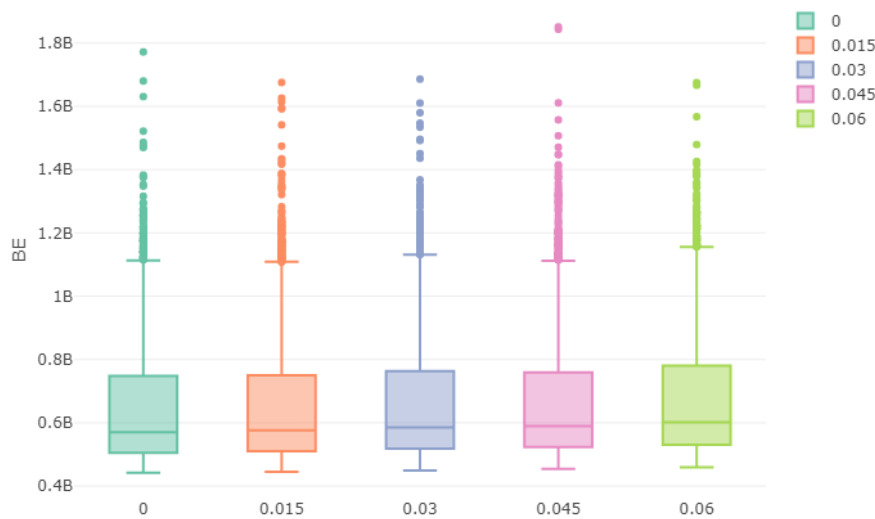
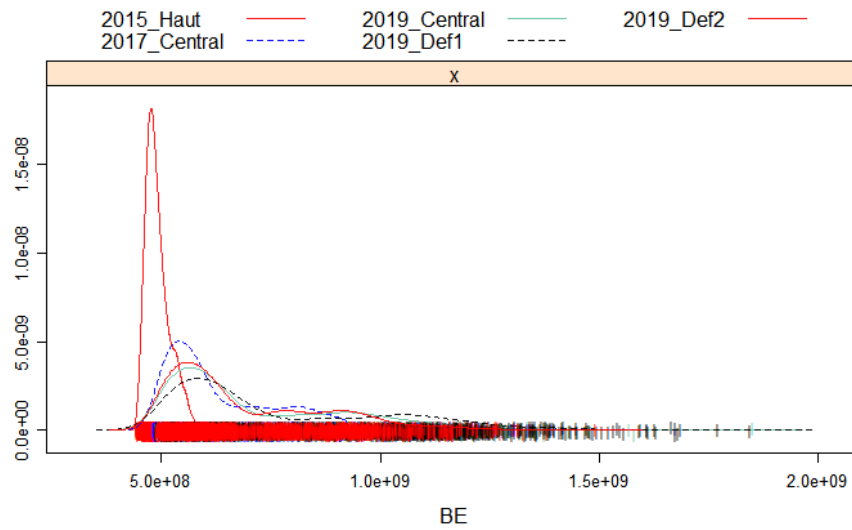


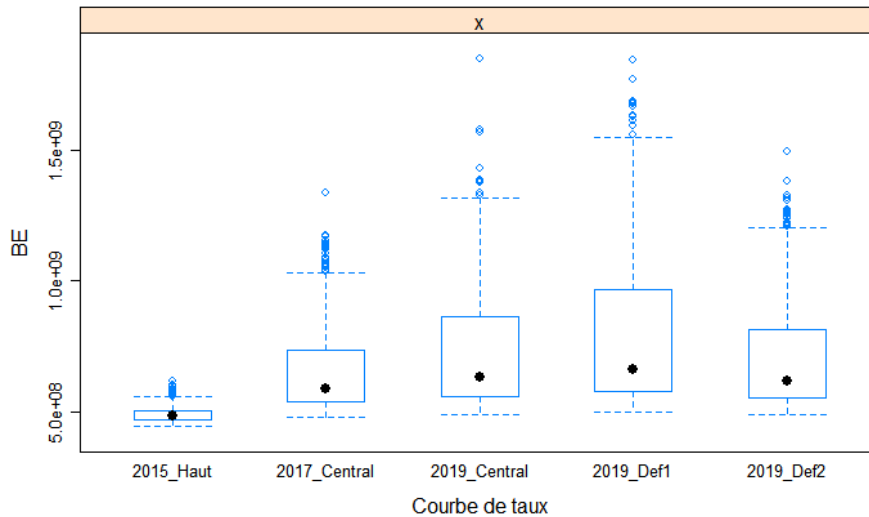
FIGURE 3.17: Boxplot des BE en fonction des PMVL

Cependant, les PMVL ont un impact mais de manière moins prononcée. Bien que cette étude ne soit que qualitative et ne permette pas de porter de conclusion définitive, elle a été effectuée pour chacune des variables et voici nos observations. Le TMG, l'âge, l'ancienneté et le pourcentage d'obligation dans l'allocation cible sont les variables pour lesquelles la distribution du BE varie en fonction de leur modalité. Inversement, les frais, le taux de PMVL et de PPE semble expliquer moins significativement la variabilité des BE. Nous devrions faire les mêmes observations au moment de notre étude *machine learning*, et c'est ce que nous étudierons.

Notre étude se tourne maintenant vers l'environnement économique et ses conséquences sur le BE.



(a) Distribution des BE en fonction de la courbe de taux



(b) Boxplot des BE en fonction de la courbe de taux

FIGURE 3.18: Description des BE en fonction des scénarios

Dans le cas de la courbe de taux de 2015 avec un choc S2 à la hausse et donc un environnement de taux positif, la répartition de nos BE est moins étalée et les valeurs de nos BE sont inférieures comparées aux autres courbes de taux. Cela s'explique premièrement par des facteurs d'actualisation plus grands mais aussi parce que cet environnement de taux est plus favorable à l'assureur qui va pouvoir contrôler les rachats dynamiques ce qui baisse la valeur des prestations.

Concernant les autres courbes de taux où nous retrouvons des taux négatifs, les densités de BE sont proches, et c'est pour la courbe de taux 2019 déformée selon la déformation 1 que nous retrouvons les BE les plus extrêmes. Cela est compréhensible car cette déformation entraîne une baisse des taux et ainsi des taux encore plus négatifs que le scénario central. Cela impacte notre actualisation et nos rendements financiers. Dans cet environnement de taux bas, l'assureur a des difficultés à servir les grands TMG et aussi atteindre la revalorisation cible, entraînant inexorablement des rachats dynamiques. Le choix de la courbe de taux en entrée va donc influencer la valeur de notre BE.



A l'issue de cette étape, nous avons nos deux bases de données pour l'apprentissage des BE et des BE par scénario. Nous les divisons en 2 afin d'obtenir la base de données d'apprentissage (80% de la base) et la base de test (20% de la base) afin de mener l'étude *machine learning*.

### 3.5 Les mesures de l'erreur de prédiction

Le choix des mesures de l'erreur dépend si le problème considéré est une régression ou une classification. Prenons un exemple simple : nous choisissons une mesure d'erreur qui calcule le pourcentage des erreurs absolues sur nos prédictions. Nous nous plaçons dans le cas d'une régression logistique où nous avons 1% de 0 et 99% de 1. Après ajustement du modèle et prédiction, nous obtenons un pourcentage d'erreur de 1%. Il s'agit d'un bon résultat au premier étant donné que nous voulons minimiser l'erreur. Cependant, en analysant nos prédictions, nous observons que notre modèle ne prédit que des 1. En d'autres termes, son pouvoir prédictif nul. C'est pour cela qu'il est conseillé d'utiliser les courbes ROC et l'AUC, c'est-à-dire l'aire sous la courbe ROC comme mesure de performance pour les problèmes de régression logistique.

Nous noterons dans la suite  $y_i$  les valeurs observées de BE et  $\hat{y}_i$  les valeurs prédites par le modèle.

#### 3.5.1 *Root Mean Square Error* ou RMSE

Cette mesure de l'erreur consiste à calculer la racine de la *Mean Square Error* ou MSE, c'est à dire à faire la moyenne arithmétique des carrés des écarts entre les prévisions et les observations et d'en prendre la racine.

Plus formellement, cela s'écrit :

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} (y_i - \hat{y}_i)^2}$$

C'est une valeur à minimiser, le MSE n'étant autre que la variance résiduelle.

#### 3.5.2 *Mean Absolute Error* ou MAE

C'est la moyenne arithmétique des valeurs absolues des écarts.

$$MAE = \frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} |y_i - \hat{y}_i|$$

Encore une fois, cette mesure est à minimiser. MAE et RMSE servent essentiellement à comparer les modèles entre eux, leurs interprétations étant ambiguës car très dépendantes de l'échelle des valeurs.

#### 3.5.3 *Mean Absolute Percentage Error* ou MAPE

$$MAPE = \frac{100}{n_{obs}} \sum_{i=1}^{n_{obs}} \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Comme son nom l'indique, c'est la moyenne des écarts en valeur absolue par rapport aux valeurs observées.

Le grand intérêt de cette mesure est que c'est un pourcentage, elle est par conséquent plus facile à interpréter. Dans notre cas, avec des BE de l'ordre de 580 M€, les mesures peuvent prendre des valeurs importantes et cela peut nous faire douter de nos modèles sans que cela ne soit nécessaire, d'où l'intérêt d'ajouter cette mesure en pourcentage qui permet de relativiser l'erreur globale.

### 3.6 Les modèles utilisés

Nous introduisons les notations que nous allons utiliser par la suite. Soit  $Y$  la variable à expliquer et  $X^1, X^2, \dots, X^p$  les  $p$  variables explicatives. On note  $f$  la fonction qui lie  $y$  à  $x = (x^1, x^2, \dots, x^p) \in \mathbb{R}^p$ . Nous disposons de  $n$  observations et on note  $z = \{(x_1, y_1), \dots, (x_n, y_n)\}$  un échantillon de loi  $F$ .

#### 3.6.1 Régression linéaire

Le premier modèle à tester est un modèle de régression linéaire. Nous commençons par les modèles les plus simples pour plusieurs raisons. Il est totalement possible que le modèle linéaire explique parfaitement la variable réponse et alors il n'est pas nécessaire d'essayer des algorithmes plus compliqués. Plus vraisemblablement, cela peut servir de repère pour les autres algorithmes et permet d'avoir une première échelle des erreurs de prédiction. Finalement, le modèle est interprétable et cela nous donne ainsi une première vision des variables plus importantes, les coefficients de la régression nous donnant directement leurs importances respectives.

Dans le cadre d'une régression linéaire, nous supposons une fonction linéaire comme fonction caractéristique. Ainsi, les prédictions  $\hat{y}$  se calculent de la manière suivante :

$$\hat{y}_i = f(x_i) = b_0 + \sum_{k=1}^p b_k x_{i,k}$$

Avec  $b = (b_1, \dots, b_p)$  déterminé par l'estimateur des moindres carrés.

Devant la complexité du modèle ALM, nous intégrons d'autres modèles ayant fait leurs preuves dans des études du même acabit.

#### 3.6.2 Le *bagging* et son application dans les forêts aléatoires

Avant d'introduire spécifiquement l'algorithme de forêts aléatoires, nous expliquons premièrement ce qu'est un arbre de décision, puis nous présenterons secondement la notion de *bagging*, pour *bootstrap aggregating* (BREIMAN, 1996), dont les forêts aléatoires sont une amélioration.

#### Les arbres de décision

Mieux connus sous l'acronyme CART (*Classification and Regression Tree*), ils font leur apparition dans les années 60 puis sont formalisés par le travail de BREIMAN et al., 1984.

Pour notre étude, nous utiliserons des arbres de régression, c'est-à-dire un arbre binaire aidant à choisir une valeur plausible pour  $Y$  dont on connaît les caractéristiques  $X_1, \dots, X_p$ . La construction de cet arbre consiste à effectuer un partitionnement récursif des données et d'estimer un modèle simple dans chaque "feuille" de l'arbre. Les solutions obtenues peuvent être représentées sous une forme graphique facile à interpréter. Le fonctionnement des arbres binaires de décision est intuitif : nous partons de la base de l'arbre appelée la racine où nous considérons encore l'ensemble des données puis nous divisons les données au niveau des nœuds pour les séparer en deux nœuds fils de manière à ce que les sous-ensembles de données à chaque nœuds soient plus homogènes selon une certaine règle. Les derniers nœuds aussi appelés nœuds terminaux sont les feuilles de l'arbre.

La construction de l'arbre s'effectue en 2 phases : la phase de *growing* et la phase de *pruning*.

#### Growing ou construction de l'arbre binaire maximal

Comme nous l'avons vu, la construction de l'arbre consiste à définir une séquence de nœuds. Chaque nœud est défini par une règle de division ou une règle d'arrêt, souvent définie comme étant un nombre minimum d'individus par nœud ou une profondeur d'arbre maximale. Concernant la règle de division, sa définition dépend du type de variable considéré. Pour une variable quantitative, la division est définie par un dépassement de seuil et pour une variable qualitative comme un partage en deux groupes de modalités.

Ainsi, l'algorithme va ajouter des nœuds à l'arbre en respectant un critère de division reposant sur la définition d'une fonction d'hétérogénéité, ayant pour but de partager les individus en deux groupes les plus homogènes possibles au sens de  $Y$ . Plaçons-nous à un nœud  $k$  et définissons  $k_d$  et  $k_g$  respectivement le nœud droit et le nœud gauche. Soit  $H_i$  la valeur de la fonction d'hétérogénéité au nœud  $i$ . Alors, parmi toutes les divisions du nœud  $k$ , l'algorithme sélectionnera la division qui minimise la somme des hétérogénéités des nœuds  $k_d$  et  $k_g$ . L'algorithme résout ainsi à chaque étape :

$$\max_{\{Divisions\ de\ X^j, j=1\dots p\}} H_k - (H_{k_g} + H_{k_d})$$

avec, par exemple, la variance comme fonction d'hétérogénéité  $H$  :

$$H_k = \frac{1}{card(k)} \sum_{i \in k} (y_i - \bar{y}_k)^2$$

Le problème revient dans ce cas à minimiser la variance intraclasse à chaque nœud.

La croissance de l'arbre est arrêtée grâce aux règles d'arrêts. Finalement, on affecte à chaque feuille une valeur, qui n'est autre que la moyenne empirique des individus présents dans la feuille.

#### Pruning ou élagage de l'arbre

L'arbre construit par la démarche présentée précédemment est très dépendant de l'échantillon avec lequel il est construit, ce qui le rend très instable. Il est sur-ajusté. Pour pallier cela, il a été mis en place une phase de *pruning* dont le but est de trouver le sous-arbre dont les performances sont les meilleurs sur un échantillon de validation. Cet arbre optimal peut notamment être obtenu par validation croisée. Nous ne développerons pas plus cette phase de *pruning*.

La notion d'arbre de décision étant introduite, nous allons maintenant détailler les notions de *bagging* et *boosting* qui nous permettront de mieux appréhender les modèles Random Forest et XGBoost. *Bagging* et *boosting* font partie de la même famille appelée "ensemble technique" ou "ensemble learning" dont le principe est d'utiliser plusieurs modèles jugés "basiques" dans le but de produire un modèle optimisé. On utilise un ensemble de *weak learners* pour créer un *strong learner*.

De nombreux exemples existent pour mieux comprendre cette idée de *weak learner* et *strong learner* et nous allons en développer un ici. Supposons que Luc, un individu souhaitant épargner dans cette période de crise sanitaire, ait l'intention de souscrire à une assurance vie. Plusieurs critères vont influencer son jugement et vont le faire décider si oui ou non il devrait souscrire. Dans le doute, il demande à son entourage leurs avis. Chacun d'entre eux va alors avoir son raisonnement, certains vont regarder le TMG, d'autres la PB contractuelle... Il aura donc finalement l'avis de plusieurs personnes qui ont tous raisonné différemment et la combinaison de ces raisonnements va lui permettre de se rapprocher du bon choix. Luc est donc parti de raisonnements personnels non stables pour arriver à un raisonnement final et une réponse finale des plus réfléchies. C'est le principe même du *boosting* ou du *bagging* : utiliser des arbres binaires qui ont pour propriété d'être instables et qui vont réduire la variance du modèle final en les agrégeant.

### Le *bagging*

Idéalement, dans le cas d'une régression, la prévision par agrégation des modèles ayant  $B$  échantillons indépendants  $\{z_b\}_{b=1\dots B}$  serait obtenue par moyenne empirique des résultats obtenus pour les modèles de chaque échantillon.

Nous aurions alors approximé  $f$  par  $\hat{f}_B$  tel que  $\hat{f}_B = \frac{1}{B} \sum_{b=1}^B \hat{f}_{z_b}$

Cependant, il n'est pas réaliste d'un point de vue quantité de données d'envisager  $B$  échantillons indépendants, c'est pour cela qu'ils sont remplacés par  $B$  échantillons *bootstrap* obtenus chacun par  $n$  tirages avec remise.

C'est le principe du *bagging*, abréviation de *bootstrap aggregating* : moyenner des prévisions sur  $B$  répliques *bootstrap* pour réduire la variance. Nous rappelons succinctement la définition du *bootstrap* : le *bootstrap* consiste à faire de l'échantillonnage par tirage aléatoire avec remise d'individus parmi une population. Sa motivation est d'approcher par simulation Monte Carlo la distribution d'un estimateur lorsque la loi de l'échantillon n'est pas connue. On effectue ainsi un grand nombre de tirage d'échantillons *bootstrap* sur lesquels on calcule l'estimateur concerné et la loi ainsi simulée est une approximation asymptotiquement convergente de la loi de l'estimateur (sous l'hypothèse que les échantillons sont indépendants de même loi et que l'estimateur soit indépendant de l'ordre des observations).

Nous pouvons résumer la démarche en détaillant le fonctionnement de l'algorithme :

---

#### Algorithm 1 Bagging

---

Soit  $z$  un échantillon et  $x_0$  à prévoir.

**for**  $b = 1$  à  $B$  **do**

    Tirage d'un échantillon *bootstrap*  $z_b$

    Estimation de  $\hat{f}_{z_b}(x_0)$

**end for**

Calculer l'estimation moyenne  $\hat{f}_B$  tel que  $\hat{f}_B = \frac{1}{B} \sum_{b=1}^B \hat{f}_{z_b}(x_0)$

---

### Forêts aléatoires

Une forêt aléatoire est l'agrégation d'une collection d'arbres aléatoires. La logique s'inscrit dans la lignée du *bagging*, mais avec une volonté de rendre plus indépendants les arbres agrégés en ajoutant une composante aléatoire dans le choix des variables qui interviennent dans le modèle. On applique donc

le *bagging* à des arbres binaires de décision en ajoutant un tirage aléatoire de  $m$  variables explicatives parmi les  $p$ . Ainsi, en se basant sur l'algorithme du bagging nous obtenons :

---

**Algorithm 2** Forêts Aléatoires
 

---

Soit  $z$  un échantillon et  $x_0$  à prévoir.

**for**  $b = 1$  à  $B$  **do**

    Tirage d'un échantillon bootstrab  $z_b$

    Estimation d'un arbre sur cette échantillon en ne tirant que  $m$  variables sur les  $p$  possibles

**end for**

Calculer l'estimation moyenne  $\hat{f}_B$  tel que  $\hat{f}_B = \frac{1}{B} \sum_{b=1}^B \hat{f}_{z_b}(x_0)$

---

Pour plus d'informations, nous vous conseillons de vous référer à l'article (BREIMAN, 2001) de Leo Breiman qui, accompagné d'Adèle Cutler, ont formellement proposé la démarche de forêts aléatoires.

### 3.6.3 Le *boosting* et son application dans XGBoost

L'algorithme adaboost de classification de variable binaire a été développé par FREUND, SCHAPIRE et al., 1996 et est à l'origine du *boosting*. S'en est suivi de nombreuses études pour adapter cet algorithme à d'autres problèmes : classification en  $k$  classes, régressions... Le *boosting* est à différencier du *bagging* vu précédemment. Contrairement au *bagging* qui itère chaque modèle "*weak learner*" indépendamment l'un de l'autre puis agrège leurs sorties à la fin sans préférence, le *boosting* favorise le "travail d'équipe". Ce que nous entendons par "travail d'équipe" est que chaque "*weak learner*" va être estimé puis va influencer le travail des autres "*weak learner*" en attribuant plus de poids aux observations mal ajustées ou mal prédites. Cela vise à réduire la variance, mais aussi le biais des prévisions. En résumé, le principe est d'agréger des modèles comme dans le cas du *bagging*, mais chaque modèle est une version adaptative du précédent, en donnant plus de poids lors de l'estimation suivante aux observations mal ajustées.

Nous avons choisi d'utiliser XGBoost (comme eXtreme Gradient Boosting) car il a brillamment fait ses preuves, notamment dans les concours Kaggle. Il présente l'intérêt d'avoir un grand nombre d'hyperparamètres, c'est-à-dire des paramètres spécifiques à l'algorithme implémenté ( par exemple la taille de l'arbre dans le cas d'un modèle CART, le nombre de neurones d'un réseau de neurones...), ce qui permet d'avoir un contrôle sur l'implémentation de l'algorithme et d'optimiser ses performances (Section 3.7). Pour une description détaillée du *boosting* et du fonctionnement de XGBoost, un lecteur intéressé pourra se référer au mémoire de GAUVILLE, 2017, grâce auquel nous avons amélioré notre compréhension de cette approche.

### 3.6.4 Les réseaux de neurones

L'Intelligence Artificielle ou IA s'applique aujourd'hui à de multiples domaines comme l'industrie, les transports (véhicules connectés) entre autres et a pour objectif la simulation du comportement humain pour fournir des réponses efficaces, fiables et personnalisées aux utilisateurs. Le réseau de neurones s'inscrit dans cette démarche et tient son nom de sa structure comparable aux réseaux de neurones du cerveau. Le concept de neurones artificiels fut d'ailleurs inventé en 1942 par le mathématicien Walter Pitts ET un neurophysicien Warren McCulloch.

Après avoir rencontré de nombreux freins technologiques (puissance de calcul insuffisante, insuffisance des références théoriques...) et après avoir été supplantés pour un temps par des apprentissages *machine learning* comme étudié précédemment, ils sont aujourd'hui sur les devants de la scène grâce à

l'utilisation répandue du *Deep Learning*. Le *Deep Learning* est un ensemble de méthodes d'apprentissage cherchant à modéliser des données d'architectures compliquées et combinant des transformations non-linéaires. Avant de présenter le réseau neuronal et sa complexité, nous allons nous intéresser au perceptron monocouche, aussi appelé neurone formel.

### Le perceptron monocouche

Commençons par un rappel succinct de biologie en expliquant le fonctionnement d'un neurone humain dont voici un schéma :

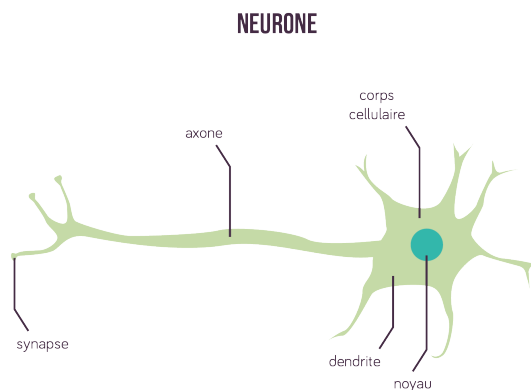


FIGURE 3.19: Neurone biologique

Ainsi, un neurone est une cellule de base du tissu nerveux composée par :

- Les synapses, les points de connexion avec les axones d'autres neurones ;
- Les dendrites permettant au neurone de recevoir l'information d'autres neurones ;
- L'axone, le canal de sortie du neurone vers d'autres neurones ;
- Le noyau donnant une réponse aux stimuli reçus en entrée.

Le principe est le suivant : le neurone reçoit des stimuli électriques par ses dendrites qui les transmettent au noyau. Ce dernier traite ces stimuli et donne une réponse en conséquence, qui est transmise dans l'axone via une impulsion électrique jusqu'à la zone d'échange d'information appelée synapse.

Par analogie, le perceptron reçoit en entrée les variables du modèle  $x = (x^1, x^2, \dots, x^p)$ , auxquelles il affecte des poids  $x = (w_1, w_2, \dots, w_p)$ . Il fait ensuite la somme de cette information et applique une fonction  $g$  appelée fonction d'activation. Le résultat de la fonction d'activation est la sortie  $y$  du perceptron.

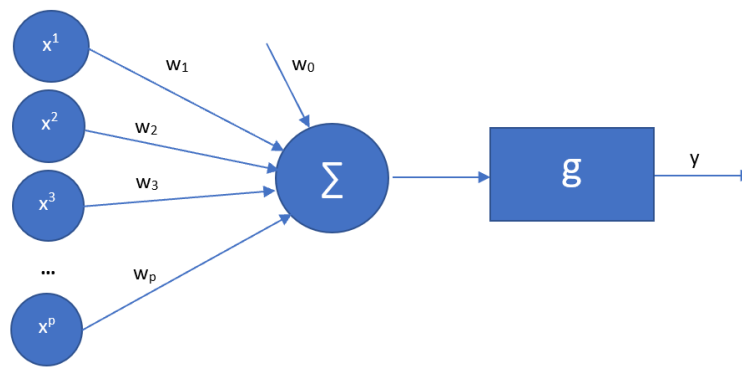


FIGURE 3.20: Neurone formel

Les neurones, utilisés à la fois pour la régression et la classification, diffèrent par leur fonction d'activation. On trouve notamment des neurones :

- linéaires avec  $g$  la fonction identité
- ReLU avec  $g(x) = \max(0, x)$
- sigmoïde avec  $g(x) = \frac{1}{1+e^x}$

Faisant face à un problème de régression, nous serons amenés à utiliser un neurone de sortie linéaire.

### Les différents types de réseaux de neurones

Un réseau de neurones combine plusieurs couches de traitement, chaque couche ayant un certain nombre de neurones de même architecture que le neurone formel étudié ci-dessus. Chaque couche utilise la sortie de la couche précédente en guise d'entrée. Les réseaux de neurones qui fonctionnent sur deux ou trois couches de neurones connectés sont appelés neurones peu profonds. À l'inverse, les réseaux sont qualifiés de *Deep Learning* lorsqu'ils comportent de nombreuses couches. Nous distinguons 3 types de réseaux de neurones : les réseaux à propagation directe (*feedforward*) aussi appelés perceptrons multicouches, les réseaux de neurones à rétroaction (*feedback*) et les réseaux de neurones convolutionnels. Nous allons détailler les spécificités des deux premiers, le deuxième étant plus adapté au traitement d'image et donc non utilisé dans ce mémoire.

Les réseaux de neurones *feedforward* ne sont pas récursifs : l'information circule de la couche d'entrée à la couche de sortie, et seulement dans ce sens. Voici ci-dessous un schéma de réseau de neurones *feedforward* à deux couches cachées, d'épaisseur quatre neurones.

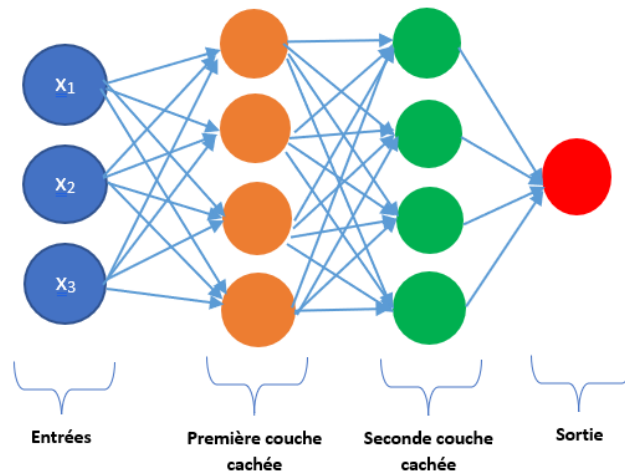


FIGURE 3.21: Réseau de neurones à propagation directe

Sur la couche de sortie, il est courant d'utiliser une fonction d'activation différente de celles utilisées sur les couches cachées, celle-ci dépendant du problème considéré.

Les réseaux de neurones *feedback* ou réseaux de neurones récurrents suivent la même logique, à l'exception qu'ils utilisent les sorties de leur couche ou des couches suivantes comme entrées en plus. Ils sont particulièrement adaptés pour le traitement des données séquentielles comme les textes et les séries temporelles.

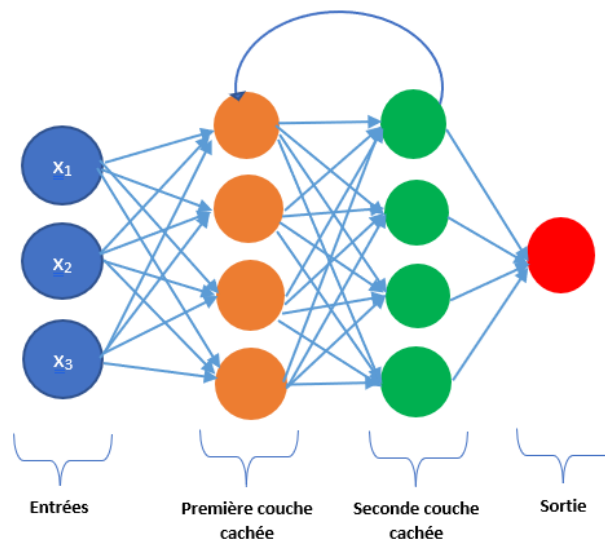


FIGURE 3.22: Réseau de neurones à rétroaction

Le succès des réseaux de neurones vient du théorème d'approximation universelle introduit par CYBENKO, 1989 puis généralisé pour toutes les fonctions d'activation par les travaux de HORNIK, 1991. Dans ce théorème, Hornik montre que n'importe quelle fonction régulière et bornée de  $\mathbb{R}^d \rightarrow \mathbb{R}$  peut être approximée aussi précisément que souhaité par un réseau de neurone à une couche cachée



contenant un nombre fini de neurones, ayant la même fonction d'activation et un neurone de sortie linéaire. Formellement, cela se réécrit :

Soit  $\phi$  une fonction d'activation bornée, continue et non décroissante. Soit  $K_d$  un ensemble compact dans  $R_d$  et  $C(K_d)$  l'ensemble des fonctions continues sur  $K_d$ . Soit  $f \in C(K_d)$ , alors pour tout  $\epsilon > 0$  il existe  $N \in \mathbb{N}$ , les réels  $v_i, b_i$  et le vecteur de  $\mathbb{R}^d$   $w_i$  tel que :

$$F(x) = \sum_{i=1}^N v_i \phi(\langle w_i, x \rangle + b_i)$$

alors nous obtenons le résultat suivant

$$\forall x \in K_d, |F(x) - f(x)| \leq \epsilon$$

Bien qu'intéressant d'un point de vue théorique, la difficulté reste la paramétrisation et l'application pratique. Nous faisons ainsi le choix pour notre étude d'utiliser un réseau de neurones à propagation directe peu profond pour notre étude sur le BE moyen seulement. Nous faisons remarquer qu'un réseau de neurones récurrent profond sur notre base de BE par scénario serait une étude très intéressante à mener : nos scénarios économiques sont des séries temporelles qui mènent à une valeur finale, le BE, et des études équivalentes ont été menées sur la prédiction de séries temporelles. Ces réseaux de neurones récurrents sont appelés LSTM pour *Long Short-Term Memory* et permettent de prendre en compte le caractère récurrent des séries temporelles. Un lecteur intéressé pourra se faire une idée du potentiel de cette méthode en se référant à l'application sur les prédictions d'inondation (LE et al., 2019).

Afin de choisir les hyperparamètres de notre réseau, nous nous basons sur le travail de STATHAKIS, 2009 dans lequel il est stipulé qu'avec un réseau de neurone à deux couches cachées, le nombre maximal de neurones cachés pour avoir une erreur suffisamment petite est :

$$\sqrt{(m+2) \times N} + 2 \times \sqrt{\frac{N}{m+2}}$$

pour la première couche et

$$m \times \sqrt{\frac{N}{m+2}}$$

pour la seconde, avec  $m$  le nombre de neurones de sortie et  $N$  la taille de l'échantillon d'apprentissage. A noter qu'il ne faut pas dépasser ces valeurs, au risque de sur-apprendre.

L'application numérique avec  $N = 9900$  et  $m = 1$  nous donne 287 neurones sur la première couche et 57 sur la seconde au maximum pour éviter l'*overfitting*.

Nous avons ainsi décidé de l'architecture de notre réseau, les poids entre les neurones sont maintenant estimés sur l'échantillon d'apprentissage en minimisant une fonction de perte par un algorithme de descente de gradient généralement. Nous choisissons comme fonction de perte et l'algorithme de descente de gradient stochastique (SGD) pour minimiser cette fonction de perte. L'algorithme SGD fonctionne de la manière suivante :

---

**Algorithm 3** SGD

---

Soit  $\theta = (w_1, b_1, \dots, w_l, b_l)$  les poids et biais du réseau de neurones

**for**  $r = 1$  à Epochs **do**

    Tirage sans remise d'un sous-ensemble de taille  $m$  dans la base d'apprentissage

$$\theta^{r+1} = \theta^r - \frac{\epsilon_r}{m} \sum_{i \in B_r} \Delta_{\theta} l(f(X_i, \theta), Y_i)$$

**end for**

---

Avec  $\epsilon_r$  le taux d'apprentissage dont le calibrage est très important pour la convergence de l'algorithme,  $l$  la fonction de perte,  $m$  le "batch" et  $N$  le nombre d'"epoch". Nous détaillons ci-dessous la signification de ces variables.

Pour des raisons de capacité de calcul, nous ne calculons pas le gradient pour la fonction de perte empirique pour chaque ligne de la base d'apprentissage : nous découpons notre base en sous-ensemble dont la taille correspond au "batch". Nous calculons ainsi la fonction de perte empirique et mettons à jour nos paramètres (poids et biais) seulement sur un sous-ensemble  $B$  tiré aléatoirement et de cardinal  $m$ , ce qui permet de gagner en temps de calcul. Trois options sont alors possibles :

1.  $m$  est égal à la taille de la base d'apprentissage c'est ce qu'on appelle le *batch mode* ;
2.  $m$  plus petit que la taille de l'échantillon mais plus grand que 1. Cette méthode est appelée *mini batch mode*. Nous choisissons  $m$  de manière à ce que la taille de la base de données soit divisible par  $m$  ;
3.  $m$  égal à 1, nous mettons alors à jour le gradient et les paramètres en même temps. C'est le *stochastic mode*.

Nous avons choisi la 2<sup>ème</sup> option dans notre modèle.

Un autre hyperparamètre est appelé "epoch". Un "epoch" est une itération de l'algorithme sur l'ensemble de la base d'apprentissage.

Nous remarquons dans le cas des réseaux de neurones, mais cela se généralise à tous nos modèles, le caractère décisif des hyperparamètres des modèles, qui doivent être choisis avec soin pour obtenir les meilleures performances possibles. Dans cette recherche d'optimisation des prédictions, nous détaillons notre démarche dans la partie qui suit.

## 3.7 Optimisation des hyperparamètres des algorithmes

Dans une démarche *machine learning* nous cherchons à minimiser notre erreur de prédiction. Pour ce faire, les différents modèles que nous utilisons ont déjà des fonctions de pertes implémentées qu'ils cherchent à minimiser.

### 3.7.1 Les problématiques de la minimisation des erreurs de prédictions

Dans l'optique de trouver les meilleurs résultats possibles, nous avons cherché à comprendre en détail les modèles que nous souhaitons utiliser afin de nous assurer de choisir, en toute connaissance de causes, les hyperparamètres adaptés et optimaux. Nous comprenons directement cette problématique dans l'implémentation d'un réseau de neurones qui nécessite de construire les différentes couches en choisissant le nombre de neurones par couche, leur fonction d'initialisation, l'initialisation des poids... Cet exemple sur les réseaux de neurones nous permet aussi de nous rendre compte que ce travail peut très rapidement être fastidieux : plus le modèle est complexe, plus il est difficile à paramétrer. A noter qu'un modèle complexe n'est pas synonyme de bons résultats, le paramétrage joue un rôle prédominant.

De plus, notre paramétrage doit prendre en compte le dilemme biais-variance soulevé au moment de l'estimation des modèles. Ce dilemme vient du fait que les modèles d'estimation commettent deux erreurs au moment de la prédiction : le biais et la variance. Le biais désigne l'erreur entre la prédiction

et la valeur observée alors que la variance est liée à la généralisation sur d'autres données. Nous parlons de dilemme car une diminution du biais va généralement augmenter la variance du modèle, d'où la nécessité de trouver un compromis pour avoir un modèle le plus performant et généralisable possible. Dans le cadre de notre mémoire cette notion est importante. Dans le cas des BE moyen par exemple, nous avons choisi 5 courbes de taux dans notre base d'apprentissage des BE et notre démarche est de pouvoir projeter les BE sur plusieurs années, avec évidemment des courbes de taux qui évoluent tout au long de notre projection temporelle. Ainsi nous souhaitons que notre modèle soit performant sur les courbes de taux implémentées, mais aussi sur les courbes de taux projetées. Pour répondre au mieux à cet enjeu, nous allons tâcher d'éviter l'*overfitting* qui consiste à avoir un modèle très, voire trop, performant sur la base d'apprentissage et qui ne s'adaptera pas à des données qui s'écartent trop des données d'apprentissage.

### 3.7.2 La *Cross Validation* en réponse à ces problématiques

#### La *k-fold Cross Validation*

Pour répondre à ces problématiques, nous utilisons la validation croisée pour paramétrer nos algorithmes. L'objectif principal de la validation croisée est de rechercher un modèle qui soit le plus généralisable possible. Plusieurs méthodes de validation croisée existent, tel que la méthode LOOCV (*Leave One Out Cross Validation*), la méthode LKOCV (*Leave K Out Cross Validation*) ou encore la *k-fold Cross Validation*.

La méthode LKOCV (LOOCV en est un cas particulier) consiste à découper notre base d'apprentissage en deux : une sous-base d'apprentissage et une base de validation contenant  $k$  éléments. L'apprentissage du modèle est alors effectué sur la sous-base d'apprentissage et un calcul d'erreur de prédiction de ce modèle est effectué sur la base de validation. L'opération est répétée jusqu'à avoir obtenu toutes les sous-bases d'apprentissage et de validation possibles.

La méthode de *k-fold Cross Validation*, elle, consiste à découper notre base d'apprentissage en  $k$  blocs et d'utiliser  $k-1$  blocs pour l'apprentissage du modèle et le dernier bloc pour validation du modèle. Nous répétons ce processus  $k$  fois, en échangeant les bases d'apprentissage et de validation. Prenons l'exemple avec  $k=3$ .

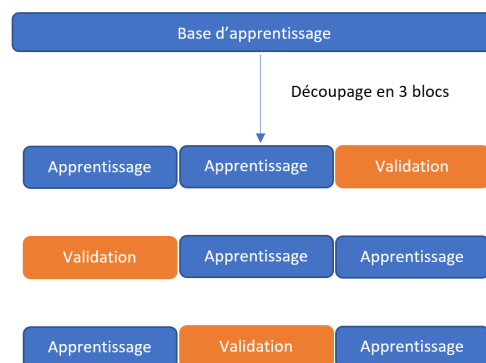


FIGURE 3.23: *3-fold Cross Validation*

Comme schématisé ci-dessus, cette méthode consiste à découper notre base d'apprentissage en 3 blocs. A chaque itération, l'apprentissage du modèle est effectué sur deux blocs et les performances

du modèle sont évaluées sur le troisième blocs. Cette démarche est effectuée à trois reprises (k fois pour généraliser) et nous obtenons ainsi trois performances pour le modèle, sur 3 bases d'apprentissage différentes. Ainsi, si nous obtenons des performances très variables sur nos différentes itérations, nous en déduirons que la variance de notre modèle est trop importante et nous chercherons d'autres paramètres.

Bien que le principe des deux méthodes présentées soit équivalent, nous notons que la méthode LKOCV est exhaustive dans le sens où la base d'apprentissage est découpée en toutes les combinaisons possibles, ce qui n'est pas le cas de la *k-fold Cross Validation* avec k combinaisons testées. Cependant, il en découle que la méthode LKOCV est plus coûteuse en temps de calcul, qui plus est dans le cadre de notre recherche des hyperparamètres les plus performants. Cela explique notre choix de la *k-fold Cross Validation*.

### La fonction *train* de Caret pour optimiser les paramètres

Cette méthode de validation croisée présentée, nous pouvons maintenant détailler notre stratégie de recherche des paramètres optimaux pour nos modèles. A cette fin, nous utilisons la library Caret sous R qui permet de chercher les paramètres les plus performants de nos modèles. La fonction *train* de cette *library* prend en entrée le type de modèle utilisé, la base d'apprentissage, la méthode de *cross-validation* retenue et les combinaisons de paramètre à tester puis renvoie finalement le modèle le plus performant. Nous détaillons ci-dessous les grandes lignes de l'algorithme implémenté de la fonction *train*.

---

#### Algorithm 4 La fonction *train* de Caret

---

```

for  $r = 1$  à  $N_{Combinaison}$  do
  for  $i = 1$  à  $k$  do
    Apprentissage sur les  $k-1$  blocs et evaluation de la performance sur le bloc de validation
  end for
  Calcul de la performance du modèle avec la  $r^{ième}$  combinaison des paramètres en faisant la moyenne
  des  $k$  performances.
end for
Le meilleur modèle est le modèle ayant la meilleure performance moyenne

```

---

Ainsi, nous parcourons les combinaisons possibles pour le modèle, et pour chaque combinaison une *k-fold Cross Validation* est effectuée. La performance du modèle est alors définie comme la performance moyenne sur les blocs de validation. Les modèles sont ainsi classés en fonction de leurs performances et le meilleur modèle est retenu.

Concernant le nombre de combinaisons possibles, deux options sont envisageables. Premièrement, nous pouvons donner les combinaisons que nous souhaitons essayer et la fonction les parcourt. Deuxièmement, nous pouvons donner le nombre de combinaisons que nous souhaitons essayer et l'algorithme génère les combinaisons de paramètres. Notre méthodologie s'articule en deux étapes. Premièrement, nous utilisons la seconde option qui permet de tester une plage large de paramètres. Une fois cela effectué, nous identifions les ordres de grandeurs des paramètres pour lesquels les performances sont les meilleures, et nous définissons des combinaisons autour de ces paramètres afin d'améliorer la performance.

Nous avons choisi le RMSE comme mesure de performance à minimiser et implémentons une *10-fold Cross Validation*. Le choix de  $k=10$  permet de tester un nombre important de combinaisons tout en évitant un temps de calcul trop grand.

Rappelons succinctement les étapes précédemment évoquées avant d'étudier les résultats du *machine learning*. Dans un premier temps, nous avons généré deux bases de données en faisant varier nos variables explicatives en entrée du modèle ALM. Une base de données est utilisée pour l'apprentissage sur les BE par scénario et l'autre pour l'apprentissage sur les BE. Nous avons sélectionné les modèles à utiliser pour notre étude et optimisé leurs hyperparamètres par validation croisée. La phase d'apprentissage effectuée, nous pouvons étudier les résultats des modèles et cela est détaillé dans la partie suivante.

### 3.8 Les résultats obtenus et la sélection du modèle

Nous utilisons les métriques introduites précédemment pour comparer les performances de nos modèles et sélectionner notre meilleur modèle.

#### 3.8.1 Les résultats obtenus par l'étude sur les BE par scénario

Nous allons dans un premier temps étudier les résultats de nos modèles sur les BE par scénario. Dans le cadre de cette étude, nous présentons les résultats de 2 modèles seulement, les résultats sur les autres n'étant pas probants : un Random Forest et un XGBoost. Nous obtenons après optimisation les résultats suivants sur la base test :

TABLE 3.5: Les résultats de nos modèles sur la base test

Modèle	RMSE (en M €)	MAE (en M €)	MAPE (en %)
<b>Random Forest</b>	584,14	80,86	0,07
<b>XGBoost</b>	920,27	630,92	0,95

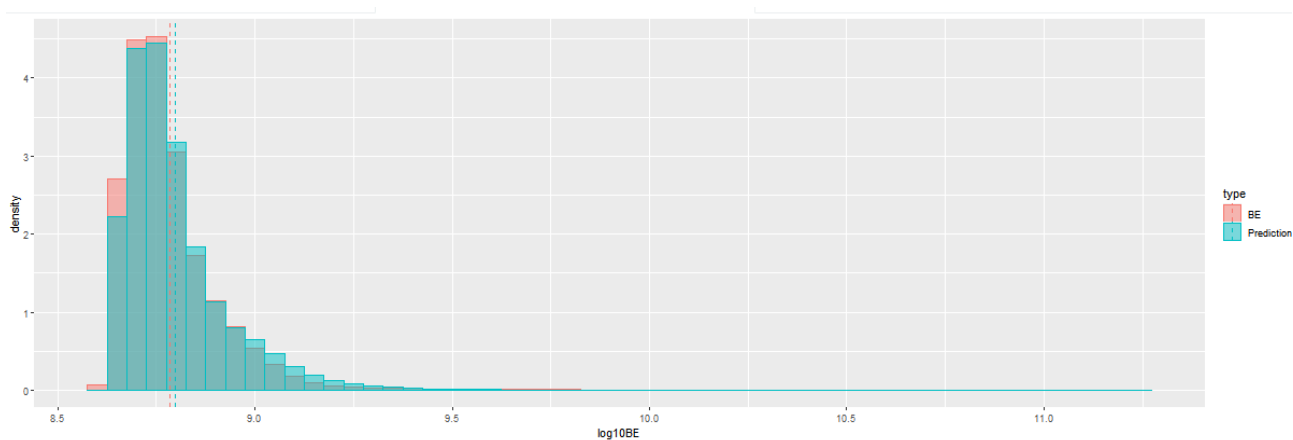
Notre analyse générale semble montrer des résultats très satisfaisants sur le Random Forest, avec une MAPE de 0.07%. RMSE et MAE sont à relativiser avec la valeur moyenne des BE par scénario de 662 M €. Alors que la MAE est plutôt satisfaisante, le RMSE est grand ce qui vient nuancer ces bons résultats. Une explication plausible est que notre modèle a du mal à prédire nos BE les plus extrêmes pouvant atteindre les 176 Md d'euros.

TABLE 3.6: Les résultats de nos modèles sur la base complète

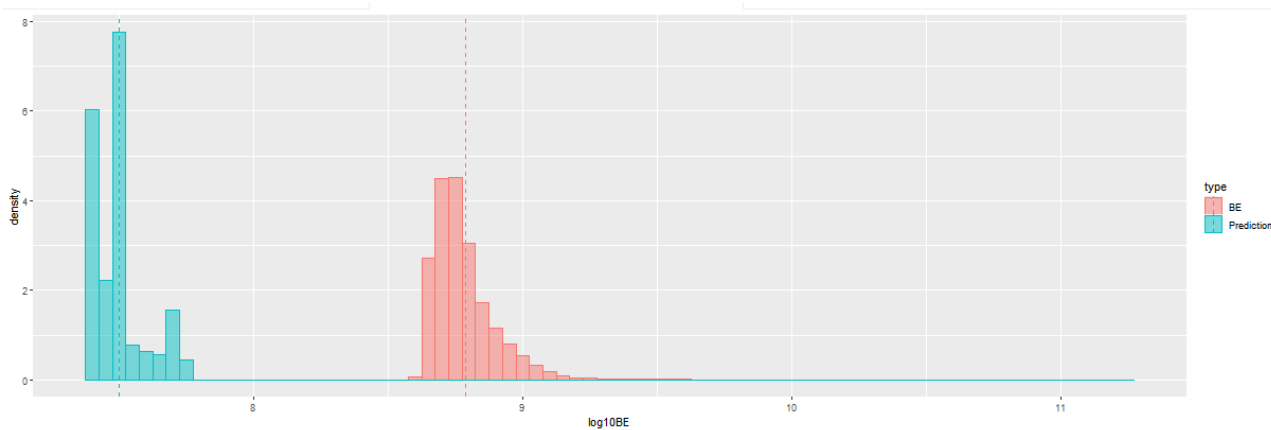
Modèle	RMSE (en M €)	MAE (en M €)	MAPE (en %)
<b>Random Forest</b>	566.82	79.82	0.07
<b>XGBoost</b>	911.25	630.46	0.95

Encore une fois, le Random Forest semble avoir mieux appris les fonctionnements de notre modèle ALM. Cependant des études complémentaires sont nécessaires pour conforter ou non les bons résultats du modèle et comprendre ce RMSE. A titre d'exemple, nous pourrions avoir un modèle qui ne prédit qu'une valeur qui serait la valeur moyenne. Il pourrait obtenir des résultats satisfaisants sur les métriques mais avec un modèle inutilisable dans le cadre de notre étude.

Nous traçons l'histogramme des BE par scénario et des prédictions sur le même graphe. Pour plus de lisibilité de nos graphiques, nous appliquons une transformation  $\log_{10}$  sur nos données et obtenons l'histogramme suivant :



(a) Histogramme du BE par scénario (en rouge) et des prédictions du Random Forest (en bleu)



(b) Histogramme du BE par scénario (en Rouge) et des prédictions du XGBoost (en bleu)

FIGURE 3.24: Histogrammes des prédictions et des BE par scénario sur la base complète

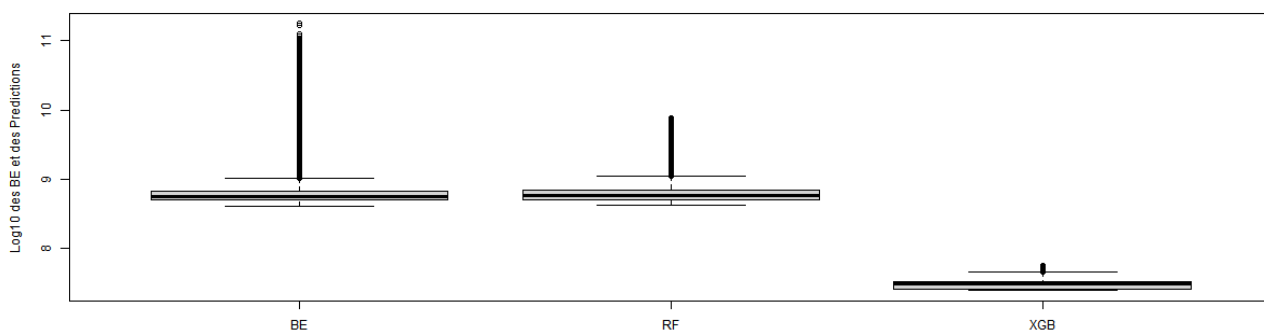


FIGURE 3.25: Boxplot des BE par scénario et des prédictions

Sur la figure 3.24a, la distribution de nos prédictions avec le Random Forest est proche de la distribution de nos BE par scénario ce qui nous conforte dans le choix de ce modèle. En revanche, la figure

3.24b indique que notre XGBoost a des prédictions largement inférieures au BE par scénario observés. Une étude complémentaire avec un nouveau calibrage des hyperparamètres permettrait potentiellement d'obtenir de meilleurs résultats, mais par manque de temps nous faisons le choix de les laisser en l'état. L'étude du boxplot confirme nos doutes liés à la grande valeur du RMSE : notre Random Forest ne parvient pas à prédire les BE par scénario extrêmes. De plus, les résultats satisfaisants sont à prendre avec précaution : nous pourrions nous retrouver dans la situation où notre modèle prédit des valeurs petites quand les BE par scénario sont grands et inversement, ce qui donnerait une comparaison d'histogrammes satisfaisante pour un modèle qui se trompe à chaque prédiction. Pour ne plus douter à ce sujet, nous traçons nos valeurs prédites en fonction du BE par scénario, ce qui devrait nous donner une tendance affine  $y = x$ . A ce stade et aux vues des résultats précédents, nous nous cantonnons à l'étude du Random Forest.

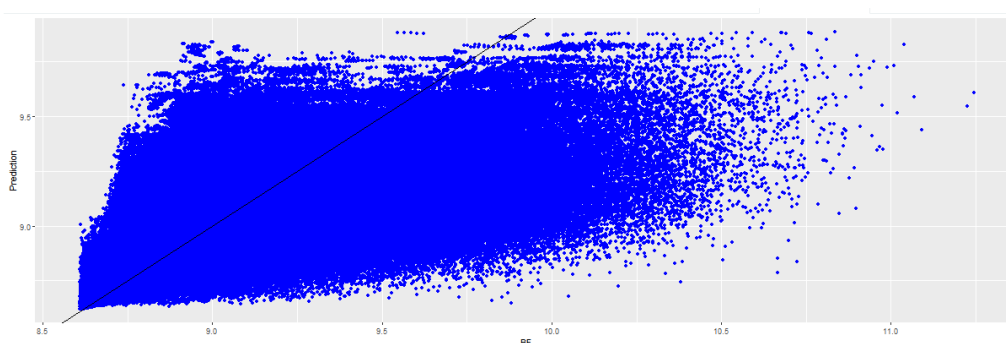


FIGURE 3.26: Graphe des prédictions en fonction des observations du BE par scénario

Le graphe 3.26 vient contredire les bons indicateurs que nous avons observés. Nos prédictions sont très dispersées et s'écartent fortement de nos BE par scénario observés et cela est vrai pour l'ensemble de nos observations et non que pour les valeurs extrêmes, comme nous le pensions.

Aux vues des résultats peu satisfaisants sur la base des BE par scénario, nous faisons le choix de nous concentrer sur l'étude des BE moyen. Une des raisons est que l'apprentissage sur cette base de données très volumineuse présente un caractère chronophage.

### 3.8.2 Les résultats obtenus pour l'étude sur les BE

Nous allons maintenant effectuer une démarche équivalente sur nos modèles implémentés sur les BE.

TABLE 3.7: Les résultats de nos modèles sur la base test

Modèle	RMSE (en M €)	MAE (en M €)	MAPE (en %)
<b>Modèle Linéaire</b>	97,24	65,08	0,1
<b>Random Forest</b>	36,82	17,34	0,02
<b>XGBoost</b>	29,41	12,03	0,01
<b>Réseau de Neurones</b>	26,05	11,52	0.013

Nous obtenons ici des résultats très satisfaisants sur l'ensemble de nos modèles, avec les meilleures performances pour le XGBoost et le Réseau de Neurones. Le RMSE est plus raisonnable, avec comme

explication probable une distribution des BE présentant moins d'*outliers* que dans le cas des BE par scénario et ainsi moins de difficulté de prédictions sur ceux-ci. Nous tachons maintenant d'implémenter les mêmes résultats sur la base complète.

TABLE 3.8: Les résultats de nos modèles sur la base complète

Modèle	RMSE (en M €)	MAE (en M €)	MAPE (en %)
<b>Modèle Linéaire</b>	93,98	65,33	0,1
<b>Random Forest</b>	31,70	16,03	0,02
<b>XGBoost</b>	24,38	10,71	0,01
<b>Réseau de Neurones</b>	25,51	10,81	0.012

Les meilleurs résultats sont obtenus avec le XGBoost et le réseau de neurones. Afin de faciliter le choix de notre modèle nous traçons nos métriques de résultat de chacun de nos modèles sur un diagramme de Kiviat comme vous pouvez le voir ci-dessous :

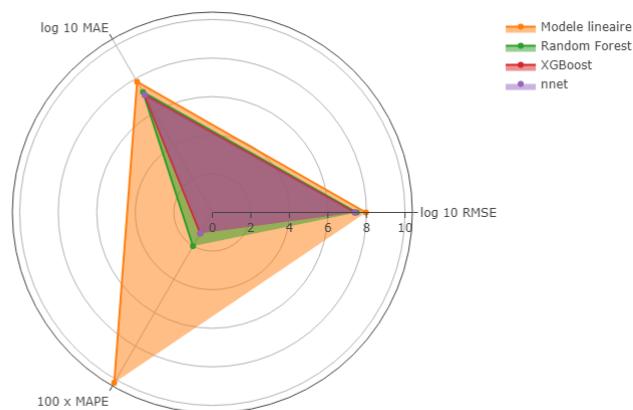
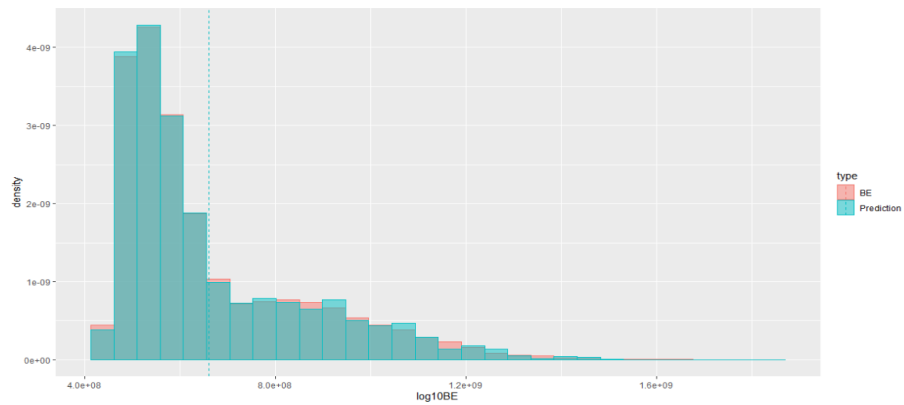


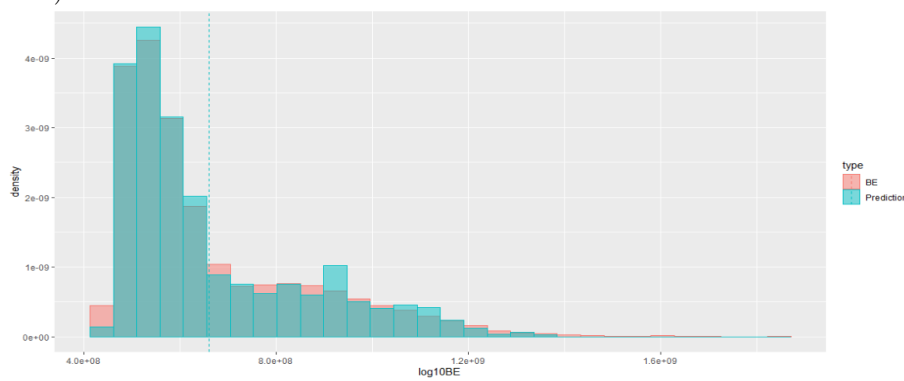
FIGURE 3.27: Diagramme de Kiviat des résultats du *machine learning*

Le modèle le plus performant minimise l'erreur et a donc l'aire la plus petite, et nous ne pouvons à ce stade pas choisir le meilleur modèle entre XGBoost et Réseau de Neurones, leurs performances étant très proches. En revanche, nous pouvons éliminer le modèle GLM de nos prétendants, aux vues de ses performances. Nous allons maintenant passer aux études complémentaires pour orienter notre choix de modèle.

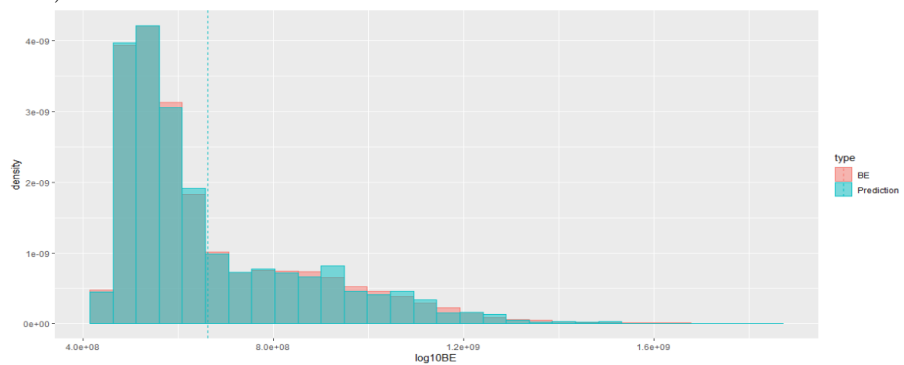




(a) Histogramme du BE (en rouge) et prédictions du Réseau de Neurons (en bleu)



(b) Histogramme du BE (en rouge) et des prédictions du Random Forest (en bleu)



(c) Histogramme du BE (en rouge) et prédictions du XGBoost (en bleu)

FIGURE 3.28: Histogrammes des prédictions et des BE sur la base complète

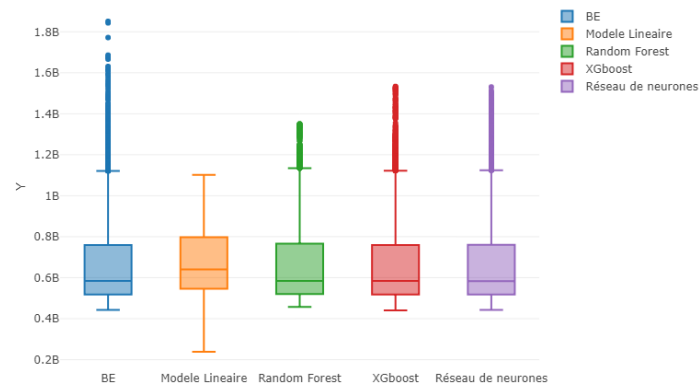
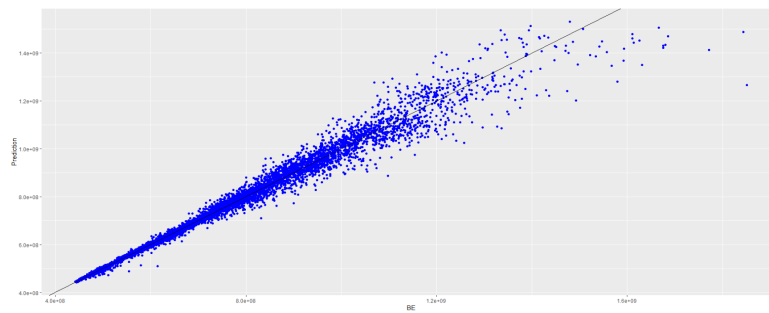
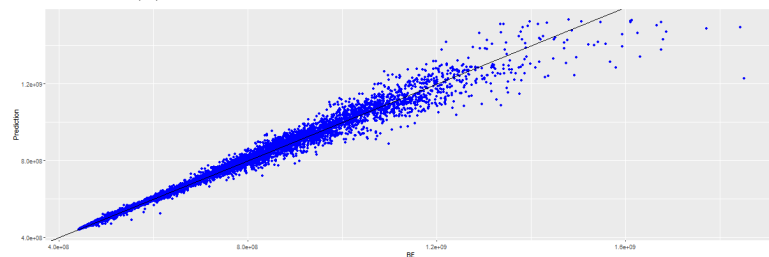


FIGURE 3.29: Boxplot des BE et des prédictions sur la base complète

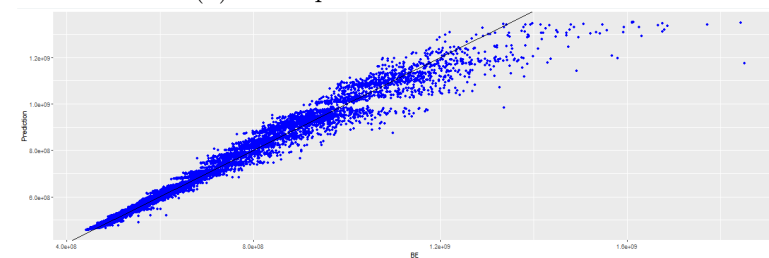
Les figures 3.28a, 3.28b et 3.28c nous montrent que les distributions des modèles sont très proches de la distribution des BE et à priori nous n'en voyons pas un sortir du lot. C'est l'étude du boxplot qui nous permet de juger les modèles XGBoost et Réseau de Neurone plus performants sur les valeurs de BE les plus extrêmes. A ce stade, le XGBoost et le Réseau de Neurones restent les meilleurs candidats et afin de clôturer notre choix de modèle, nous représentons les prédictions en fonction du BE.



(a) BE et prédictions du Réseau de Neurones



(b) BE et prédictions du XGBoost



(c) BE et prédictions du Random Forest

FIGURE 3.30: Graphe des prédictions en fonction du BE

Ces graphes nous permettent finalement d'affirmer que les modèles XGBoost et Réseau de Neurones sont plus performants que le modèle Random Forest, sans qu'un ne se distingue clairement. Nous remarquons aussi que les erreurs de prédictions augmentent avec le BE, en restant moindre comparées à nos erreurs sur l'étude des BE par scénario. Les résultats du XGBoost étant légèrement meilleurs sur la base complète (cf 3.8) nous le choisirons pour la suite de notre étude.

Nous clôturons cette étude par l'identification des variables importantes pour notre modèle retenu. À noter que chaque modèle a sa propre métrique d'importance implémentée. Nous rappelons en premier lieu que le XGBoost s'apparente à une somme d'arbres. L'importance des variables du XGBoost va donc se faire en calculant l'importance des variables pour chaque arbre, puis agréger l'information de chacun des arbres pour définir globalement le score d'importance de chaque variable. En pratique, pour chaque arbre, le MSE est calculé sur les données dites *out-of-bag* (OOB) en permutant les variables. Pour définir les données OOB, nous rappelons que les méthodes de *bagging* ou *boosting* s'appuient sur des méthodes de *bootstrap* consistant à sélectionner un échantillon parmi les données d'apprentissage. Ainsi, les données *Out-of-bag* sont comme, leur nom l'indique, les données qui n'ont pas été sélectionnées par le *bootstrap*. Ainsi, les variables qui permettent d'avoir un meilleur MSE verront leur score d'importance augmenter.

Cette méthode nous permet d'obtenir le diagramme d'importance ci-dessous :

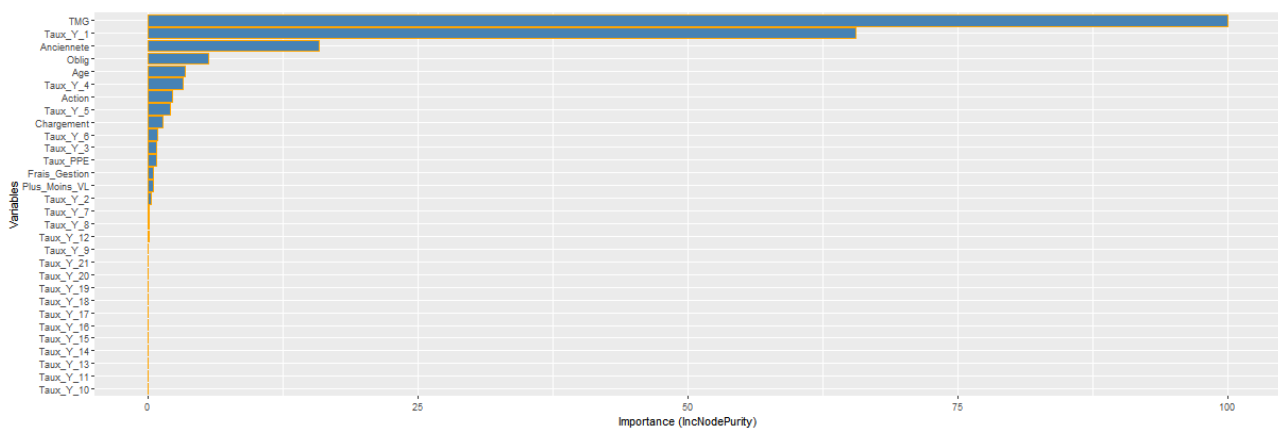


FIGURE 3.31: Classement des variables explicatives

Nous allons maintenant tâcher d'étudier ces résultats en les mettant en relation avec nos observations effectuées lors de l'étude préliminaire de nos BE et leur visualisation dans la partie 3.4. La variable ayant le plus d'impact sur nos prédictions est le TMG ce qui est en accord avec nos observations sur la figure 3.15 sur laquelle nous avons remarqué que la distribution des BE variait sensiblement en fonction du TMG et que notamment les valeurs de BE extrêmes qui surviennent pour des TMG grands. La seconde variable importante est le Taux ZC observé la première année. Les taux étant très corrélés, l'importance de cette variable est à mettre en relation avec la figure 3.18 sur laquelle nous avons jugé de l'importance de la courbe de taux sur la distribution de nos BE et du grand potentiel de celle-ci pour expliquer les variations du BE. Cette variable est déterminante au niveau du rendement de l'actif de l'assureur et a par conséquent une influence considérable sur sa stratégie. Les trois variables suivantes : "Ancienneté", "Age" et "Oblig" avaient aussi été repérées comme ayant un fort potentiel explicatif.

Le modèle XGBoost sur les BE moyen est donc le plus performant et ses résultats sont cohérents avec nos études préliminaires sur le BE, c'est donc le modèle que nous sélectionnons pour l'application à l'ORSA que nous détaillerons dans la partie suivante.



## Chapitre 4

# Interprétation du modèle à posteriori et applications à l'ORSA

Nous allons dans ce chapitre final mettre à contribution notre étude *machine learning* pour une étude de type ORSA de notre compagnie fictive introduite précédemment. Dans un premier temps, nous allons utiliser une méthode d'interprétabilité à posteriori sur notre modèle XGBoost afin de piloter la stratégie de notre compagnie et voir l'apport de telles méthodes au pilotage stratégique de la compagnie d'assurance. L'étude et l'identification des variables discriminantes au niveau du BE fait partie intégrante de la politique de gestion des risques d'une compagnie. Identifier la sensibilité du passif à certains paramètres est un enjeu majeur lors du lancement de nouveaux produits par exemple. Dans un second temps, nous projeterons nos BE selon un *business plan* de 3 ans et des données de marchés que nous spécifierons comme il est demandé dans l'ORSA.

### 4.1 Interprétation de notre modèle de *machine learning*

L'utilisation des modèles de *machine learning* reste limitée dans la projection et le calcul de variables bilantielles et réglementaires comme le SCR ou le BE. Une des raisons à cela est l'effet "boîte noire" et le manque d'interprétabilité de certains modèles de *machine learning*. Cependant, nous voulons montrer qu'en plus d'être plus rapides, des outils ont été mis au point pour améliorer l'interprétabilité et dont l'utilisation pour piloter la stratégie de la compagnie est très prometteuse. Nous jugeons ainsi pertinent de placer cette étude dans ce chapitre dédié à l'ORSA.

#### 4.1.1 Définitions et enjeux de l'interprétabilité des algorithmes

Nous commençons par définir certaines notions. Un algorithme est qualifié familièrement de "boîte noire" s'il suffit de lui donner les entrées nécessaires pour qu'il fournisse la réponse voulue sans que l'utilisateur n'ait accès au fonctionnement et aux étapes qui ont menés à cette réponse. À l'inverse, comme défini dans l'article de MILLER, 2019 nous pouvons définir l'interprétabilité comme le degré qu'un humain peut comprendre des causes d'une décision. Plus un modèle est interprétable, plus il est facile de comprendre pourquoi la décision ou la prédiction a été faite. Il n'est pas suffisant de connaître le "quoi", il faut aussi connaître le "pourquoi".

Ainsi, nous distinguons les modèles interprétables des modèles "boîtes noires". À titre d'exemple

et en référence aux modèles utilisés dans notre étude, les modèles linéaires généralisés, les arbres de décision sont des algorithmes faciles à interpréter. Dans le cas des modèles linéaires, nous savons directement comment va réagir le modèle aux variables par le biais des coefficients devant les différentes variables. Ainsi, un coefficient de grande valeur devant une variable va faire qu'une variation même infime de la variable va avoir un impact significatif sur la valeur prédite. Pour les arbres de classification, l'importance d'une variable se lit directement à son placement dans l'arbre. Un lecteur intéressé pourra se référer à l'article de BAEHRENS et al., 2010 à ce sujet. A l'inverse, les modèles comme XGBoost, les réseaux de neurones ou encore Random Forest sont qualifiés de "boîtes noires". XGBoost et Random Forest sont des modèles qui somment des arbres ce qui rend nos remarques sur les arbres de décisions non applicables.

Comme le spécifie l'article de RIBEIRO et al., 2016b, trois raisons viennent expliquer l'engouement autour des recherches sur cette notion d'interprétabilité en *machine learning*. La première est que la confiance de l'utilisateur est directement impactée par l'interprétabilité du modèle. Un utilisateur sera moins réticent à utiliser un modèle pour lequel il peut anticiper les réactions et les sorties en fonction des données qu'il reçoit. C'est ce problème de confiance qui limite l'utilisation du *machine learning* dans les cadres réglementaires comme Solvabilité 2 où l'assureur doit être en capacité de pouvoir expliquer tous les résultats fournis. La seconde raison est que la personne qui a mis au point et/ou utilise le modèle sera plus à même à le réparer ou l'auditer s'il comprend son fonctionnement en profondeur. Finalement, un modèle interprétable est plus facile à utiliser ou à vendre. Dans le cadre de notre étude, si notre compagnie projette son bilan sur 3 ans et que le modèle prédit une augmentation significative du BE la dernière année, nous devons pouvoir expliquer que c'est, par exemple, un changement du pourcentage d'obligation dans l'allocation cible de 70% à 60% qui induit plus de risque dans nos actifs et impacte à la hausse le BE.

Alors deux solutions sont envisageables. La première est de se limiter à des modèles interprétables, au détriment peut-être des performances du modèle. La seconde est celle que nous plébiscitons et est une approche qualifiée généralement de "*model-agnostic*" qui consiste à expliquer à posteriori les modèles "boîtes noires", c'est-à-dire après qu'ils aient appris sur les observations. Plusieurs arguments viennent appuyer notre choix pour cette étude. Le modèle ALM met en place des mécanismes très complexes avec des interactions multiples entre les variables. Ainsi, des modèles interprétables mais moins flexibles n'ont pas permis d'en capter toute la complexité. Des modèles plus complexes sont alors nécessaires, ce qui mène la plupart du temps à des modèles "boîtes noires". C'est ce que notre étude a d'ailleurs montré avec le XGBoost comme modèle sélectionné. De plus, les méthodes d'interprétabilité peuvent être appliquées indifféremment des modèles ce qui rend les résultats comparables tout d'abord, mais qui permet aussi à l'utilisateur de changer de modèle à sa guise.

Cependant une interprétabilité dite globale est difficile à atteindre et c'est pour cela que ces études sont le plus souvent effectuées localement, pour certaines observations qui nous intéressent. C'est une des raisons qui freine l'utilisation de ces méthodes dans le cadre du pilier I de Solvabilité 2 pour lesquelles l'assureur a besoin de pouvoir expliquer tous les résultats.

Plusieurs méthodes existent et diffèrent par leur philosophie : là où certaines méthodes vont chercher à approximer localement les données par un modèle interprétable (CRAVEN et SHAVLIK, 1996), d'autres vont choquer les entrées du modèle et voir sa réaction (RIBEIRO et al., 2016a, KRAUSE et al., 2016). Nous avons déjà vu une implémentation de la seconde méthode lors de notre classement des variables importantes figure 3.31. Pour plus de détails et d'exemples concrets, MOLNAR, 2020 liste et explique précisément les différentes approches possibles.

Nous allons maintenant appliquer une des méthodes d'interprétabilité à posteriori, que nous prendrons soin d'introduire auparavant.

#### 4.1.2 *Local Interpretable Model-agnostic Explanations* (LIME) : remplacer localement le modèle "boîte noire" par un modèle interprétable.

##### Introduction à la méthode LIME

Premièrement nous allons faire le distinguo entre les méthodes *global surrogate* et les méthodes *local surrogate* comme LIME. Les méthodes *global surrogate* vont chercher à apprendre un modèle interprétable à partir de l'ensemble des prédictions du modèle "boîte noire". La méthode LIME, introduite par RIBEIRO et al., 2016a, va localement, sur un ensemble de prédictions individuelles, entraîner un modèle interprétable. Le but est ainsi de comprendre pourquoi le modèle a fait cette prédiction en particulier.

Pour ce faire, LIME va faire varier localement les variables en entrée du modèle et évaluer l'impact sur les prédictions. LIME génère ainsi un nouveau jeu de données sur lequel il va prédire avec le modèle "boîte noire", puis va apprendre un modèle interprétable comme par exemple une régression de LASSO ou un arbre de décision. Le modèle interprétable doit être une bonne approximation locale des prédictions et non des observations. Il doit être **localement fidèle** au modèle "boîte noire". Nous résumons ci-dessous les étapes de la méthode LIME :

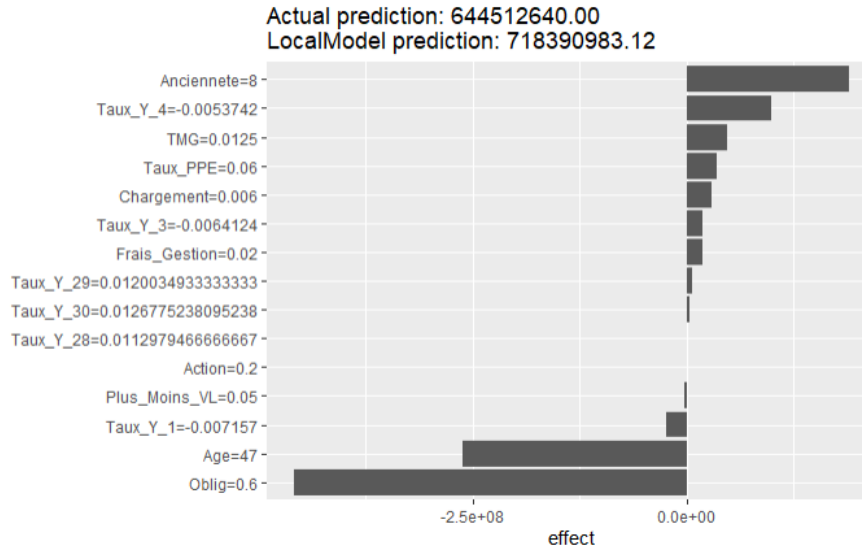
1. Nous sélectionnons les observations que nous voulons expliquer ;
2. LIME perturbe les variables ayant mené à ces prédictions pour créer un nouveau jeu de variables sur lequel il effectue les prédictions avec le modèle "boîte noire". Une nouvelle base d'apprentissage locale est ainsi créée ;
3. LIME va donner un poids aux nouvelles prédictions en fonction de leurs proximités aux prédictions initiales ;
4. LIME apprend un modèle interprétable sur la base d'apprentissage locale ;
5. LIME utilise l'interprétabilité du modèle local pour expliquer nos observations.

Cette méthode, encore au stade de développement, présente néanmoins des limites. La première principale limite est le choix du voisinage (étapes 2 et 3 ci-dessus) qui reste empirique et impact directement les explications fournies par la méthode. La seconde limite est que deux points très proches peuvent avoir des explications très différentes, sans que cela ne soit justifié. Ainsi, les résultats que nous présentons sont à analyser avec précaution et en gardant ces points en tête.

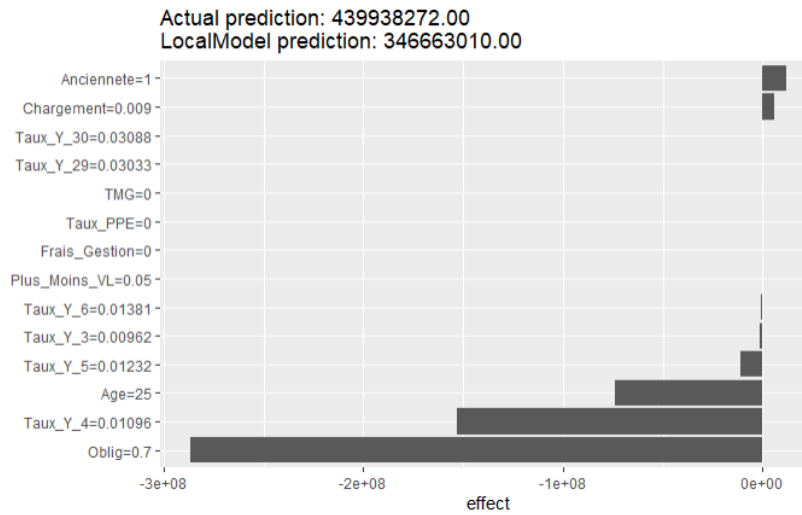
##### Application de cette méthode dans un cadre ORSA

Dans le cadre de l'ORSA, nous allons utiliser cette méthode afin d'expliquer le BE minimal et un BE de valeur centrale afin d'identifier quelles variables en sont à l'origine et en quelles proportions. L'assureur pourra ainsi adapter sa stratégie en fonction des résultats observés.

Nous n'effectuerons pas d'études sur les valeurs de BE les plus grandes car nous remarquons sur la figure 3.30b que notre modèle commet des erreurs non négligeables sur les valeurs maximales. Or, le modèle interprétable va apprendre sur les valeurs prédites par XGBoost, qui n'a pas décelé les mécanismes menant à ces valeurs extrêmes. Nous risquerions donc d'obtenir des résultats erronés avec la méthode LIME. Nous faisons donc le choix de mettre de côté les observations maximales et nous concentrer sur la valeur minimale et une valeur centrale choisie arbitrairement.



(a) Explications d'un BE de valeur centrale



(b) Explications pour le BE minimal

FIGURE 4.1: Explication de BE par la méthode LIME

Nous avons choisi d'effectuer l'étude LIME en retenant 15 variables explicatives pouvant intervenir dans l'apprentissage de notre modèle interprétable. Cela nous permet de limiter les variables intervenant, tout en ayant un modèle interprétable suffisamment proche nos prédictions "boîte noire" et mener une étude la plus juste possible. Pour atteindre 15 variables explicatives, l'algorithme met en place les méthodes "*Backward*", "*Forward*" (voire "*Stepwise*") des régressions linéaires, consistant à partir respectivement du modèle complet ou du modèle avec une seule variable explicatives, et ajouter ou enlever une à une des variables explicatives de manière à réduire l'erreur de prédiction (généralement, une minimisation du critère d'Akaike AIC est implémentée). Une régression de Lasso peut aussi être utilisée. Ainsi, dans les figures 4.1a et 4.1b, les impacts ne sont autres que les coefficients de la régression linéaire relatifs aux 15 variables expliquant le modèle.

Nous commençons par l'étude de la figure 4.1a. Les variables "Anciennete", "Taux\_Y\_4" et "TMG" sont les variables qui vont avoir un impact positif majeur sur la valeur de notre BE. A l'inverse, les variables "Age" et "Oblig" vont impacter négativement le BE. Voici l'interprétation que



nous en tirons :

- Une augmentation du TMG va venir augmenter les engagements de l'assureur envers les assurés qui, dans un environnement de taux négatifs comme celui dans lequel se trouve cette observation, sont difficiles à honorer ce qui peut entraîner des pertes et des difficultés pour atteindre la revalorisation cible ce qui peut mener à une augmentation des rachats conjoncturels et donc une augmentation des BE.
- En remarquant que l'ancienneté est de 8 ans, son augmentation fait que le MP entre dans la vague de rachat structurelle intervenant à 9 ans d'ancienneté ce qui entraîne des prestations de rachat et une augmentation du BE.
- Une augmentation de la variable "Taux\_Y\_4" qui est négative, vient diminuer le facteur d'actualisation et donc augmenter notre BE.
- L'impact négatif de la part obligation peut traduire le fait qu'un portefeuille plus prudent limite les risques liés au marché pour l'assureur. Un portefeuille d'actif comportant plus d'obligations est moins volatile et il y a ainsi moins de BE par scénario hauts. Si nous augmentons la part d'action, nous subissons les scénarios aux rendements actions défavorables qui peuvent enclencher les rachats dynamiques.
- Concernant l'impact négatif de l'âge, l'interprétation est plus subtile car nous aurions tendance à penser qu'un accroissement de l'âge augmenterait les prestations de décès et par conséquent le BE. Cependant, nous avons fait l'hypothèse d'un TMG viager, les engagements liés au TMG sont donc à honorer jusqu'à la mort de tous les assurés du MP. Ainsi, notre modèle capte le fait qu'une augmentation de l'âge rapproche du décès des assurés, mettant fin par conséquent aux engagements.

L'assureur qui se trouverait dans la situation de l'observation va pouvoir orienter son choix d'allocation d'actif en augmentant sa part d'obligation afin de baisser son BE.

Concernant la figure 4.1b, peu de variables ont un impact positif, seule la variable "Anciennete" a un impact significatif sur le BE et les variables "Age", "Taux\_Y\_4" et "Oblig" vont avoir un impact négatif. En voici notre interprétation :

- L'impact positif de l'ancienneté est dû au fait qu'une augmentation de l'ancienneté (de valeur 1 pour cette observation) fait que les vagues de rachats structurels à 5 et 9 ans d'ancienneté sont plus proches et donc le facteur d'actualisation en est diminué (car nous sommes en environnement de taux positifs) ce qui va augmenter le BE ;
- Les impacts négatifs de la part obligation et de l'âge trouve la même interprétation que précédemment ;
- L'impact négatif du "Taux\_Y\_4" va venir traduire qu'une augmentation du taux d'actualisation va ainsi diminuer le facteur d'actualisation et ainsi diminuer le BE.

Encore une fois, l'assureur peut se servir des observations sur la part d'obligation pour piloter sa stratégie d'allocation d'actif.

Cette étude met en exergue que l'analyse LIME d'une observation peut être utile pour le pilotage de la stratégie et l'anticipation des risques. Nous devons cependant l'utiliser comme un outil d'aide à la compréhension des BE et de pilotage et non comme une science exacte.

## 4.2 Application à l'ORSA : la projection de BE

Au moment de l'introduction du pilier II de solvabilité (1.1.2), nous évoquons l'aspect prospectif et pluriannuel de l'ORSA qui pousse les compagnies d'assurance à projeter leur bilan et des indicateurs sur plusieurs années, et cela à chaque changement stratégique.

Dans ce cadre, nous allons mettre à contribution notre étude *machine learning* afin de projeter le bilan économique de notre compagnie sur un horizon temporel que nous avons fixé à 3 ans. Nous nous concentrerons, comme tout au long de ce mémoire, sur le calcul du BE et tâcherons de comparer une projection de bilan en utilisant XGBoost et une projection de bilan en utilisant l'outil ALM.

Nous tâcherons ainsi de projeter le bilan selon plusieurs scénarios "Monde Réel", en variant la stratégie de notre compagnie. Cette application nous permettra aussi de tester les performances de notre modèle qui sera mis en relation avec le temps de calcul gagné.

### 4.2.1 La démarche mise en place

Nous allons maintenant spécifier la démarche que nous mettons en place, qui ressemble à une méthode des simulations dans les simulations pour le calcul du SCR, aux différences près que nous évoluons selon un seul scénario "Monde Réel" mais sur plusieurs années de projection.

Nous partons de la situation initiale de notre compagnie comme spécifié dans la section 2.6. Nous définissons ensuite un scénario d'évolution économique "Monde Réel" sur 3 ans dans lequel notre compagnie est susceptible d'évoluer. Ainsi, nous définissons une évolution des taux courts, des taux ZC, des rendements immobiliers et actions sur ces trois années. Finalement, nous modélisons de nouveaux entrants dans notre portefeuille de passif chaque année que nous appellerons *New Business* (NB) que nous supposons arriver en début d'année. Cet ensemble d'informations nous permet de faire évoluer notre compagnie selon un scénario que nous modélisons. Nous récupérerons ainsi au 31 décembre de chaque année la situation de l'entreprise : son bilan, les PM et nombre d'individus de chaque MP, l'évolution de son portefeuille obligataire... mais aussi la situation économique. Cette dernière nous permet de générer les 1000 scénarios économiques, grâce au GSE, nécessaires au calcul du BE avec l'outil ALM. Notre démarche peut ainsi être schématisée comme suit :

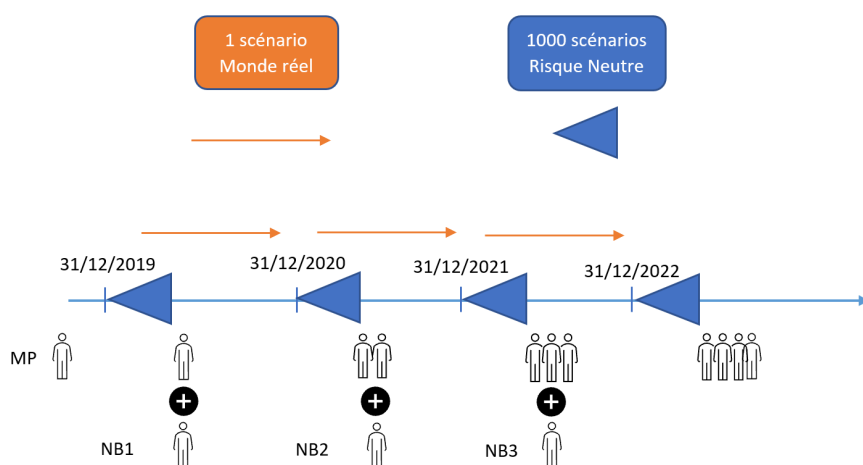


FIGURE 4.2: Démarche ORSA de projection de nos BE sur 3 ans

Etant donné que notre XGBoost est modélisé pour calculer un BE par MP, nous supposons que les BE de chaque MP s'additionnent et négligerons ainsi les interactions entre les MP et les effets de mutualisation, comme expliqué 3.2.2. De plus, pour le calculer, nous sommerons nos BE par MP, en les pondérant par leur part de PM dans la PM globale de la compagnie. Nous nommerons par la suite  $BE_{outil}$  le BE fourni par l'outil ALM et  $BE_{ML}$  le BE prédit par XGBoost.

En plus du BE, nous introduisons un autre indicateur : le ratio  $BE/(PPE+PM)$ . Ce ratio permet de comparer les provisions comptables mises en face des engagements envers les assurés avec leurs valeurs *best estimate*. La PPE est un engagement envers les assurés, comme stipulé dans le code des assurances. De plus, le modèle ALM suppose qu'à l'issue des 30 ans de projection, notre compagnie d'assurance stoppe ces activités. A ce moment-là, l'assureur liquide ses actifs et les redistribue à ses actionnaires et à ses assurés. Les actifs en face de la PPE et de la PM sont alors redistribués aux assurés et les PVL effectuées sont partagées entre les *Model Points*, au prorata de leur part dans la PM globale. Cette redistribution au terme de la projection est appelée allocation du *Terminal Surplus*.

#### 4.2.2 Première application : projection du BE de notre assureur fictif

Nous supposons ici que l'évolution des taux courts en "Monde Réel" se fait selon la courbe de taux EIOPA au 31/12/2019 et nous supposons un rendement annuel de 4% de nos actions et de notre immobilier sur les trois années de projection. De plus, notre compagnie prévoit l'entrée en portefeuille de nouveaux contrats chaque année de projection. Nous supposons que notre produit d'épargne est attractif et concurrentiel, ce qui entraîne des entrées en portefeuille en augmentation sur les 3 prochaines années. Nous aurons ainsi une entrée de NB à hauteur de 4% de l'encours du stock de 2019 pour l'année 2020, 5% pour l'année de 2021 et 6% pour l'année 2022. De plus, notre produit respecte les conditions de marché actuelles, un TMG à 0% est donc implémenté dans nos NB et nous fixons leurs âges moyens à 40 ans. Nous ne modélisons pas de primes pour les NB. Finalement, le MP en stock n'est autre que le MP moyen défini dans le tableau 2.3. Nous l'appellerons *In Force* par la suite et le noterons MP dans nos schémas et tableaux.

Nous avons ainsi introduit l'ensemble des éléments nécessaires à cette étude. La première étape va être de projeter notre bilan sur les 3 années futures puis nous récupérerons ensuite les PM par MP.

Nous détaillons année par année l'évolution de notre bilan et les calculs du BE qui en découlent.

#### 31/12/2019

Nous nous trouvons au 31/12/2019 et la situation de la compagnie est la situation précédemment décrite. Nous utilisons la courbe de taux EIOPA comme entrée de notre GSE et de notre modèle XGBoost, ce qui nous permet de calculer les  $BE_{outil}$  et le  $BE_{ML}$ .

TABLE 4.1: BE au 31/12/2019

	31/12/2019	Temps de calcul
$BE_{outil}$	612 578 993 €	14 min
$BE_{ML}$	611 512 000 €	3 s
<b>Ecart</b>	0,17%	13,95 min

Au 31/12/2019, nous implémentons une PVL de 10% pour nos actions. Notre ratio  $BE/(PPE+PM)$  s'élève à 123%. Nous expliquons le caractère élevé de ce ratio par le TMG d'1,85% servi contractuellement à l'*In Force*, qui représente une garantie de grande valeur dans un environnement de taux bas tel qu'implémenté dans cette étude.

Nous notons que l'écart entre le BE et la prédiction est légèrement plus élevé comparé à la MAPE. Cela s'explique par le fait que notre *In Force* a une combinaison de paramètres sur laquelle notre modèle n'a pas appris. Cela est un premier indicateur à garder en tête quand nous concluons sur la variance de notre modèle.

### 31/12/2020

L'outil ALM permet de projeter notre bilan en valeur nette comptable au 31/12/2020, que nous détaillons ci-dessous en prenant soin d'indiquer la part de la PM relative à l'*In Force* et au NB1.

Actif	Passif	
Obligations 399 470 118€ Actions 17 231 875€ Immo 80 413 467€ Monétaire 67 464 490€	FP 52 403 189€ PM 488 190 506€ PPE 14 280 000€ RKPI 9 706 255€	MP : 469 343 727€ NB1 : 18 846 779€

FIGURE 4.3: Bilan en VNC au 31/12/2020

La PVL s'élève alors à 14%. Nous calculons finalement nos BE comme précédemment, à l'exception que le calcul du  $BE_{ML}$  est cette fois-ci une agrégation des BE par MP.

TABLE 4.2: BE, erreurs et temps de calcul au 31/12/2020

	31/12/2020	Temps de calcul
$BE_{outil}$	611 224 806 €	11,8 min
$BE_{ML}$	605 187 848 €	7 s
<b>Ecart</b>	0,99%	11,75 min

Notre écart augmente légèrement. Cela s'explique par l'ajout d'un MP qui apporte un biais supplémentaire, mais aussi par le fait que le modèle prédise sur une courbe de taux forwardée d'un an et sur laquelle il n'a pas appris.

TABLE 4.3: PM, PPE et Ratio BE/(PPE+PM) par MP au 31/12/2020

	PM	PPE ( en % de PM)	Ratio BE/(PPE+PM)
<b>MP</b>	469 343 727 €	2,81%	124%
<b>NB1</b>	18 846 779 €	0,11%	109%
<b>Global</b>	488 190 506 €	2,92%	120%

Nous notons alors que les ratio BE/(PPE+PM) restent supérieurs à 1. Le TMG non nul pour l'*In Force* explique que son ratio soit plus élevé que celui de NB1. De plus, malgré un TMG à 0, le ratio du NB1 est différent de 1 et cela s'explique par l'effet de frais et le *Terminal Surplus*.

**31/12/2021**

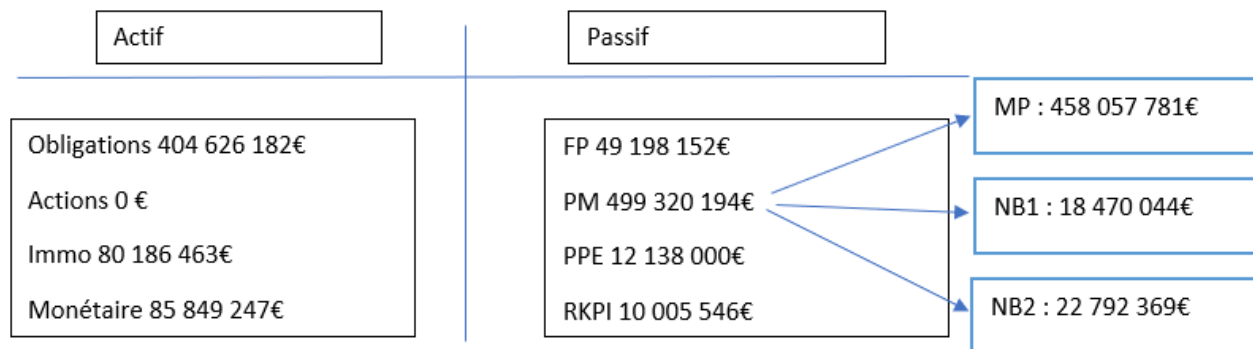


FIGURE 4.4: Bilan en VNC au 31/12/2021

Nous remarquons que nous projetons de liquider toutes nos actions au courant de l'année 2021. Notre compagnie va donc être amenée à réaliser les PVL actions afin d'atteindre soit la revalorisation cible soit la revalorisation contractuelle. L'assureur va aussi devoir retrouver son allocation cible l'année 2022.

TABLE 4.4: BE, erreurs et temps de calcul au 31/12/2021

	31/12/2021	Temps de calcul
$BE_{outil}$	613 018 382 €	12,1 min
$BE_{ML}$	609 603 012 €	10 s
<b>Ecart</b>	0,56%	11,9 min

Notre erreur ne se dégrade pas malgré l'avancée temporelle de notre projection et l'ajout d'un MP. Le modèle s'adapte donc plutôt bien aux nouvelles variables d'entrées, et notre hypothèse d'agrégation des BE ne semble pas induire d'erreurs significatives.

Nous reportons dans le tableau ci-dessous les indicateurs par MP :

TABLE 4.5: PM, PPE et Ratio BE/(PPE+PM) par MP au 31/12/2021

	PM	PPE ( en % de PM)	Ratio BE/(PPE+PM)
<b>MP</b>	458 057 781 €	2,23%	123%
<b>NB1</b>	18 470 044 €	0,09%	109%
<b>NB2</b>	22 792 369 €	0,11%	109%
<b>Global</b>	499 320 194 €	2,43%	119%

Nous remarquons alors que le ratio est égal entre NB1 et NB2. Cela s'explique par le fait que le BE prédit par XGBoost est le même pour ces 2 *model points*. La maille de variation des variables "Anciennete" et "Age" dans notre base de données n'étant pas suffisamment fine, XGBoost ne différencie pas ces 2 MP n'ayant qu'une année de différence et prédit le même BE.

### 31/12/2022

Nous finissons par projeter notre bilan en 2022, tout en simulant l'entrée en portefeuille du NB3 au début d'année.

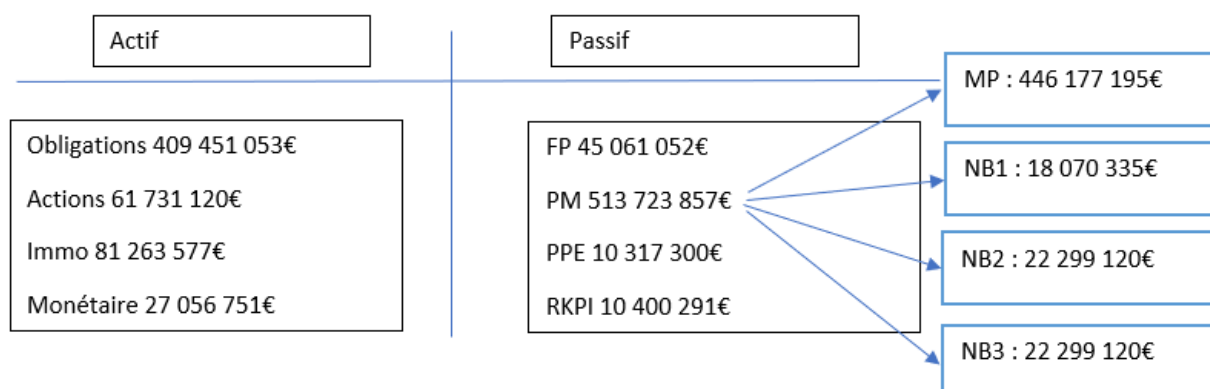


FIGURE 4.5: Bilan en VNC au 31/12/2022

Nous obtenons ainsi les résultats suivants au terme de notre projection :

TABLE 4.6: BE, erreurs et temps de calcul au 31/12/2022

	31/12/2022	Temps de calcul (min)
$BE_{outil}$	625 342 860 €	13 min
$BE_{ML}$	622 165 693 €	15 s
<b>Ecart</b>	0,51 %	12,75 min

Nous obtenons encore des résultats satisfaisants cette dernière année de projection.

Nous reportons dans le tableau ci-dessous les indicateurs par *Model Point* :

TABLE 4.7: PM, PPE et Ratio BE/(PPE+PM) par MP au 31/12/2022

	PM	PPE ( en % de PM)	Ratio BE/(PPE+PM)
<b>MP</b>	446 177 195 €	1,74%	123%
<b>NB1</b>	18 070 335 €	0,07%	107%
<b>NB2</b>	22 299 120 €	0,09%	107%
<b>NB3</b>	27 177 208 €	0,10%	107%
<b>Global</b>	513 723 857 €	2,01%	119%

### Bilan de nos projections

L'erreur moyenne entre le BE calculé par le modèle ALM et le BE prédit par XGBoost est de 0,56%. Celle-ci se dégrade peu avec l'augmentation de l'année de projection, ce qui indique une bonne adaptation du modèle. Cette erreur nous semble être encourageante dans le sens où ces prédictions sont faites sur des variables nouvelles, XGBoost n'ayant pas effectué l'apprentissage sur celles-ci. La variance de notre modèle semble plutôt faible, compte tenu des limites de notre base de données sur les BE moyens.

Faisant abstraction du biais de nos prédictions, nous voulons vérifier que la tendance des BE prédits avec XGBoost d'une année à l'autre est la même que celle des BE calculés via l'outil. Cette analyse a pour but de s'assurer que le modèle XGBoost décèle les situations engendrant une hausse de BE et inversement.

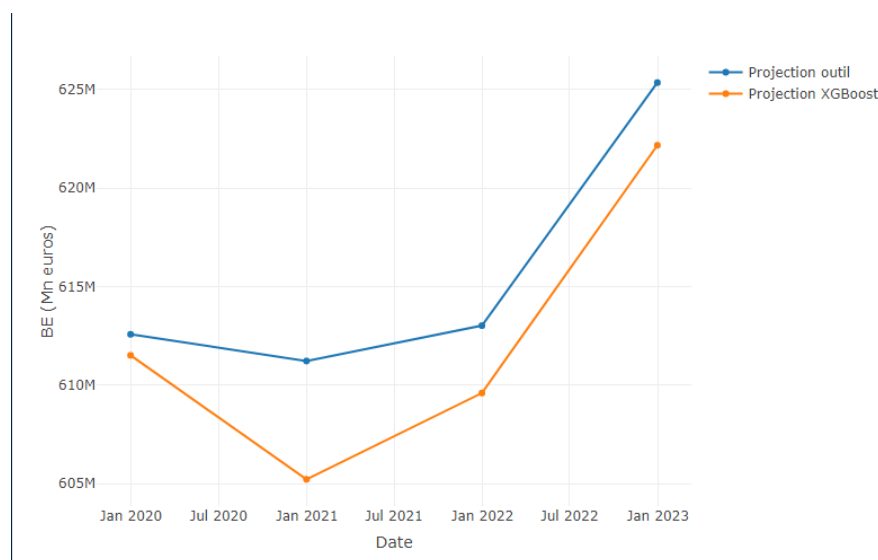


FIGURE 4.6: Comparaison des projections de BE avec l'outil et avec XGBoost

Nous observons que, malgré une tendance à la baisse plus marquée pour le BE XGBoost au 31/12/2020, les tendances sont respectées. De plus, nous notons que XGBoost a tendance à sous-évaluer le BE. Dans cet exemple, XGBoost décèle les mouvements de baisse et de hausse du BE, ce qui est un point positif. De plus la tendance à la hausse de BE s'explique essentiellement par deux raisons. La première est l'ajout de New Business chaque année, augmentant les engagements de l'assureurs et

par conséquent le BE. La seconde est le vieillissement de l'*In force*, entraînant une augmentation des prestations décès.

Nous voyons cependant tout l'intérêt de cette approche dans cet exemple : une projection sur seulement 3 années, selon un seul scénario et avec des hypothèses simplificatrices fortes sur le nombre de MP par exemple, a demandé près d'une heure de calcul en comptabilisant la génération des scénarios chaque fin d'année. En comparaison, l'approche *machine learning* nous permet de faire un travail équivalent en seulement 5 minutes ce qui est un gain considérable.

### 4.2.3 Seconde application : projection du BE dans le cadre de la crise sanitaire Covid-19

Les résultats satisfaisants de l'application précédente nous encouragent à compléter notre étude. Nous allons ainsi projeter notre BE sur 3 années (horizon d'un plan stratégique ORSA), mais dans l'environnement de crise que nous connaissons actuellement.

Pour ce faire, nous commençons par supposer que l'environnement économique au 31/12/2020 peut être approché par celui du 17/11/2020, date à laquelle nous menons l'étude. Ainsi, nous nous replaçons au 31/12/2019, et projetons notre compagnie dans un environnement monde réel. Concrètement, cela se traduit par des modifications à plusieurs niveaux. Pour commencer, notre approche concernant l'évolution des taux diffère. Contrairement à l'étude précédente, nous ne supposons plus que notre évolution des taux suit la courbe EIOPA du 31/12/2019 sur l'intégralité de la projection. A compter de 2020, nous utilisons plutôt la courbe de taux EIOPA au 17/11/2020, puis projetons nos taux de la même manière que précédemment. Nous modélisons l'évolution Monde Réel de nos taux comme suit :

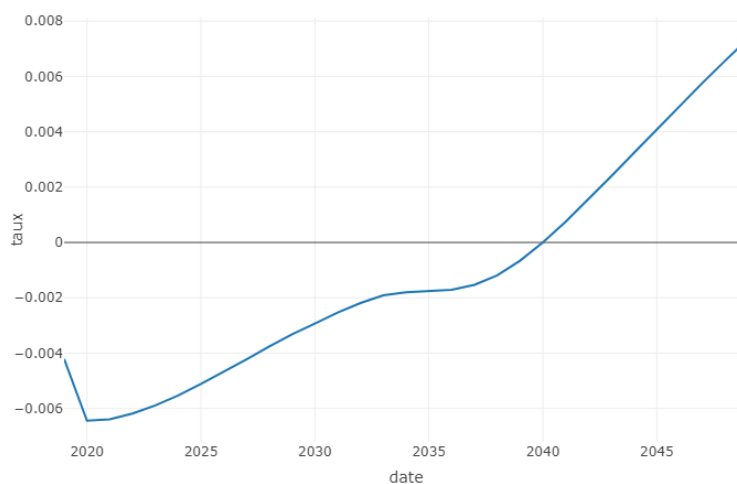


FIGURE 4.7: Evolution Monde Réel de nos taux.

Nous remarquons une forte baisse des taux entre 2019 et 2020 s'expliquant par la crise sanitaire et économique.

A ce stade, nous avons défini l'évolution Monde Réel de nos taux pour la projection. En plus de cela, nous devons modifier l'évolution de nos rendements actions entre le 31/12/2019 et le 31/12/2020 afin qu'ils soient représentatifs des conditions de marché observées actuellement. Pour ce faire, nous nous



basons sur l'évolution de l'indice CAC 40\*, en supposant notre portefeuille action composé seulement d'entreprises françaises cotées. Ainsi, nous observons sur la figure 4.8 que l'indice s'élevait à 5978 au 31/12/2019, sur la figure 4.9 qu'il a atteint 5483 au 17/12/2020 et finalement sur la figure 4.10 que la valeur minimale du CAC 40 au plus fort de la crise était à 3755. Voici le détail de ces valeurs ci-dessous :

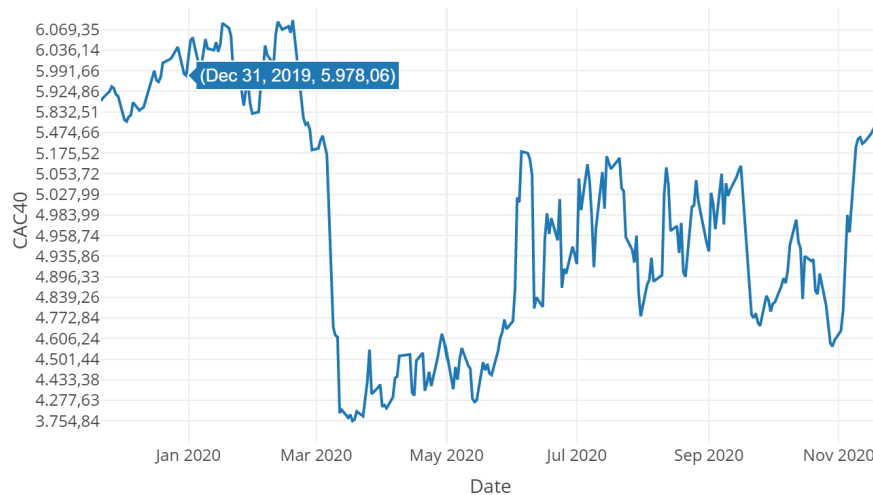


FIGURE 4.8: Valeur du CAC 40 le 31/12/2019

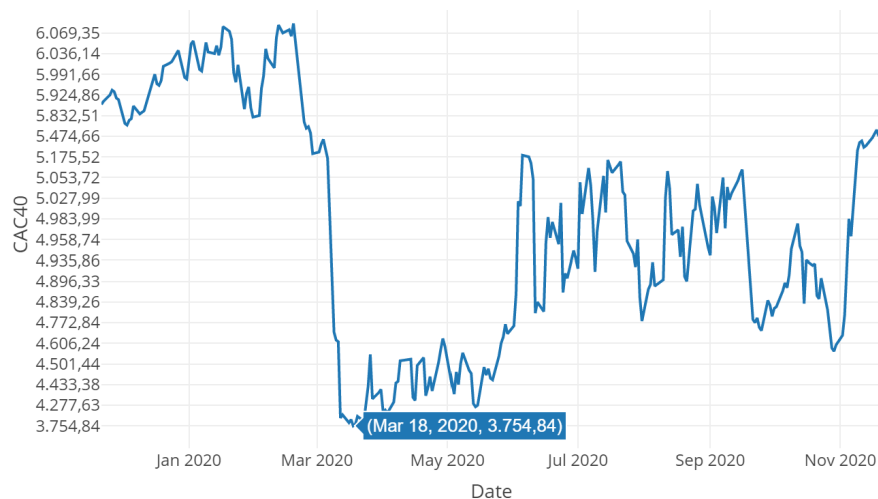


FIGURE 4.9: Valeur minimale du CAC 40 en 2020

---

\*FCHISET : CAC 40 Settlement

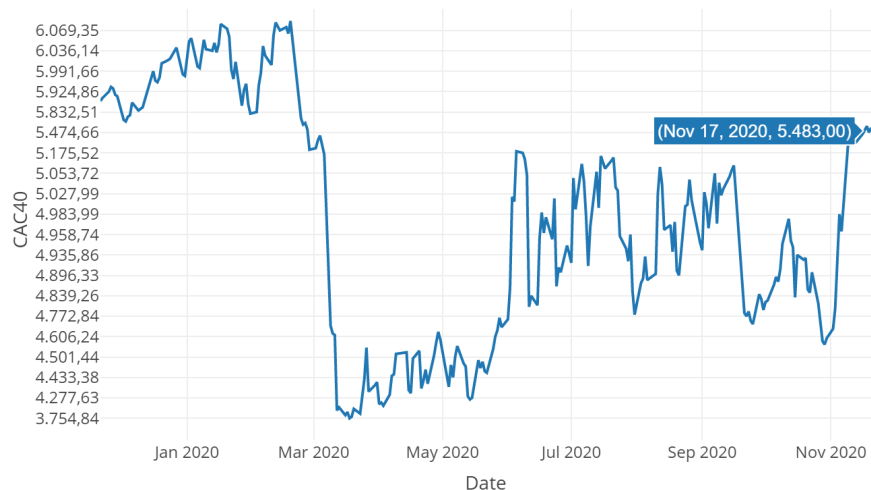


FIGURE 4.10: Valeur du CAC 40 le 17/11/2020

Ces valeurs introduites, nous allons maintenant expliquer comment nous les avons utilisées pour projeter les rendements actions de notre compagnie. Pour commencer, le rendement actions de l'année 2020 est défini comme étant la baisse de l'indice CAC 40 entre le 31/12/2019 et la date de l'étude, soit un rendement s'élevant à  $-8,5\%$ . Les autres rendements projetés nécessitent de faire des hypothèses sur l'évolution future de l'environnement économique. Nous allons ainsi implémenter 4 scénarios se distinguant par leurs formes d'évolution et leurs points de chute au 31/12/2022 (date de fin de projection) :

- Scénario d'évolution en U symétrique, c'est à dire avec un point de chute au même niveau qu'avant la crise. Un scénario en U est un scénario pour lequel l'indice CAC 40 (que nous rappelons représentatif des variations de notre rendement actions) est à la baisse la première année, ce qui est le cas pour nous en 2020, puis reste au même niveau l'année suivante et finit par remonter à son niveau du 31/12/2019 la dernière année. Nous appellerons ce scénario U1.
- Scénario d'évolution en U avec un point de chute à un niveau inférieur qu'avant la crise. Ce scénario est le même que le précédent, à l'exception près que nous supposons que l'indice du CAC 40 remonte à un niveau inférieur à celui du 31/12/2019 la dernière année. Nous appellerons ce scénario U2.
- Scénario d'évolution en V avec un point de chute au même niveau qu'avant la crise. Un scénario en V est un scénario pour lequel l'indice CAC 40 est à la baisse la première année, puis baisse encore en 2021 pour atteindre la valeur minimale atteinte pendant la crise et finit par remonter à son niveau du 31/12/2019 la dernière année. La crise économique ferait donc écho à la crise sanitaire, avec un an de décalage. Nous appellerons ce scénario V1.
- Scénario d'évolution en V avec un point de chute à un niveau inférieur qu'avant la crise. C'est un scénario semblable au précédent en supposant que l'indice CAC 40 au 31/12/2022 ne remonte qu'à la valeur du 31/12/2020. Nous appellerons ce scénario V2.

Nous résumons cela graphiquement ci-dessous :

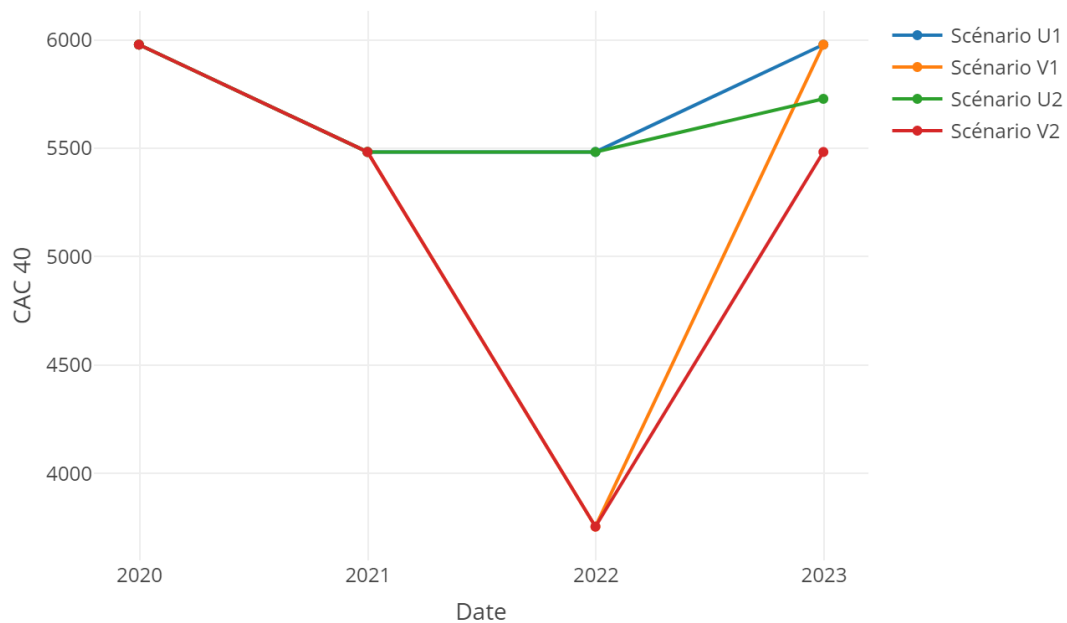


FIGURE 4.11: Les scénarios d'évolution économique envisagés

Nous partons de ces évolutions du CAC 40 pour définir nos rendements financiers. Pour définir le rendement financier de l'année  $N$ , nous faisons la différence entre la valeur de l'indice CAC 40 attendue l'année  $N+1$  et la valeur de l'année  $N$ , que nous relativisons par la valeur de l'année  $N$ . Nous obtenons finalement les rendements actions suivants :

TABLE 4.8: Rendements actions projetés pour chaque scénario

	31/12/2020	31/12/2021	31/12/2022
<b>U1</b>	-8,5%	0%	9,3 %
<b>U2</b>	-8,5%	0%	4,6%
<b>V1</b>	-8,5%	-31,4%	59,3 %
<b>V2</b>	-8,5%	-31,4%	45,8%

Finalement, nous considérons que l'effet sur l'immobilier n'est pas direct et les rendements restent inchangés.

Nous avons défini 4 scénarios Monde Réel dans lesquels nous avons fait évoluer notre compagnie d'assurance vie fictive en calculant les BE chaque année selon la même démarche décrite dans le schéma 4.2. Nous allons maintenant tâcher d'en détailler les résultats.

### Projection des BE pour le scénario U1

Nous résumons ici les résultats des projections dans le cas du scénario U1.

TABLE 4.9: BE projetés sur 3 ans pour le scénario U1

	31/12/2019	31/12/2020	31/12/2021	31/12/2022
$BE_{outil}$	612 578 993 €	664 585 597 €	688 565 988 €	714 372 424 €
$BE_{ML}$	611 512 000 €	615 830 136 €	645 273 190 €	663 728 965 €
<b>Ecart relatif</b>	0,17 %	7,34%	6,29%	7,09%

Nous remarquons que les BE sont plus élevés en comparaison à l'application précédente que nous qualifions de centrale. Cela s'explique par l'environnement de taux plus négatifs comparé à la situation centrale, venant augmenter les facteurs d'actualisation et par conséquent les BE. A cela s'ajoute l'environnement économique défavorable à l'assureur qui se trouve en moins-values sur les actions durant cette projection. Il ne peut ainsi plus effectuer de plus-values sur les actions pour servir les taux de revalorisation contractuel et cible et limiter les rachats conjoncturels par ce biais.

### Projection des BE pour le scénario U2

TABLE 4.10: BE projetés sur 3 ans pour le scénario U2

	31/12/2019	31/12/2020	31/12/2021	31/12/2022
$BE_{outil}$	612 578 993 €	664 585 597 €	688 565 988 €	714 665 887 €
$BE_{ML}$	611 512 000 €	615 830 136 €	645 273 190 €	663 732 897 €
<b>Ecart relatif</b>	0,17 %	7,34%	6,29%	7,13%

Ici, seul le rendement actions de la dernière année diffère avec le scénario U1, ce qui explique que nous trouvions les mêmes valeurs de BE jusqu'au 31/12/2021. Concernant le BE au 31/12/2022, nous notons que celui-ci est très légèrement supérieur à celui observé dans le scénario U1. Cela est cohérent car nos rendements actions sont inférieurs en 2022 dans le scénario U2, et donc moins de PVL sont disponibles pour atteindre la revalorisation cible ce qui entraîne une augmentation des prestations de rachats. Nous avons tâché d'étudier cette variation de BE la dernière année de projection. Pour ce faire, nous avons extrait les projections de PPE, de PVL et de rachats sur les 1000 scénarios économiques et nous en avons calculé la moyenne.

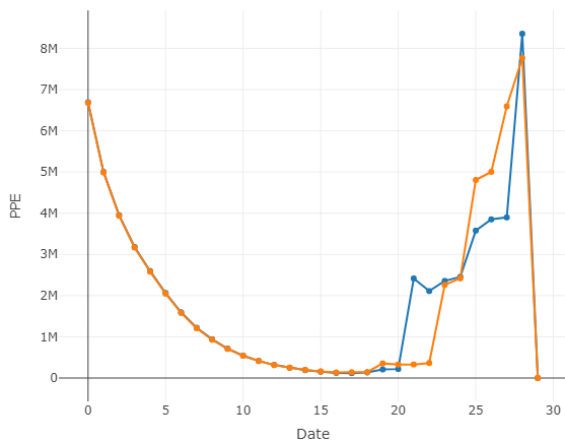


FIGURE 4.12: Evolution de la PPE moyenne sur les 1000 scénarios

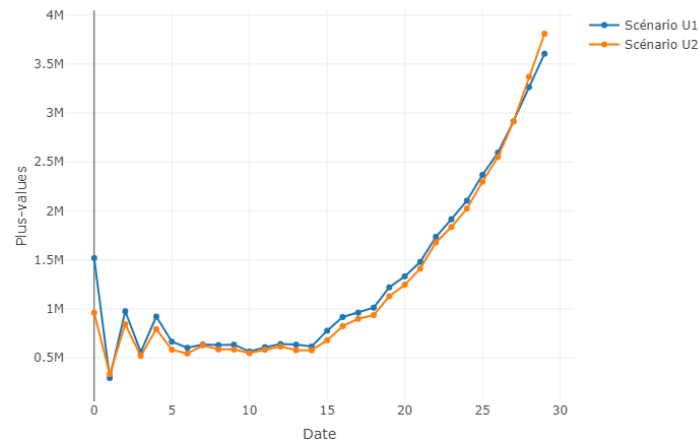


FIGURE 4.13: Evolution des plus-values actions moyennes sur les 1000 scénarios

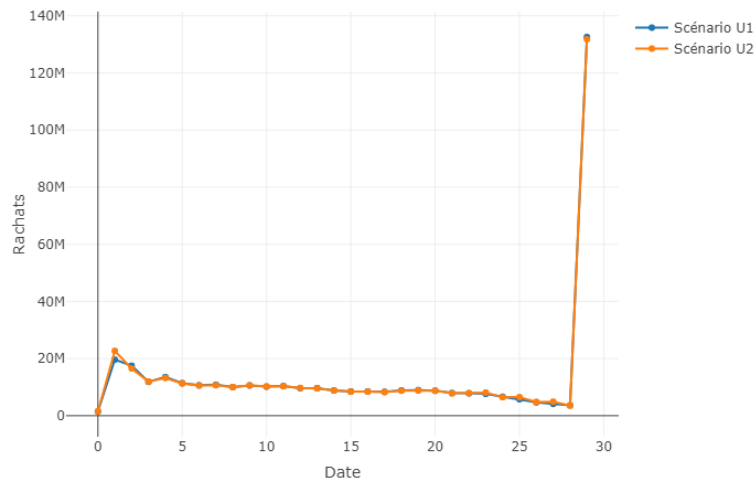


FIGURE 4.14: Evolution du montant des rachats moyens sur les 1000 scénarios

Nous observons sur la figure 4.12 que les montants moyens de PPE sont à un niveau comparable sur les deux scénarios, excepté en fin de projection où le montant moyen de PPE dans le scénario U1 augmente à partir de la 20ème année de projection, avant d'être dépassé par le scénario U2. Sur la figure 4.13, nous observons que les plus-values des actions au 31/12/2022 du scénario U1 s'élèvent à 1,5 M€ contre 0,9 M€ pour U2. Cela est expliqué par le rendement actions plus élevé au 31/12/2022 dans le scénario U1. Bien que les niveaux de plus-values soient proches, ils sont légèrement plus élevés dans le cas du scénario U1 sur l'ensemble de la projection. Nous observons une chute du niveau de plus-values action la première année de projection, ainsi que la 3ème et la 5ème. Finalement, la figure 4.14 nous montre que les prestations de rachat sont très proches. A noter que le pic de rachat en fin de projection vient retranscrire le *terminal surplus*. Les niveaux des prestations de rachat sont différents la première année de projection, pour le scénario U1 elles sont légèrement supérieures en comparaison avec le scénario U2 (19,5 M€ contre 22,5 M€).

Ces observations nous permettent de mieux comprendre cette légère variation de BE. L'effondrement des plus-values action et la baisse de la réserve de PPE nous font supposer que les plus-values ont

été liquidées afin d'atteindre le taux de revalorisation cible. En effet, comme nous l'avons étudié sur le schéma 2.6, une reprise de PPE et de plus-values peut être effectuée afin d'atteindre la revalorisation cible. Cela étant dit, dans le cas du scénario U1, le niveau de plus-values est plus élevé, l'assureur a donc plus de ressources pour essayer d'atteindre la revalorisation cible. Ainsi, cela explique que nous ayons des prestations de rachats moins grandes sur le scénario U1, l'assureur ayant pu limiter les rachats dynamiques plus efficacement que dans le cas du scénario U2. Ainsi, nous supposons que ce sont les prestations de rachats, supérieures dans le cas du scénario U2, qui expliquent que la valeur du BE de U2 soit légèrement supérieure.

### Projection des BE pour le scénario V1

TABLE 4.11: BE projetés sur 3 ans pour le scénario V1

	31/12/2019	31/12/2020	31/12/2021	31/12/2022
$BE_{outil}$	612 578 993 €	664 585 597 €	712 699 047 €	707 743 272 €
$BE_{ML}$	611 512 000 €	615 830 136 €	660 794 896 €	674 301 670 €
<b>Ecart relatif</b>	0,17 %	7,34%	7,28%	4,73%

Ici l'augmentation des BE est plus significative que les scénarios en U. Les explications sont les mêmes, mais avec des moins-values plus grandes. La hausse significative des rendements en fin de projection vient inverser la tendance des BE qui diminuent. Nous remarquons cependant que cette variation de tendance n'est pas captée par XGBoost.

### Projection des BE pour le scénario V2

TABLE 4.12: BE projetés sur 3 ans pour le scénario V2

	31/12/2019	31/12/2020	31/12/2021	31/12/2022
$BE_{outil}$	612 578 993 €	664 585 597 €	712 699 047 €	710 829 395 €
$BE_{ML}$	611 512 000 €	615 830 136 €	660 794 896 €	663 820 828 €
<b>Ecart relatif</b>	0,17 %	7,34%	7,28%	6,61%

Le BE en fin de projection est légèrement supérieur à celui du scénario V1, et cela est expliqué par un rendement moins important des actions en 2022. Nous allons encore une fois tâcher d'expliquer cette différence de BE.

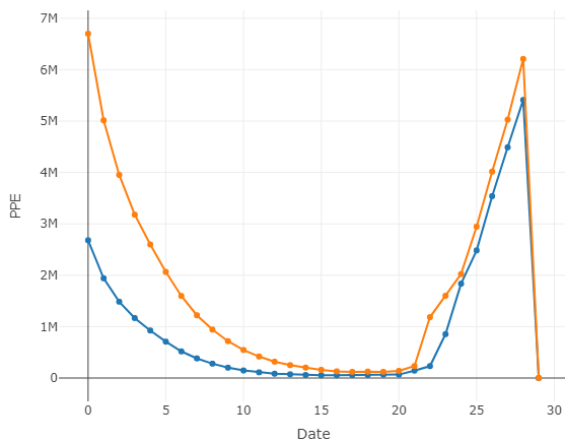


FIGURE 4.15: Evolution de la PPE moyenne sur les 1000 scénarios

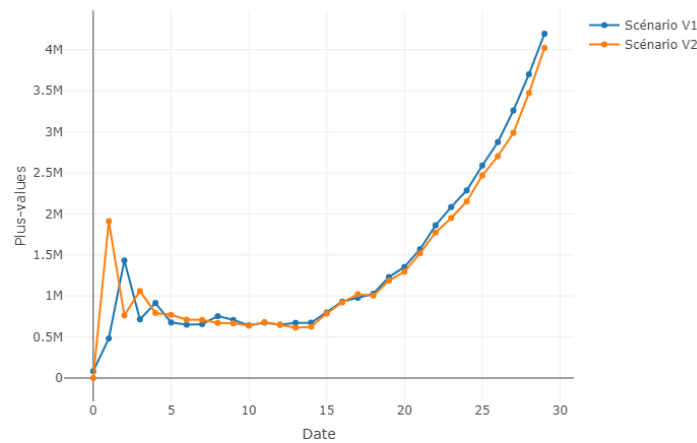


FIGURE 4.16: Evolution des plus-values actions moyennes sur les 1000 scénarios

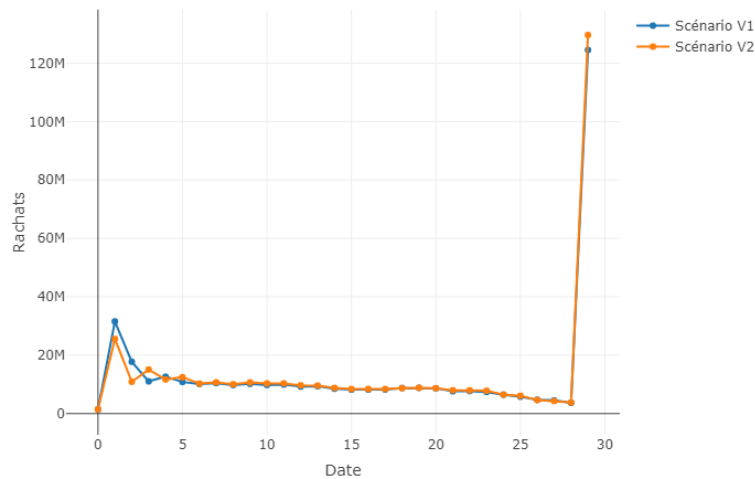


FIGURE 4.17: Evolution du montant des rachats moyens sur les 1000 scénarios

Nous observons sur la figure 4.15 que le montant moyen de PPE est à un niveau plus élevé pour le scénario V2. Nous remarquons aussi une baisse importante du niveau de PPE sur les 15 premières années de projection, avec un niveau de PPE proche de 0 à compter de la 15ème année de projection. Sur la figure 4.16, nous observons que les plus-values des actions au 31/12/2022 sont proches de 0, malgré une forte augmentation des rendements actions cette année-là. Nous supposons que les actions se trouvaient en moins-values les années précédentes, étant donné l'environnement action très défavorable. Bien que les niveaux de plus-values soient proches, ils sont légèrement plus élevés dans le cas du scénario V2 sur la première année de projection, avant de chuter et retrouver le niveau de V1. Finalement, la figure 4.17 nous montre que les prestations de rachat sont très proches. Les niveaux de prestations de rachat sont différents de la première à la cinquième année de projection, pour le scénario V1 elles sont légèrement supérieures comparé au scénario V2 (31,6 M€ pour V1 contre 25,6 M€ pour V2 la première année de projection).

Ces observations nous permettent à nouveau de mieux comprendre cette légère variation de BE. L'effondrement de la réserve de PPE nous fait supposer que la PPE est liquidée afin d'atteindre le

taux de revalorisation cible dans les 2 scénarios. De plus, les plus-values amassées la première année de projection sont directement liquidées la seconde année sur le scénario V2. Ainsi, dans le cas du scénario V2, le niveau de PPE et de plus-values étant plus élevé, l'assureur a donc plus de ressources pour essayer d'atteindre la revalorisation cible. Cela explique que nous ayons des prestations de rachats moins grandes sur le scénario V2 les première et seconde années de projection, l'assureur ayant pu limiter les rachats dynamiques plus efficacement que dans le cas du scénario V1. Ainsi, nous supposons que ce sont les prestations de rachats, supérieures dans le cas du scénario V2, qui expliquent que la valeur du BE de V1 soit supérieure.

Les erreurs de prédiction à compter de 2020 sont plus grandes que lors de notre première application. Les erreurs moyennes pour U1, U2, V1 et V2 sont respectivement de 5,22%, 5,23%, 4,88% et 5,35%, soit près de 10 fois supérieures à celle que nous observions la première application.

Nous allons tâcher d'expliquer l'augmentation de ces erreurs. Premièrement, à partir du 31/12/2020, c'est une nouvelle courbe de taux sur laquelle XGBoost effectue des prédictions, qui plus est introduisant un environnement de taux très bas. Or jusqu'à maintenant nous n'avions pas utilisé de courbe totalement nouvelle, nous avons simplement déformé la courbe au 31/12/2019. Ainsi, la variance du modèle explique en partie l'erreur croissante. En lien avec cela, nous introduisons des scénarios de crise très adverses, qui amènent à des variations extrêmes de nos variables que peine à capter notre modèle XGBoost. En effet, non sans liens avec la remarque précédente sur la courbe de taux, nos projections du bilan montrent notamment des moins-values conséquentes non prises en compte dans nos plages de variation de la variable "Plus\_Moins\_VL". En plus de cela, nous prédisons fréquemment des BE identiques entre les *New Business* d'une même année (voire d'une année à l'autre dans le cas d'évolution des paramètres proches) et cela s'explique encore une fois par le manque de variabilité dans notre base.

#### 4.2.4 Limites de notre étude et pistes d'amélioration

Bien que nos applications soient encourageantes, avec un gain de temps considérable sur les projections effectuées pour des erreurs moyennes minimales sur le scénario central, notre démarche reste à ce stade pédagogique et peut être améliorée. En effet, nous avons rencontré des contraintes logistiques de temps et coût de calcul qui nous ont amené à faire des hypothèses lors de la création de notre base de données réduisant les champs d'application de notre approche, notamment sur les scénarios plus adverses.

Au niveau du passif, la principale limite vient du fait de ne considérer qu'un unique MP et qu'un produit d'épargne en euros. Nous jugeons cependant ce choix justifié car il nous permet de faciliter les analyses de sensibilités liées aux paramètres du contrat et d'avoir une étude plus complète de leurs impacts sur le BE, bien qu'il nous éloigne de la situation réelle d'un assureur.

Cependant, nous pensons que le plus problématique dans la construction de la base de données est d'avoir peu de sensibilité de nos variables, plus spécialement à l'allocation d'actifs et aux paramétrages du GSE. En effet, dans le cadre de l'assurance vie, une des sources principales de variations du BE vient de l'environnement économique et de la capacité de l'assureur à tirer profit de son actif. Concernant l'environnement économique, les scénarios économiques en sortie du GSE sont très corrélés malgré les perturbations de la courbe de taux EIOPA en input comme nous avons pu le voir sur la figure 3.12. Une piste d'amélioration consisterait donc à compléter la base de données en jouant sur le paramétrage du GSE, apportant ainsi plus de variabilité dans l'environnement économique. Nous améliorerions ainsi nos erreurs de projection, avec un modèle qui s'adapterait certainement mieux aux déformations économiques. De plus, notre allocation d'actif reste très simplifiée, nous pourrions notamment considérer un *rating* moyen des obligations moins élevé afin de nous rapprocher des conditions réelles du marché.



Ainsi, compléter notre base de données améliorerait avec certitude les performances de notre modèle sur la projection selon des scénarios plus adverses. Une approche alternative serait d'effectuer un recalibrage de notre modèle chaque année de projection, nous aurions ainsi un modèle adapté à la situation de l'assureur et à la situation économique. Cependant, il faudrait prendre en compte le temps de recalibrage, qui pourrait réduire le gain en temps de calcul qui motive cette approche.



# Conclusion

Ce mémoire se place dans le cadre d'une étude de type ORSA d'un assureur-vie souhaitant pouvoir projeter à tout moment son *Best Estimate* sur plusieurs années, sans utiliser le modèle ALM trop gourmand en temps et puissance de calcul. A cette fin, nous avons sélectionné des variables explicatives avec lesquelles nous avons généré deux bases d'apprentissage : une pour le BE par scénario et une pour le BE.

Nous avons ensuite testé plusieurs algorithmes, dont les hyperparamètres avaient été préalablement optimisés par validation croisée afin d'obtenir les meilleures performances possibles. Alors que notre approche sur les BE par scénario n'a pas donné de résultats satisfaisants, notre modèle XGBoost sur les BE montre des performances remarquables avec une MAPE (*Mean Absolute Percentage Error*) de 0,01%. Il a ainsi été sélectionné pour mener à bien notre étude ORSA.

Dans le cadre de l'étude de type ORSA, notre première application a été d'utiliser la méthode LIME (*Local Interpretable Model-agnostic Explanations*) permettant d'interpréter a posteriori nos prédictions de BE. Nous avons montré qu'un tel outil présentait des intérêts pour permettre à l'assureur de piloter sa stratégie et analyser ses risques dans des scénarios donnés.

Nous avons ensuite montré que notre modèle XGBoost répondait à notre problématique en permettant de réduire à 5 minutes le temps de projections de notre BE sur 3 ans, contre une heure nécessaire dans le cas de l'outil ALM. De plus, nous jugeons son erreur moyenne de 0,56% encourageante, l'étude ayant été menée sur une combinaison de variables sur laquelle XGBoost n'a pas effectué l'apprentissage. Le modèle XGBoost a réussi à capter les variations de tendance du BE sur ce scénario central.

Finalement, nous avons effectué la même projection sur des scénarios stressés simulant les effets de la crise sanitaire de la Covid-19 et la crise économique qui en découle. Cette étude montre des résultats plus mitigés avec une erreur moyenne 10 fois supérieure à celle observée sur le scénario central. Néanmoins, cela nous a permis de juger des limites de notre approche.

Ces limites résident dans les hypothèses effectuées et notre base de données d'apprentissage que nous jugeons incomplète. Le fait de ne considérer qu'un *Model Point* au moment de l'apprentissage et de considérer que le BE puisse être approximer en agrégeant des BE par *Model Point* est une hypothèse forte qui n'est pas envisageable pour un assureur. De plus, des contraintes opérationnelles et temporelles ont limité notre choix de plage de variations de nos variables. Nous pensons notamment qu'il serait intéressant de faire varier les paramètres de notre GSE ce qui permettrait de limiter la corrélation détectée entre nos scénarios projetés et apporter plus de variabilité à notre base d'apprentissage. Les limites de variabilité dans notre base de données est la cause principale de l'augmentation de nos erreurs dans l'application à l'ORSA. Finalement, nous ne projetons que le BE, ce qui limite l'exhaustivité de notre étude dans le cadre de l'ORSA. Des travaux équivalents sur le SCR seraient donc à envisager, encouragés par les résultats satisfaisants sur le BE.

Nous jugeons tout de même que l'intégration d'approches *machine learning* sont à envisager par

les assureurs dans leurs études ORSA. Bien que notre approche soit simplifiée, nous avons montré avec les méthodes d'interprétabilité à posteriori de type LIME que le *machine learning* peut être digne de confiance et met à disposition des outils d'analyse puissants, qui, nous en sommes persuadés, peuvent trouver leurs places dans une stratégie de pilotage de BE d'une compagnie d'assurance vie. De plus, les assureurs ont les moyens financiers et logistiques permettant de dépasser les contraintes qui ont limité notre étude.

Afin d'encourager la poursuite de cette étude nous proposons des pistes qui, selon nous, méritent d'être explorées. Notre étude pourrait être poursuivie en approfondissant l'apprentissage sur les BE par scénario avec une approche *Deep Learning*. Nous pensons notamment aux réseaux de neurones récurrents LSTM, dont les résultats sur les séries temporelles sont très prometteurs. De plus, en gardant une approche sur les BE, il pourrait être intéressant de considérer un modèle par année de projection, qui prendrait en considérations l'information des années précédentes. Finalement, nous pensons qu'il serait très intéressant de combiner nos deux applications ORSA : nous pourrions projeter nos BE sur 3 années en effectuant chaque année une étude de type LIME permettant d'orienter la stratégie de l'assureur. Cela lui permettrait de comprendre encore mieux l'évolution de son BE et de faire les choix qui lui paraissent les plus pertinents.

# Bibliographie

- BAEHRENS, D., SCHROETER, T., HARMELING, S., KAWANABE, M., HANSEN, K. et MÜLLER, K.-R. (2010). How to explain individual classification decisions. *The Journal of Machine Learning Research* 11, p. 1803-1831.
- BREIMAN, L. (1996). Bagging predictors. *Machine learning* 24.2, p. 123-140.
- BREIMAN, L. (2001). Random forests. *Machine learning* 45.1, p. 5-32.
- BREIMAN, L., FRIEDMAN, J., STONE, C. J. et OLSHEN, R. A. (1984). Classification and regression trees. CRC press.
- CRAVEN, M. et SHAVLIK, J. W. (1996). Extracting tree-structured representations of trained networks. *Advances in neural information processing systems*, p. 24-30.
- CYBENKO, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems* 2.4, p. 303-314.
- DECUPÈRE, S. (2011). Agrégation des risques et allocation de capital sous Solvabilité II. Mémoire d'actuaire. ENSAE.
- DEVINEAU, L. et LOISEL, S. (2009). Construction d'un algorithme d'accélération de la méthode des «simulations dans les simulations» pour le calcul du capital économique Solvabilité II. *Bulletin Français d'Actuariat* 10.17, p. 188-221.
- FREUND, Y., SCHAPIRE, R. E. et al. (1996). Experiments with a new boosting algorithm. *icml*. T. 96. Citeseer, p. 148-156.
- GAUVILLE, R. (2017). Projection du ratio de solvabilité : des méthodes de machine learning pour contourner les contraintes opérationnelles de la méthode des Sds. Mémoire d'actuaire. EURIA.
- HAGUET, E. (2013). Mise en place d'indicateurs de suivi du risque dans un cadre d'ORSA Epargne. Mémoire d'actuaire. ENSAE.
- HORNIK, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural networks* 4.2, p. 251-257.
- KRAUSE, J., PERER, A. et NG, K. (2016). Interacting with predictions: Visual inspection of black-box machine learning models. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, p. 5686-5697.
- LE, X.-H., HO, H. V., LEE, G. et JUNG, S. (2019). Application of long short-term memory (LSTM) neural network for flood forecasting. *Water* 11.7, p. 1387.
- MATTEI, J. (2019). Apports de la data science dans l'analyse des résultats d'un modèle de gestion actif-passif en assurance-vie. Mémoire d'actuaire. ENSAE.
- MILLER, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence* 267, p. 1-38.
- MOLNAR, C. (2020). Interpretable Machine Learning. Lulu. com.
- RIBEIRO, M. T., SINGH, S. et GUESTRIN, C. (2016a). " Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, p. 1135-1144.

- RIBEIRO, M. T., SINGH, S. et GUESTRIN, C. (2016b). Model-agnostic interpretability of machine learning. *arXiv preprint arXiv:1606.05386*.
- STATHAKIS, D (2009). How many hidden layers and nodes? *International Journal of Remote Sensing* 30.8, p. 2133-2147.
- SYLVAIN DETROUILLEAU, S. M. (2013). Modèle ALM : Apport de la Logique Floue dans la modélisation des comportements. Mémoire d'actuaire. ENSAE.



## Annexe A

## Annexes

### A.1 Chocs sur les taux S2

Échéance (en années)	Diminution
1	75 %
2	65 %
3	56 %
4	50 %
5	46 %
6	42 %
7	39 %
8	36 %
9	33 %
10	31 %
11	30 %
12	29 %
13	28 %
14	28 %
15	27 %
16	28 %
17	28 %
18	28 %
19	29 %
20	29 %
90	20 %

FIGURE A.1: Chocs S2 à la baisse de la courbe de taux ( $ChocDown_t$ )



Échéance (en années)	Augmentation
1	70 %
2	70 %
3	64 %
4	59 %
5	55 %
6	52 %
7	49 %
8	47 %
9	44 %
10	42 %
11	39 %
12	37 %
13	35 %
14	34 %
15	33 %
16	31 %
17	30 %
18	29 %
19	27 %
20	26 %
90	20 %

FIGURE A.2: Chocs S2 à la hausse de la courbe de taux ( $ChocUp_t$ )

Les autres valeurs sont obtenues par interpolation linéaire.

## A.2 Matrices de corrélations

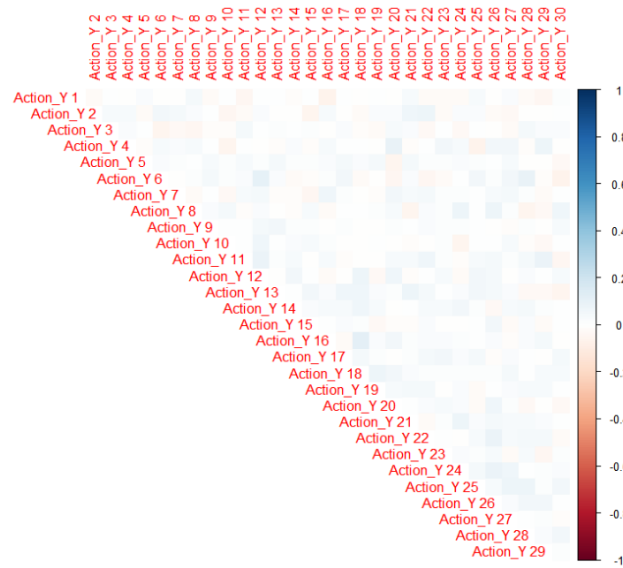


FIGURE A.3: Matrice des corrélations des rendements actions

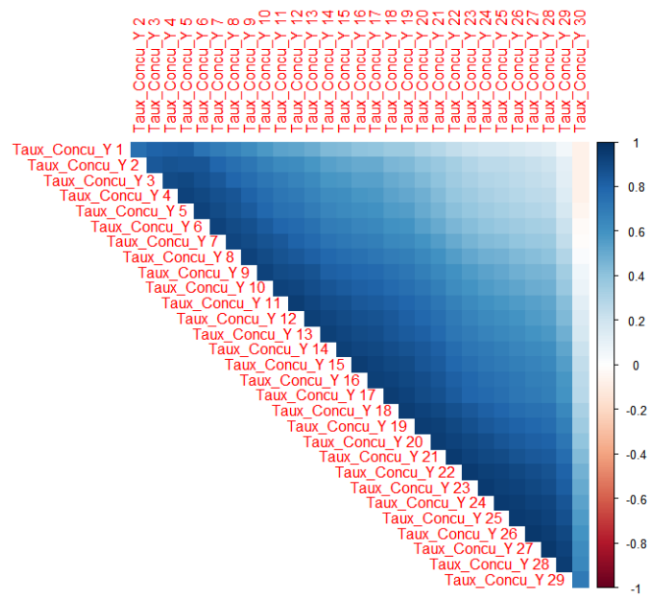


FIGURE A.4: Matrice des corrélations du taux concurrent

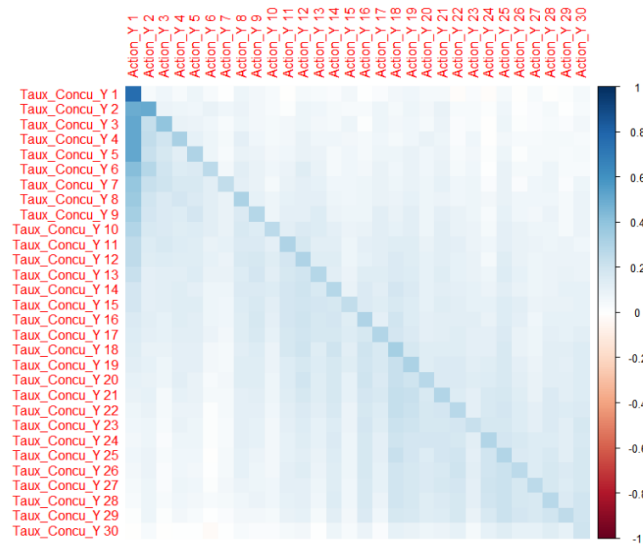


FIGURE A.5: Matrice des corrélations des rendements actions et des taux concurrents

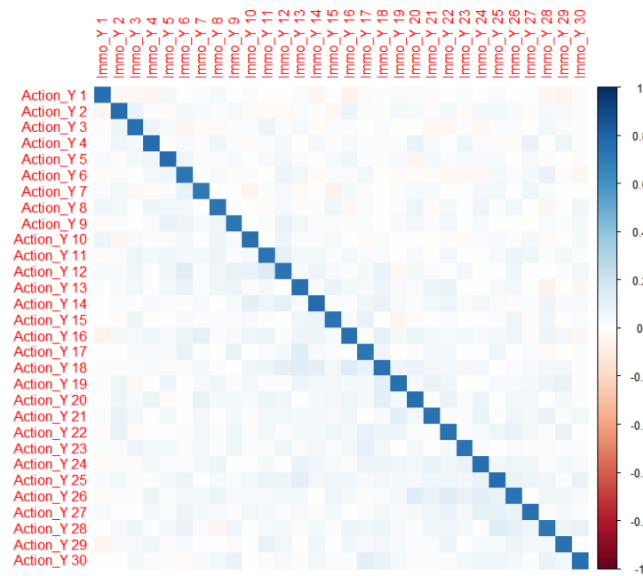


FIGURE A.6: Matrice des corrélations des rendements actions et immobiliers

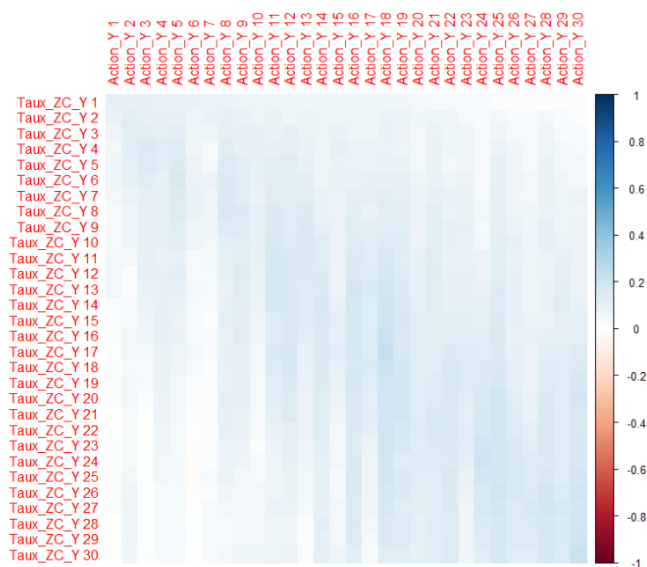


FIGURE A.7: Matrice des corrélations des rendements actions et des taux ZC

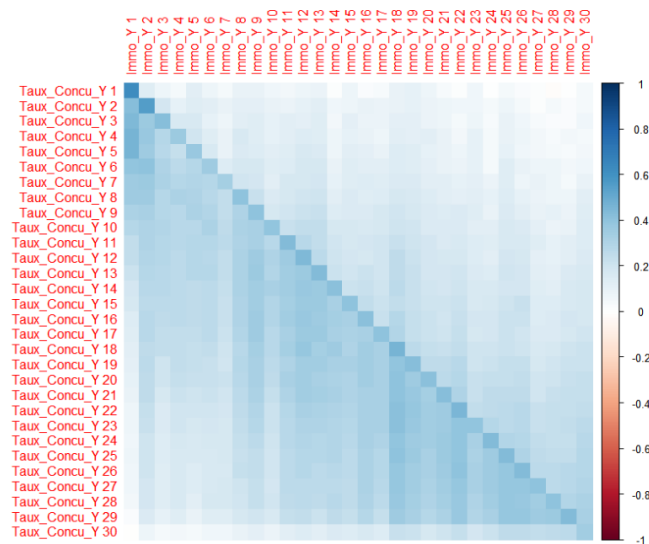


FIGURE A.8: Matrice des corrélations des rendements immobiliers et des taux concurrent

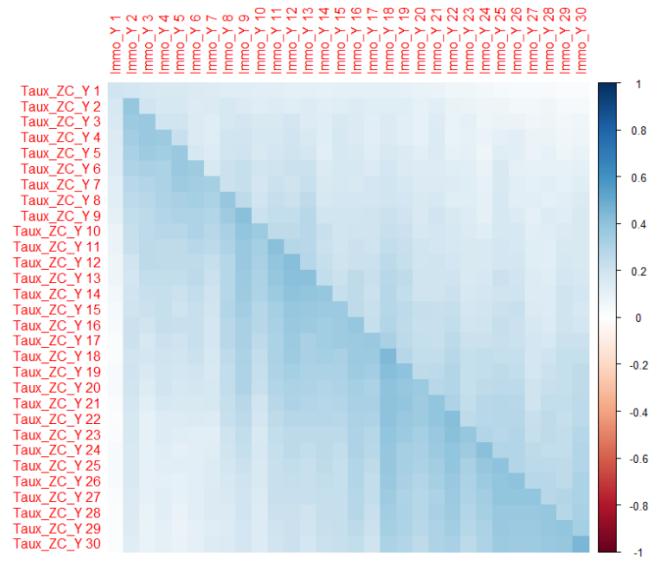


FIGURE A.9: Matrice des corrélations des rendements immobiliers et des taux ZC

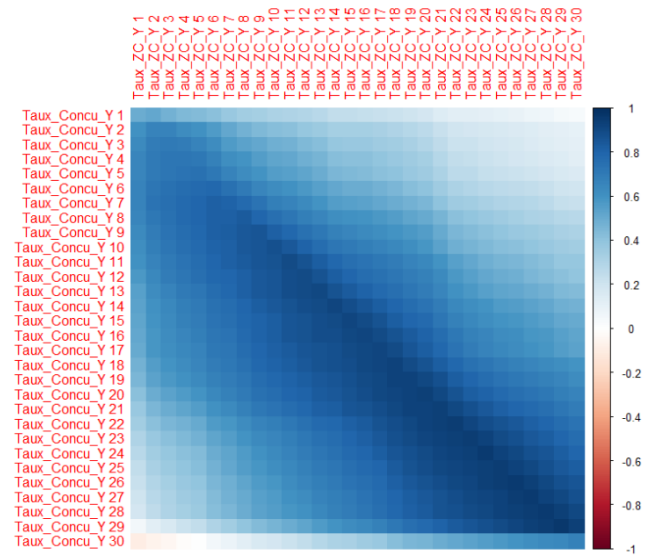


FIGURE A.10: Matrice des corrélations des taux concurrents et des taux ZC

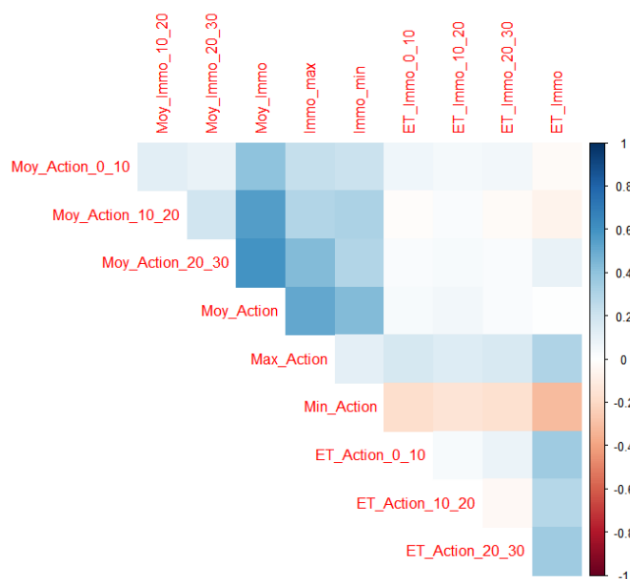


FIGURE A.11: Matrice des corrélations des variables de complétion Action

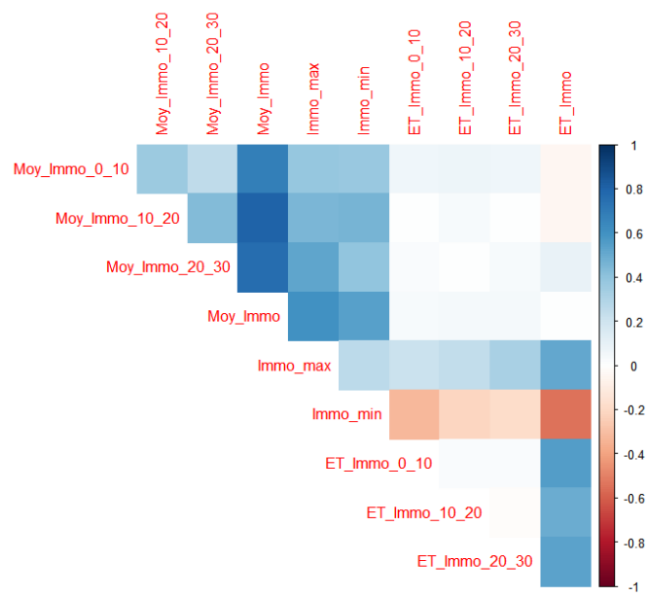


FIGURE A.12: Matrice des corrélations des variables de complétion Immobilier

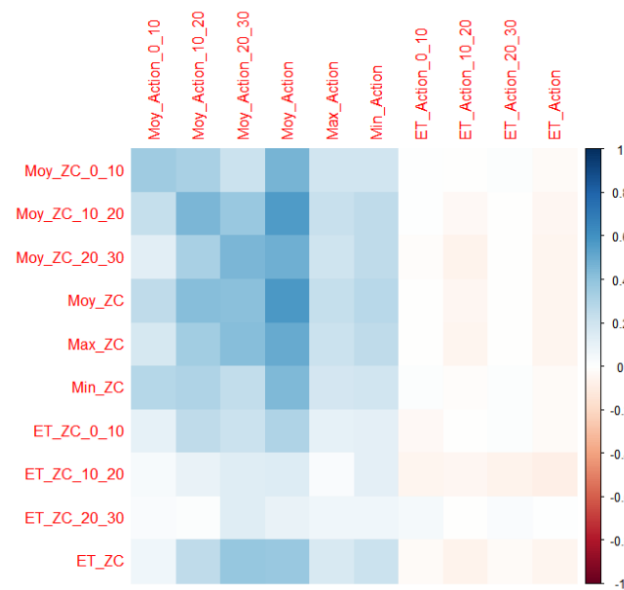


FIGURE A.13: Matrice des corrélations des variables de complétion Action et ZC

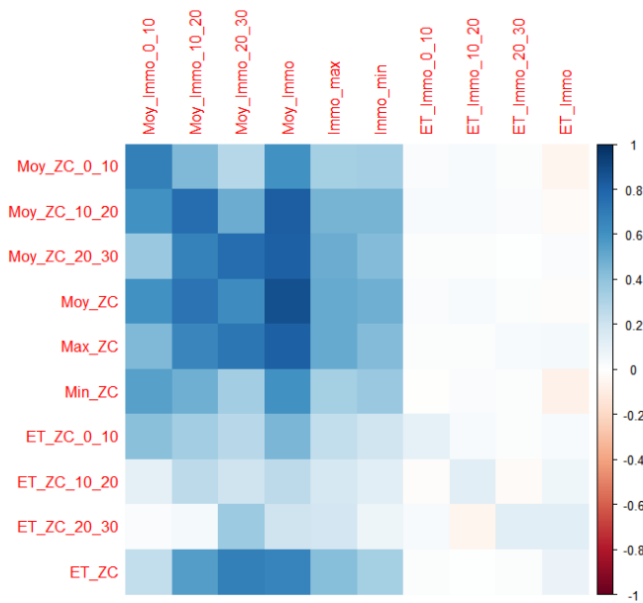


FIGURE A.14: Matrice des corrélations des variables de complétion Immobilier et ZC