





**Mémoire présenté le :
pour l'obtention du diplôme
de Statisticien Mention Actuariat
et l'admission à l'Institut des Actuaraires**

Par : Axel Elbaz	
Titre du mémoire : Prédiction du SCR marché par machine learning et application aux tests de sensibilité	
Confidentialité : <input checked="" type="checkbox"/> NON <input type="checkbox"/> OUI (Durée : <input type="checkbox"/> 1 an <input type="checkbox"/> 2 ans)	
Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus.	
<u>Membres présents du jury de la filière :</u>	Signature : <u>Entreprise : Deloitte Conseil</u> Nom : Cyril Chalin Signature : CYRIL CHALIN 
	<u>Directeur de mémoire en entreprise</u>
<u>Membres présents du jury de l'Institut des Actuaraires :</u>	Signature : Nom : Guillaume Limbert Signature : 
	<u>Invité :</u> Nom : 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) <u>Signature du responsable entreprise :</u> CYRIL CHALIN  <u>Signature du candidat :</u> 

REMERCIEMENTS

Mes remerciements vont en premier lieu vers Cyril Chalin, Claude Chassain et Baptiste Brechot, associés en charge de l'équipe Actuariat, qui m'ont permis de rejoindre leur équipe.

Un grand merci à Guillaume Limbert, mon tuteur d'alternance, et à Aurélien Suray et Jeffrey Calvez, qui m'ont accompagné tout au long de ce mémoire et m'ont accordé un soutien immense lorsque les obstacles apparaissaient sur la voie.

Merci à l'ensemble de l'équipe Actuariat, qui ont montré un intérêt constant pour ce projet et qui se sont rendus disponibles pour me faire bénéficier de leur expertise.

Et enfin, je remercie ma famille qui a fait l'effort de comprendre ce sujet et qui m'a permis, en l'expliquant, de mieux le comprendre moi-même.

TABLE DES MATIÈRES

REMERCIEMENTS	2
TABLE DES MATIÈRES	3
RÉSUMÉ	5
ABSTRACT	6
NOTE DE SYNTHÈSE	7
EXECUTIVE SUMMARY.....	12
INTRODUCTION	17
I/ L'ASSURANCE VIE EN FRANCE.....	18
1. Contrat d'épargne en euros	18
2. Contrat d'épargne en UC.....	19
3. Autres versions du contrat d'épargne : multisupport et Eurocroissance	19
4. Les avantages fiscaux	19
II/ CONTEXTE RÉGLEMENTAIRE.....	21
1. L'introduction de Solvabilité I puis II	21
2. Les trois piliers.....	22
a. Pilier 1 : Exigences quantitatives	22
b. Pilier 2 : Exigences qualitatives.....	22
c. Pilier 3 : <i>Reporting</i> et information du public.....	23
3. Focus sur le pilier 1 : Le bilan prudentiel.....	23
a. Le Best Estimate	23
b. La Risk Margin	24
c. Les fonds propres	24
d. SCR et MCR.....	25
e. Retour sur le SCR Marché.....	29
III/ PRINCIPES DE MODÉLISATION ALM.....	35
1. Evaluation des actifs.....	35
2. Evaluation des passifs.....	35
3. Le générateur de scénarios économiques.....	36
4. Hypothèses assurés et assureur	36
5. Fonctionnement général du modèle.....	39
a. Fonctionnement de l'algorithme de réinvestissement	40
b. Méthode de <i>flexing</i>	40
c. Une limite à considérer : la non-modélisation des fonds propres	42
IV/ ÉLABORATION DE LA MÉTHODE ET MISE EN PLACE DE L'OUTIL.....	43

1. Rappel de l'objectif.....	43
2. Génération des données	43
a. Variabilisation de l'action et de l'immobilier	44
b. Variabilisation des obligations.....	44
c. Variabilisation des taux	49
3. Calibration du modèle de Machine Learning	50
a. Préparation des données (normalisation, sélection de variables).....	50
b. Choix du modèle.....	53
c. Tuning des hyperparamètres	54
4. Etudes des résultats	55
5. Explicabilité du modèle	60
V/ ÉTUDE DE CAS : SCR ET SENSIBILITÉS	63
1. Rôle et déroulement d'un test de sensibilité	63
2. Présentation du portefeuille fictif.....	64
3. Prédiction de SCR dans différents scénarios.....	66
a. Sélection des scénarios pertinents pour l'étude de cas.....	66
b. Résultats	67
4. Limites et prochaines étapes.....	75
CONCLUSION	76
BIBLIOGRAPHIE.....	77
ANNEXES.....	78
Annexe I – Table des figures.....	78
Annexe II – Présentation du Gradient Boosting.....	80
Annexe III – Résultats des prédictions par sous-module et par test de sensibilité pour le BEL et la PVFP.....	81

RÉSUMÉ

Les assureurs ont un besoin constant d'évaluer leur sensibilité au risque. Plus précisément, dans le cas du risque de marché, il est essentiel de comprendre comment se comporterait son portefeuille d'actifs et comment serait impacté son rendement en cas de changement de l'environnement économique. L'interaction entre l'actif et le passif inhérente aux contrats d'assurance vie implique que ces impacts potentiels sur le portefeuille d'actifs entraînent des répercussions directes sur le taux servi et le comportement des assurés, et donc sur le bilan et la solvabilité de l'assureur.

Ce dernier a donc intérêt à explorer un large éventail de sensibilités afin d'avoir la connaissance la plus exhaustive possible de ses expositions les plus critiques. Cependant, mener un grand nombre de ces tests de sensibilité peut s'avérer informatiquement coûteux et humainement laborieux.

L'objectif de ce mémoire est de créer un outil capable de prédire avec une précision satisfaisante le SCR marché d'un assureur vie dans différents contextes économiques. Un tel outil permettrait à l'assureur d'explorer quasi-instantanément un ensemble complet de scénarios, pouvant croiser des chocs affectant simultanément différentes parts du portefeuille. Pour y parvenir, nous nous servons du modèle ALM pour générer une base de données puis d'un modèle d'apprentissage supervisé pour pouvoir prédire le SCR marché à partir des variables économiques retenues.

Ce mémoire rappelle d'abord le cadre réglementaire et les principes de fonctionnement des outils employés. Il s'attache ensuite à présenter la méthode mise en place, les obstacles rencontrés et les solutions trouvées afin de générer une base de données pertinente vis-à-vis de l'objectif fixé puis de paramétrer un algorithme de machine learning sur celle-ci. Enfin, notre processus est testé au cours d'une étude de cas impliquant un assureur vie fictif souhaitant explorer la sensibilité de son SCR à divers chocs de marché.

ABSTRACT

Insurers have a constant need to assess their risk sensitivity. Specifically, in the case of market risk, it is essential to understand how their portfolio of assets would behave and how its performance would be impacted by changes in the economic environment. The interaction between assets and liabilities inherent in life insurance contracts implies that these potential impacts on the asset portfolio have direct repercussions on the distributed rate, the behavior of policyholders, and therefore on the insurer's balance sheet and solvency.

Therefore, it is in the insurer's interest to explore a wide range of sensitivities to have a comprehensive understanding of its most critical exposures. However, conducting a large number of these sensitivity tests can be computationally expensive and laborious.

The aim of this paper is to create a tool capable of predicting an insurer's market Solvency Capital Requirement (SCR) with satisfactory accuracy in different economic contexts. Such a tool would allow the insurer to quasi-instantaneously explore a complete set of scenarios, potentially crossing shocks affecting different parts of the portfolio simultaneously. To achieve this, we will use the Asset and Liability Management model to generate a database and then utilize a supervised learning model to predict the market SCR based on selected economic variables.

This paper first describes the regulatory framework and operational principles of the tools used. It then focuses on presenting the method implemented, the obstacles encountered, and the solutions found to generate a relevant database to reach our objective and to fine-tune a machine learning algorithm on it. Finally, our process is tested through a case study involving a fictional life insurer who wants to explore the sensitivity of its SCR to various market shocks.

NOTE DE SYNTHÈSE

Dans ce mémoire, nous cherchons à créer un outil capable de prédire via *machine learning*, pour une compagnie d'assurance à un instant donné, le montant du SCR marché en fonction du portefeuille d'actifs et de l'environnement économique.

Le SCR (*Solvency Capital Requirement*), défini dans la norme Solvabilité II, peut se calculer via une formule standard qui correspond pour les modules considérés en une agrégation de différences entre la NAV (*Net Asset Value*) de l'assureur en conditions normales et la NAV après application de chocs prédéfinis. Ces chocs s'appliquent directement sur les valeurs d'entrée du modèle ALM et c'est à ce titre qu'ils impactent le montant de NAV en sortie.

La méthode employée pour prédire le SCR marché consiste dans un premier temps à prédire les NAV en fonction des variables affectées par les chocs réglementaires, c'est-à-dire les valeurs de marché des actifs, les scénarios économiques considérés et le spread du portefeuille obligataire. En y parvenant avec une performance suffisante, nous pouvons prédire presque instantanément les NAV centrales et choquées, y compris pour des tests de sensibilité faisant varier l'état initial de l'actif.

En parallèle, nous prédirons également le BEL (*Best Estimate of Liabilities*) afin d'obtenir un bilan complet.

Le processus mis en place pour y parvenir se décompose en quatre phases :

- 1) La variabilisation des conditions économiques en entrée du modèle ALM ;
- 2) La production d'une base de données contenant une information synthétique et nécessaire pour pouvoir expliquer les variations de la NAV et du BEL calculées en sortie du modèle ALM à partir des données d'entrée ;
- 3) La calibration d'un modèle de *machine learning* sur la base des données générées pour prédire NAV et BEL, ainsi que la mise en place de méthode d'interprétabilité ;
- 4) La création des *inputs* adaptés pour prédire les NAV centrales et choquées pour chaque sous-module de SCR marché et pour chaque test de sensibilité considéré, puis l'analyse des résultats.

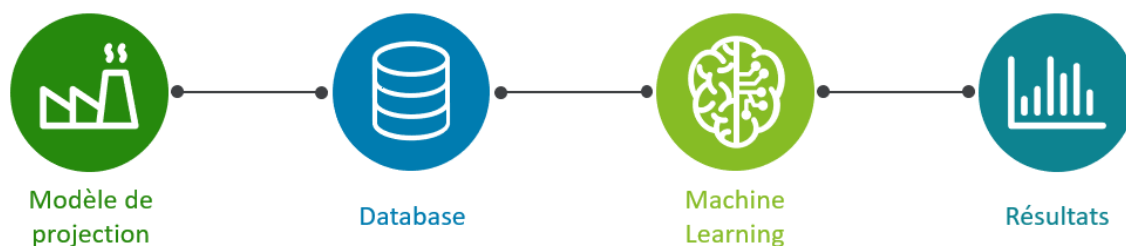


Schéma récapitulatif du processus général

1. Variabilisation des données d'entrée au sein du modèle de projection

Le premier objectif est de construire une base de données. Celle-ci doit contenir un grand nombre de combinaisons de variables décrivant l'actif et les BEL et NAV associés. Ces combinaisons doivent être assez diverses et exhaustives pour s'assurer que les tests de sensibilités et l'application des chocs réglementaires sur l'actif aboutiront à des *inputs* cohérents avec notre base d'apprentissage.

Deux problématiques se dégagent à ce niveau : quels *inputs* choisit-t-on de perturber pour générer cette base et comment les perturbe-t-on, pour chaque simulation.

Le choix des éléments variabilisés à l'actif est justifié par l'objectif de pouvoir prédire les NAV après chocs de SCR marché pour les modules de taux, d'action, d'immobilier et de spread. Le choc de taux implique une modification de la courbe des taux employée pour la génération des scénarios économiques, les chocs action et immobilier correspondent à une baisse de valeur de marché des actifs concernés et le choc de spread correspond à une baisse de la valeur de marché des obligations en conséquence à l'augmentation du spread associé.

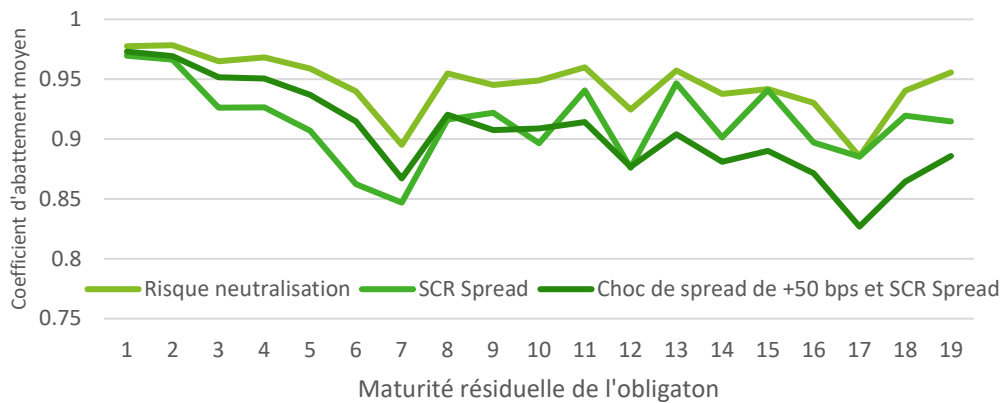
Des trois premiers chocs, il est déduit immédiatement que **les variables incluront les taux *spots* et *forward* et les valeurs de marché des actions et de l'immobilier.** Pour le choc de spread, le choix a été effectué en fonction de la méthode choisie pour risque-neutraliser les obligations. En effet, des abattements sont appliqués aux coupons et nominaux afin de faire correspondre la valeur de marché calculée au sein du modèle à celle observée sur le marché. **Ce sont donc les montants des coupons et des nominaux des obligations qui seront perturbés pour générer la base.**

Vient ensuite la question des modalités de cette variabilisation. Les données d'entrée du modèle doivent être perturbées de façon à inclure suffisamment de combinaisons cohérentes avec l'usage souhaité. Il faut cependant éviter de produire un nombre trop important de combinaisons aberrantes qui s'avèreraient avoir un effet nocif sur le pouvoir prédictif du modèle de *machine learning*.

Pour l'action et l'immobilier, des coefficients aléatoires sont tirés entre des bornes définies au préalable pour chaque simulation, puis appliqués aux valeurs de marché. Nous utilisons une loi uniforme et paramétrons nos bornes de façon à être sûrs d'inclure les chocs SCR dans l'amplitude choisie. Les bornes sont fixées entre 50% et 130% pour l'action et entre 60% et 120% pour l'immobilier.

Pour les taux, nous considérons que la variabilisation s'effectue naturellement via la génération aléatoire de scénarios stochastiques. Chaque scénario propose une structure différente de taux en input. Pour s'assurer de l'exhaustivité de la base, plusieurs tables de scénarios économiques sont produites, centrés autour de translations de la courbe des taux sans risques au 31/12/2022.

Si les solutions trouvées pour les données ci-dessus sont plutôt intuitives, plusieurs obstacles ont cependant été rencontrés lors de la variabilisation des obligations, qui est pourtant nécessaire pour prédire efficacement les effets de la risque-neutralisation et des chocs de spread. La logique adoptée pour les actions et l'immobilier, c'est-à-dire dans notre cas l'application de chocs tirés indépendamment aux coupons et nominaux, produisait des inputs incohérents, trop diffus et ne permettant pas une prédiction correcte des chocs de spread. **Nous nous sommes donc penchés sur le type d'abattement que nous devons anticiper pour pouvoir le reproduire au mieux lors de la génération des coefficients de bruit.** La figure ci-dessous montre, par année de maturité, l'abattement appliqué à notre portefeuille d'actifs de référence lors de la risque neutralisation, du SCR de spread et du SCR de spread combiné à une hausse de 50 bps du spread.



Représentation des abattements moyens appliqués au nominaux dans différents scénarios en fonction de la maturité résiduelle

Suite à ces observations, quelques contraintes supplémentaires sur la variabilisation des obligations ont été adoptées, sans perte de généralisation :

- Les coefficients sont tirés par année de maturité et appliqués à tous les coupons et nominaux des obligations de la maturité concernée.
- Les coefficients ne sont plus tirés indépendamment dans des bornes fixées. Pour chaque année de maturité (exceptée la première), le choc est tiré uniformément dans des bornes relatives centrées sur le coefficient choisi pour la maturité précédente.
- Des plafonds sont mis en place pour s'assurer que les coefficients restent compris entre deux bornes absolues (entre 80% et 100%)

Cette méthode s'avère avoir par la suite un impact très positif sur la performance de prédiction des BEL et NAV, nous permettant ainsi de progresser dans la prédiction du SCR Marché.

L'ensemble des coefficients aléatoires utilisés pour perturber les *inputs* et générer les données sont calculés directement au sein du modèle Prophet puis appliqués aux *inputs* à chaque simulation avant le début des calculs actuariels.

2. Génération de la base de données

Nous sommes capables de perturber nos données d'entrée afin de générer une base exhaustive. L'ajout d'une fonction dans le code Prophet permet ainsi, pour chaque simulation, d'inscrire une ligne contenant les informations sur l'actif et le BEL et la NAV associés au format CSV. Il reste encore à **définir quelles variables doivent être affichées dans la base de données pour qu'elle contienne une information complète et exploitable**. Une fois encore, la solution s'avère intuitive pour l'action, l'immobilier et les taux. Les valeurs de marché et les taux *spot* et *forward* contiennent toute l'information pertinente vis-à-vis du modèle et dans un format synthétique. Pour les obligations, il a fallu réfléchir à une méthode efficace pour décrire le portefeuille obligataire sans proposer un nombre trop important de variables. Par exemple, deux obligations peuvent avoir une valeur de marché similaire mais des coupons, nominaux et maturités radicalement différentes, ce qui a un impact majeur sur le déroulement des projections. La valeur de marché n'est donc pas jugée pertinente. Finalement, nous avons opté pour **l'utilisation d'une variable représentant la production financière annuelle anticipée du portefeuille obligataire**.

Au total, ce sont 750 000 simulations qui sont effectuées, pour un montant équivalent de lignes produites au sein de la base de données consolidée.

3. Calibration et interprétabilité d'un algorithme de *machine learning*

Ayant à disposition une base de données exploitable, l'apprentissage supervisé a été mis en place, avec en variables à expliquer le BEL et la NAV. Des travaux menés préalablement ont montré que l'algorithme XGBoost était le plus performant pour accomplir nos objectifs, c'est donc celui utilisé dans la suite de ce mémoire.

Nous commençons par présenter conceptuellement le fonctionnement d'XGBoost, soit celui d'un algorithme d'apprentissage supervisé utilisant des arbres de décisions, prédisant à chaque itération les erreurs de prédictions de l'itération précédente en fonction des variables explicatives puis agrégeant l'ensemble des classifieurs construits en un classifieur final.

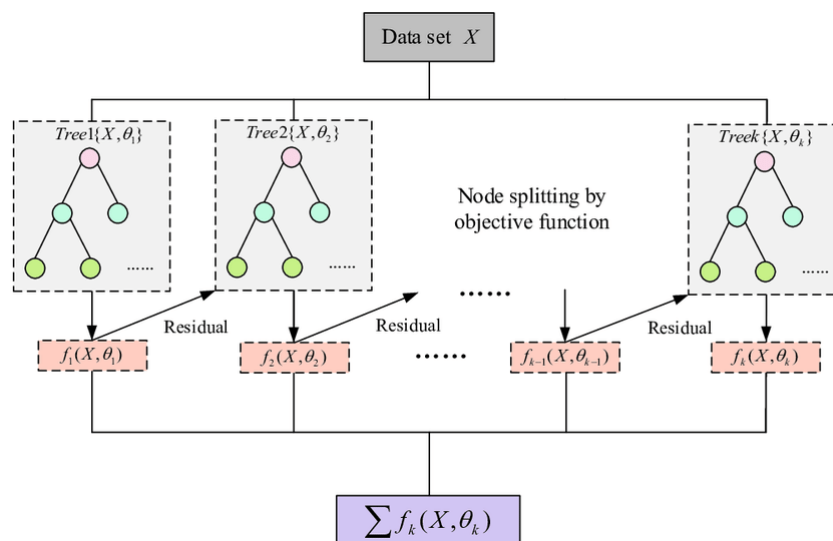


Schéma explicatif du fonctionnement de XGBoost

Après les étapes classiques de normalisation des variables et de séparation de la base entre jeu d'apprentissage, jeu de test, variables explicatives et variables à expliquer, nous procédons à une sélection de variables. Celle-ci s'opère via une pré-sélection manuelle dans un premier temps, puis dans un second temps, la *feature importance* déterminée en faisant apprendre le modèle sur un échantillon réduit est utilisée pour affiner cette sélection. Des modèles intégrant un nombre incrémentalement grand de variables par ordre décroissant de *feature importance* sont entraînés et l'évolution de leur performance est étudiée afin de choisir un nombre plus faible de variables sans perte de performance. D'autres manipulations d'agrégation de variables ont lieu afin d'amplifier leur significativité. **Ainsi, des 1732 variables initialement présentes, une centaine seulement est retenue après l'ensemble des retraitements.** Les hyperparamètres du modèle sont ensuite calibrés un à un sur un échantillon de ce jeu de données raffiné, puis le modèle finalement obtenu est entraîné sur l'ensemble du jeu d'apprentissage.

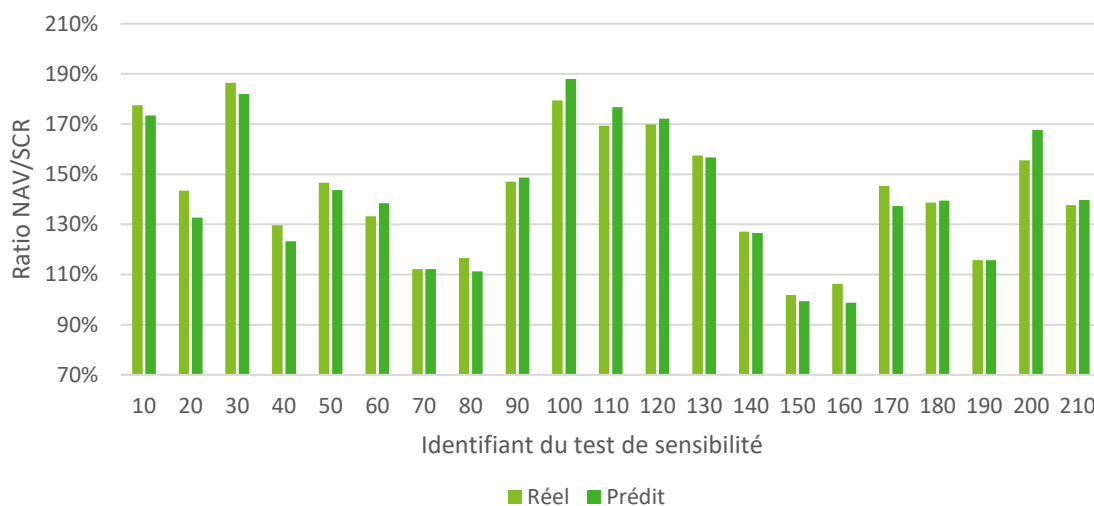
Le processus de génération des données puis de calibration du modèle est itératif. A chaque tentative, la RMSE du modèle sur le jeu de test est contrôlée pour avoir un indicateur de sa performance, puis le modèle est utilisé pour prédire des BEL et NAV dans des cas de tests de sensibilités classiques pour vérifier sa pertinence.

Ayant obtenu un modèle satisfaisant, nous nous sommes penchés sur son interprétabilité. Le XGBoost souffre d'un effet « boîte noire » rendant opaque le processus derrière l'obtention de ses résultats. Des méthodes permettant d'explicitier ces modèles, telles que LIME et SHAP, existent. Nous nous sommes notamment servis de cette dernière pour avoir un regard sur le rôle et l'impact des variables essentielles, et s'assurer que le modèle est cohérent avec notre intuition. Une autre possibilité offerte par SHAP est, pour une prédiction donnée, avoir une explication de l'impact de chaque variable sur l'écart à la moyenne de la prédiction.

4. Obtention des résultats et application pratique

La dernière étape consiste à définir les tests de sensibilités à mener pour notre étude, pour un portefeuille fictif créé pour l'exercice. Il est alors nécessaire de produire et formater les *inputs* de façon qu'ils reflètent les tests de sensibilités et l'application des chocs réglementaires de SCR. Chaque NAV est prédite en faisant la moyenne des prédictions pour 1000 simulations stochastiques. Cette méthode garantit l'obtention d'une NAV prenant en compte la TVOG, mais permet également à des mécanismes de compensation entre les erreurs de prédiction de se produire. **L'erreur de prédiction sur les NAV est ainsi de l'ordre du pourcent.**

Les NAV centrales et choquées obtenues pour chaque test de sensibilité sont utilisées pour déterminer le SCR marché, puis pour prédire le comportement d'un indicateur rendement/risque, ici le ratio NAV/SCR. Dans un premier temps, les résultats des prédictions de SCR marché sont d'une précision moyenne mais tolérables, avec plusieurs pistes d'améliorations laissées ouvertes. Cependant, **les résultats pour l'indicateur NAV/SCR marché sont très satisfaisants, avec une capacité pour notre modèle à prédire efficacement les tendances à la hausse ou à la baisse de l'indicateur**, dans des scénarios croisant des chocs simultanés sur plusieurs axes.



Impact des tests de sensibilités sur le ratio NAV/SCR marché

Pouvoir prédire un SCR marché à un instant donné et en fonction d'évolutions possibles de l'environnement économique est un succès important en vue du développement d'un outil complet et utilisable pour accompagner l'actuaire dans une étude ALM. Les prochaines étapes pour y parvenir seront notamment la variabilisation du passif et des *management actions* (telles que l'allocation d'actifs cible) et l'intégration de ces éléments au modèle.

EXECUTIVE SUMMARY

In this paper, we're creating a tool capable of predicting, with the help of machine learning and for an insurance company at a given time, the amount of market SCR based on the asset portfolio and the economic environment.

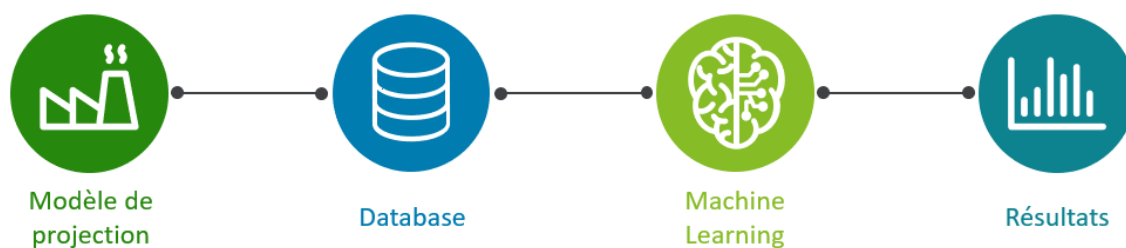
The Solvency Capital Requirement (SCR), defined in the Solvency II standard, can be computed using a standard formula which, for the considered modules, corresponds to an aggregation of differences between the insurer's NAV (Net Asset Value) under normal conditions and his NAV after the application of predefined shocks. These shocks directly affect the input values of the ALM (Asset and Liability Management) model, thereby impacting the NAV amount as output.

The method used to predict the market SCR first involves predicting the NAV based on variables affected by regulatory shocks, namely, the market values of assets, the considered economic scenarios, and the spread of the bond portfolio. By achieving this with sufficient performance, we can almost instantly predict the central and shocked NAV, even for sensitivity tests varying the initial state of the asset portfolio.

Simultaneously, we will also predict the BEL (Best Estimate of Liabilities) to obtain a complete balance sheet.

The process implemented to achieve this is divided into four phases:

- 1) Perturbation of economic conditions as input to the ALM model;
- 2) Production of a database containing synthetic and necessary information to explain variations in NAV and the BEL calculated at the output of the ALM model based on input data;
- 3) Calibration of a machine learning model based on the generated data to predict NAV and BEL, as well as the implementation of interpretability methods;
- 4) Creation of suitable inputs to predict central and shocked NAV for each sub-module of the Market SCR and for each considered sensitivity test, followed by an analysis of results.



Summary diagram of the overall process

1. Variabilization of input data within the projection model

The first objective is to build a database. It must contain a large number of combinations of variables describing the assets and the associated BEL and NAV. These combinations must be diverse and exhaustive enough to ensure that the application of sensitivity tests and regulatory shocks to the assets result in inputs consistent with our training database. Two issues emerge at this point: which inputs are chosen to be perturbed to generate this database and how are they perturbed for each simulation.

The choice of variabilized elements must be coherent with the objective of predicting NAV after Market SCR shocks for the interest rate, equity, real estate, and spread modules. The interest rate shock involves a modification of the yield curve used for generating economic scenarios, equity and real estate shocks correspond to a decrease in market value of the respective assets, and the spread shock corresponds to a decrease in market value of bonds due to an increase in the associated spread.

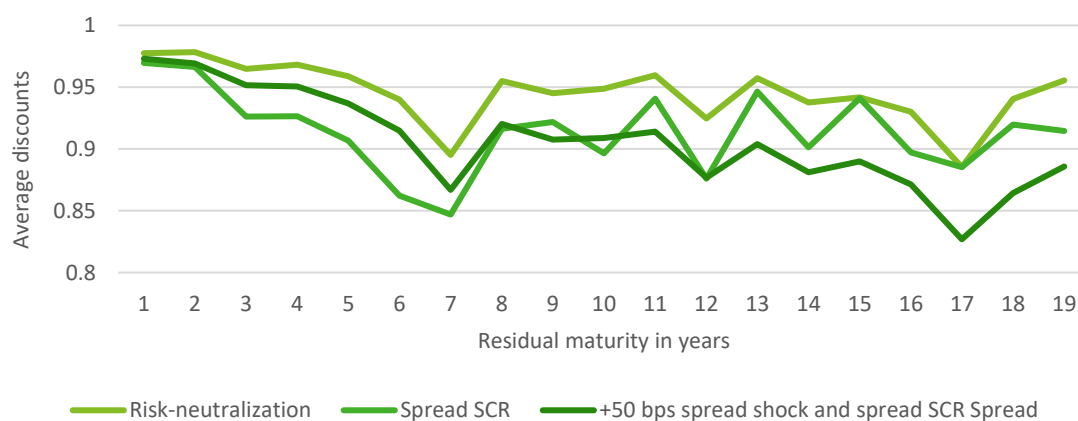
From the first three shocks, **it is immediately deduced that the variables will include spot and forward rates and market values of equities and real estate.** For the spread shock, the choice was made based on the method chosen to risk-neutralize bonds. Indeed, discounts are applied to coupons and principals to match the calculated market value within the model to that observed in the real-world market. **Therefore, the amounts of coupons and principals of bonds will be perturbed to generate the database.**

Next comes the question of the modalities of this variabilization. The input data of the model must be perturbed in a way that includes enough combinations consistent with the desired use. However, it is necessary to avoid producing too many aberrant combinations that would have a harmful effect on the predictive power of the machine learning model.

For equities and real estate, random coefficients are drawn between pre-defined bounds for each simulation, then applied to market values. We use a uniform distribution and parameterize our bounds to ensure that the SCR shocks are included within the chosen range. The bounds are set between 50% and 130% for equities and between 60% and 120% for real estate.

For interest rates, we consider that variabilization occurs naturally through the random generation of stochastic scenarios. Each scenario proposes a different rate structure as input. To ensure the completeness of the database, several tables of economic scenarios are produced, centered around translations of the risk-free yield curve as of 31/12/2022.

While the solutions found for the above variables are rather intuitive, several obstacles were encountered in the variabilization of bonds, which is necessary to predict the effects of risk-neutralization and spread shocks. The logic adopted for equities and real estate, i.e., in our case the application of independently drawn shocks to coupons and principals, produced inconsistent, too diffuse inputs, and did not allow for a correct prediction of spread shocks. **We therefore looked into the type of discount that we should anticipate in order to reproduce it as closely as possible during the generation of noise coefficients.** The figure below shows, by maturity year, the discount applied to our reference asset portfolio during risk neutralization, spread SCR, and spread SCR combined with a 50 bps spread increase.



Representation of the average discounts applied to principals in different scenarios based on residual maturity.

Following these observations, some additional constraints on the variabilization of bonds have been adopted, without loss of generalization:

- Coefficients are drawn per residual maturity year and applied to all coupons and principals of bonds of the respective maturity.
- Coefficients are no longer independently drawn within fixed bounds. For each maturity year (except the first), the shock is uniformly drawn within relative bounds centered on the coefficient chosen for the previous maturity.
- Caps are implemented to ensure that coefficients remain between two absolute bounds (between 80% and 100%).

This method subsequently proves to have a very positive impact on the prediction performance of BEL and NAV, allowing us to make progress in predicting the Market SCR.

All random coefficients used to perturb inputs and generate data are calculated directly within the Prophet model and applied to inputs at each simulation before the start of actuarial calculations.

2. Generation of the database

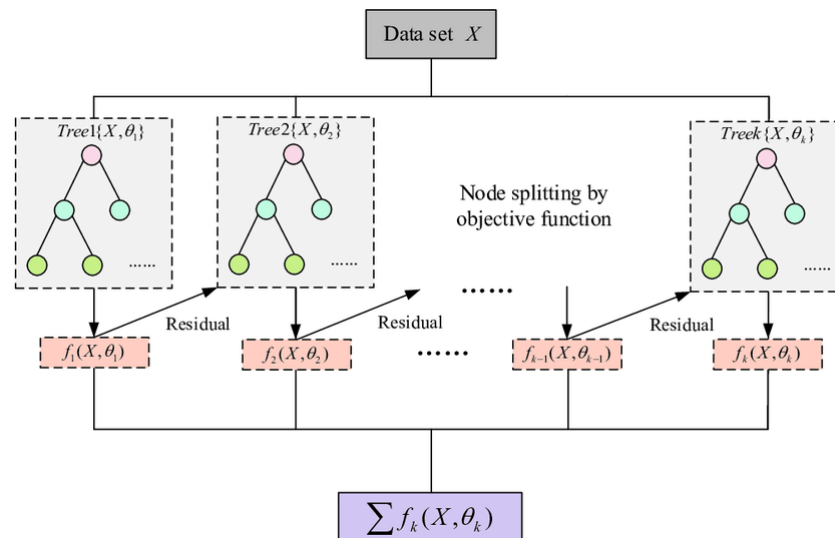
We are able to randomize our input data to generate a comprehensive database. Adding a function in the Prophet code allows, for each simulation, to record a row containing information about the asset and the associated BEL and NAV in CSV format. There is left to **define which variables should be displayed in the database to ensure it contains complete and usable information**. Once again, the solution proves intuitive for equities, real estate, and interest rates. Market values and spot and forward rates contain all the relevant information regarding the model and in a synthetic format. For bonds, we had to think of an efficient method to describe the bond portfolio without creating too many variables. For example, two bonds may have a similar market value but radically different coupons, principals, and maturities, which have a major impact on the projections. Therefore, the market value is not considered relevant. Finally, **we opted for the use of a variable representing the anticipated annual financial production of the bond portfolio**.

In total, 750,000 simulations are performed, resulting in an equivalent number of observations produced within the consolidated database.

3. Calibration and interpretability of a machine learning algorithm

With an exploitable database available, supervised learning has been implemented, with BEL and NAV as variables to be explained. Previous work has shown that the XGBoost algorithm was the most effective in achieving our objectives, so it is the one used in the rest of this paper.

We begin by conceptually presenting XGBoost, which is a supervised learning algorithm using decision trees, predicting at each iteration the prediction errors of the previous iteration based on the explanatory variables and then aggregating all the classifiers built into a final classifier.



Summary diagram of the XGBoost algorithm

After the usual steps of variable normalization and separation of the database into training set, test set, explanatory variables, and target variables, we proceed with feature selection. This is done through a first manual pre-selection, followed by using the feature importance determined by training the model on a reduced sample to refine this selection. Models incorporating an incrementally large number of variables starting with those with the highest feature importance are trained, and the evolution of their performance is studied to select a smaller number of variables without loss of performance. Other variable aggregation manipulations take place to amplify their significance. **Thus, out of the initially 1732 variables present, only around a hundred are retained after all the processing steps.** The model's hyperparameters are then calibrated one by one on a sample of this refined dataset, and the final model obtained is trained on the entire training set.

The process of data generation and model calibration is iterative. At each attempt, the RMSE of the model on the test set is checked to assess its performance, and then the model is used to predict BELs and NAVs in classic sensitivity test cases to verify its relevance.

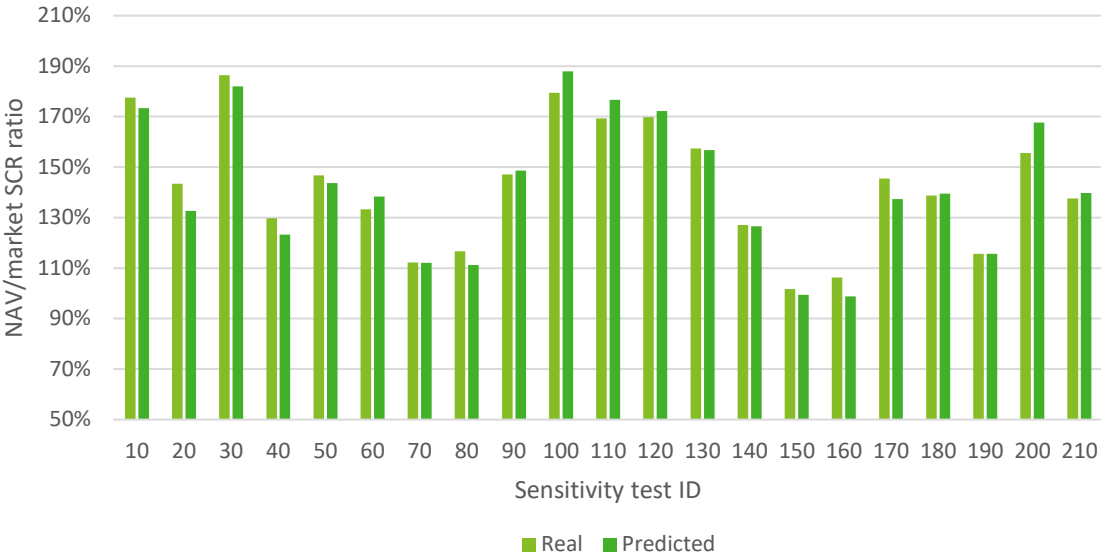
Having obtained a satisfactory model, we turned our attention to its interpretability. XGBoost suffers from a "black box" effect that makes the process behind its results opaque. Methods to explain these models, such as LIME and SHAP, exist. We have used the latter to examine the role and impact of essential variables and ensure that the model is consistent with our intuition. Another possibility offered by SHAP is, for a given prediction, to have an explanation of the impact of each variable on the deviation from the average prediction.

4. Obtaining results and practical application

The final step involves defining sensitivity tests for our study, for a fictional portfolio created for the exercise. It is then necessary to produce and format the inputs so that they reflect the sensitivity tests and the application of regulatory shocks to the SCR. Each NAV is predicted by averaging predictions for 1000 stochastic simulations. This method ensures obtaining a NAV accounting for the TVOG (Time Value of Options and Guarantees), but also allows for mechanisms of error compensation between predictions to occur. **The prediction error on NAVs is thus in the order of a percent.**

The central and shocked NAVs obtained for each sensitivity test are used to determine the Market SCR, and then to predict the behavior of a risk-return indicator, here the NAV/market SCR ratio. Initially, the results of Market SCR predictions are of mid but tolerable accuracy, with several ways to improve

left open. However, **the results for the NAV/Market SCR indicator are very satisfactory, with our model's ability to effectively predict upward or downward trends in the indicator**, in scenarios involving simultaneous shocks on multiple axes.



Impact of sensitivity tests on the NAV/Market SCR ratio

Being able to predict a Market SCR at a given moment and based on possible developments of the economic environment is a significant success towards the development of a comprehensive and usable tool to assist the actuary in Asset and Liability Management studies. The next steps to achieve this will notably include the variabilization of liabilities and management actions (such as target asset allocation) and the integration of these elements into the model.

INTRODUCTION

Les assureurs ont un besoin constant d'évaluer leur exposition au risque. Il est essentiel de connaître les scénarios les plus risqués, dans lesquels la solvabilité de la compagnie peut être dangereusement affectée, afin de pouvoir mettre en place des stratégies de mitigation ou de mieux communiquer sur les risques de la compagnie. Un des outils à leur disposition pour y parvenir est l'exécution de tests de sensibilités, c'est-à-dire l'analyse du comportement de métriques face à des perturbations de paramètres en entrée du modèle de calcul prudentiel. Dans ce mémoire, nous nous concentrerons sur l'exposition au risque de marché d'un assureur vie, via la sensibilité du SCR marché.

L'assureur a intérêt à explorer un grand nombre de scénarios possibles pour avoir la vision la plus exhaustive de son exposition. Cependant, les calculs de SCR et les tests de sensibilité sont des processus lents, coûteux informatiquement et parfois laborieux, requérant une importante part d'intervention humaine. En effet, calculer un SCR marché selon la formule standard requiert le paramétrage, l'exécution et l'agrégation de plusieurs *runs* au sein des modèles de projection afin de déterminer les variations de NAV associées à chaque sous-module. Cet effort est multiplié par le nombre de tests de sensibilité à mener.

C'est pourquoi, il peut être intéressant pour un assureur d'avoir accès à un outil capable d'approximer efficacement le SCR marché pour chaque scénario de test, sans avoir à utiliser à chaque fois le modèle ALM. Les NAV à calculer pour déterminer le SCR sont le produit d'une fonction très complexe prenant en entrée des données sur les engagements envers les assurés, l'état du portefeuille d'actifs, l'environnement économique, la stratégie de l'assureur et des assurés... C'est pourquoi, il est pertinent de chercher du côté des algorithmes d'apprentissage supervisé qui, après un apprentissage sur une base de données contenant les données d'entrée du modèle ALM et les NAV associées, pourrait nous permettre d'approximer presque instantanément le SCR marché pour tous les tests de sensibilité proposés.

Dans un premier temps, nous présenterons le cadre réglementaire de l'assurance vie puis de Solvabilité II. Nous détaillerons ensuite le fonctionnement du modèle de calcul, en décrivant les inputs, l'algorithme sous-jacent et les spécificités de notre modèle, notamment l'usage d'une méthode *flexing*. Enfin, nous exposerons les différentes phases du processus mis en place pour atteindre l'objectif fixé, soit la prédiction du SCR marché pour différents tests de sensibilités, ainsi que les obstacles rencontrés et la façon dont ils ont été surmontés pour atteindre un niveau de précision satisfaisant. L'efficacité de notre méthode sera évaluée dans une étude de cas impliquant un assureur vie fictif souhaitant évaluer son exposition au risque de marché.

I/ L'ASSURANCE VIE EN FRANCE

Avant de rentrer dans le détail de Solvabilité II et de la modélisation ALM, il est essentiel de rappeler les éléments caractéristiques des produits d'assurance vie avec lesquels nous allons travailler.

Un contrat d'assurance vie est un produit financier qui permet à un individu (appelé souscripteur) de constituer une épargne à long terme via le versement de primes tout en bénéficiant d'une couverture d'assurance en cas de survie ou de décès. L'assurance en cas de vie pure n'existe quasiment pas en France, les contrats d'épargne contiennent généralement à la fois une garantie en cas de vie et une en cas de décès. L'assuré a également le droit d'exercer une option de rachat avant l'échéance du contrat. Dans ce cas, l'assureur lui verse un capital correspondant à une part ou toute la mise initiale ainsi que les intérêts générés. Certaines pénalités de rachat peuvent être appliquées.

Il existe différents types de contrats d'épargne possédant chacun des risques et des rémunérations différentes pour l'assureur et l'assuré. Les deux types de contrats d'épargne principaux sont les produits en euros et les produits en Unités de Compte (UC). Il existe aussi des produits multi-supports, combinant les deux types précédents, et depuis peu des produits Eurocroissance. Nous allons présenter ces différents supports et leurs spécificités.

1. Contrat d'épargne en euros

Sur un produit en euros, les sommes versées par l'assuré sont garanties et revalorisées chaque année par un taux minimum garanti (TMG) et possiblement augmenté d'une participation aux bénéfices de l'assureur. Le rendement est donc dépendant des résultats techniques et financiers de l'assureur. Les intérêts sont définitivement acquis par l'assuré et se composent. La garantie du capital acquis par l'assuré implique que le risque de perte est intégralement supporté par l'assureur.

- Le Taux Minimum Garanti

Le TMG est donc un taux minimum de revalorisation garanti par l'assureur sur une période donnée. Celui-ci peut être fixe ou basé sur un indice, par exemple le taux moyen des emprunts de l'Etat français (TME). Pour éviter que les assureurs ne promettent un taux trop élevé mettant en risque leurs activités en cas de contexte défavorable, le code des Assurances (Article A132-1) limite celui-ci en fonction de la durée d'engagement. Il ne peut pas dépasser :

- 75% du TME pour un contrat à prime unique ou à versements libres si l'engagement est inférieur à 8 ans
- Le minimum entre 60% du TME et 3,5% sinon.

- La Participation aux Bénéfices

L'assureur a l'obligation de verser aux assurés une part ou la totalité de ses bénéfices techniques et financiers annuels liés aux contrats en euros à travers la revalorisation des contrats. Il est légalement tenu de reverser sous forme de participation aux bénéfices au moins 85% du résultat technique et 90% du résultat financier.

Il a cependant la possibilité de ne pas la verser immédiatement et de la différer dans le temps via la constitution d'une Provision pour Participation aux Bénéfices (PPB) ou aux Excédents (PPE). Les fonds et les intérêts produits par ceux-ci sont la propriété des assurés. Les sommes dotées sur cette provision

doivent nécessairement être reprises et versées dans les 8 années qui suivent. L'intérêt de la PPE est d'offrir à l'assureur la possibilité de lisser son résultat pour éviter des variations trop fortes de taux servi d'une année à l'autre.

2. Contrat d'épargne en UC

Dans un contrat en unités de comptes, les assurés placent leur argent dans une variété d'instruments financiers tels que des actions, des obligations, des fonds communs de placement ou d'autres actifs, appelés "unités de compte". Contrairement aux produits en euros, les rendements sur un contrat d'épargne en unités de compte n'ont pas de minimum garanti, mais dépendent des performances et des fluctuations des marchés financiers. Par conséquent, les assurés assument un certain niveau de risque, car la valeur des unités de compte peut augmenter ou diminuer en fonction des conditions économiques et des tendances du marché. La flexibilité des contrats d'épargne en unités de compte réside dans la diversification des investissements possibles, permettant aux assurés de participer aux opportunités de croissance tout en reconnaissant le potentiel de perte de capital.

Des garanties peuvent néanmoins être ajoutées à ces contrats afin de fournir un minimum de protection au capital investi par l'assuré, via des garanties plancher par exemple.

3. Autres versions du contrat d'épargne : multisupport et Eurocroissance

Il est possible de combiner ces deux types de contrats en un seul : on parle alors de multisupport. Une part de l'épargne sera investie sur des supports en Euros et l'autre partie sur des unités de compte. Les parts d'euros et d'UC sont fixés selon le profil de risque de l'assuré (prudent, équilibré, dynamique), et des arbitrages sont possibles pour transférer des fonds entre les deux supports.

Similairement, les produits eurocroissance combinent des caractéristiques des produits d'épargne en euros, qui offrent un rendement garanti et une sécurité du capital, avec des éléments de potentiel de croissance liés aux performances des marchés financiers. Contrairement aux produits d'épargne en euros, les produits eurocroissance permettent une croissance potentielle du capital en fonction de la performance d'un support d'investissement, tel qu'un fonds en unités de compte. Cependant, cette croissance potentielle est assortie d'une garantie partielle ou différée du capital, ce qui signifie que la protection du capital peut dépendre de la durée de détention et des performances du support d'investissement. Les produits eurocroissance peuvent offrir une certaine flexibilité pour les retraits, mais des pénalités peuvent s'appliquer en cas de sortie anticipée. Comme ils sont liés aux marchés financiers, ces produits comportent un niveau de risque plus élevé que les produits d'épargne en euros, mais offrent en contrepartie un rendement potentiellement plus élevé.

4. Les avantages fiscaux

L'un des principaux avantages du contrat d'assurance vie est la possibilité de bénéficier d'une fiscalité avantageuse sur les gains générés par le contrat au fil du temps. Les intérêts et les plus-values

accumulés dans un contrat d'assurance vie peuvent généralement être exonérés d'impôt sur le revenu jusqu'à un certain seuil ou dans certaines conditions, par exemple, si le rachat a lieu plus de 8 ans après la souscription du contrat.

Un autre avantage fiscal est lié à la transmission de l'assurance vie en cas de décès de l'assuré. Les bénéficiaires désignés dans le contrat peuvent recevoir les fonds sans avoir à payer de droits de succession, jusqu'à un certain montant défini par la législation fiscale. Cela peut permettre de transmettre un capital à ses héritiers de manière plus avantageuse sur le plan fiscal que d'autres formes de transmission de patrimoine.

II/ CONTEXTE RÉGLEMENTAIRE

Nous avons donc présenté le cadre de l'assurance vie dans lequel se déroule notre étude. Nous allons maintenant définir le capital de solvabilité requis, sa méthode de calcul, et comment il répond à une logique prudentielle.

1. L'introduction de Solvabilité I puis II

Commençons par rappeler ce que l'on entend par un organisme « solvable ». Un organisme d'assurance est solvable s'il est capable de faire face à ses engagements, c'est-à-dire s'il est en mesure de verser à chacun de ses assurés les prestations qui leur sont dues.

De nombreux risques peuvent remettre en cause la solvabilité d'un tel organisme, tels qu'une sous-estimation des engagements, une sous-tarification des garanties, une dérive de la sinistralité, des taux servis trop élevés, des pertes importantes sur les marchés financiers, etc. Différents processus et mesures permettent de limiter l'impact de ces risques, en mettant en avant un principe de prudence lors du provisionnement, de la tarification, ou dans la gestion d'actif. De manière plus générale, c'est une véritable culture de suivi et de contrôle du risque qui est attendue de la part des assureurs.

En 1973, l'Union Européenne adopte la directive Solvabilité 1 ayant pour but de réglementer ces pratiques afin de protéger au mieux les assurés. Déjà à cette époque, celle-ci se présente sous la forme de trois piliers :

- Le Pilier I porte sur la nécessité d'avoir des provisions techniques suffisantes
- Le Pilier II ajoute des contraintes sur les actifs pouvant représenter ces provisions techniques : elle définit les actifs admissibles en fonction de catégories ou des limitations de volumes détenus par émetteurs
- Le Pilier III définit une exigence de marge de solvabilité, soit un montant minimal de fonds propres à détenir. En non-vie, celle-ci correspondait au maximum entre un pourcentage (16% à 18%) des primes et un pourcentage (23% à 26%) des sinistres. En vie, elle correspondait à la somme de 4% des provisions en euros, 1% des provisions en UC et entre 0.1% et 0.3% des capitaux sous risques

La directive S1 souffrait cependant de nombreux écueils. Les calculs forfaitaires n'étaient pas adaptés aux profils de risques des organismes d'assurance et la directive se limitait à la régulation des aspects quantitatifs.

Le 1^{er} janvier 2016, la directive Solvabilité II entre en application, avec pour objectif de pallier ces lacunes. Celle-ci apporte un renforcement des exigences en termes de solvabilité, de gouvernance et de reporting. Cette fois encore, elle se décompose en trois piliers : exigences quantitatives, exigences qualitatives et reporting prudentiel/information du public.

2. Les trois piliers

a. Pilier 1 : Exigences quantitatives

Le Pilier 1 porte sur les exigences quantitatives. Il impose une vision économique du bilan prudentiel, avec l'actif et le passif présentés en valeur de marché plutôt qu'en valeur comptable. Aussi il définit deux nouvelles métriques d'exigence de capital : le SCR (*Solvency Capital Requirement*) et le MCR (*Minimum Capital Requirement*), ainsi que leur méthode de calcul. Ce pilier est central pour ce mémoire dans la mesure où il regroupe les règles de calculs pour l'ensemble des éléments que nous souhaiterons prédire. Nous le détaillerons par la suite.

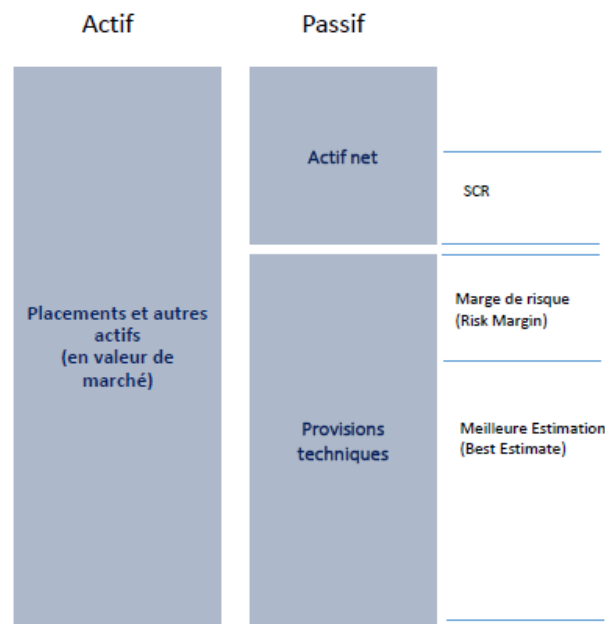


Figure 1 – Schéma simplifié d'un bilan prudentiel

b. Pilier 2 : Exigences qualitatives

Le Pilier 2 de Solvabilité II regroupe notamment les règles de gouvernance et de gestion des risques, ainsi que l'évaluation de ses propres risques de solvabilité au travers de l'ORSA (*Own Risk and Solvency Assessment*). La norme impose aux organismes la mise en place d'un système de gouvernance efficace dont le but est de garantir une gestion saine et prudente de l'activité. Ce système doit comprendre au minimum « 4 yeux », soit deux dirigeants effectifs, et 4 responsables de fonctions clés : fonction actuarielle, gestion des risques, audit interne et conformité. En parallèle à ce contrôle interne, le Pilier 2 codifie le contrôle externe exercé par le superviseur européen. Il énonce notamment que le superviseur peut imposer une marge additionnelle de capital de solvabilité requis (*Capital Add-On*) s'il perçoit une inadéquation entre la formule standard ou le modèle interne de l'organisme et le profil de risque de celui-ci, ou s'il détecte une défaillance au niveau de la gouvernance ou de la gestion des risques. Enfin, l'organisme doit mener un processus interne d'évaluation des risques et de la solvabilité via l'ORSA. Il doit illustrer la capacité de l'organisme ou du groupe à identifier, mesurer et gérer les éléments de nature à modifier sa solvabilité ou sa situation financière. Aussi, sa déclinaison

opérationnelle en fait un outil stratégique de premier plan qui doit être appréhendé par l'organisme comme un outil de pilotage de l'activité en fonction des risques.

c. Pilier 3 : *Reporting* et information du public

Le Pilier 3 de Solvabilité II concerne la communication d'informations au public et aux autorités de contrôle. Il vise à harmoniser au niveau européen les informations publiées par les organismes d'assurance ainsi que celles remises aux superviseurs.

Cette publication est destinée au marché ainsi qu'au superviseur, et s'organise autour de trois principes : Les publications comptables, réglementaires et celles destinées au superviseur doivent être cohérentes ; La publication d'informations pour le régulateur européen doit être harmonisée à l'échelle de l'UE ; Les mêmes règles doivent s'appliquer à tous les assureurs. Ces informations, à la fois quantitatives et qualitatives, doivent être remises à une fréquence annuelle et, pour certaines, trimestrielle.

3. Focus sur le pilier 1 : Le bilan prudentiel

a. Le Best Estimate

Le *Best Estimate* correspond au montant que doit détenir un assureur pour être en mesure d'indemniser ses assurés « selon une vision moyenne ». Il se calcule comme la somme des *cash flows* futurs liés à l'exécution des contrats en portefeuille, en tenant compte de leur probabilité de survenance et actualisés via la courbe des taux sans risques fournie par l'EIOPA.

$$BEL = \sum_i \frac{CF_i \times p_i}{(1 + r_i)^i}$$

Avec :

- CF_i les *cash flows* attendus au temps i
- p_i leur probabilité de survenance
- r_i le taux sans risque pour la maturité i

Cette courbe est construite à partir des taux swaps « euro contre EURIBOR 6 mois ». Un ajustement de risque de crédit (CRA) est appliqué aux taux observés jusqu'au *Last Liquid Point* (LLP), fixé à 20 ans. Les taux entre le LLP et la dernière maturité observée sont interpolés, et ceux allant au-delà du LLP sont extrapolés jusqu'à l'*Ultimate Forward Rate*, le taux vers lequel la courbe converge. Aussi, un *Volatility Adjustment* est appliqué sur l'ensemble de la courbe pour limiter l'impact de la volatilité du marché sur le bilan.

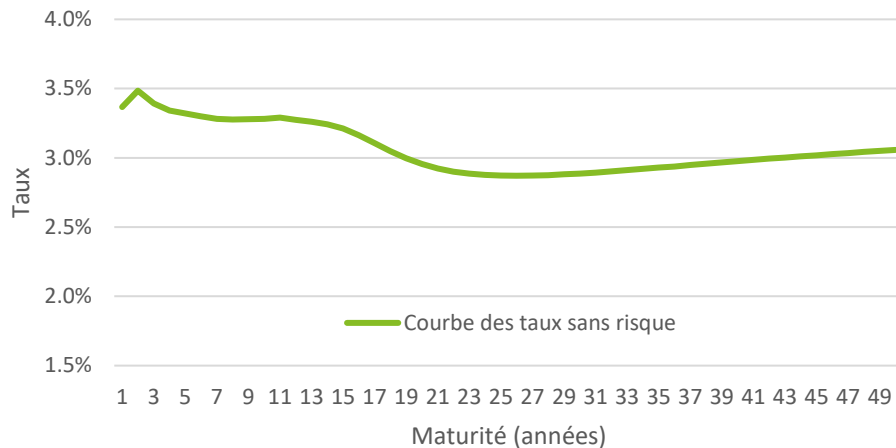


Figure 2 – Courbe des taux sans risques avec VA de l'EIOPA au 31/12/2022

Le *Best Estimate* doit prendre en compte toutes les options et garanties offertes par le contrat en anticipant le comportement probable des assurés, l'ensemble des versements discrétionnaires futurs (notamment la participation aux bénéfices), et ce sur l'ensemble des contrats existants que l'assureur ne peut plus unilatéralement refuser.

b. La Risk Margin

La Marge pour Risque est une partie complémentaire au *Best Estimate*. Elle représente le montant requis pour couvrir le coût du transfert du portefeuille d'assurance à un autre organisme et s'assurer que celui-ci pourrait reprendre et honorer les engagements d'assurance et de réassurance. Elle est calculée en actualisant le coût du capital annuel généré par l'immobilisation du SCR, estimé à 6% par an, sur la durée de vie résiduelle des engagements utilisés pour le calcul du *Best Estimate*.

$$RM = CoC \times \sum_{t>0} \frac{\mathbb{E}[SCR(t)]}{(1 + r_{t+1})^{t+1}}$$

Avec :

- *RM* la Risk Margin
- *CoC* le coût du capital, fixé à 6%
- *SCR(t)* le capital de solvabilité requis à la date *t*
- r_{t+1} le taux sans risque de maturité *t+1*

c. Les fonds propres

Les fonds propres désignent les ressources financières d'une compagnie d'assurance qui sont disponibles pour absorber les pertes éventuelles et maintenir sa stabilité financière. Sous Solvabilité II, les fonds propres correspondent à l'actif net (ou NAV), c'est-à-dire la différence entre la valeur économique de l'actif et la valeur des provisions techniques (soit l'engagement envers les assurés). L'actif peut être vu comme la combinaison des fonds propres en vision historique et de la *Value in Force*, cette dernière correspondant à la valeur actuelle des profits futurs retirés de la marge de risque.

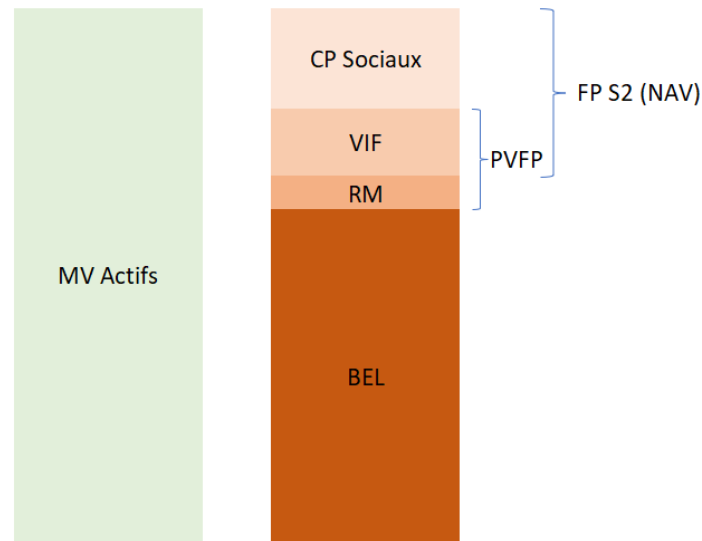


Figure 3 – Schéma d'un bilan prudentiel avec décomposition de la NAV

d. SCR et MCR

Formule standard et modèle interne

Le SCR représente le capital nécessaire pour absorber des pertes inattendues, entraînées par exemple par des événements extrêmes et imprévisibles (catastrophes naturelles, terrorisme, chute des marchés boursiers, etc.). Il correspond à la Value at Risk à 99.5% de l'actif net économique de l'organisme d'assurance à l'horizon d'un an. Pour rappel, la VaR indique la perte maximale probable que l'on pourrait subir avec un certain degré de certitude sur une période donnée. En d'autres termes, une ruine ne devrait avoir lieu dans l'année qu'en cas d'événement bicentenaire.

Il peut être évalué de deux façons : via une formule standard fournie par l'EIOPA avec un calibrage par défaut, ou à partir d'un modèle interne développé spécifiquement par et pour l'organisme. La formule standard propose une approche modulaire. La charge en capital est calculée pour chaque source de risque que l'on nomme module de base. Chaque module peut être divisé en plusieurs sous-modules, correspondant chacun à un choc à appliquer sur des inputs de notre modèle ALM. Les montants retenus sont les différences entre la NAV du scénario central et les NAV choquées.

$$SCR_{\text{sousmodule}} = \Delta NAV = NAV_{\text{centrale}} - NAV_{\text{choquée}}$$

Ils sont finalement agrégés par des matrices de corrélations pour obtenir des SCR par module puis un SCR total.

Une telle approche présente ses avantages et ses inconvénients. D'une part, l'investissement humain et financier nécessaire pour l'appliquer est relativement faible. Aussi, cet aspect standardisé facilite la comparaison entre les organismes d'assurance. D'autre part, le calibrage par défaut à appliquer peut ne pas refléter les spécificités du profil de risque de l'organisme, en plus de présenter un aspect très arbitraire. De telles approximations peuvent aboutir à un capital requis surestimé. C'est pourquoi les organismes les plus importants peuvent opter pour le modèle interne.

Il existe deux catégories de modèles internes : une approche par module inspirée par la formule standard mais avec un calibrage adapté (on parle alors de modèle interne partiel) ou une approche intégrée qui vise à reconstruire complètement la VaR à 99.5% d'actifs nets.

Au sein de l'approche intégrée, l'impact des risques est considéré dans son ensemble et non pas risque par risque. On retrouve alors la modélisation de toutes les interdépendances entre les différentes sources de risques.

L'utilisation d'un modèle interne est soumise à l'approbation de l'ACPR. Celle-ci contrôle notamment le calibrage pour s'assurer de l'équivalence avec une VaR à 99.5% à un an, de la qualité statistique du modèle (données fiables, hypothèses cohérentes, sensibilité du résultat sous contrôle, ...), et de la qualité de la documentation sur la conception, les hypothèses et les données qui doit être tenue à jour.

Cette méthode présente comme avantages notables de mesurer une image fidèle du profil de risque et de fournir des indicateurs plus nombreux et pertinents sur la santé financière de l'entreprise. Cependant, elle s'avère très lourde et coûteuse aussi bien humainement que financièrement ou informatiquement.

Dans ce mémoire, nous avons choisi de nous concentrer sur la prédiction des résultats obtenus via la formule standard. Nous allons donc revenir plus en détail sur les différents modules, et plus spécifiquement sur le module de marché.

Présentation de la formule standard

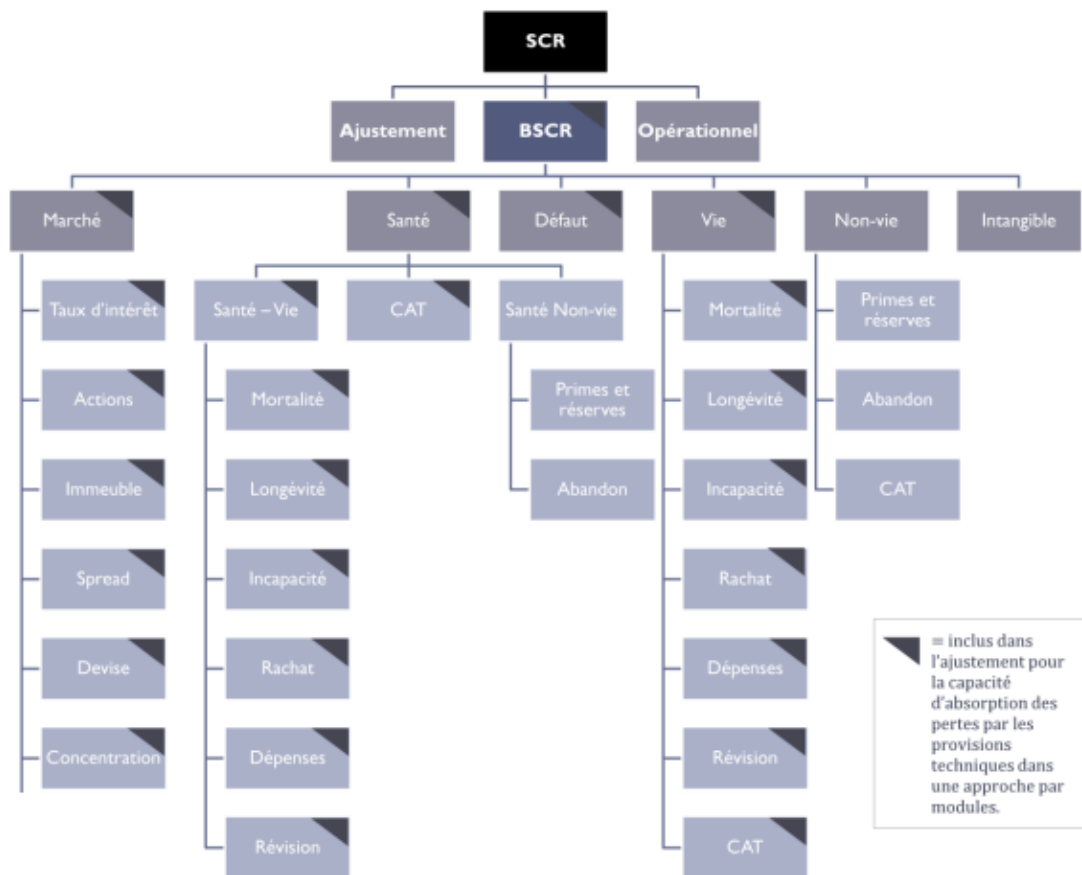


Figure 4 – Décomposition du SCR en modules et sous-modules

En formule standard, six modules sont compris dans le calcul du BSCR : Marché, Santé, Défaut (ou Contrepartie), Vie, Non-Vie et Intangible. Commençons par présenter brièvement chacun des modules.

Le module de SCR Marché représente le risque lié à la volatilité des marchés qui génèrent des variations de prix des instruments financiers. Ces risques auront donc dans un premier temps un impact sur l'actif de l'organisme, en affectant sa valeur de marché, son rendement et ses opportunités de réinvestissement. Mais le passif sera également affecté, de façon plus indirecte, via l'exercice des options et garanties des assurés notamment dans le cadre des contrats en euros. Celui-ci étant central pour notre mémoire, nous le détaillerons par la suite.

Le module de Défaut de contrepartie évalue le risque lié à la possibilité que les contreparties avec lesquelles la compagnie d'assurance traite ne puissent pas remplir leurs obligations contractuelles. Le risque de contrepartie peut provenir de différentes sources, telles que les transactions de produits dérivés, les opérations de réassurance, les prêts ou autres engagements financiers. Lorsqu'une compagnie d'assurance s'engage dans de telles transactions, elle est exposée au risque que la contrepartie ne puisse pas honorer ses engagements, ce qui peut entraîner des pertes financières.

Le module de souscription en Vie est conçu pour évaluer les risques spécifiques liés aux contrats d'assurance vie et aux engagements envers les assurés. Ce module prend en compte divers risques associés à la nature particulière des produits d'assurance vie, dont la tarification repose sur des hypothèses de mortalité, de longévité, de durée d'incapacité ou de morbidité, de frais ou de rachats.

Le module de souscription en Non-Vie représente le risque lié à l'activité de l'assurance non-vie (pour tout type d'entreprise du secteur) provenant directement des sinistres couverts ou d'une mauvaise gestion de l'activité. Ce risque comprend aussi le risque lié à l'incertitude sur le comportement de l'assuré (reconduction de contrat par exemple) et le risque Catastrophe (CAT) : risque de survenance d'une catastrophe d'origine naturelle (ex : tsunami, cyclone, tempête, ...) ou d'origine humaine (attaque terroriste par exemple).

Le module de souscription en santé représente des risques similaires à ceux considérés en vie et en non-vie mais spécifiquement pour les Lines of Business afférentes à la santé. Une distinction est effectuée entre la santé SLT qui incluent les Lines of Business pouvant être traitées via des méthodes similaires à celles utilisées pour le SCR Vie et la Santé non-SLT ou assimilable à la non-vie (Frais médicaux, indemnisation des travailleurs, protection du revenu, ...).

Le module Intangible correspond au risque de baisse de la valeur des actifs incorporels présents dans le bilan économique, comme par exemple la marque de l'entreprise (ou Goodwill).

Les résultats obtenus pour chacun de ces modules de risques élémentaires sont ensuite agrégés via la formule suivante :

$$BSCR = \sqrt{\sum_{i,j} (Corr_{ij} \times SCR_i \times SCR_j)} + SCR_{intangibles}$$

avec les coefficients de corrélations présentés dans la matrice ci-dessous :

	SCR Marché	SCR défaut	SCR vie	SCR santé	SCR non vie
SCR Marché	1	0,25	0,25	0,25	0,25
SCR défaut	0,25	1	0,25	0,25	0,5
SCR vie	0,25	0,25	1	0,25	0
SCR santé	0,25	0,25	0,25	1	0
SCR non vie	0,25	0,5	0	0	1

Tableau 1 – Coefficients de corrélation pour l'agrégation des modules de SCR

Pour obtenir le SCR, il faut prendre en compte les effets d'absorption des provisions techniques et des impôts différés : la somme de ces éléments correspond au montant d'ajustement à déduire du BSCR.

Cela reflète d'une part la capacité d'absorption des pertes par provisions techniques et par les impôts différés. En effet, en cas de perte, c'est un moyen pour l'assureur de transférer une partie des pertes aux assurés et à l'État. La section 9 de la directive Solvabilité II et plus spécifiquement les articles 206 et 207 détaillent les calculs.

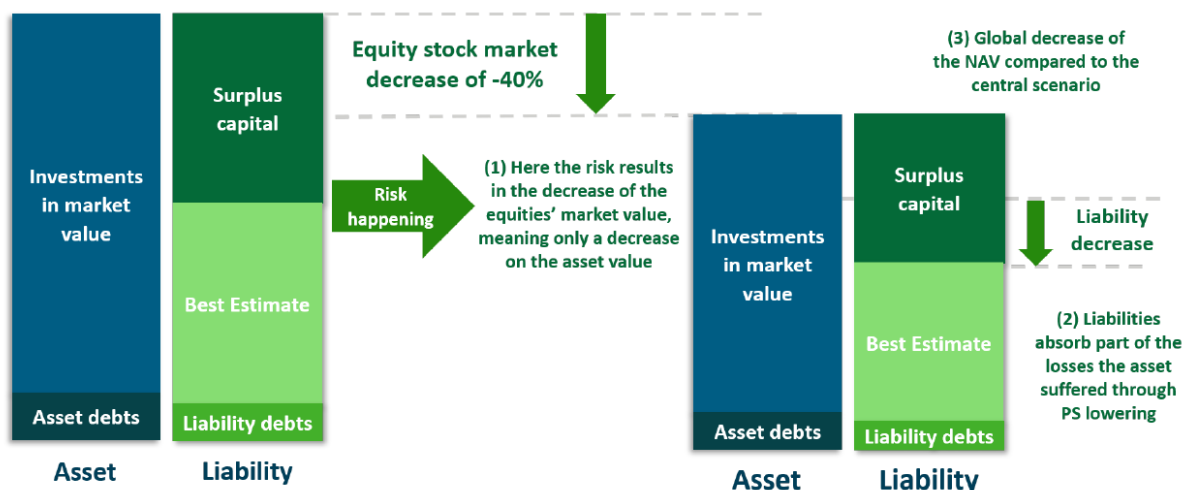


Figure 5 – Exemple d'absorption des pertes par les provisions techniques

La capacité de l'assureur à partager ses pertes avec les assurés provient de sa capacité à gérer la participation aux bénéfices distribuée chaque année. Si, en théorie, ce levier permet à l'assureur de partager une perte avec les assurés, cette capacité à partager les pertes est limitée par la présence de taux techniques, de taux minimums garantis ou même du mécanisme réglementaire de participation aux bénéfices.

L'ajustement par les impôts différés suit la même logique. Lorsqu'une entreprise réalise une perte, cela peut donner lieu à un crédit d'impôt. Ensuite, on peut considérer que cette perte n'impactera les fonds propres de l'assureur que dans la mesure de ce montant de perte net d'impôts. Dans ce cas, une exigence de capital peut être partiellement absorbée par l'impôt en appliquant au SCR le taux d'imposition auquel l'entreprise est soumise. Cependant, afin de pouvoir justifier une telle capacité d'absorption via les impôts, l'assureur doit démontrer que le crédit d'impôt généré par la perte peut être attribué à des impôts futurs. En d'autres termes, il faut justifier que la perte peut être utilisée pour générer une économie d'impôt pour l'assureur.

Détermination du MCR

Le MCR (ou *Minimal Capital Requirement*), quant à lui, représente le montant minimal de fond propre à détenir en dessous duquel l'intervention de l'autorité de contrôle est quasi-systématique. Son calcul est défini dans les articles 248 à 251 de la norme. Il correspond au maximum entre un montant plancher (compris entre 1M€ et 3,2M€ selon le type d'activité de l'organisme d'assurance) et une valeur comprise entre 25% et 45% du SCR.

e. Retour sur le SCR Marché

Le SCR Marché est divisé en 6 sous-modules :

- Le risque de taux : Ce risque est lié à l'évolution de la courbe des taux. Celui-ci consiste à appliquer des coefficients de majoration ou de minoration à la courbe des taux sans risques en fonction de la maturité. Les chocs à appliquer aux maturités non explicitement mentionnées peuvent être déduits par interpolation linéaire. Le règlement précise que l'écart de taux lors du choc à la hausse doit être d'au moins 1%, tandis que lors du choc à la baisse, les taux négatifs ne sont pas affectés. Le choc générant la plus grosse perte de fonds propres est retenu. Concrètement, ce

changement brutal des taux impacte à l'actif la valeur de marché des instruments financiers sensibles aux taux (obligations, swaps, ...) et au passif le facteur d'actualisation des flux.

Maturité (en années)	Hausse	Baisse
1	70%	75%
2	70%	65%
3	64%	56%
4	59%	50%
5	55%	46%
6	52%	42%
7	49%	39%
8	47%	36%
9	44%	33%
10	42%	31%
11	39%	30%
12	37%	29%
13	35%	28%
14	34%	28%
15	33%	27%
16	31%	28%
17	30%	28%
18	29%	28%
19	27%	29%
20	26%	29%
90	20%	20%

Tableau 2 – Coefficients de chocs appliqués à la courbe des taux centrale pour le SCR de taux

Une fois ces chocs appliqués à la courbe, deux courbes choquées sont obtenues et serviront de base au GSE pour la génération de scénarios stochastiques.

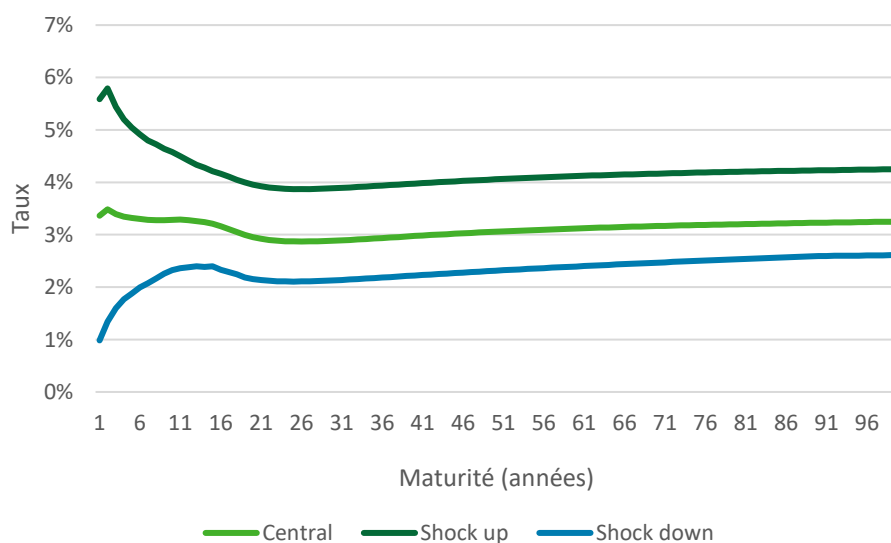


Figure 6 – Courbe de taux centrales et choquées au 31/12/2022

- Le risque actions : Ce sous-module concerne les produits sensibles aux variations de cours des actions. Un coefficient de minoration est appliqué à la valeur de marché des actions et autres actifs liés au marché actions (options, fonds actions, ...). Ce coefficient dépend notamment du type d'action, parmi lesquels on distingue :
 - les actions type 1, cotées sur un marché régulé dans un pays de l'OCDE ou de l'EEE
 - les actions type 2, non cotées ou cotées dans un pays hors de l'OCDE ou de l'EEE

La diminution du cours de l'action à appliquer est de 39% pour les actions de type 1 et de 49% pour les actions de type 2. Précisons que le choc n'est que de 22% pour les placements en actions dans des entreprises liées et de nature stratégique. À ce premier choc s'ajoute un ajustement symétrique, permettant de compenser les effets cycliques du marché. Celui-ci est compris entre -10% et 10% :

$$AS = \frac{1}{2} \left(\frac{CI - AI}{CI} - 8\% \right)$$

Avec :

- *CI* le montant de l'indice actions au moment du calcul
- *AI* la moyenne pondérée des niveaux quotidiens de l'indice sur les trois dernières années.

A la date de l'écriture de ce mémoire, celui-ci est d'environ 2%.

Les actions de chaque type sont choquées séparément, puis les impacts sur les fonds propres S2 sont agrégés via la formule suivante :

$$SCR_{equity} = \sqrt{SCR_{type1}^2 + 2 \times 0,75 \times SCR_{type1} \times SCR_{type2} + SCR_{type2}^2}$$

- Le risque immobilier : Le concept est similaire à celui du risque actions. Le montant retenu est la baisse de l'actif net à la suite d'une diminution de 25% de la valeur de marché des actifs immobiliers.
- Le risque de spread : Rappelons d'abord que le spread correspond à l'écart entre le rendement d'une obligation et le taux sans risque pour une même maturité. Cet écart rémunère le risque pris en acquérant une obligation potentiellement moins liquide et avec un risque de défaut plus important. Le risque de spread représente donc le risque d'une hausse soudaine de l'incertitude liée à la détention de certaines obligations, prêts, positions de titrisation ou dérivés de crédits. La hausse du spread d'une obligation implique un coefficient d'actualisation plus élevé et donc une baisse de sa valeur de marché. La réglementation propose un tableau définissant le choc à appliquer à la valeur de marché des obligations en fonction de leur durée modifiée (ou sensibilité) et de leur notation.

Échelon de qualité de crédit		0		1		2		3		4		5 et 6	
Duration (dur_i)	$stress_i$	a_i	b_i	a_i	b_i	a_i	b_i	a_i	b_i	a_i	b_i	a_i	b_i
Jusqu'à 5 ans	$b_i \cdot dur_i$	—	0,9 %	—	1,1 %	—	1,4 %	—	2,5 %	—	4,5 %	—	7,5 %
Supérieure à 5 et inférieure ou égale à 10 ans	$a_i + b_i \cdot (dur_i - 5)$	4,5 %	0,5 %	5,5 %	0,6 %	7,0 %	0,7 %	12,5 %	1,5 %	22,5 %	2,5 %	37,5 %	4,2 %
Supérieure à 10 et inférieure ou égale à 15 ans	$a_i + b_i \cdot (dur_i - 10)$	7,0 %	0,5 %	8,5 %	0,5 %	10,5 %	0,5 %	20,0 %	1,0 %	35,0 %	1,8 %	58,5 %	0,5 %
Supérieure à 15 et inférieure ou égale à 20 ans	$a_i + b_i \cdot (dur_i - 15)$	9,5 %	0,5 %	11 %	0,5 %	13,0 %	0,5 %	25,0 %	1,0 %	44,0 %	0,5 %	61,0 %	0,5 %
Plus de 20 ans	$\min[a_i + b_i \cdot (dur_i - 20); 1]$	12,0 %	0,5 %	13,5 %	0,5 %	15,5 %	0,5 %	30,0 %	0,5 %	46,6 %	0,5 %	63,5 %	0,5 %

Tableau 3 – Coefficients de choc à appliquer à la valeur de marché des obligations en fonction de la maturité et la sensibilité pour le SCR de spread

Les obligations ou prêts pour lesquels une évaluation de crédit par un Organisme Externe d'Évaluation du Crédit n'est pas disponible, et qui n'ont pas fourni de sûreté conforme aux critères énoncés à l'article 214 du règlement délégué, se voient attribuer un facteur de stress en fonction de leur sensibilité et durée conformément au tableau suivant :

duration (dur_i)	$stress_i$
Jusqu'à 5 ans	$3\% \cdot dur_i$
Supérieure à 5 et inférieure ou égale à 10 ans	$15 + 1,7\% \cdot (dur_i - 5)$
Supérieure à 10 et inférieure ou égale à 20 ans	$23,5\% + 1,2\% \cdot (dur_i - 10)$
Plus de 20 ans	$\min(35,5\% + 0,5\% \cdot (dur_i - 20); 1)$

Tableau 4 - Coefficients de choc à appliquer à la valeur de marché des obligations non notées

Il est important de souligner que les obligations souveraines ne sont pas affectées par le choc de spread.

- Le risque de concentration : Ce risque vient du fait que la volatilité d'un portefeuille augmente avec sa concentration, que ce soit sur un même émetteur, un même domaine d'activité ou une même zone géographique. Le sous-module de concentration ne concerne que le risque de concentration sur un même émetteur. Il consiste en l'application d'un choc g_i sur la valeur de marché de l'excédent d'actifs issus d'un même émetteur.

Notation	0	1	2	3	4	5	6
g_i	12%	12%	21%	27%	73%	73%	73%

Tableau 5 – Choc appliqué à la valeur de marché de l'excédent en fonction de la notation

Cette exposition en excès est définie à partir d'un seuil CT_i fixé selon l'échelon de qualité moyen des actifs visés, de même que le choc à appliquer sur cet excédent :

$$XS_i = \text{Max}(0; E_i - CT_i \times \text{Assets})$$

Avec :

- XS_i l'exposition en excès
- E_i l'exposition à la signature unique i
- CT_i le seuil relatif d'exposition en excès
- Assets l'assiette totale d'actifs considérés pour le calcul du sous-module

Moyenne pondérée des échelons de qualité de l'exposition sur signature unique i	0	1	2	3	4	5	6
Seuil relatif d'exposition en excès CT_i	3%	3%	3%	1.5%	1.5%	1.5%	1.5%

Tableau 6 – Seuil relatif d'exposition en excès en fonction de la notation pondérée

Les actifs dont l'émetteur est la BCE ou un Etat avec une notation de 0 ou 1 ne sont pas sensibles à ce choc. De plus, le choc est réduit pour les actifs issus d'Etat noté 2 ou plus.

Le montant retenu pour ce sous-module correspond à $\sqrt{\sum_i (XS_i \times g_i)^2}$.

- Le risque de change : ce sous-module consiste en l'application d'une hausse ou d'une baisse du taux de change de 25% entre l'euro et chacune des devises présentes. Le choc retenu est le plus important suite à la hausse ou la baisse du taux. Le montant total, quant à lui, correspond à la somme des montants retenus pour chaque devise.

Une fois que tous les sous-modules ont été calculés, leurs résultats sont agrégés via une matrice de corrélation pour obtenir le SCR marché total.

i \ j	Taux d'intérêt	Actions	Actifs immobiliers	Marge	Concentration	Devise
Taux d'intérêt	1	A	A	A	0	0,25
Actions	A	1	0,75	0,75	0	0,25
Actifs immobiliers	A	0,75	1	0,5	0	0,25
Marge	A	0,75	0,5	1	0	0,25
Concentration	0	0	0	0	1	0
Devise	0,25	0,25	0,25	0,25	0	1

Le paramètre A est égal à 0 lorsque l'exigence de capital pour risque de taux d'intérêt fixée à l'article 165 est l'exigence de capital visée au point a) dudit article. Dans tous les autres cas, le paramètre A est égal à 0,5.

Tableau 7 – Coefficients de corrélation utilisés pour l'agrégation des sous-modules du SCR marché

Dans ce mémoire, nous nous concentrerons sur la prédiction des sous-modules taux, actions, immobilier et spread.

III/ PRINCIPES DE MODÉLISATION ALM

Nous avons présenté le cadre réglementaire de Solvabilité II et la méthode de calcul à appliquer pour calculer le SCR marché en formule standard. Cependant, la détermination des bilans centraux et choqués nécessaires requiert l'utilisation d'un modèle de gestion actif/passif capable de représenter l'ensemble des interactions complexes pouvant affecter le résultat.

1. Evaluation des actifs

Selon l'article 75 de la Directive Solvabilité II, « les actifs sont valorisés au montant pour lequel ils pourraient être échangés dans le cadre d'une transaction conclue, dans des conditions de concurrence normales, entre des parties informées et consentantes ». Les actifs sont donc évalués en valeur de marché, soit en obtenant directement les données sur les marchés financiers, ou, si celles-ci ne sont pas accessibles ou existantes, en reconstituant les flux financiers de notre actif à partir de différents actifs au prix connu. Cette approche est dite *market consistent*, elle repose sur l'hypothèse d'absence d'opportunité d'arbitrage, c'est-à-dire la possibilité de faire un gain de façon certaine sans capital initial.

Dans notre version de modèle, les actifs sont représentés par des model points regroupant des actifs aux caractéristiques similaires. Les informations essentielles dont nous disposons sur nos model points se résument à la valeur de marché et la valeur comptable pour nos actions, obligations et immobilier, ainsi que les maturités, coupons, nominaux, notations et type d'émetteur (*Government* ou *Corporate*) pour nos obligations. Par mesure de simplification, nous ne considérerons que des obligations à taux fixe.

2. Evaluation des passifs

Toujours d'après l'article 75 de la directive, les passifs doivent être valorisés dans les mêmes conditions que l'actif. Le calcul du *Best Estimate* doit ainsi faire appel à des hypothèses statistiques et actuarielles pertinentes. En présence d'options et de garanties dans le contrat d'assurance, une approche stochastique doit être adoptée afin de prendre en compte le caractère aléatoire des risques. En effet, l'environnement économique a un impact direct sur le comportement des assurés via la participation aux bénéfices.

Le coût de ces options et garanties est difficilement calculable via une formule fermée ou une approximation par un portefeuille d'actif, c'est pourquoi une approche de type Monté Carlo est nécessaire.

Les données sur notre portefeuille d'épargne sont également regroupées sous la forme de model points. Les principales caractéristiques permettant de regrouper les contrats d'assurance au sein d'un model point sont :

- L'âge
- Le sexe
- La date de souscription
- Le nombre de contrats

- Le montant moyen d'encours (ou Provision Mathématiques) au début de la projection
- Le taux minimum garanti
- Les primes programmées

A ces model points sur les assurés s'ajoutent également des hypothèses de frais et de chargements pour les commissions et la gestion et acquisition des contrats.

3. Le générateur de scénarios économiques

Comme évoqué précédemment, calculer la meilleure estimation de nos engagements envers les assurés requiert une approche modélisant les aléas du marché et les comportements des assurés qui en découlent. C'est ici qu'intervient le Générateur de Scénarios Economiques (GSE).

Un GSE a pour objectif de modéliser les comportements des marchés financiers. Il est donc essentiel dans la mesure où le résultat de l'assureur dépend directement de l'évolution de ses actifs et des taux d'intérêts. Parmi les variables modélisées on retrouve :

- Les taux d'intérêts, indexés ou non sur l'inflation
- L'inflation
- Les rendements des actions et de l'immobilier

Dans le cadre de Solvabilité II, le GSE est structuré selon une approche « risque neutre ». Dans l'univers risque neutre, il n'y a pas de rémunération supplémentaire pour le risque, tous les actifs rapportent en moyenne le taux sans risque. Seule la volatilité des actifs a donc un impact sur les variations de résultats.

La modélisation stochastique des taux et des indices n'étant pas centrale dans le but de ce mémoire, nous ne détaillerons pas celle-ci.

4. Hypothèses assurés et assureur

Au-delà des données sur l'actif, le passif et l'environnement économique, pour qu'une projection soit complète, il faut également anticiper les comportements aussi bien de l'assureur que des assurés.

D'abord, les hypothèses sur les assurés portent notamment sur la sinistralité (dans notre cas la mortalité), les rachats conjoncturels et les arbitrages.

Pour la sinistralité, il n'existe pas de table règlementaire obligatoire à utiliser dans le cadre de la modélisation. Les assureurs peuvent utiliser des tables d'expériences basées sur leur portefeuille ou des tables règlementaires – ces dernières étant potentiellement moins adaptées à une vision Best Estimate dans la mesure où elles incluent déjà une marge de prudence. Dans tous les cas, l'organisme doit justifier que les tables utilisées sont adaptées à son risque.

Le comportement des assurés en matière de rachats (partiels ou totaux) et d'arbitrages (euros vers UC ou vice versa) peuvent être modélisés en distinguant deux types de phénomènes : les phénomènes

structurels, liés à la nature des produits d'assurance, et les phénomènes conjoncturels, qui sont des conséquences de changement de l'environnement économique.

Les rachats structurels sont en partie liés aux avantages fiscaux des contrats d'assurance vie, notamment la défiscalisation des rendements après 8 ans. Une des méthodes souvent employées pour mesurer ces rachats est d'observer la moyenne mobile des rachats sur une période déterminée, en fonction de l'ancienneté des contrats.

Les rachats et arbitrages conjoncturels, quant à eux, peuvent avoir différentes causes. La principale reste l'insatisfaction des assurés. En effet, ces derniers comparent les taux servis par l'assureur à des taux concurrentiels de produits similaires (livret A, obligations d'état, autres organismes d'assurance, ...). Dans le modèle, la différence entre ce taux servi et le taux concurrentiel devient alors l'argument d'une fonction déterminant le taux de rachats dynamiques attendu. L'ACPR fournit une fonction définie ainsi :

$$RC(\Delta Tx) = \begin{cases} TxMax & \text{si } \Delta Tx \leq \alpha \\ TxMax \times \frac{\Delta Tx - \beta}{\alpha - \beta} & \text{si } \alpha \leq \Delta Tx \leq \beta \\ 0 & \text{si } \beta \leq \Delta Tx \leq \gamma \\ TxMin \times \frac{\Delta Tx - \gamma}{\delta - \gamma} & \text{si } \gamma \leq \Delta Tx \leq \delta \\ TxMin & \text{si } \Delta Tx \geq \delta \end{cases}$$

Avec :

- $\Delta Tx = \text{Taux servi} - \text{Taux référence}$;
- $TxMax$: taux maximum de rachats conjoncturels ;
- $TxMin$: taux minimum de rachats conjoncturels ;
- α, β, γ et δ les différents seuils de la loi.

L'ACPR a donnée dans les Orientations Nationales Complémentaires aux Spécifications Techniques pour l'exercice 2013 un plafond maximum et minimum pour les paramètres des lois de rachats conjoncturels. Voici les valeurs retenues dans notre modèle :

- $TxMax = 30\%$
- $TxMin = -5\%$
- $\alpha = -5\%$
- $\beta = -1\%$
- $\gamma = 1\%$
- $\delta = 3\%$

Et voici la loi représentée graphiquement :

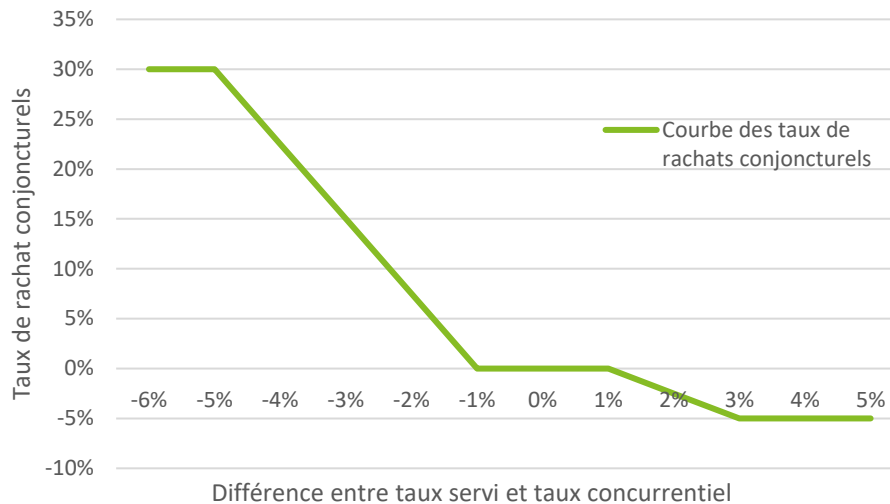


Figure 7 – Loi de rachats conjoncturels

Une logique similaire existe pour les arbitrages, avec une tendance à comparer les performances des contrats euro et UC pour déterminer de potentiels arbitrages.

Dans notre cas d'étude, nous ne modéliserons qu'un produit euro. En effet, un produit UC ne présente pas les mêmes options et garanties, ce qui implique que l'impact des variations de l'environnement économique sont plus aisément prévisible. Aussi, par simplification, nous ne modéliserons pas d'arbitrage.

Pour l'assureur, nous souhaiterons modéliser les montants initiaux de réserves ainsi que les futures décisions de gestion, telles que la gestion de la participation aux bénéficiers (détermination du taux cible), la stratégie d'allocation d'actif ou encore la stratégie de réinvestissement. Dans le cadre de la modélisation, celles-ci doivent être cohérentes avec la stratégie actuelle de l'assureur ou des évolutions prévues.

Est d'abord modélisée la stratégie de participation aux bénéficiers via le taux servi. Comme nous l'avons vu, ce taux cible détermine quelle part des profits financiers (au-delà de 85%) doit être reversée aux assurés, et a donc un impact direct sur le comportement de ces derniers. Dans le modèle, ce taux est calculé comme un pourcentage du TME (taux moyen d'emprunt d'Etat) pondéré par un coefficient de lissage pour éviter des écarts de rendements trop brutaux d'une année à l'autre. Est également définie la politique de reprise et de dotation de la Provision pour Participation aux Bénéficiers.

Ensuite, le modèle doit représenter la stratégie d'allocation d'actif, c'est-à-dire, la composition de notre portefeuille selon le type d'actif. Cette allocation peut être définie en valeur comptable ou en valeur de marché, fixe ou dynamique au cours du temps. Notre modèle utilise une stratégie fixe prédéfinie : la stratégie d'allocation est définie via une table Prophet au sein de laquelle l'utilisateur indique pour chaque pas de temps le pourcentage d'obligations, d'actions, d'immobilier et de *cash* à détenir après réinvestissements.

Concernant le portefeuille obligataire, notre modèle ne permet cependant pas de dynamiser la stratégie de réinvestissement en fonction du type d'obligation. Afin d'obtenir le pourcentage d'obligations souhaité, le modèle puise au sein d'une table définissant par avance un portefeuille obligataire disponible à l'achat. Ce portefeuille contient donc des obligations classiques dans des proportions fixes et définies en début de simulation.

5. Fonctionnement général du modèle

Ayant ainsi présenté l'ensemble des inputs nécessaire à la modélisation actif/passif, nous pouvons rentrer plus en profondeur dans le fonctionnement du modèle en lui-même.

Nous disposons à Deloitte d'un modèle ALM codé sur Prophet. Le processus de projections peut être décomposé en quatre phases principales :

- La définition des hypothèses et la construction des tables de données en entrée
- L'importation des données et le calibrage du modèle ALM
- La projection des cash-flows sur la durée des contrats
- Le reporting des indicateurs souhaités pour analyse

Pour être plus précis quant au fonctionnement de la projection, le pas de temps des calculs est annuel, et il est considéré que les prestations interviennent en milieu d'année et les revalorisations en fin d'année. Le détail de ces étapes est présenté dans le diagramme suivant :

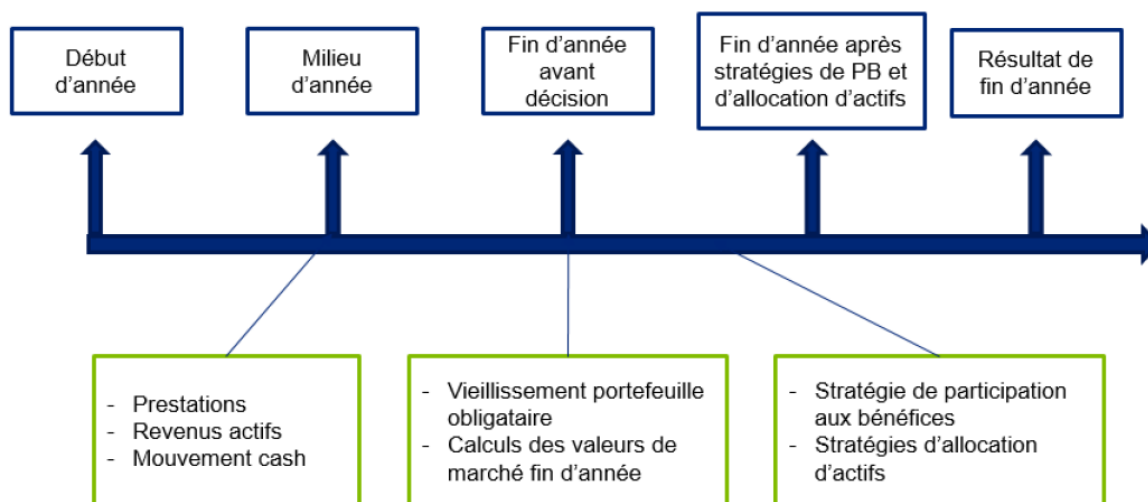


Figure 8 – Déroulement d'une année de projection au sein du modèle ALM

a. Fonctionnement de l'algorithme de réinvestissement

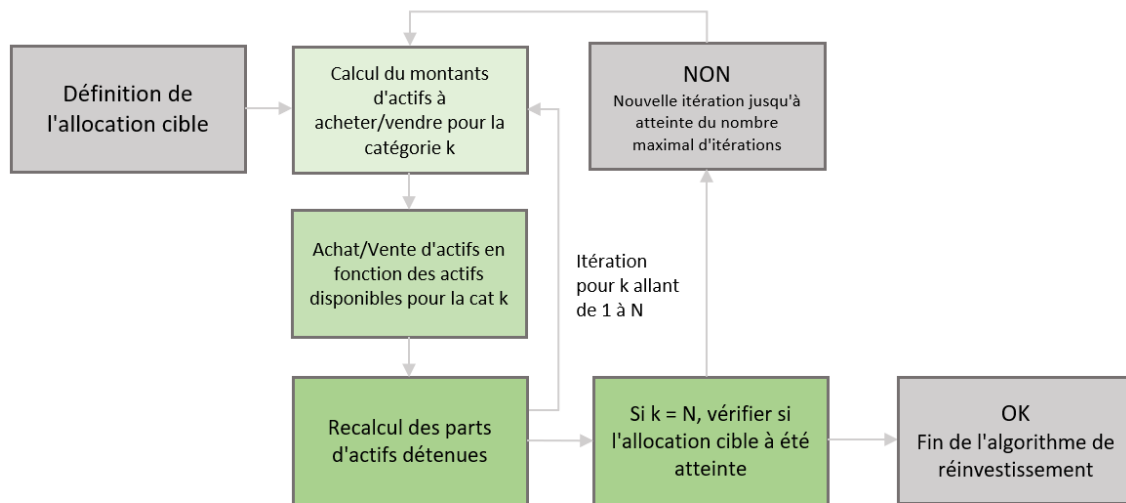


Figure 9 – Déroulement de l'algorithme de réinvestissement

Le modèle effectue jusqu'à un nombre maximum d'itérations (dans notre cas 25) au cours desquelles il va séquentiellement ajuster chacune des N classes d'actifs pour atteindre l'allocation cible. Pour chaque classe d'actif, il vérifie si la part détenue est conforme à la stratégie d'investissement. Dans le cas contraire, il calcule le montant d'actif à acheter ou vendre pour rééquilibrer l'allocation. S'il y a nécessité d'acheter, il puise dans la table d'obligations à disposition dans des proportions prédéfinies. Il procède ainsi classe par classe d'actif, en recalculant à chaque fois la part détenue par rapport au portefeuille obtenu après ajustement des classes d'actifs précédentes, jusqu'à atteindre la trésorerie, traitée en dernier.

b. Méthode de *flexing*

Nous utilisons des méthodes de *flexing* pour la modélisation actif/passif. Nous allons présenter cette méthode et ses avantages et inconvénients.

Le *flexing* repose sur une projection en deux temps : tout d'abord, le passif est projeté seul pour une trajectoire déterministe, à partir des Model Points présentés plus tôt et d'hypothèses de mortalité, de rachats structurels, de frais et de chargements sur les contrats. Seul le TMG est pris en compte pour la revalorisation en l'absence de données sur la production financière. Les flux projetés sont synthétisés dans une table (intitulée DET_CF sur Prophet) qui sera utilisée comme input pour la modélisation actif/passif. Cette table contient notamment des projections de la provision mathématique, du nombre de contrats, du montant de primes reçu, des différentes prestations pour décès ou rachats, et des frais et chargements.

	prod_name	IDX_CFLOW	VARNAME	2022	2023	2024	2025
1	A_EURO	1	NO_POLS_IF	4000	3862.258409	3727.62972976132	3595.92322036236
2	A_EURO	1	MATH_RES	173088478.552	170198863.193826	167254320.060716	164250937.260329
3	A_EURO	1	PREM_INC	0	1700000	1646266.081675	1594029.61114883
4	A_EURO	1	NET_PREMS	0	1666000	1613340.7600415	1562149.01892585
5	A_EURO	1	SUM_ASSURED	0	0	0	0
6	A_EURO	1	FV_SUM_ASSD	173088478.552	170198863.193826	167254320.060716	164250937.260329
7	A_EURO	1	FV_NET_PREM	0	0	0	0
8	A_EURO	1	FV_ANNUIITY	0	0	0	0
9	A_EURO	1	FEED_AMOUNT	0	1488081.27937843	1456968.07351345	1425227.21967891
10	A_EURO	1	DTH_OUT	0	788627.309293624	849480.029778944	916796.693096992
11	A_EURO	1	MAT_OUT	0	0	0	0
12	A_EURO	1	SURR_OUT	0	5255069.32825925	5165371.93688519	5073962.34589513
13	A_EURO	1	ANN_OUT	0	0	0	0
14	A_EURO	1	RIDERC_OUTGO	0	0	0	0
15	A_EURO	1	P_SURR_OUT	0	0	0	0
16	A_EURO	1	INVEXP	0	0	0	0
17	A_EURO	1	REXP	0	859410.427845525	846243.232068867	832892.499319625
18	A_EURO	1	ICOMM	0	34000	32925.3216335	31880.5922229766
19	A_EURO	1	RCOMM	0	0	0	0
20	A_EURO	1	IEXP	0	0	0	0
21	A_EURO	1	DET_MATH_RES_INT	0	1488081.27937843	1456968.07351345	1425227.21967891
22	A_EURO	1	CASHB_OUT	0	0	0	0
23	A_EURO	1	INT_BON_OUT	0	0	0	0
24	A_EURO	1	VAL_INT_PC	0	0	0	0

Tableau 8 – Extrait de la table DET_CF

Pendant la projection ALM, ces flux vont être ajustés, ou « flexés », pour prendre en compte les interactions avec l'actif, les décisions de gestions et l'environnement économique. Cette phase est essentielle dans la mesure où ces interactions déterminent le montant de participation aux bénéficiaires reversés aux assurés et donc le comportement de ces derniers via les rachats conjoncturels.

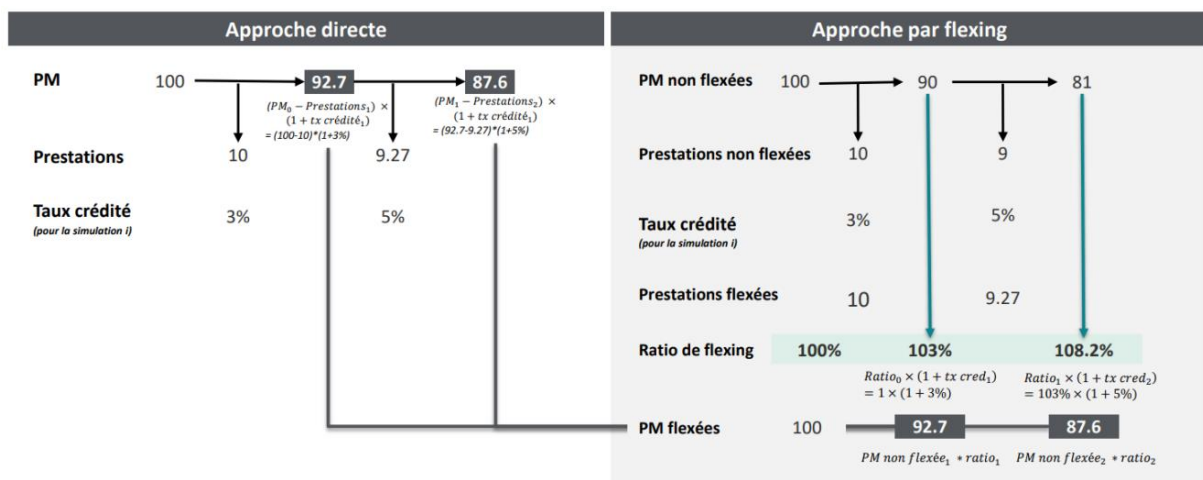


Figure 10 – Schéma simplifié de flexing

Le flexing présente comme avantage majeur une simplification des calculs et donc une réduction du temps de projection, au dépend d'une meilleure précision offerte par les modélisations intégrées.

Le modèle utilise deux ratios de flexing « de base », le premier pour représenter l'impact de la participation aux bénéficiaires et le second pour les rachats conjoncturels. Ces deux ratios sont recalculés puis appliqués aux cash flows déterministes à chaque pas de temps.

c. Une limite à considérer : la non-modélisation des fonds propres

Un élément essentiel à mentionner vis-à-vis de notre étude est la non-modélisation des fonds propres au sein de notre modèle actif/passif. Celui-ci ne modélise au passif que les engagements envers les assurés. En pratique, afin que le bilan en début de simulation reste équilibré, les actifs sont ajustés proportionnellement pour être égaux en valeur comptable au passif modélisé.

L'implication directe est que la différence entre l'actif modélisé et le *Best Estimate* obtenu correspond en fait à la PVFP (*Present Value of Future Profits*). Tout au long de ce mémoire, nous ne considérerons pas non plus la marge pour risque. Les sous-modules de SCR que nous calculons correspondent donc à la différence non pas entre les NAV mais entre les PVFP centrale et choquées.

Cette limite n'est pas considérée comme un obstacle à l'objectif fixé. Les méthodes développées par la suite restent valides si utilisées avec un modèle ALM modélisant les fonds propres.

IV/ ÉLABORATION DE LA MÉTHODE ET MISE EN PLACE DE L'OUTIL

Maintenant que le cadre de notre étude et les outils à notre disposition ont été présentés, nous pouvons aborder la partie essentielle de ce mémoire : la prédiction d'un SCR marché.

1. Rappel de l'objectif

Les sous-modules du SCR marché que nous souhaitons prédire correspondent à des différences entre une NAV centrale et une NAV choquée selon des modalités propres à chaque sous-module. Chacun de ces chocs affecte directement les données à l'actif ou les scénarios économiques en entrée du modèle ALM. Prédire un SCR marché équivaut donc à prédire les NAV dans le scénario central et pour chacun des sous-modules.

Nous optons donc pour des méthodes d'apprentissage supervisé prenant en entrée, pour un passif et certains paramètres fixés, des informations sur l'actif et le scénario économique, et aboutissant en sortie à l'obtention d'un BE et d'une PVFP. En d'autres termes, nous allons générer une DET_CF via le modèle Prophet déterministe, fixer les hypothèses assurés et assureurs dans le modèle actif/passif puis prédire le bilan en fonction du reste des *inputs*.

Une fois que nous sommes capables de prédire avec une précision suffisante un bilan déterministe, il en découle immédiatement le bilan stochastique qui correspond à la moyenne des prédictions effectuées sur le même portefeuille d'actif mais pour des scénarios économiques différents.

2. Génération des données

La première étape consiste à construire une base de données. Nous souhaitons ainsi générer une étendue de combinaisons *market values*/scénarios économiques assez vaste pour être capable de prédire l'impact de combinaisons de chocs réglementaires assez lourds, imposés par la formule standard, et de chocs plus subtils appliqués lors d'un test de sensibilité ou de quelque étude interne.

Pour ce faire, nous allons profiter de la fonctionnalité de calcul stochastique de Prophet. En temps normal, Prophet prend en entrée une table de scénarios économiques et, lorsqu'un run stochastique est lancé, calcule l'ensemble des indicateurs économiques désirés pour chacun de ces scénarios. Lors du *reporting*, l'ensemble de ces résultats sont agrégés en moyenne, permettant ainsi d'obtenir une projection stochastique conformément à la méthode de Monte-Carlo.

Nous pouvons tirer profit de fonctionnement séquentiel de Prophet, qui effectue l'ensemble de ces calculs scénario par scénario. Il nous suffit d'ajouter une fonction au modèle permettant, pour chaque simulation stochastique, de sortir l'ensemble des *inputs* ainsi que les outputs (BEL et PVFP) associés dans un fichier plat qui servira de base d'apprentissage. Il reste donc à définir quelles informations sur l'actif nous voulons retenir, et comment nous souhaitons les faire varier en entrée du modèle ALM pour couvrir l'ensemble des possibles.

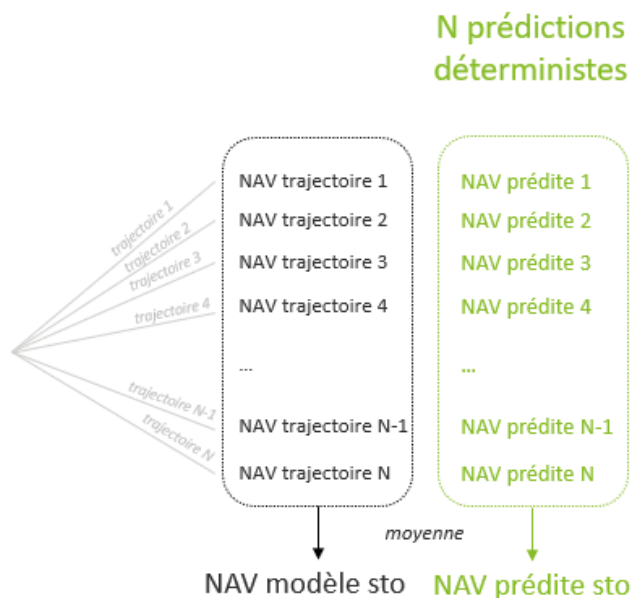


Figure 11 – Récapitulatif de la méthode pour l'obtention d'un bilan stochastique

a. Variabilisation de l'action et de l'immobilier

Dans le cas des actions et de l'immobilier, la seule métrique pertinente du point de vue du modèle et impactée par les stress tests ou les chocs SCR est la *market value* totale. C'est celle-ci qui va déterminer le montant de plus ou moins-values latentes et qui aura donc un impact sur le rendement et la stratégie de l'assureur.

Pour chaque simulation effectuée, des coefficients aléatoires sont tirés dans des bornes définies au préalable, puis appliqués aux valeurs de marché des actions et de l'immobilier. Nous utilisons une loi uniforme et paramétrons nos bornes de façon à être sûrs d'inclure les chocs SCR dans l'amplitude choisie.

Nous fixons donc les bornes entre 50% et 130% pour l'action et entre 60% et 120% pour l'immobilier.

b. Variabilisation des obligations

Être capable de variabiliser correctement le portefeuille obligataire est essentiel que ce soit pour la risque-neutralisation des obligations, pour les chocs de spread, ou pour le cas particulier de ces derniers qu'est le SCR de spread. Commençons par clarifier ces concepts.

La nécessité de « risque-neutraliser » réside dans le fait que sous Solvabilité II, les actifs doivent être évalués sous l'hypothèse de *market consistency*. En d'autres termes, dans un univers risque-neutre, leur valeur de marché dans le modèle doit être égale à celle observable sur le marché, tout en servant le taux sans risque.

Rappelons que la *market value* d'une obligation se calcule comme étant la somme des cash flows actualisés au taux de rendement actuariel.

$$MV = \sum_{t=1}^T \frac{CF_t}{(1+r)^t}$$

Avec :

- CF_t les *cash flows* au temps t (coupon ou coupon et nominal)
- r le taux de rendement actuariel
- MV la valeur de marché
- T la maturité résiduelle

Une des possibilités pour parvenir à la *market consistency* est d'implémenter dans le modèle un module de défaut. Celui-ci doit être calibré de façon à ce qu'en moyenne les obligations aient un rendement égal au taux sans risque, c'est-à-dire que l'on garantit l'absence d'opportunité d'arbitrage en s'assurant que toute rémunération du risque est associée à un risque de défaut équivalent.

$$MV = \sum_{t=1}^T \frac{CF_t \times (1 - p_t)}{(1 + r_{rf(t)})^t}$$

Avec :

- $r_{rf(t)}$ le taux sans risque pour une obligation zéro-coupon de maturité t
- p_t la probabilité de défaut au temps t

En l'absence d'un module de défaut, nous décidons de nous orienter vers une autre option consistant à ajuster directement les *cash flows*. Nous commençons par recalculer dans le modèle la valeur de marché de nos obligations actualisées au taux sans risque, puis nous appliquons le rapport $MV_{market} / MV_{recalculée}$ à l'ensemble de nos *cash flows*, c'est-à-dire au montant de coupon et au nominal.

$$MV_{recalculée} = \sum_{t=1}^T \frac{CF_t}{(1 + r_{rf})^t}$$

Et mécaniquement :

$$\sum_{t=1}^T \frac{CF_t \times \frac{MV_{market}}{MV_{recalculée}}}{(1 + r_{rf(t)})^t} = MV_{market}$$

Plutôt que de compenser la prime de risque avec un risque de défaut, nous choisissons donc d'éliminer simplement la prime de risque en diminuant proportionnellement nos coupons et nominaux.

Durant cette phase de risque-neutralisation, la valeur des coefficients appliqués est dépendante de la courbe des taux sans risque utilisée. Si nous voulons être capable de faire varier la courbe des taux risques en input de nos prédictions sans avoir à recommencer tout le processus de génération de données et d'apprentissage, nous devons variabiliser notre portefeuille obligataire. Ainsi, les chocs que celui-ci subit lors de la risque-neutralisation seront pris en compte par le modèle de machine learning.

Cette variabilisation répond également à notre besoin de pouvoir effectuer des chocs de spread sur notre portefeuille.

Le spread de crédit fait référence à l'écart entre le rendement d'un titre obligataire risqué et le taux sans risque. Ce spread est utilisé pour évaluer le niveau de risque perçu d'un émetteur de dette par rapport à un émetteur plus sûr. En d'autres termes, le spread représente la rémunération supplémentaire lors de la détention d'un titre obligataire risqué.

Pour obtenir le spread implicite de nos obligations, nous calculons leur taux de rendement actuariel (TRA) puis faisons la différence entre celui-ci et le taux sans risque. Le TRA correspond au taux permettant d'actualiser les *cash flows* programmés d'une obligation afin de retrouver son prix de marché.

Calculer le TRA requiert l'aide d'un solveur car il n'existe pas de formule fermée permettant de l'obtenir. Il existe cependant une formule permettant de l'approximer avec une précision satisfaisante :

$$TRA = \frac{C + \frac{N - MV}{T}}{\frac{1}{2}(N + MV)}$$

Avec :

- C le coupon
- N le nominal
- MV la valeur de marché
- T la maturité résiduelle

Ayant clarifié ce qu'est le spread de crédit, nous pouvons expliquer comment appliquer un stress test sur celui-ci. Un choc de spread correspond à une hausse soudaine du risque lié à la détention de certaines obligations, soit une perte de confiance envers leurs émetteurs. Etant plus risquées, le spread de crédit de ces obligations augmente, ce qui correspond mécaniquement à une baisse de son prix sur le marché (les *cash flows* étant actualisés à un taux plus important). Un choc de spread équivaut donc à un choc sur la *market value* avant risque-neutralisation, et donc en pratique à un abattement plus important des *cash-flows* si nous avons opté pour cette méthode de risque-neutralisation.

Nous pouvons remarquer que dans les deux cas, les mécanismes sont très similaires dans la mesure où ils correspondent à une modification des coupons et nominaux. En fin de compte, les seules informations prises en compte par le modèle et impactées par cette étape de risque-neutralisation sont le coupon et le nominal. Ce sont donc les deux inputs que nous avons décidés de variabiliser.

Il a fallu ensuite réfléchir à quelles données nous voulions utiliser dans le cadre de la conception des algorithmes de *machine learning*. Dans le cas des actions et de l'immobilier, l'ensemble de l'information était synthétisable dans une seule variable représentant leur valeur de marché. Pour les obligations, la complexité était beaucoup plus importante. La *market value* seule ne permettait pas de convoquer l'ensemble de l'information sur la structure du portefeuille obligataire, et de fait, deux portefeuilles présentant des *market values* identiques peuvent avoir des structures extrêmement différentes. La répartition de nos obligations parmi les différentes maturités et notations possibles, ainsi que l'importance relative du nominal et du coupon, sont autant de données que nous devons représenter et qui ont un impact significatif sur le résultat. Nous avons donc décidé de classer les *market values*, les montants de coupons et les nominaux associés à chaque obligation dans des *buckets* de notation X maturité.

Vient enfin la problématique de la randomisation des *inputs*. Nous savons que nous devons impacter les nominaux et les coupons, il reste à savoir comment. De nombreuses possibilités ont été explorées au cours de ce mémoire, avec plus ou moins de succès. Nous allons brièvement les présenter.

A l'origine, nous avons voulu adopter une approche agnostique : nous tirions aléatoirement un coefficient entre 60% et 110% par *bucket* de notation X maturité, que nous appliquions à l'ensemble des nominaux des obligations contenues, puis nous répétions l'opération pour les montants de coupon. La logique était simple : quel que soit le choc appliqué à notre portefeuille obligataire, celui-ci serait compris dans les bornes de notre base d'apprentissage. Cette stratégie s'est avérée avoir de multiples conséquences adverses.

La première, c'est qu'une fois la risque neutralisation effectuée, la notation n'a plus d'impact sur le déroulement des calculs du modèle. Choisir un coefficient aléatoire par *bucket* a donc pour effet de créer une « fausse » variance au sein des inputs obligataires, voire permettre des mécanismes de compensation entre différents *buckets* de même maturité. Nous avons donc décidé de tirer une valeur unique par année de maturité.

La seconde, c'est que choisir de randomiser nos variables à une maille aussi fine ne nous permettait pas d'avoir des variables suffisamment significatives pour les algorithmes de *machine learning*. Une solution immédiate faisant écho à notre problématique précédente fut d'agréger nos *buckets* par maturité. Cependant, l'algorithme peinait à considérer comme significative les variables représentant les coupons et ignorait simplement cette information. Un moyen d'y remédier fut de créer une variable synthétisant les coupons et nominaux à travers une estimation de la production financière obligataire annuelle :

$$prodoblign_t = N_t + \sum_{i \geq t} C_i$$

soit la somme des nominaux des obligations arrivées à maturité et des coupons des obligations encore détenues.

Enfin, cette approche s'avérait trop décorrélée de la réalité, dans la mesure où cette façon de choquer nos *inputs* aboutissait à des situations où certains *buckets* voyaient leurs coupons et nominaux presque coupés en deux tandis que les *buckets* directement adjacents voyaient ces derniers augmentés. Nous produisons donc des scénarios irréalistes et non pertinents vis-à-vis de nos objectifs, ce qui limitait la quantité de données réellement exploitable à notre disposition, en plus de forcer l'apprentissage des algorithmes sur des données anormales. Nous nous sommes donc penchés plus en détail sur la nature des chocs que nous souhaitons prédire. Voici les abattements moyens appliqués aux nominaux de nos obligations en fonction de leur maturité dans différents cas :

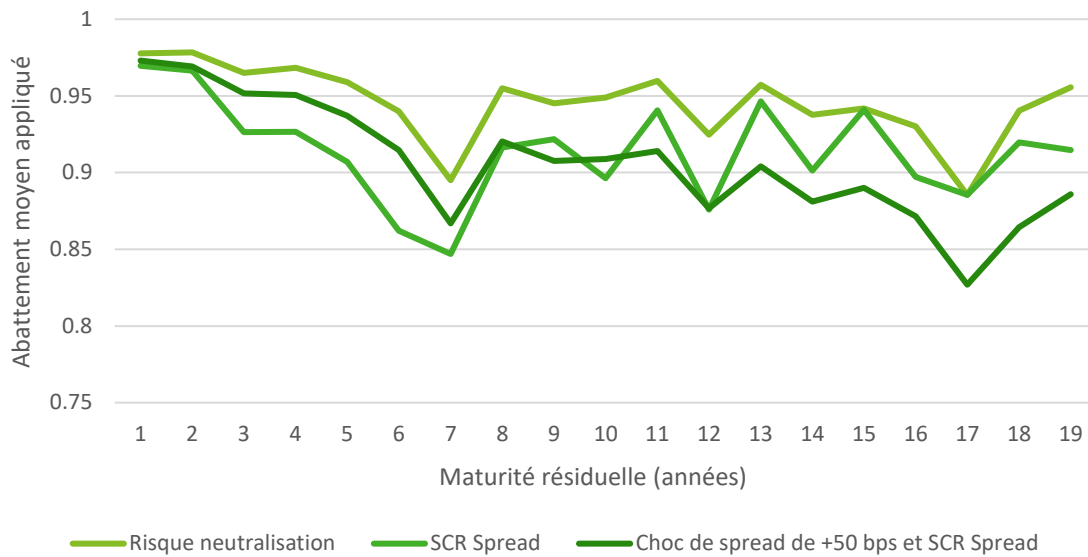


Figure 12 – Représentation des abattements moyens appliqués au nominaux dans différents scénarios en fonction de la maturité résiduelle

Plusieurs conclusions émanent de cette analyse. Tout d'abord, en voulant balayer l'éventail le plus large possible de combinaisons de chocs, nous avons au départ largement exagéré l'amplitude des bornes de nos tirages aléatoires. Ensuite, on observe qu'effectivement les abattements ne sont pas indépendants d'une maturité à l'autre ce qui fait sens : en général, on vise une certaine harmonie dans le portefeuille obligataire. Aussi, on remarque qu'appliquer un choc de spread de 50 bps sur l'ensemble du portefeuille aboutit à des abattements similaires à ceux du scénario central mais s'amplifiant pour les maturités plus longues.

Nous décidons donc de laisser de côté les tirages uniformes et de viser plutôt la construction d'un cône de diffusion. Ainsi, une valeur sera tirée aléatoirement pour représenter l'abattement des obligations de maturité 1 an, puis pour chaque année qui suit, une nouvelle valeur sera tirée puis additionnée à l'abattement précédent. Pour le coefficient initial, nous nous permettons une amplitude assez faible car l'actualisation ayant un effet presque négligeable, un choc de spread ou l'utilisation d'une courbe différente pour la risque-neutralisation ont un impact réduit. Nous choisissons donc d'initialiser nos coefficients avec un premier abattement entre 95% et 100%. Les différences observées entre les abattements appliqués sur deux années consécutives sont comprises entre -4% et +4%. Nous choisissons donc d'ajouter à chaque coefficient une valeur tirée uniformément entre -5% et 5%. Enfin, pour s'assurer de ne pas se retrouver avec des chocs absurdes, nous définissons un plafond à 100% et un plancher à 80%.

Afin de nous assurer que la masse de données générée inclut bien des coefficients similaires à ceux appliqués sur les coupons et nominaux, nous avons choisi de représenter les coefficients d'abattement appliqués au cours de la risque-neutralisation de notre portefeuille fictif et au cours du choc règlementaire pour SCR de spread. Sur le même graphique, nous présentons, parmi 5000 courbes générées selon la méthode indiquée ci-dessus, les trois courbes de coefficients d'abattements les plus proches en termes de RMSE de nos abattements observés dans le portefeuille fictif.

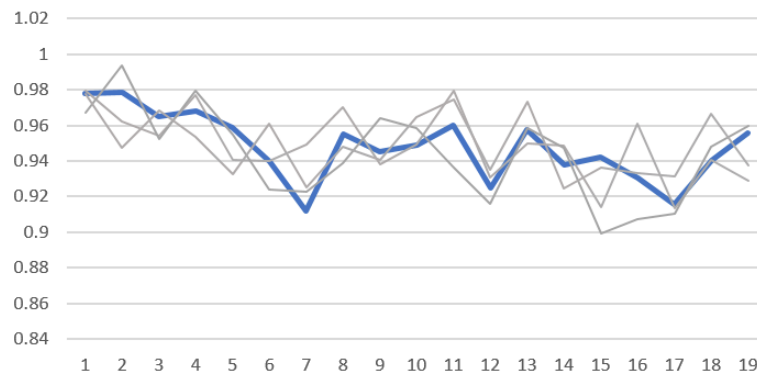


Figure 13 – Représentation de l'abattement des nominaux pour risque-neutralisation et de trois combinaisons d'abattements proches générés par notre processus

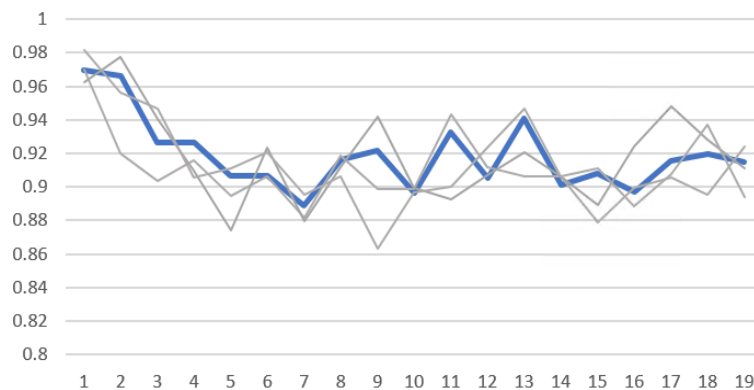


Figure 14 - Représentation de l'abattement des nominaux pour SCR de spread et de trois combinaisons d'abattements proches générés par notre processus

Dans les deux cas, nous constatons que notre nouvelle méthode de génération de chocs sur les coupons et nominaux produit bien des scénarios similaires et cohérents avec ce qui est observable en pratique. Nous limitons ainsi le nombre de scénarios incohérents introduisant un bruit superflu au sein de notre base de données pour nous concentrer sur des données crédibles.

c. Variabilisation des taux

La variabilisation des taux s'effectue via deux procédés. D'une part, le générateur de scénarios économiques produit un grand nombre de simulations aléatoires à partir de la courbe de taux centrale et d'autres paramètres. Cette masse de données permet déjà d'obtenir une grande variété de taux possibles au sein de notre base de données. Cependant, ne se servir que d'une seule courbe centrale produit peu de diversité pour les taux de faibles maturités, ce qui va à l'encontre de notre perspective d'effectuer des tests de sensibilité, et plus encore de prédire des chocs de taux règlementaires qui affectent particulièrement les premières maturités. C'est pourquoi, d'autre part, nous avons choisi un ensemble de courbes choquées obtenues par translations de la courbe centrale de l'EIOPA au 31/12/2022 et qui serviront de courbes centrales pour la génération de scénarios économiques.

Les courbes de taux sélectionnées à ce titre sont la courbe centrale de l'EIOPA avec VA au 31/12/2022 et des translations par ajustements *flat* de 25, 50, 75, 100, 125, 150 et 200 bips, additionnés et

soustraits. Ce sont donc 15 tables de scénarios économiques de 50 000 lignes chacune qui sont produites.

Nous obtenons ainsi 750 000 lignes d'inputs pour le *machine learning*, présentant chacune une combinaison différente de valeurs de marché, de production obligatoire et de taux, ainsi que le BEL et la PVFP qui lui sont associés.

3. Calibration du modèle de Machine Learning

a. Préparation des données (normalisation, sélection de variables)

Normalisation

L'étape suivante consiste à normaliser nos variables. En effet, nous nous retrouvons avec des variables étant pour certaines de l'ordre du million (*market values*, production financière), tandis que d'autres sont de l'ordre de 1 (taux, indices). Afin que l'algorithme accorde une importance équivalente à l'ensemble des variables, et pour réduire le temps de calcul, nous avons choisi d'effectuer une normalisation min/max afin de ramener toutes nos variables entre 0 et 1, et qu'ainsi les seules informations transmises soient les variations de leurs valeurs.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Séparation du data set

Ensuite, conformément à la méthodologie habituelle à un processus de *machine learning*, nous séparons les variables à expliquer (PVFP, BEL) des variables explicatives. Nous les stockons respectivement dans les data frames y et X.

Enfin, nous divisons le jeu de données en un jeu d'apprentissage et un jeu de test. Nous optons pour du 80%/20%.

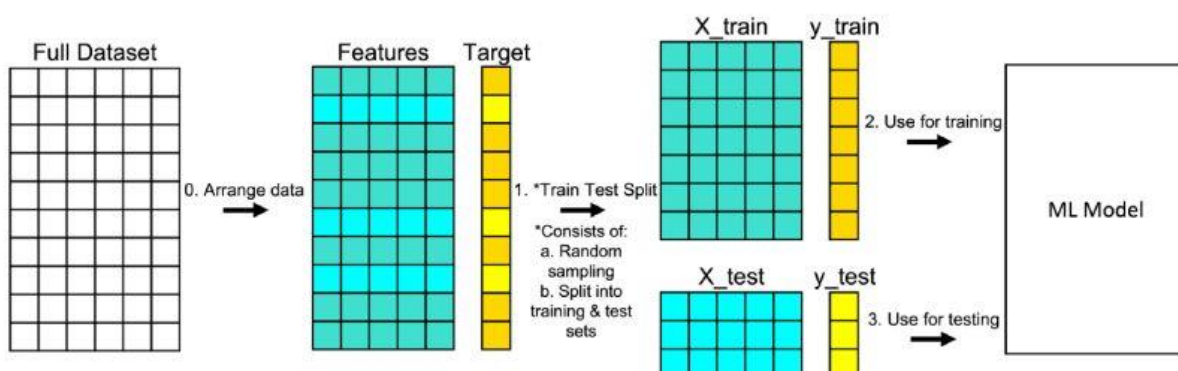


Figure 15 – Schéma récapitulatif des divisions du data set

Sélection de variables

La sélection de variables s'est opérée en deux parties.

La première est une pré-sélection manuelle des variables que nous laissons à l'algorithme. En effet, notamment dans le cas des obligations, nous avons un ensemble important de variables portant une

information très similaire. Pour rappel, à ce stade du projet, nous disposions de la *market value* totale des obligations, des *market values* de l'action et de l'immobilier, des nominaux et des coupons obligataires pour chaque année de maturité et de la production financière obligataire annuelle. Il s'est avéré que laisser l'ensemble de ces données à disposition des algorithmes pouvait fausser la qualité de prédiction : par exemple, nous avons illustré plus tôt pourquoi la *market value* totale des obligations était un mauvais indicateur de l'état de notre portefeuille obligataire ; néanmoins, cette variable reste fortement corrélée avec la PVFP et le BEL. De ce fait, les algorithmes utilisant des arbres de décision avaient tendance à lui donner une forte importance puis à ignorer les variables plus précises mais fortement corrélées (telles que les variables par année de maturité), ce qui limitait la captation des nuances du portefeuille.

Après de multiples tentatives, nous avons décidé de ne laisser que les variables de production financière obligataire comme représentantes du portefeuille obligataire.

Après cette première phase, il restait plus de 1300 variables, dont un très grand nombre de taux et d'indices. Nous avons donc laissé un modèle XGBoost calibré par défaut tourner sur un échantillon représentatif de 45 000 lignes afin de déterminer la *feature importance* de nos variables. La *feature importance* correspond au gain, c'est-à-dire à l'amélioration de la performance provoqué par l'ajout d'une variable explicative aux arbres de décisions. L'ajout de cette variable aboutit à la création de nouveaux nœuds et donc de branches plus précises.

Une fois cette étape menée à bien, nous avons fait tourner le modèle avec un nombre croissant de variables sélectionnées par ordre d'importance. Chaque *fitting* aboutit à un score, la RMSE, afin de pouvoir constater si celle-ci continue de décroître (démontrant ainsi une amélioration de la qualité de prédiction) ou si elle atteint un plateau. Une fois ce plateau atteint, nous pouvons retenir les variables sélectionnées.

Pour rappel, la RMSE ou *Root Mean Squared Error* (Racine de la moyenne des erreurs quadratiques) est une mesure statistique utilisée pour évaluer la précision d'un modèle de prédiction ou d'estimation par rapport aux valeurs réelles observées.

$$RMSE = \frac{1}{n} \sqrt{\sum (y_i - \hat{y}_i)^2}$$

Avec y_i les valeurs réelles et \hat{y}_i nos prédictions.

Plus la RMSE est faible, plus le modèle est en mesure de prédire avec précision les valeurs réelles.

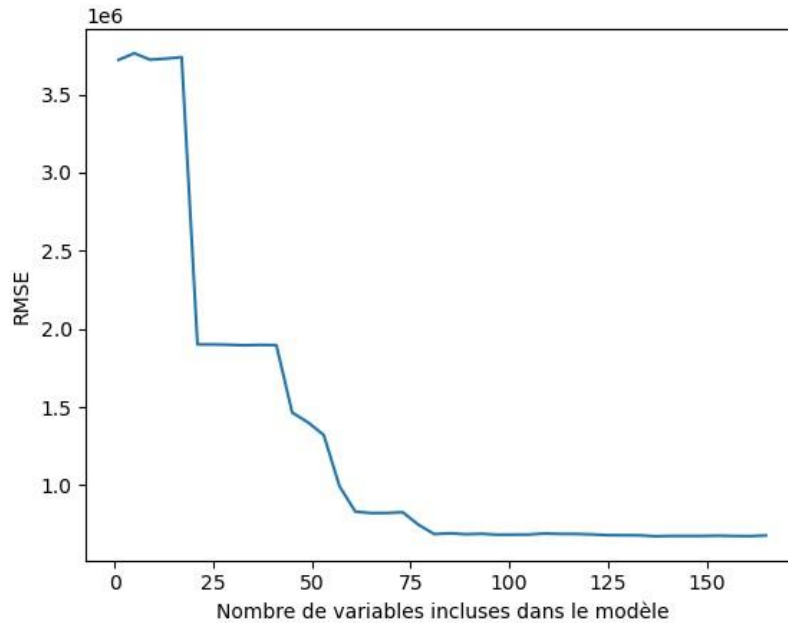


Figure 16 – RMSE du modèle de ML en fonction du nombre de variables retenues

Ce graphique présente la RMSE en fonction du nombre de variables retenues pour entraîner le modèle.

Sur 1732 variables, nous en gardons 1342 après agrégation, traitements et présélection, pour finalement n'en conserver qu'une centaine sans perte de performance et en limitant le risque de surapprentissage. Parmi elles, on retrouve nos *market values* actions et immobilier, les productions financières du portefeuille obligataire, et un ensemble d'indices actions, immobilier, de taux *spot* et *forward* jugés pertinents par notre algorithme. Voici les variables avec la plus grande *feature importance*.

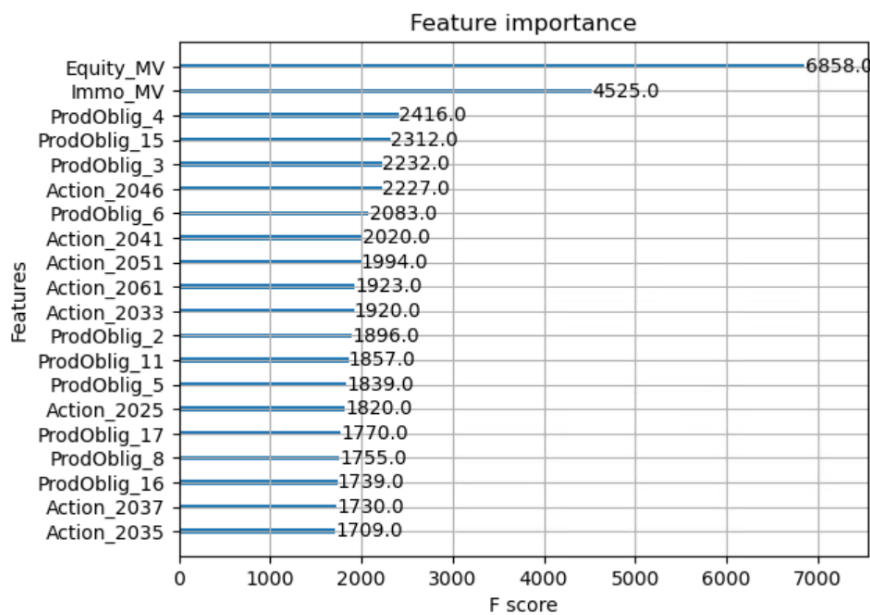


Figure 17 – Score de feature importance des variables les plus significatives retenues

Sans surprise, les variables avec la plus grande importance sont les valeurs de marché des actions et de l'immobilier. Ces deux variables représentent à elles-seules une part très importante du portefeuille

d'actifs et leurs variations ont un impact fort sur la PVFP. Viennent ensuite les variables obligataires et les indices actions.

Cette méthodologie est appliquée successivement pour l'algorithme de prédiction de la PVFP puis pour celui du BEL.

b. Choix du modèle

Il existe de nombreux algorithmes de *machine learning* capables de traiter notre problématique. Nous pouvons notamment mentionner le *Support Vector Machine*, le *Convolutional Neural Network* ou encore le *Random Forest*.

Les études préliminaires menées en interne dans un cadre plus restreint, notamment par Clément Gounel dans son mémoire ont montré qu'aucun d'entre eux ne présentait de meilleures performances que le *eXtreme Gradient Boosting*, ou XGBoost, face à la très importante quantité de données et de dimensions à laquelle nous faisons face.

Comme son nom l'indique, XGBoost repose sur des méthodes de *boosting*. Le *boosting* est une technique d'ensemble qui combine plusieurs modèles plus faibles (appelés "*weak learners*" ou "*base learners*") pour créer un modèle plus puissant. Il se concentre sur les erreurs commises par les modèles faibles et cherche à les corriger dans les itérations suivantes, en mettant davantage l'accent sur les exemples mal classés. Cela permet d'améliorer progressivement la performance du modèle global.

Similairement au *Random Forest*, XGBoost utilise des arbres de décision comme *base learners*. Les arbres de décision sont des structures hiérarchiques où chaque nœud interne représente une décision basée sur une caractéristique (ou une combinaison de caractéristiques) et chaque feuille représente une prédiction.

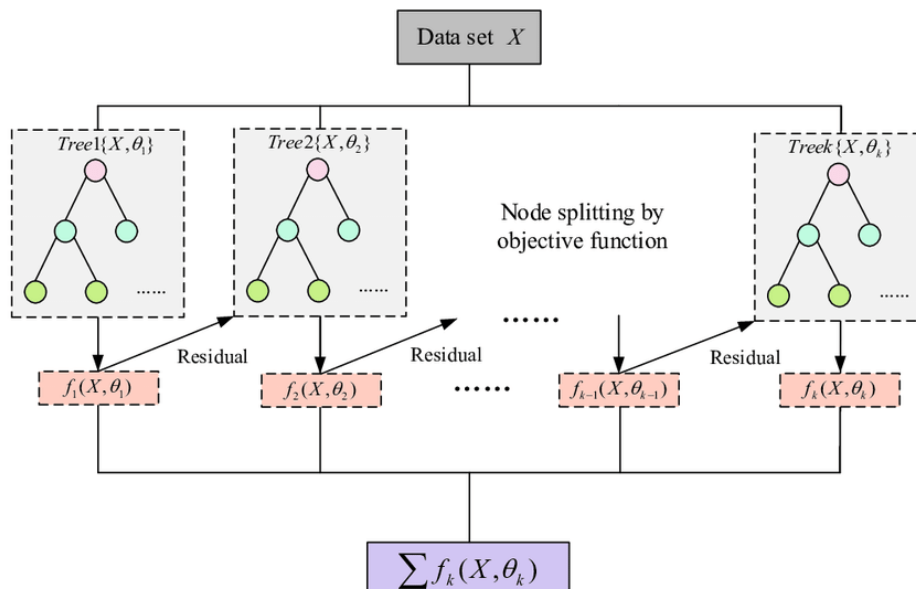


Figure 18 – Schéma explicatif du fonctionnement de XGBoost

c. Tuning des hyperparamètres

Nous avons tenté une méthode de recherche par grille pour trouver le set d'hyperparamètres optimal, c'est-à-dire établir une matrice d'hyperparamètres à explorer, en comparant pour chaque combinaison possible le score obtenu. Cependant, au vu du nombre important d'hyperparamètres à ajuster et de valeurs possibles pour chacun d'entre eux, nous avons finalement opté pour une stratégie plus linéaire : nous optimiserons chaque paramètre tour à tour, par ordre d'importance. De plus, pour limiter le temps de calcul, nous conserverons l'échantillon utilisé pour la sélection de variables lors du tuning.

Pour chaque valeur possible d'un hyperparamètre, nous effectuerons une *4-fold cross validation*. En pratique, cela signifie que le jeu de données sur lequel notre modèle tourne sera découpé en quatre sous-jeux. Il y aura alors quatre itérations au cours desquelles le modèle apprendra sur trois des sous-jeux de données puis sera testé sur le quatrième et obtiendra un score de performance (nous avons choisi d'observer la RMSE). Le score final pour la valeur testée correspond alors à la moyenne des quatre scores retenus. Une telle méthode permet de limiter fortement le risque d'*overfitting* en s'assurant que le score obtenu, la métrique dont nous nous servons pour déterminer le paramètre optimal, n'est pas dépendant du jeu de test arbitrairement déterminé.



Figure 19 – Illustration d'un 4-fold cross validation

Nous allons présenter ici les hyperparamètres à optimiser.

- Le taux d'apprentissage : Chaque nouvel arbre sera pondéré par ce taux afin de permettre une convergence vers un optimal. Une valeur trop élevée empêchera le modèle de converger vers un optimal en accordant trop d'importance aux arbres essayant de prédire les résidus les plus faibles, tandis qu'une valeur trop faible provoquera une convergence trop lente qui n'atteindra pas cet optimal.
- La profondeur maximale : Ce paramètre définit la distance maximale entre la racine et une feuille qu'un arbre peut atteindre, déterminant à quel point l'algorithme est sensible aux variations. Une profondeur trop élevée peut causer du surapprentissage.
- Le poids minimal des enfants : Ce paramètre contrôle le nombre d'enfants déterminés à chaque nœud de nos arbres en imposant un nombre minimal d'individus dans chaque enfant. En d'autres termes, l'algorithme ne considérera pas un critère de classification si celui-ci ne permet pas

d'aboutir à des nœuds contenant suffisamment de données. L'augmenter réduit donc le risque de surapprentissage.

- La proportion de colonnes et de lignes prélevées pour chaque arbre
- Le paramètre de régularisation L2 : Ce paramètre contrôle l'importance relative de la pénalité par rapport à l'ajustement aux données d'entraînement. Une valeur plus élevée de λ conduira à une plus grande pénalisation des coefficients, ce qui peut conduire à un modèle plus simple mais avec potentiellement une performance de prédiction réduite sur les données d'entraînement. En revanche, une valeur trop faible de λ pourrait permettre un surajustement excessif.

L'ensemble du processus d'optimisation pour la PVFP prend environ 2 heures, puis l'apprentissage environ 1h. La durée pour le BEL est similaire. Voici les évolutions de la RMSE pour la PVFP et le BEL, démontrant le gain de performance après optimisation et utilisation du jeu de données complet.

RMSE	Avant optimisation (échantillon)	Après optimisation (échantillon)	Après optimisation (jeu d'apprentissage complet)
PVFP	635 k€	490 k€	350 k€
BEL	980 k€	735 k€	520 k€

Tableau 9 – Evolution de la RMSE après optimisation du modèle et usage du jeu complet

Une fois ce processus terminé, les prédictions sont quasi-instantanées.

4. Etudes des résultats

Une fois notre modèle optimisé et entraîné sur les données qui lui sont fournies, nous pouvons analyser la qualité des résultats obtenus.

Nous vérifions d'abord à quel point le modèle parvient à prédire avec précision les PVFP et les BEL du jeu de test fourni. En exportant les prédictions dans un fichier Excel, nous pouvons étudier la répartition des erreurs et récupérer les scénarios les moins bien prédits pour identifier de potentielles failles dans le processus. Cette étape permet notamment de s'assurer que le *machine learning* se déroule correctement et que le modèle identifie bien les corrélations présentes dans nos données.

Dans la dernière version mise en place, après avoir été optimisé correctement, XGBoost parvient à une RMSE de 350k€ pour une PVFP de l'ordre de 10 millions d'euros dans le scénario central.

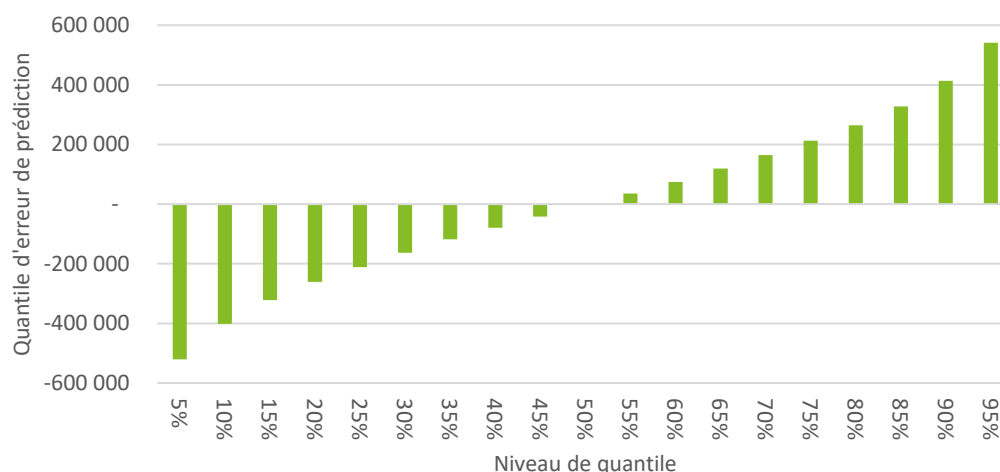


Figure 20 – Montants des erreurs de prédiction de la PVFP par quantile sur le jeu de test

Nous pouvons ainsi observer que dans le cas de la PVFP, l'erreur est symétriquement distribuée, centrée en 0 et sa valeur absolue reste en dessous de 400k€ pour 80% des prédictions, soit 4% de la PVFP.

La seconde étape consiste à s'assurer que nos prédictions restent cohérentes vis-à-vis de scénarios spécifiques, correspondant au type de stress tests que nous pourrions appliquer lors d'une étude ALM ou pour un calcul de SCR.

Premières études – Variabilisation des actions, de l'immobilier et des taux

Durant nos tests préliminaires, nous n'avons introduit de la variabilité que sur les actions, l'immobilier et les taux. Il s'est avéré qu'adopter une approche « stochastique », en considérant la moyenne des PVFP prédites pour chacune des 1000 simulations stochastiques, permettait à des mécanismes de compensation de se produire : les erreurs de prédiction pour chaque simulation individuelle se compensaient en moyenne, pour ainsi obtenir des erreurs inférieures à 100 k€, soit de l'ordre de 1%.

Scénario	Taux	Action	Immo	Résultats (en M€)		Ecart	
				PVFP Modèle	PVFP Prédite	Absolu (€)	Relatif (%)
Ref 31.12.22	-	-	-	10.92	10.96	38 306	0.35%
2	-	-20 %	-10 %	9.18	9.24	62 032	0.68%
3	-	+20 %	+10 %	12.70	12.60	-98 220	-0.77%
4	+25 bps	-	-	11.72	11.72	-4 414	-0.04%
5	+25 bps	-20 %	-10 %	10.06	10.00	-61 043	-0.61%
6	-25 bps	-	-	10.46	10.49	30 022	0.29%
7	-25 bps	+20 %	+10 %	12.22	12.15	-62 726	-0.51%
8	+50 bps	-	-	12.04	12.05	956	0.01%
9	+50 bps	-20 %	-10 %	10.33	10.31	-18 154	-0.18%
10	-50 bps	-	-	10.09	10.05	-34 548	-0.34%
11	-50 bps	+20 %	+10 %	11.76	11.75	-5 800	-0.05%
12	+100 bps	-	-	12.61	12.58	-30 489	-0.24%
13	+100 bps	-20 %	-10 %	10.81	10.80	-6 407	-0.06%
14	-100 bps	-	-	9.05	8.97	-77 767	-0.86%
15	-100 bps	+20 %	+10 %	10.44	10.55	108 946	1.04%

Tableau 10 – Résultats des prédictions de PVFP après variabilisation de l'action, de l'immobilier et des taux

Scénario	Taux	Action	Immo	Résultats (en M€)		Ecart	
				BEL Modèle	BEL Prédit	Absolu (€)	Relatif (%)
Ref 31.12.22	-	-	-	160.21	160.24	31 564	0.02%
2	-	-20 %	-10 %	154.38	154.21	-168 380	-0.11%
3	-	+20 %	+10 %	166.01	165.92	-87 576	-0.05%
4	+25 bps	-	-	157.13	157.40	269 764	0.17%
5	+25 bps	-20 %	-10 %	151.22	151.43	214 083	0.14%
6	-25 bps	-	-	163.05	163.12	75 462	0.05%
7	-25 bps	+20 %	+10 %	168.87	168.85	-13 276	-0.01%
8	+50 bps	-	-	154.54	154.71	165 064	0.11%
9	+50 bps	-20 %	-10 %	148.68	148.81	131 781	0.09%
10	-50 bps	-	-	165.87	165.98	110 030	0.07%
11	-50 bps	+20 %	+10 %	171.77	171.82	46 472	0.03%
12	+100 bps	-	-	149.70	149.70	-3 384	0.00%
13	+100 bps	-20 %	-10 %	143.92	144.01	87 463	0.06%
14	-100 bps	-	-	171.95	172.03	74 842	0.04%
15	-100 bps	+20 %	+10 %	178.13	178.01	-114 241	-0.06%

Tableau 11 - Résultats des prédictions de PVFP après variabilisation de l'action, de l'immobilier et des taux

	PVFP	BEL
Satisfaisant	< 1%	< 0.1%
Acceptable	1% - 2%	0.1% - 0.25%
Limitant	> 2%	> 0.25%

Pour nos tests préliminaires de sensibilités, nous nous sommes concentrés sur des scénarios types, en prenant en compte par exemple la corrélation inverse entre taux et actions ou immobilier observée sur le marché. Chaque valeur correspond à la moyenne de 1000 prédictions, prenant chacune en entrée un des scénarios économiques produits par le GSE.

Bien qu'il soit d'un ordre de grandeur plus de 16 fois supérieur à la PVFP, nous remarquons que le BEL ne souffre pas d'erreurs de prédictions proportionnellement accrues. Les erreurs sont d'un niveau similaire pour la PVFP et le BEL, voire légèrement supérieures en valeur absolu pour ce dernier.

Cependant, l'introduction de la variabilité sur les obligations afin de pouvoir prédire l'impact de la risque-neutralisation et des chocs de spread a largement complexifié les relations entre les inputs et les résultats, ce qui a fortement détérioré la qualité de nos prédictions.

Premières tentatives d'introduction du portefeuille obligataire en variable explicative

Pour les tests suivants, nous avons gardé le même ensemble de scénarios combinant chocs de taux, d'actions et d'immobilier, à la différence que cette fois-ci, nous avons fait varier aléatoirement et indépendamment les montants de coupons et de nominaux de nos obligations pour chaque simulation. Au sein du modèle de *machine learning*, le portefeuille obligataire était représenté par sa *market value* totale et par les montants de coupon et de nominaux par maturité et par notation des

obligations. Dans cette première version, nous faisons varier le portefeuille obligataire *real world* pour créer la base de données, et nous utilisons le même portefeuille obligataire mais risque-neutralisé pour les tests de sensibilités. Le résultat fut pour le moins décevant.

ID	Taux	Action	Immo	PVFP			
				Réel (k€)	Prédit (k€)	Ecart k€	Ecart %
100	-	-	-	12 875	12 277	-598	-4.64%
101	-	-20%	-10%	11 355	10 986	-368	-3.24%
102	-	+20%	+10%	14 910	13 674	-1236	-8.29%
103	+25 bps	-	-	13 427	12 670	-757	-5.64%
104	+25 bps	-20%	-10%	11 913	11 392	-520	-4.37%
105	-25 bps	-	-	12 435	12 084	-351	-2.82%
106	-25 bps	+20%	+10%	14 536	13 487	-1049	-7.21%
107	+50 bps	-	-	13 743	12 960	-783	-5.70%
108	+50 bps	-20%	-10%	12 164	11 694	-470	-3.87%
109	-50 bps	-	-	12 101	12 052	-49	-0.40%
110	-50 bps	+20%	+10%	14 195	13 434	-761	-5.36%
111	+100 bps	-	-	14 313	13 443	-869	-6.07%
112	+100 bps	-20%	-10%	12 625	12 216	-409	-3.24%
113	-100 bps	-	-	11 559	11 582	23	0.20%
114	-100 bps	+20%	+10%	13 272	12 981	-290	-2.19%
115	Test Contrôle			8 248	8 220	-28	-0.34%

Tableau 12 - Résultats des prédictions de PVFP après ajout de la variabilisation du portefeuille obligataire

ID	Taux	Action	Immo	BEL			
				Réel (M€)	Prédit (M€)	Ecart k€	Ecart %
100	-	-	-	164.8	158.9	-5 976	-3.63%
101	-	-20%	-10%	158.8	152.7	-6 076	-3.83%
102	-	+20%	+10%	170.4	164.6	-5 793	-3.40%
103	+25 bps	-	-	161.9	156.1	-5 807	-3.59%
104	+25 bps	-20%	-10%	155.8	150.1	-5 749	-3.69%
105	-25 bps	-	-	167.8	161.9	-5 944	-3.54%
106	-25 bps	+20%	+10%	173.3	167.6	-5 691	-3.28%
107	+50 bps	-	-	159.1	153.2	-5 940	-3.73%
108	+50 bps	-20%	-10%	153.2	147.3	-5 802	-3.79%
109	-50 bps	-	-	170.7	164.9	-5 870	-3.44%
110	-50 bps	+20%	+10%	176.2	170.6	-5 635	-3.20%
111	+100 bps	-	-	154.0	147.7	-6 349	-4.12%
112	+100 bps	-20%	-10%	148.1	142.1	-6 048	-4.08%
113	-100 bps	-	-	176.6	171.4	-5 223	-2.96%
114	-100 bps	+20%	+10%	182.5	177.0	-5 547	-3.04%
115	Test Contrôle			153.6	153.4	-247	-0.16%

Tableau 13 - Résultats des prédictions de PVFP après ajout de la variabilisation du portefeuille obligataire

Ces tests ont révélé que la façon dont nous faisons varier le portefeuille obligataire introduisait de très forts biais, facilement observable dans le cas du BEL, ainsi qu'une variance très importante. En effet, presque toutes nos erreurs de prédictions étaient dans le rouge, et celle qui ne l'étaient pas semblaient plutôt relever de la coïncidence.

Un test de sensibilité « contrôle » a été introduit. Les données d'entrée de celui-ci ont été construites selon la même méthode que celle appliquée pour le jeu d'apprentissage, c'est-à-dire en appliquant un coefficient tiré aléatoirement et indépendamment aux coupons et nominaux de chaque *bucket* Notation X Maturité. Nous pouvons remarquer que ce jeu de test a fait l'objet de prédictions correctes, aussi bien pour le BE que pour la PVFP, ce qui exclut un problème qui viendrait du modèle de *machine learning*. Pour identifier la source du biais, il a donc fallu se tourner vers la génération de la base de données.

Focus sur le portefeuille obligataire et résultats avec la méthode retenue

Nous avons déjà exposé précédemment les différentes évolutions qu'a suivi le procédé de génération de données pour s'approcher au mieux de la réalité et ainsi réduire le bruit superflu.

Nous savions que le problème provenait de la façon dont nous choquions les obligations. Nous avons donc choisi de nous concentrer sur des tests de sensibilités affectant le spread. Lors des itérations qui ont suivi, seuls les taux et les obligations ont été variabilisés, à chaque fois de façon différente, jusqu'à atteindre un niveau de performance acceptable. Le procédé final utilisant l'estimation de la production financière annuelle du portefeuille obligataire a donné lieu à des résultats satisfaisants.

Nous avons finalement réintégré l'action et l'immobilier puis testé à nouveau le modèle, cette fois-ci spécifiquement sur des scénarios impliquant chocs de taux et de spread.

ID	Taux	Spread	BEL				PVFP			
			Réel (M€)	Prédit (M€)	Ecart k€	Ecart %	Réel (k€)	Prédit (k€)	Ecart k€	Ecart %
2	-	-	154.5	154.5	-3	0.00%	10 059	10 147	87	0.87%
4	-	+50 bps	151.3	150.9	-336	-0.22%	8 856	8 959	103	1.16%
5	+50 bps	-	148.9	149.1	154	0.10%	11 166	11 076	-91	-0.81%
6	+50 bps	+50 bps	146.0	145.8	-144	-0.10%	9 911	9 852	-59	-0.59%
7	+50 bps	SCR	146.5	146.4	-125	-0.09%	10 246	10 221	-25	-0.24%
8	-50 bps	-	159.9	159.9	55	0.03%	9 434	9 417	-17	-0.18%
9	-50 bps	+50 bps	156.9	156.7	-129	-0.08%	7 794	7 835	42	0.53%
10	-50 bps	SCR	157.5	157.3	-205	-0.13%	8 268	8 322	54	0.66%
11	+100 bps	-	144.2	144.4	210	0.15%	11 636	11 539	-97	-0.83%
12	+100 bps	+50 bps	141.4	141.4	33	0.02%	10 475	10 362	-113	-1.08%
13	+100 bps	SCR	141.9	141.8	-37	-0.03%	10 796	10 725	-72	-0.66%
14	-100 bps	-	165.4	165.6	147	0.09%	8 859	8 740	-120	-1.35%
15	-100 bps	+50 bps	163.4	163.4	42	0.03%	5 977	5 988	11	0.18%
16	-100 bps	SCR	163.6	163.5	-72	-0.04%	6 957	7 052	95	1.36%
17	-	+50 bps et SCR	148.8	148.5	-236	-0.16%	8 081	8 195	114	1.40%
20	SCR Action		145.2	145.5	294	0.20%	6 658	6 496	-162	-2.43%
21	SCR Immo		152.6	152.5	-172	-0.11%	9 228	9 382	154	1.67%
22	SCR Spread		151.9	151.7	-258	-0.17%	9 170	9 284	114	1.25%
23	SCR Taux Up		142.4	143.3	896	0.63%	11 324	11 393	69	0.61%
24	SCR Taux Down		165.5	164.0	-1 438	-0.87%	8 962	9 247	285	3.18%

Tableau 14 - Résultats des prédictions de PVFP et BEL après application de la nouvelle méthode de variabilisation du portefeuille obligataire

Nous remarquons immédiatement une très nette amélioration de la performance du modèle, avec la plupart des erreurs de prédiction se situant dans des plages acceptables.

C'est donc cette version du modèle, allant de la génération de la base de données jusqu'à l'apprentissage, dont nous nous servirons pour mener nos prédictions de PVFP centrales et choquées et ainsi prédire nos indicateurs pour les scénarios de tests choisis.

Cependant, vérifier la performance de quelques points ne donne qu'un faible aperçu du fonctionnement général et de la fiabilité du modèle. C'est pourquoi des méthodes ont été mises en place pour mieux comprendre les tendances identifiées par celui-ci et la logique derrière les résultats qu'il produit.

5. Explicabilité du modèle

Une des limites des modèles de Machine Learning tels que XGBoost ou le *Random Forest* est l'effet « boîte noire », c'est-à-dire le manque d'explicabilité derrière les résultats.

XGBoost est considéré comme une « boîte noire » en raison de sa complexité et de sa non-linéarité. Il utilise plusieurs arbres de décision pour former un modèle puissant, mais cette approche complexe rend souvent difficile la compréhension directe de son fonctionnement. Chaque arbre corrige les erreurs des arbres précédents, créant ainsi un modèle difficile à interpréter. De plus, XGBoost est capable de capturer des interactions complexes entre les caractéristiques, ce qui rend l'impact de chaque caractéristique sur les prédictions globales plus difficile à expliquer, surtout dans les cas où il en existe un grand nombre.

A contrario, dans le cas d'une régression linéaire, les coefficients retenus par le modèle permettent de comprendre l'impact marginal de chaque variable sur les résultats fournis, et ainsi d'établir une confiance plus ou moins importante en ces derniers.

Des méthodes existent pour tenter de remédier à ce manque d'explicabilité, aussi bien au niveau global qu'au niveau local, notamment les méthodes LIME et SHAP.

LIME (*Local Interpretable Model-agnostic Explanations*) est une méthode d'interprétation en apprentissage automatique. Elle vise à expliquer les prédictions d'un modèle pour une observation donnée. L'idée centrale de LIME est de créer localement des modèles explicables autour d'une instance spécifique en perturbant les données et en observant l'impact de ces perturbations sur les prédictions du modèle. Cette méthode génère des versions altérées de l'observation choisie, puis crée un modèle simple et interprétable sur ces données altérées pour expliquer la prédiction du modèle complexe. L'avantage de LIME réside dans sa capacité à fournir des explications locales pour des prédictions individuelles, même pour des modèles complexes, ce qui peut aider à comprendre les décisions prises par le modèle. Cependant, nous avons choisi de nous tourner plutôt vers SHAP, qui en plus de remplir les mêmes fonctions que LIME en termes d'interprétabilité locale, permet également d'avoir une meilleure compréhension du modèle dans son ensemble.

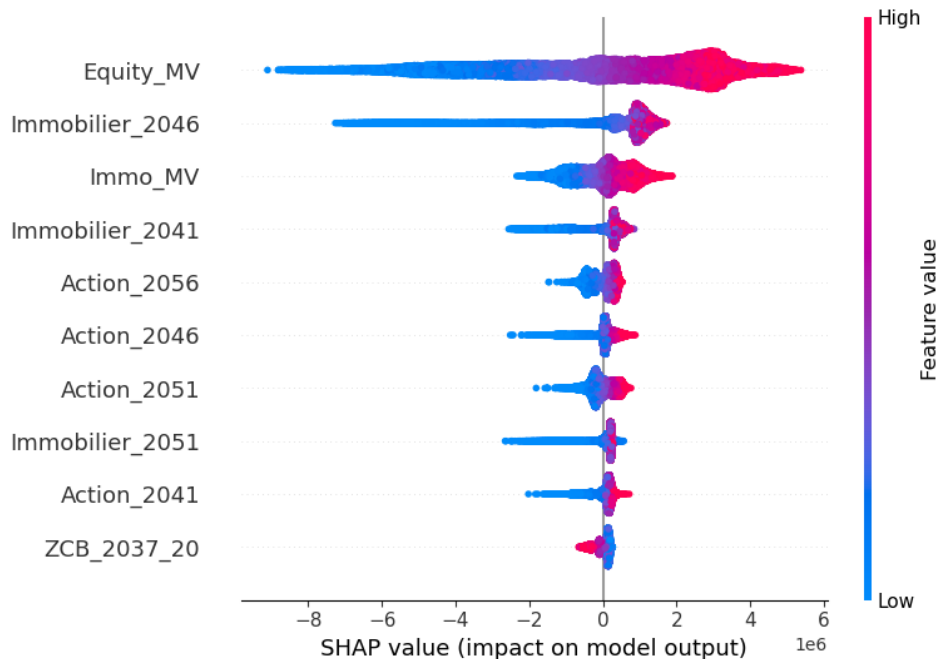
SHAP (SHapley Additive exPlanations) est une approche d'interprétabilité se basant sur la théorie des jeux pour attribuer des valeurs d'importance à chaque attribut dans la prédiction d'un modèle.

Cette méthode cherche à expliquer les prédictions de manière globale en évaluant l'impact de chaque attribut sur les prédictions du modèle. Elle calcule la contribution marginale de chaque attribut à la

différence entre la prédiction du modèle pour une observation spécifique et la prédiction moyenne sur toutes les observations.

SHAP attribue des valeurs d'importance individuelles à chaque caractéristique en considérant toutes les combinaisons possibles d'attributs et en mesurant la contribution moyenne de chaque attribut sur ces combinaisons.

Son application sur notre modèle aboutit aux résultats suivants.



Cette figure indique comment les variables jugées les plus importantes influent sur les prédictions en fonction de leur valeur. Plusieurs éléments permettent de penser que le modèle est effectivement cohérent quant aux relations entre les variables explicatives et la PVFP. La variable la plus influente est la valeur de marché des actions, suivie de près par celle de l'immobilier. Aussi, la corrélation positive entre ces variables et la PVFP a bien été relevée par le modèle. Cela en va de même pour les indices actions et immobiliers. D'autre part, la corrélation est inverse pour les taux (comme montré par le déflateur ZCB_2037_20). Cela peut s'expliquer par une sensibilité plus forte du portefeuille de référence à une baisse des taux.

La forme des graphiques semble aussi indiquer que les options et garanties offertes par le contrat euro ont été correctement relevées par le modèle. Nous observons une asymétrie forte entre les impacts positifs et négatifs de nos variables sur la NAV : les impacts négatifs peuvent atteindre des niveaux bien plus importants que ceux positifs. Cette observation est cohérente avec le fait que dans un contrat euro, l'assureur porte la totalité des pertes mais doit partager ses bénéfices, et que l'assuré a la possibilité de racheter son contrat en cas de rendement insatisfaisant.

Un point cependant qui attire notre attention est l'importance accordée par SHAP à l'indice immobilier de 2046. Lors de la sélection des variables par *feature importance*, cet indice ne s'est pas vu attribué un score justifiant cet exceptionnalisme. La méthode LIME lui accorde également une importance

significative. Il est probable que ce résultat soit une anomalie propre à la base de données générée aléatoirement.

Ayant une meilleure vision globale du fonctionnement du modèle et des relations déterminées par celui-ci, SHAP permet également de comprendre le résultat derrière une prédiction unique.

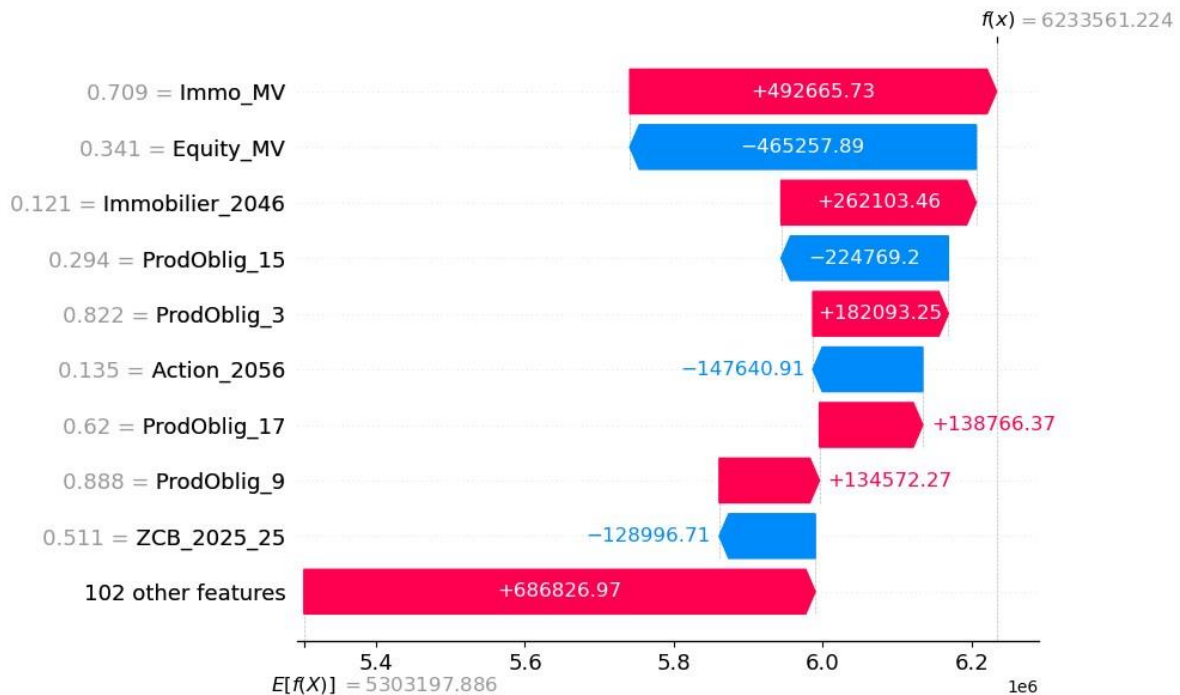


Figure 22 – Diagramme en cascade d'explication de l'écart à la moyenne pour une prédiction choisie aléatoirement dans le jeu de test

Nous avons pris une observation aléatoire pour comprendre comment SHAP explique les écarts à la moyenne déterminé par le modèle. Les résultats vont dans le même sens que ce qui était décrit dans la figure précédente. Des valeurs de marché ou des montants de production financière annuelle du portefeuille obligataire élevés vont de pair avec une PVFP élevée. Nous pouvons nous interroger sur la valeur en apparence faible de l'indice immobilier de 2046 (0,121) qui impacte positivement le résultat ; une analyse rapide du jeu de données après normalisation révèle que la moyenne de l'indice immobilier est d'environ 0,1. Certains scénarios économiques présentent des valeurs très élevées pour cet indice ce qui biaise les montants après normalisation.

Dans cette partie, nous avons mis en place un processus de génération de données et d'apprentissage supervisé permettant de prédire BEL et PVFP en fonction de variables économiques. Ces prédictions sont d'une précision satisfaisante pour l'usage souhaité et l'utilisation de méthodes permettant de mitiger l'effet « boîte noire » de l'algorithme de *machine learning* nous permet d'avoir une meilleure confiance envers les résultats.

V/ ÉTUDE DE CAS : SCR ET SENSIBILITÉS

Dans cette dernière partie, nous commencerons par présenter notre portefeuille d'épargne et d'actif actuel, notre stratégie d'allocation et d'investissement, et notre bilan économique en conditions normales.

Ensuite, nous définirons un ensemble de tests de sensibilités pertinents pour lesquels évaluer l'évolution du bilan et du SCR marché.

1. Rôle et déroulement d'un test de sensibilité

Les assureurs entreprennent des tests de sensibilité sur leurs actifs pour diverses raisons fondamentales liées à la gestion des risques, à la conformité réglementaire, à la planification stratégique et à la communication avec les parties prenantes.

En premier lieu, la gestion des risques constitue l'une des principales motivations. Les assureurs sont exposés à une multitude de risques, tels que les fluctuations des marchés financiers, les variations des taux d'intérêt et les changements dans l'économie. Les tests de sensibilité leur permettent d'évaluer comment leur portefeuille d'investissements réagirait à ces divers scénarios, identifiant ainsi les points de vulnérabilité potentiels. Cela offre aux assureurs la possibilité de mettre en œuvre des stratégies d'atténuation des risques afin de renforcer leur résilience financière.

Un autre aspect important est la conformité réglementaire. Au travers de l'ORSA notamment, les autorités de régulation exigent que les assureurs réalisent des stress tests pour s'assurer de leur capacité à faire face à des situations économiques adverses sans compromettre leur stabilité financière. Ces tests contribuent à garantir la solidité du secteur de l'assurance et à protéger les intérêts des assurés.

Sur le plan de la planification stratégique, les tests de sensibilité sont des outils essentiels. Ils aident les assureurs à élaborer des stratégies d'investissement en anticipant divers scénarios économiques. La capacité à ajuster rapidement l'allocation d'actifs en fonction des conditions changeantes peut permettre aux assureurs d'optimiser leur rendement-risque, assurant ainsi une performance financière optimale.

En outre, ces tests jouent un rôle crucial dans la communication avec les parties prenantes. Les résultats des stress tests peuvent être utilisés pour démontrer la solidité financière de l'entreprise aux actionnaires, clients, régulateurs et autres parties intéressées. Une transparence accrue sur la capacité d'une compagnie d'assurance à résister à des chocs économiques renforce la confiance du public et maintient la crédibilité de l'entreprise.

Un assureur initie des tests de sensibilité en identifiant d'abord les facteurs de risque cruciaux auxquels il est exposé. Sur cette base, il sélectionne des scénarios représentatifs de conditions économiques diverses, allant des situations de crise aux périodes de prospérité.

L'assureur évalue ensuite l'impact potentiel de chaque scénario sur son bilan. L'analyse des résultats permet à l'assureur de comprendre comment chaque scénario pourrait influencer son portefeuille d'investissements, ses revenus, sa solvabilité et d'autres indicateurs financiers.

En fonction des résultats des tests de sensibilité, l'assureur développe des stratégies d'atténuation des risques. Ces stratégies peuvent impliquer des ajustements dans l'allocation d'actifs, des modifications de politiques d'investissement, ou la mise en place de mesures de couverture pour se protéger contre les risques identifiés.

Les conclusions des tests de sensibilité sont ensuite documentées dans des rapports détaillés, qui servent à communiquer avec les parties prenantes, telles que les autorités de régulation, les actionnaires et les clients.

Ces tests ne sont pas des événements isolés, mais font partie d'un processus continu. Les assureurs les répètent périodiquement pour tenir compte des changements dans l'environnement économique, des ajustements de portefeuille, et pour garantir que leurs stratégies d'atténuation restent appropriées. En somme, les tests de sensibilité sont une composante intégrale de la gestion des risques, permettant aux assureurs d'anticiper et de gérer les impacts potentiels de conditions économiques adverses sur leurs actifs.

2. Présentation du portefeuille fictif

Afin de tester l'efficacité et les limites de notre modèle prédictif, nous avons mis en place une société d'assurance vie fictive qui souhaite mener des tests de sensibilités pour les raisons ci-dessus. Nous décrivons ici les caractéristiques de cette société.

Nous avons créé quatre *model points* pour que la composition globale du portefeuille reflète celui d'un assureur vie en France. Ces *model points* couvrent différentes tranches d'âge, de 40 à 70 ans, avec une moyenne d'âge de 55 ans, représentant approximativement l'âge moyen des détenteurs de contrats d'assurance vie en France. Chaque *model point* présente un encours moyen de 44 000 euros par contrat pour 1000 contrats, assurant ainsi une similitude entre eux. Les taux minimums garantis sont de 0.5% pour le *model point* des assurés de 40 ans, de 1% pour les assurés de 50 ans et de 2% pour les assurés de 60 et 70 ans. Aussi, pour ces deux derniers *model points*, nous considérons un versement de primes périodiques annuelles de 850 euros.

Le stock de provision pour participation aux bénéfices et la réserve de capitalisation s'élèvent respectivement à 6,0 millions d'euros et 0,9 million d'euros, représentant 3,5% et 0,5% des provisions mathématiques. Pour des raisons de simplicité, d'autres provisions comptables telles que la provision globale de gestion sont considérées comme négligeables.

La structure du portefeuille actif (*model point* actif) est représentative d'un portefeuille d'un assureur vie au 31 décembre 2022, notamment en ce qui concerne l'allocation d'actifs, la durée, le rating des obligations et le niveau de plus ou moins-values latentes.

L'allocation en valeur comptable choisie est 75% en obligations, 15% en actions, 5% en immobilier et 5% en cash. La stratégie d'allocation d'actifs est fixe et consiste à rester proche de l'allocation initiale. Les plus-values latentes sont de 20% sur les actions et l'immobilier.

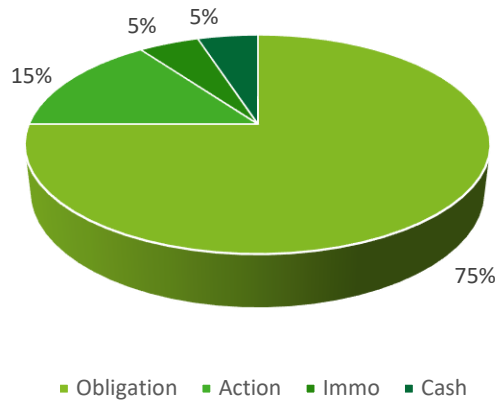


Figure 23 – Allocations d'actifs initiale

Le portefeuille obligataire a un taux de rendement actuariel moyen de 2%, une duration moyenne d'environ 8 ans, et est composé de 36% d'obligations *Corporate* pour 64% d'obligations souveraines.

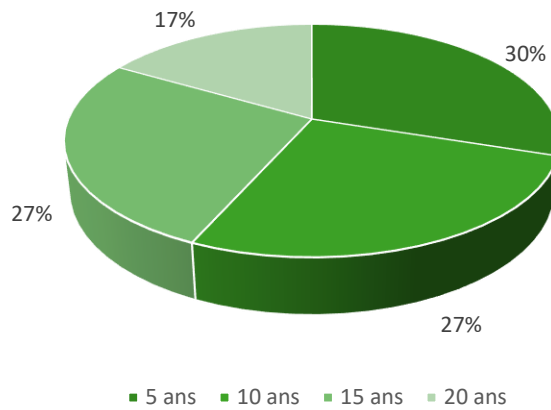


Figure 24 – Répartition des obligations par maturité

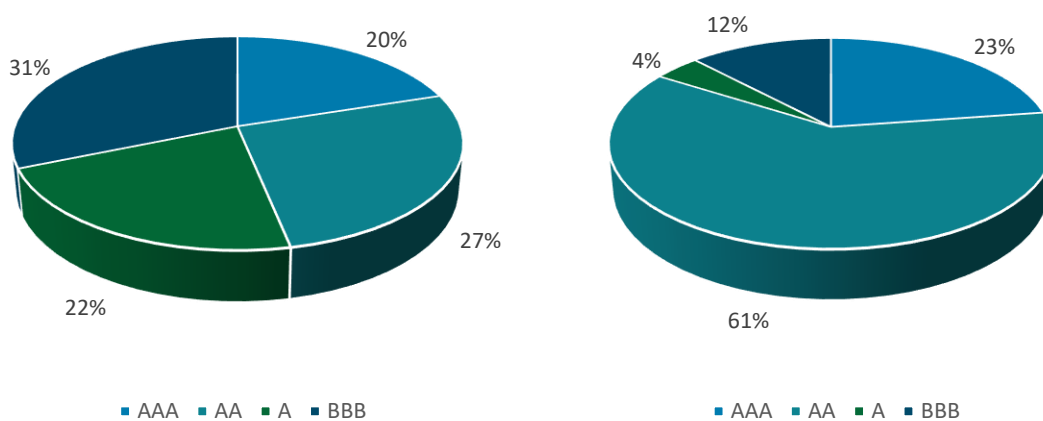


Figure 25 – Répartition par notation des obligations corporate et souveraines

La courbe de taux utilisée est celle de l'EIOPA avec VA au 31 décembre 2022.

Pour rappel, nous ne modélisons pas les fonds propres de la société. Nous n'analyserons donc que l'impact des chocs sur la PVFP prédite par le modèle.

3. Prédiction de SCR dans différents scénarios

a. Sélection des scénarios pertinents pour l'étude de cas

Pour notre étude de cas, nous considérons plusieurs scénarios de *stress tests* affectant l'exposition de notre assureur fictif au risque de marché. Le choix des scénarios testés répond à certains critères afin de couvrir un éventail large de possibilités tout en conservant des situations vraisemblables.

Nous nous limitons donc à des sensibilités correspondant à des ajustements *flat* de -50 bps à +50 bps sur la courbe des taux, et des chocs de -10% à +10% sur les actions, de -5% à +5% sur l'immobilier et jusqu'à +50 bps sur le spread obligataire.

Parmi les scénarios retenus, dans un objectif de vraisemblance, nous nous assurons de retrouver la corrélation positive entre les indices actions et immobilier ainsi que la corrélation nulle ou négative entre l'action et les taux.

Finalement, voici les *stress tests* conservés pour notre étude.

Scénario	Taux	Action	Immo	Spread
10	-	-	-	-
20	-	-10%	-5%	+25 bps
30	-	+10%	+5%	+25 bps
40	-	-5%	-	+50 bps
50	-	+5%	-	+50 bps
60	+25 bps	-	-	+25 bps
70	+25 bps	-	-	+50 bps
80	+25 bps	-10%	-5%	+25 bps
90	+25 bps	-5%	-	-
100	+50 bps	-	-	-
110	+50 bps	-	-	+25 bps
120	+50 bps	-10%	-5%	-
130	+50 bps	-10%	-5%	+25 bps
140	-25 bps	-	-	+25 bps
150	-25 bps	-	-	+50 bps
160	-25 bps	+10%	+5%	+25 bps
170	-25 bps	+5%	-	-
180	-50 bps	-	-	-
190	-50 bps	-	-	+25 bps
200	-50 bps	+10%	+5%	-
210	-50 bps	+10%	+5%	+25 bps

Tableau 15 – Scénarios retenus pour les tests de sensibilité

b. Résultats

Chacun de ces *stress tests* implique normalement la réalisation de six *runs* sur Prophet (central, choc action, choc immobilier, choc de spread, hausse des taux et baisse des taux). Notre machine met environ 20 minutes pour faire tourner 1000 simulations et pour produire le tableau de résultats stochastiques nécessaire à la détermination du bilan stochastique. Réaliser une telle série de *stress tests* via un outil classique prendrait environ 42 heures de temps machine, entrecoupées de nombreuses manipulations par l'utilisateur.

Notre algorithme de *machine learning* obtient ses prédictions en quelques minutes.

En premier lieu, observons la qualité des prédictions sur la PVFP et le BEL centraux pour les scénarios considérés.

Scénario	PVFP				BEL			
	Réel (k€)	Prédit (k€)	Ecart k€	Ecart %	Réel (M€)	Prédit (M€)	Ecart k€	Ecart %
10	10 059,1	10 152,9	93,8	0,93%	154,5	154,6	106,9	0,07%
20	8 324,0	8 413,8	89,8	1,08%	150,2	150,1	(132,9)	-0,09%
30	10 253,5	10 355,4	101,9	0,99%	155,8	155,7	(106,1)	-0,07%
40	8 377,3	8 580,6	203,3	2,43%	150,1	150,0	(124,9)	-0,08%
50	9 218,4	9 271,8	53,4	0,58%	152,5	152,3	(274,7)	-0,18%
60	10 211,7	10 023,2	(188,5)	-1,85%	149,9	150,1	199,3	0,13%
70	9 591,8	9 529,4	(62,4)	-0,65%	148,4	148,6	182,7	0,12%
80	9 169,4	9 024,5	(145,0)	-1,58%	147,2	147,3	144,4	0,10%
90	10 432,0	10 343,6	(88,4)	-0,85%	150,3	150,5	207,9	0,14%
100	11 166,0	11 074,2	(91,8)	-0,82%	148,9	149,0	97,4	0,07%
110	10 501,8	10 379,0	(122,7)	-1,17%	147,4	147,5	70,0	0,05%
120	10 083,2	10 053,5	(29,7)	-0,29%	146,2	146,3	120,4	0,08%
130	9 464,3	9 398,5	(65,8)	-0,69%	144,7	144,7	23,9	0,02%
140	8 992,1	8 998,1	6,0	0,07%	155,6	155,5	(150,6)	-0,10%
150	8 331,2	8 432,3	101,0	1,21%	154,0	154,0	(76,3)	-0,05%
160	7 889,7	7 887,5	(2,2)	-0,03%	152,9	152,8	(105,3)	-0,07%
170	9 167,5	9 358,2	190,8	2,08%	156,1	156,2	38,3	0,02%
180	9 434,4	9 463,7	29,3	0,31%	159,9	160,0	134,0	0,08%
190	8 631,6	8 595,6	(36,0)	-0,42%	158,3	158,1	(257,4)	-0,16%
200	10 515,8	10 476,4	(39,3)	-0,37%	162,6	162,8	207,7	0,13%
210	9 536,5	9 569,7	33,2	0,35%	161,2	161,1	(140,5)	-0,09%

Tableau 16 - Résultats des prédictions de PVFP et BEL pour les scénarios choisis

	PVFP	BEL
Satisfaisant	< 1%	< 0.1%
Acceptable	1% - 2%	0.1% - 0.25%

Limitant	> 2%	> 0.25%
-----------------	------	---------

Deux des scénarios (40 et 170) présentent un niveau d'erreur de prédiction de la PVFP supérieur au reste. Ces deux scénarios ne partagent pas de point commun particulier. Une observation plus fine révèle que 81% et 82% des prédictions sur les 1000 scénarios économiques sont surestimées pour respectivement les *stress tests* 40 et 170. Ces imprécisions témoignent des faiblesses des algorithmes « Boîte noire » tels que le XGBoost, et des difficultés à expliquer de tels biais. Des solutions possibles peuvent consister en l'agrandissement de jeu de données d'entraînement autour des points les moins bien prédits.

Pour les autres *stress tests*, le niveau de précision atteint pour la prédiction des scénarios centraux est très satisfaisant malgré l'incorporation du spread en variable explicative. Le tableau ci-dessous présente les moyennes, écarts-type et maximum des écarts relatifs et en milliers d'euros.

	PVFP		BEL	
	Ecart réel (k€)	Ecart relatif	Ecart réel (k€)	Ecart relatif
Moyenne	1,5	0,06%	7,8	0,01%
Ecart-type	102,3	1,09%	151,9	0,10%
Ecart absolu max	203,3	2,43%	274,7	0,18%

Tableau 17 – Statistiques sur les erreurs de prédiction des PVFP et BEL pour les scénarios choisis

Notons que les écarts sont centrés aussi bien pour la PVFP que pour le BEL. L'ordre de grandeur très supérieur du BEL par rapport à la PVFP n'affecte que peu la performance de prédiction, ce qui est cohérent dans la mesure où dans notre modèle ALM l'un est le complémentaire de l'autre.

Pour chacun des scénarios, nous avons lancé les simulations nécessaires à l'obtention des résultats pour les sous-module choisis.

Voici les résultats obtenus sur les prédictions de SCR marché :

Scénario	Sous-module Action		
	Réel	Prédit	Ecart k€
10	3 700	4 165	465
20	4 092	4 267	176
30	3 627	3 824	196
40	4 374	4 590	216
50	4 209	4 328	119
60	4 063	4 073	10
70	4 296	4 298	2
80	4 137	4 089	-49
90	3 947	4 062	114
100	3 905	3 877	-28
110	3 965	3 889	-75

120	3 817	3 769	-48
130	3 890	3 859	-32
140	4 321	4 594	274
150	4 812	5 053	241
160	4 682	4 788	106
170	3 988	4 480	492
180	4 427	4 854	427
190	4 842	5 042	200
200	4 337	4 326	-11
210	4 386	4 508	122

Tableau 18 – Résultats des prédictions pour le sous-module de SCR action

Scénario	Sous-module Immobilier		
	Réel	Prédit	Ecart k€
10	831	627	-204
20	635	645	10
30	670	754	84
40	667	642	-25
50	643	542	-101
60	715	566	-149
70	695	600	-95
80	678	608	-70
90	725	675	-50
100	747	589	-158
110	713	567	-146
120	645	582	-64
130	656	583	-73
140	781	642	-138
150	735	625	-110
160	671	690	19
170	793	801	9
180	900	746	-154
190	821	706	-115
200	977	884	-93
210	797	878	81

Tableau 19 - Résultats des prédictions pour le sous-module de SCR immobilier

Scénario	Sous-module Spread		
	Réel	Prédit	Ecart k€
10	889	899	10
20	609	681	72

30	819	816	-3
40	705	804	99
50	774	783	9
60	859	750	-109
70	867	851	-17
80	827	735	-92
90	911	889	-23
100	920	912	-8
110	796	708	-88
120	804	794	-10
130	734	678	-56
140	895	840	-54
150	1 004	1 028	23
160	837	810	-27
170	923	927	3
180	1 166	1 094	-73
190	1 066	1 019	-47
200	1 427	1 317	-109
210	1 225	1 165	-60

Tableau 20 - Résultats des prédictions pour le sous-module de SCR spread

Scénario	Sous-module Taux Up		
	Réel	Prédit	Ecart k€
10	-1 265	-1 261	4
20	-1 214	-1 441	-227
30	-1 299	-1 329	-30
40	-1 324	-1 331	-7
50	-1 344	-1 402	-59
60	-265	-350	-85
70	-297	-370	-73
80	-264	-363	-99
90	-259	-283	-23
100	-742	-222	520
110	-809	-381	428
120	-647	-353	294
130	-720	-499	222
140	-1 531	-1 393	137
150	-1 586	-1 484	101
160	-1 527	-1 517	10
170	-1 546	-1 284	262
180	-1 760	-1 597	163
190	-1 897	-1 786	111

200	-1 522	-1 527	-5
210	-1 797	-1 709	88

Tableau 21 - Résultats des prédictions pour le sous-module de SCR taux à la hausse

Sous-module Taux Down			
Scénario	Réel	Prédit	Ecart k€
10	1 097	915	-182
20	1 293	1 743	450
30	1 234	1 132	-102
40	1 713	2 057	344
50	1 646	1 829	183
60	3 522	3 219	-303
70	4 365	4 398	33
80	3 749	4 244	495
90	2 878	2 632	-246
100	1 703	1 450	-253
110	1 763	1 658	-105
120	1 654	1 667	13
130	1 725	1 888	163
140	2 340	2 230	-110
150	3 152	3 335	183
160	2 514	3 140	626
170	1 649	1 684	35
180	1 339	920	-418
190	1 942	1 797	-145
200	978	380	-597
210	1 659	1 319	-341

Tableau 22 - Résultats des prédictions pour le sous-module de SCR taux à la baisse

SCR Marché				
Scénario	Réel	Prédit	Ecart k€	Ecart %
10	5 667	5 855	187	3,31%
20	5 806	6 341	535	9,22%
30	5 501	5 692	191	3,48%
40	6 459	6 958	500	7,73%
50	6 285	6 452	167	2,66%
60	7 662	7 242	-421	-5,49%
70	8 550	8 496	-54	-0,63%
80	7 863	8 113	250	3,18%
90	7 094	6 959	-135	-1,90%
100	6 221	5 892	-330	-5,30%

110	6 201	5 874	-327	-5,27%
120	5 936	5 840	-96	-1,62%
130	6 012	5 998	-14	-0,23%
140	7 078	7 112	34	0,47%
150	8 186	8 481	295	3,61%
160	7 420	7 983	563	7,59%
170	6 305	6 814	509	8,08%
180	6 800	6 784	-16	-0,23%
190	7 459	7 429	-29	-0,39%
200	6 758	6 249	-509	-7,53%
210	6 929	6 848	-81	-1,17%

Tableau 23 – Résultat final sur les prédictions de SCR marché après agrégation

	Satisfaisant	Acceptable	Limitant
Ecart	< 200k	200k - 400k	> 400k

Rappelons tout d'abord que prédire un SCR avec une précision satisfaisante est un exercice difficile. Le SCR étant une différence entre une PVFP centrale et une PVFP choquée, si les erreurs de prédictions pour le scénario central et le scénario choqué sont de sens contraire, l'écart entre le SCR réel et le SCR prédit peut s'accroître fortement.

	Action	Immobilier	Spread	Taux Up	Taux Down	Marché
Moyenne	139	-73	-27	83	-13	58
Ecart-type	168	77	53	179	309	307

Tableau 24 – Statistiques sur les erreurs de prédiction des PVFP choquées pour les sous-modules de SCR marché

Le succès le plus notable est sans doute la précision avec laquelle nous prédisons le sous-module de spread. Introduire le spread en tant que variable explicative a été l'un des défis les plus importants de ce mémoire, et obtenir des écarts pour la plupart inférieurs à 100 k€ est une amélioration très satisfaisante. Le SCR de spread est ainsi le sous-module le mieux prédit, avec une erreur moyenne centrée de -27 k€ et un écart-type faible de 53 k€.

Cependant, les résultats sont plus mitigés quant aux prédictions du SCR Marché total. Après agrégation des sous-modules considérés, environ la moitié des prédictions atteignent des écarts importants voire très importants, suffisamment significatifs pour remettre en question leur utilité pour des tests de sensibilités.

Une part importante de ces écarts peut être expliquée par une sous-performance des prédictions pour les sous-modules d'actions et de taux.

En effet, les erreurs de prédiction du SCR action soulignent un biais important du modèle. Un examen plus détaillé des PVFP prédites lors de l'application du choc action révèle une sous-estimation de la PVFP choquée avec une erreur moyenne de -137 k€, ce qui explique la différence accrue entre PVFP centrale et choquée. Cette conjecture de biais est aussi appuyée par le fait que 20 des 21 prédictions de PVFP choquées pour le sous-module action sont sous-estimées. Ce phénomène ne dépend donc pas du scénario de *stress test* étudié.

A ce jour, il existe peu de pistes d'explication quant à ce biais. Les actions du portefeuille sont représentées par une variable conceptuellement simple puisqu'il ne s'agit que de la valeur de marché totale. Ce paramètre est le seul à varier lors de la création de nos bases de données et lors de l'application du choc de la formule standard. Une solution potentielle serait d'augmenter le nombre de lignes de la base d'apprentissage. Il est plus probable que la solution, si elle existe, est à trouver dans le calibrage du modèle XGBoost.

Les erreurs de prédictions du sous-module de taux, quant à elles, bien que plus centrées que pour le sous-module action, souffrent d'une variance très forte. Une solution qui mitigera largement le problème, en plus d'être aisément implémentable lors de la prochaine itération du modèle, est d'introduire plus de variance dans le générateur de scénarios économiques. En effet, nous avons entraîné notre modèle sur des tables de scénarios produites à partir de translation de la courbe de taux sans risque. Cependant, le choc réglementaire de taux affecte significativement la forme de la courbe des taux, notamment pour les premières maturités. Il est donc probable que ces formes spécifiques ne soient pas correctement identifiées par le modèle de *machine learning*.

Passé ces axes d'amélioration, le potentiel de notre modèle réside dans sa capacité à prédire efficacement l'impact de nos tests de sensibilité sur le SCR. Le graphique ci-dessous présente côte-à-côte l'impact relatif réel et l'impact relatif prédit.

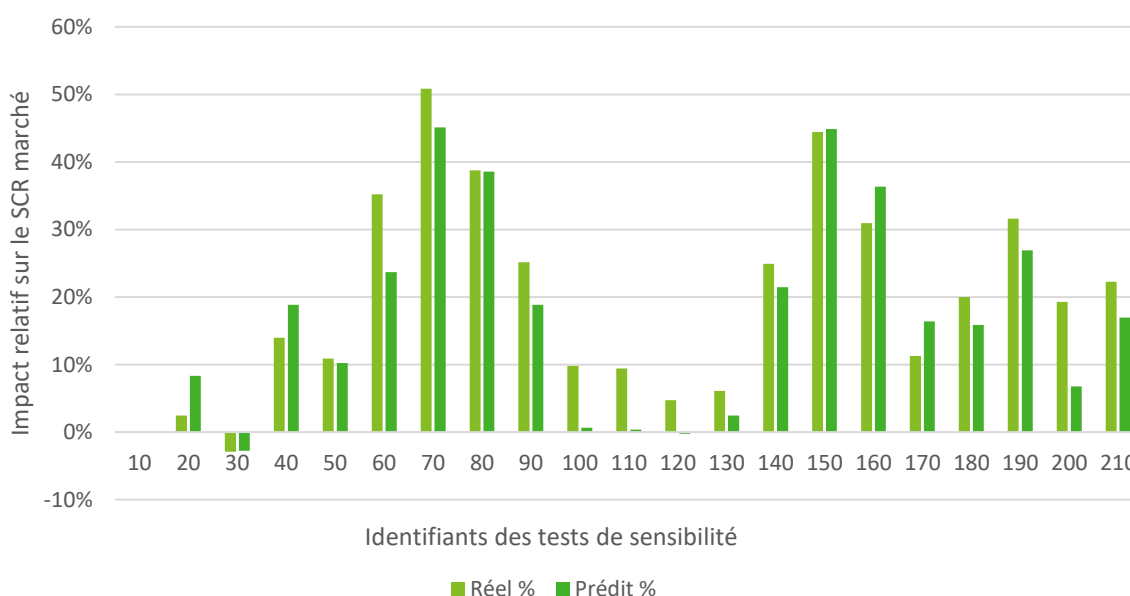


Figure 26 – Impact des tests de sensibilités sur le SCR marché

Nous observons que les prédictions d'impact des tests de sensibilité sur le SCR Marché suivent avec une précision satisfaisante dans la plupart des cas les impacts réels. Des écarts pouvant aller jusqu'à 10% sur l'impact peuvent cependant exister. Nous avons évoqué précédemment des pistes possibles

pour améliorer plus encore ces résultats, notamment au travers d'une meilleure prédiction du sous-module de taux.

Une métrique exploitable pour évaluer son exposition au risque en fonction des scénarios est le ratio PVFP/SCR. Se servir de cette métrique rendement/risque est plus pertinent que de simplement s'intéresser à l'évolution du SCR dans la mesure où une hausse de SCR peut être tolérée si la PVFP augmente suffisamment pour la compenser.

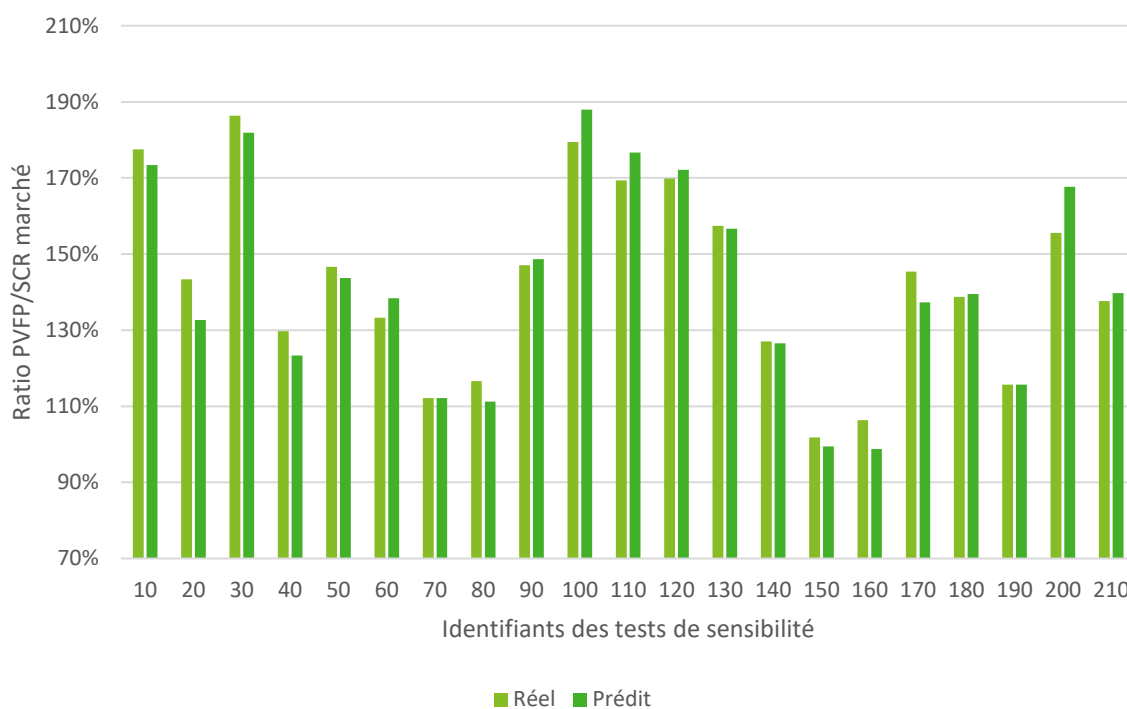


Figure 27 – Impact des tests de sensibilités sur le ratio PVFP/SCR marché

L'indicateur prédit est calculé à partir de la PVFP et du SCR prédits. En scénario central, celui-ci vaut 177% (prédit à 173%). Il diminue le plus fortement dans les scénarios 70, 80, 150, 160 et 190.

Voici pour rappel les scénarios en question :

Scénario	Taux	Action	Immo	Spread
70	+25 bps	-	-	+50 bps
80	+25 bps	-10%	-5%	+25 bps
150	-25 bps	-	-	+50 bps
160	-25 bps	+10%	+5%	+25 bps
190	-50 bps	-	-	+25 bps

Tableau 25 – Scénarios retenus pour leur impact critique sur l'indicateur

Ces scénarios sont de façon surprenante plutôt diversifiés. Ils présentent des hausses et des baisses de taux, d'actions, d'immobilier, et des hausses de spread plus ou moins fortes. Mais des combinaisons de chocs faibles, parfois même de sens différents, peuvent avoir des impacts significatifs sur la solvabilité de l'assureur.

C'est ici que transparait l'intérêt de notre méthode. L'obtention de ces prédictions étant quasi-instantanées, nous pouvons mener autant de tests de sensibilités que nous pouvons produire d'inputs. Dans ce cas d'étude, nous avons utilisé le modèle en parallèle des prédictions pour pouvoir attester de leur performance. En pratique, une maquette Excel ou un algorithme sur Python peut produire suffisamment d'inputs au format adapté pour produire une nappe de sensibilités. L'assureur serait ainsi capable d'identifier rapidement l'ensemble des scénarios à haut risque nécessitant une étude plus approfondie.

4. Limites et prochaines étapes

Dans cette étude, nous nous sommes concentrés sur la prédiction en fonction des paramètres qui varient lors d'un calcul de SCR ou d'une étude ALM classique, c'est-à-dire les *market values*, les taux et le spread.

Les premiers résultats obtenus, bien qu'améliorables, sont très encourageants. Plusieurs axes d'améliorations sont envisageables afin de tendre vers un outil plus complet.

Tout d'abord, l'absence du passif au sein des variables explicatives a pour conséquences que notre modèle est figé dans le temps. Il ne peut obtenir ses prédictions qu'à la date à laquelle il est paramétré. En effet, la méthode de *flexing* utilisée ne permet d'appliquer directement notre méthodologie de variabilisation de l'actif au passif. Si les paramètres affectant l'actif sont en input direct du modèle Prophet actif/passif, les inputs affectant le passif sont introduits dans le modèle Prophet déterministe, puis les outputs de celui-ci deviennent l'input du modèle actif/passif via la table DET_CF.

Cette problématique ne se poserait pas dans un modèle intégré. Mais dans notre situation, cela nécessiterait par exemple de trouver un moyen de variabiliser la table DET_CF tout en s'assurant que son contenu reste cohérent au cours du temps.

Intégrer le passif au sein des variables explicatives permettraient de projeter son SCR sur les années à venir, en adaptant les inputs pour prendre en compte le vieillissement du portefeuille. Cela ouvrirait également la voie pour des tests de sensibilités plus complet, incluant le SCR Souscription vie en plus du SCR Marché.

Une autre limite significative du processus est la non-intégration de variables explicatives représentant les potentielles *management actions*. Une évolution possible à ce titre serait d'introduire l'allocation initiale ou l'allocation cible en variable explicative. Ainsi, l'assureur pourrait étudier l'impact d'une variation de son allocation d'actifs initiale ou de sa stratégie de réallocation dans le temps sur son indicateur rendement/risque, et ainsi trouver des allocations possibles pour optimiser cet indicateur.

D'autres éléments variabilisables au titre des *management actions* seraient le taux de participation aux bénéfices et la politique de dotation/reprise de PPE.

Toutes ces améliorations permettraient à ce projet de tendre un peu plus vers un objectif plus ambitieux : la création d'un véritable outil d'aide au reporting et à la prise décision dans le cadre d'études ALM.

CONCLUSION

L'objectif de cette étude était la mise en place d'un processus viable permettant de prédire l'impact de variations de l'environnement économique sur le SCR marché d'un assureur vie à un instant donné. Nous avons commencé par rappeler le cadre réglementaire et les principes de modélisation actifs/passifs, avant de présenter la méthode adoptée.

Les étapes du processus ont été détaillées, de la perturbation des *inputs* afin de produire une base de données à l'application aux tests de sensibilité, en passant par le calibrage du modèle de *machine learning*. Ce mémoire a ainsi permis de retracer l'évolution de notre réflexion face aux obstacles qui ont pu se présenter, notamment lors de la phase de génération de la base de données. Trouver une solution viable pour variabiliser le portefeuille obligataire et transmettre efficacement et synthétiquement l'information associée a sans aucun doute été le défi le plus important que nous avons relevé.

Après un grand nombre d'itérations, de développement et d'améliorations, la conclusion est qu'il est effectivement possible de se servir de l'apprentissage supervisé, en l'occurrence avec l'algorithme XGBoost, pour prédire le SCR marché tel que défini par la formule standard. Des pistes d'améliorations possibles ont également été proposées pour pouvoir continuer le projet dans un objectif d'application concrète en entreprise.

BIBLIOGRAPHIE

Guo, Rui & Zhao, Zhiqian & Wang, Tao & Liu, Guangheng & Zhao, Jingyi & Gao, Dianrong. (2020). *Degradation state recognition of piston pump based on ICEEMDAN and XGBoost*. Applied Sciences. 10. 6593. 10.3390/app10186593.

Cross Validation, Math Works, <https://fr.mathworks.com/discovery/cross-validation.html>

GALARNYK Michael (2022), *Understanding train test split*, BuiltIn, <https://builtin.com/data-science/train-test-split>

Amjad Abu-Rmileh (2019), *The Multiple faces of 'Feature importance' in XGBoost*, Towards Data Science, <https://towardsdatascience.com/be-careful-when-interpreting-your-features-importance-in-xgboost-6e16132588e7>

XGBoost : Tout savoir sur le Boosting, blent, XG<https://blent.ai/blog/a/xgboost-tout-comprendre>, 2022

Règlement délégué (UE) 2015/35 de la Commission du 10 octobre 2014 complétant la directive 2009/138/CE du Parlement européen et du Conseil sur l'accès aux activités de l'assurance et de la réassurance et leur exercice (solvabilité II)

Ofi Asset Management (2016), *Méthodologie Ofi AM de calcul du SCR marché*

MICHEL Loïc (2015), Mémoire d'actuariat - *Impacts de management actions sur des indicateurs de risque et de rentabilité dans le cadre d'un contrat d'épargne Euro*

GOUNEL Clément (2023), Mémoire d'actuariat - *Optimisation du processus de pilotage du risque de Spread par des méthodes de Machine Learning*

GAUVILLE Rémi (2017), Mémoire d'actuariat - *Projection de ratio de solvabilité : des méthodes de machine learning pour contourner les contraintes opérationnelles de la méthode des SdS*

DESIRE Hortense (2021), Mémoire d'actuariat - *Euro fund asset and liability interaction modelling in a low interest rate environment: impact on solvency and profitability*

OUADEIH Mohamed Amine (2018), Mémoire d'actuariat - *Implémentation d'une nouvelle méthode de modélisation des flux de passif d'un contrat Épargne dans un modèle ALM et comparaison avec la méthode "Flexing"*

DELCAILLAU Dimitri (2019), Mémoire d'actuariat – *Contrôle et transparence des modèles complexes en actuariat*

ANNEXES

Annexe I – Table des figures

Figure 1 – Schéma simplifié d'un bilan prudentiel.....	22
Figure 2 – Courbe des taux sans risques avec VA de l'EIOPA au 31/12/2022.....	24
Figure 3 – Schéma d'un bilan prudentiel avec décomposition de la NAV	25
Figure 4 – Décomposition du SCR en modules et sous-modules.....	27
Figure 5 – Exemple d'absorption des pertes par les provisions techniques.....	29
Figure 6 – Courbe de taux centrales et choquées au 31/12/2022.....	30
Figure 7 – Loi de rachats conjoncturels.....	38
Figure 8 – Déroulement d'une année de projection au sein du modèle ALM.....	39
Figure 9 – Déroulement de l'algorithme de réinvestissement.....	40
Figure 10 – Schéma simplifié de flexing	41
Figure 11 – Récapitulatif de la méthode pour l'obtention d'un bilan stochastique	44
Figure 12 – Représentation des abattements moyens appliqués au nominaux dans différents scénarios en fonction de la maturité résiduelle.....	48
Figure 13 – Représentation de l'abattement des nominaux pour risque-neutralisation et de trois combinaisons d'abattements proches générés par notre processus	49
Figure 14 - Représentation de l'abattement des nominaux pour SCR de spread et de trois combinaisons d'abattements proches générés par notre processus	49
Figure 15 – Schéma récapitulatif des divisions du data set	50
Figure 16 – RMSE du modèle de ML en fonction du nombre de variables retenues	52
Figure 17 – Score de feature importance des variables les plus significatives retenues.....	52
Figure 18 – Schéma explicatif du fonctionnement de XGBoost.....	53
Figure 19 – Illustration d'un 4-fold cross validation.....	54
Figure 20 – Montants des erreurs de prédiction de la PVFP par quantile sur le jeu de test	56
Figure 21 – Diagramme en violon indiquant la valeur de SHAP en fonction de la valeur de la variable pour les 10 variables les plus impactantes.....	61
Figure 22 – Diagramme en cascade d'explication de l'écart à la moyenne pour une prédiction choisie aléatoirement dans le jeu de test	62
Figure 23 – Allocations d'actifs initiale.....	65
Figure 24 – Répartition des obligations par maturité	65
Figure 25 – Répartition par notation des obligations corporate et souveraines	65
Figure 26 – Impact des tests de sensibilités sur le SCR marché.....	73
Figure 27 – Impact des tests de sensibilités sur le ratio PVFP/SCR marché.....	74
Tableau 1 – Coefficients de corrélation pour l'agrégation des modules de SCR	28
Tableau 2 – Coefficients de chocs appliqués à la courbe des taux centrale pour le SCR de taux.....	30
Tableau 3 – Coefficients de choc à appliquer à la valeur de marché des obligations en fonction de la maturité et la sensibilité pour le SCR de spread	32
Tableau 4 - Coefficients de choc à appliquer à la valeur de marché des obligations non notées	32
Tableau 5 – Choc appliqué à la valeur de marché de l'excédent en fonction de la notation	32
Tableau 6 – Seuil relatif d'exposition en excès en fonction de la notation pondérée.....	33
Tableau 7 – Coefficients de corrélation utilisés pour l'agrégation des sous-modules du SCR marché	33
Tableau 8 – Extrait de la table DET_CF.....	41

Tableau 9 – Evolution de la RMSE après optimisation du modèle et usage du jeu compet.....	55
Tableau 10 – Résultats des prédictions de PVFP après variabilisation de l’action, de l’immobilier et des taux	57
Tableau 11 - Résultats des prédictions de PVFP après variabilisation de l’action, de l’immobilier et des taux.....	57
Tableau 12 - Résultats des prédictions de PVFP après ajout de la variabilisation du portefeuille obligataire	58
Tableau 13 - Résultats des prédictions de PVFP après ajout de la variabilisation du portefeuille obligataire	58
Tableau 14 - Résultats des prédictions de PVFP et BEL après application de la nouvelle méthode de variabilisation du portefeuille obligataire.....	59
Tableau 15 – Scénarios retenus pour les tests de sensibilités	67
Tableau 16 - Résultats des prédictions de PVFP et BEL pour les scénarios choisis.....	67
Tableau 17 – Statistiques sur les erreurs de prédiction des PVFP et BEL pour les scénarios choisis....	68
Tableau 18 – Résultats des prédictions pour le sous-module de SCR action.....	69
Tableau 19 - Résultats des prédictions pour le sous-module de SCR immobilier	69
Tableau 20 - Résultats des prédictions pour le sous-module de SCR spread	70
Tableau 21 - Résultats des prédictions pour le sous-module de SCR taux à la hausse.....	71
Tableau 22 - Résultats des prédictions pour le sous-module de SCR taux à la baisse.....	71
Tableau 23 – Résultat final sur les prédictions de SCR marché après agrégation	72
Tableau 24 – Statistiques sur les erreurs de prédiction des PVFP choquées pour les sous-modules de SCR marché.....	72
Tableau 25 – Scénarios retenus pour leur impact critique sur l’indicateur	74

Annexe II – Présentation du Gradient Boosting

Dans ce mémoire, nous nous sommes servis de XGBoost en tant qu'algorithme d'apprentissage supervisé pour atteindre nos objectifs. Cette annexe offre une explication plus détaillée du fonctionnement de l'algorithme de *boosting* de gradient, dans le cas d'une régression avec des variables quantitatives.

Considérons L notre fonction de perte et f_i le classifieur retenu après la i -ème itération. La fonction de coût choisie dans notre cas est celle de l'écart quadratique : $L(y_i, f(x_i)) = \frac{1}{2}(y_i - f(x_i))^2$.

Cette fonction présente l'avantage d'être sensible aux écarts mais peu robuste face aux valeurs aberrantes. Ayant un contrôle important sur la génération de notre base de données, cette fonction semble appropriée.

Voici l'algorithme de *gradient boosting* sur lequel repose notre modèle d'apprentissage supervisé :

- Initialisation avec un modèle constant : $f_0(x) = \operatorname{argmin}_{\theta} \sum_{i=1}^n L(y_i, \theta)$ (dans le cas d'une fonction de perte quadratique, cela correspond à une fonction constante égale à la moyenne des y_i)
- Pour m allant de 1 à M :
 - Calcul des pseudo-résidus (partie résiduelle non-expliquée lors de l'itération précédente) : $r_{i,m} = -\frac{\partial L(y_i, f_{m-1}(x_i))}{\partial f_{m-1}(x_i)}$
 - Ajustement d'un *weak learner* (dans notre cas, un arbre de décision) $h_m(x)$ entraîné sur les données $(x_i, r_{i,m})_{1 \leq i \leq n}$ avec x_i les variables explicatives et $r_{i,m}$ les pseudo-résidus associés
 - Calcul du paramètre de régularisation θ_m en résolvant le problème d'optimisation : $\theta_m = \operatorname{argmin}_{\theta} \sum_{i=1}^n L(y_i, f_{m-1}(x_i)) + \theta h_m(x_i)$
 - Calcul du nouveau classifieur : $f_m(x) = f_{m-1} + \theta_m h_m(x)$
- A la sortie de cet algorithme, on obtient $f_M(x)$, notre classifieur final.

Annexe III – Résultats des prédictions par sous-module et par test de sensibilité pour le BEL et la PVFP

Choc	ID	Taux	Action	Immo	Spread	BEL Réel (M€)	BEL prédit (M€)	Ecart réel BEL (k€)	Ecart relatif BEL	PVFP réelle (k€)	PVFP prédite (k€)	Ecart réel PVFP (k€)	Ecart relatif PVFP
Central	10	-	-	-	-	154.5	154.6	106	0.07%	10 059	10 153	94	0.93%
Action	11					144.9	145.1	220	0.15%	6 359	5 988	-371	-5.84%
Immo	12					152.6	152.7	25	0.02%	9 228	9 526	298	3.23%
Spread	13					151.9	152.0	101	0.07%	9 170	9 254	84	0.92%
Up	14					142.4	143.1	692	0.49%	11 324	11 414	90	0.79%
Down	15					165.5	163.8	-1 644	-0.99%	8 962	9 238	276	3.08%
Central	20	-	-10%	-5%	+25 bps	150.2	150.1	-132	-0.09%	8 324	8 414	90	1.08%
Action	21					142.3	142.4	98	0.07%	4 232	4 146	-86	-2.03%
Immo	22					148.3	148.0	-223	-0.15%	7 689	7 769	80	1.04%
Spread	23					147.5	147.2	-293	-0.20%	7 715	7 733	17	0.23%
Up	24					138.4	138.9	481	0.35%	9 538	9 855	317	3.32%
Down	25					161.1	160.0	-1 127	-0.70%	7 031	6 671	-360	-5.13%
Central	30	-	+10%	+5%	+25 bps	155.8	155.7	-106	-0.07%	10 254	10 355	102	0.99%
Action	31					144.8	144.9	65	0.05%	6 626	6 532	-94	-1.43%
Immo	32					153.7	153.6	-119	-0.08%	9 583	9 601	18	0.19%
Spread	33					153.3	153.3	-27	-0.02%	9 434	9 540	105	1.12%
Up	34					143.9	144.7	716	0.50%	11 553	11 685	132	1.14%
Down	35					166.7	165.1	-1 629	-0.98%	9 019	9 223	204	2.26%
Central	40	-	-5%	-	+50 bps	150.1	150.0	-124	-0.08%	8 377	8 581	203	2.43%
Action	41					141.9	141.8	-50	-0.04%	4 003	3 991	-13	-0.31%
Immo	42					148.1	147.9	-202	-0.14%	7 710	7 938	228	2.96%
Spread	43					147.5	147.4	-163	-0.11%	7 673	7 777	104	1.36%
Up	44					138.5	139.3	886	0.64%	9 701	9 912	211	2.17%
Down	45					161.3	160.0	-1 253	-0.78%	6 664	6 523	-141	-2.12%

Central	50	-	+5%	-	+50 bps	152.5	152.3	-274	-0.18%	9 218	9 272	53	0.58%
Action	51					142.8	142.7	-58	-0.04%	5 010	4 944	-66	-1.31%
Immo	52					150.5	150.2	-260	-0.17%	8 575	8 729	154	1.80%
Spread	53					150.0	149.7	-349	-0.23%	8 444	8 489	45	0.53%
Up	54					140.8	141.6	738	0.52%	10 562	10 674	112	1.06%
Down	55					163.6	162.1	-1 489	-0.91%	7 573	7 443	-130	-1.71%
Central	60	+25 bps	-	-	+25 bps	149.9	150.1	199	0.13%	10 212	10 023	-189	-1.85%
Action	61					140.7	140.8	161	0.12%	6 149	5 951	-199	-3.23%
Immo	62					147.9	148.0	100	0.07%	9 496	9 457	-39	-0.41%
Spread	63					147.5	147.6	172	0.12%	9 352	9 273	-80	-0.85%
Up	64					145.8	146.7	843	0.58%	10 477	10 373	-104	-0.99%
Down	65					169.4	167.5	-1 887	-1.11%	6 690	6 804	114	1.71%
Central	70	+25 bps	-	-	+50 bps	148.4	148.6	182	0.12%	9 592	9 529	-62	-0.65%
Action	71					139.4	139.4	-9	-0.01%	5 296	5 232	-64	-1.21%
Immo	72					146.4	146.4	32	0.02%	8 897	8 930	33	0.37%
Spread	73					146.0	146.1	28	0.02%	8 725	8 679	-46	-0.53%
Up	74					144.4	145.4	1 001	0.69%	9 889	9 899	10	0.11%
Down	75					168.3	166.8	-1 511	-0.90%	5 227	5 131	-95	-1.83%
Central	80	+25 bps	-10%	-5%	+25 bps	147.2	147.3	144	0.10%	9 169	9 024	-145	-1.58%
Action	81					139.3	139.5	156	0.11%	5 032	4 936	-96	-1.91%
Immo	82					145.3	145.2	-36	-0.03%	8 492	8 416	-75	-0.89%
Spread	83					144.7	144.5	-169	-0.12%	8 343	8 290	-53	-0.64%
Up	84					143.1	143.8	737	0.52%	9 434	9 388	-46	-0.49%
Down	85					166.9	166.0	-896	-0.54%	5 420	4 780	-640	-11.80%
Central	90	+25 bps	-5%	-	-	150.3	150.5	207	0.14%	10 432	10 344	-88	-0.85%
Action	91					141.6	141.6	57	0.04%	6 485	6 282	-203	-3.13%
Immo	92					148.3	148.4	146	0.10%	9 707	9 669	-39	-0.40%
Spread	93					147.8	148.1	285	0.19%	9 521	9 455	-66	-0.69%
Up	94					146.1	147.1	1 016	0.70%	10 691	10 626	-65	-0.61%
Down	95					169.5	168.1	-1 414	-0.83%	7 554	7 712	158	2.09%

Central	100	+50 bps	-	-	-	148.9	149.0	97	0.07%	11 166	11 074	-92	-0.82%
Action	101					139.5	139.5	-13	-0.01%	7 261	7 198	-64	-0.88%
Immo	102					146.9	147.1	168	0.11%	10 420	10 485	66	0.63%
Spread	103					146.5	146.8	279	0.19%	10 246	10 162	-84	-0.82%
Up	104					136.1	137.5	1 390	1.02%	11 908	11 296	-612	-5.14%
Down	105					162.3	160.8	-1 482	-0.91%	9 463	9 624	161	1.70%
Central	110	+50 bps	-	-	+25 bps	147.4	147.5	70	0.05%	10 502	10 379	-123	-1.17%
Action	111					138.1	138.2	89	0.06%	6 537	6 490	-47	-0.73%
Immo	112					145.4	145.5	37	0.03%	9 789	9 812	23	0.24%
Spread	113					145.0	145.1	129	0.09%	9 706	9 671	-35	-0.36%
Up	114					134.8	136.4	1 556	1.15%	11 311	10 760	-551	-4.87%
Down	115					160.6	158.9	-1 687	-1.05%	8 739	8 722	-17	-0.20%
Central	120	+50 bps	-10%	-5%	-	146.2	146.3	120	0.08%	10 083	10 053	-30	-0.29%
Action	121					138.0	138.0	-47	-0.03%	6 266	6 285	19	0.30%
Immo	122					144.3	144.4	88	0.06%	9 438	9 472	34	0.36%
Spread	123					143.7	143.8	82	0.06%	9 279	9 260	-19	-0.21%
Up	124					133.5	134.9	1 423	1.07%	10 731	10 407	-324	-3.02%
Down	125					159.5	158.1	-1 384	-0.87%	8 429	8 387	-42	-0.50%
Central	130	+50 bps	-10%	-5%	+25 bps	144.7	144.7	23	0.02%	9 464	9 399	-66	-0.69%
Action	131					136.6	136.7	94	0.07%	5 574	5 540	-34	-0.61%
Immo	132					142.8	142.7	-71	-0.05%	8 808	8 815	7	0.08%
Spread	133					142.2	142.0	-140	-0.10%	8 730	8 720	-10	-0.12%
Up	134					132.2	133.6	1 452	1.10%	10 185	9 897	-288	-2.82%
Down	135					157.8	156.5	-1 309	-0.83%	7 740	7 511	-229	-2.96%
Central	140	-25 bps	-	-	+25 bps	155.6	155.5	-150	-0.10%	8 992	8 998	6	0.07%
Action	141					146.6	146.7	119	0.08%	4 671	4 404	-268	-5.73%
Immo	142					153.7	153.5	-203	-0.13%	8 212	8 356	144	1.76%
Spread	143					153.1	152.9	-181	-0.12%	8 097	8 158	60	0.74%
Up	144					145.8	146.6	800	0.55%	10 523	10 391	-131	-1.25%
Down	145					169.5	167.6	-1 895	-1.12%	6 652	6 768	116	1.75%

Central	150	-25 bps	-	-	+50 bps	154.0	154.0	-76	-0.05%	8 331	8 432	101	1.21%
Action	151					145.5	145.6	91	0.06%	3 519	3 379	-140	-3.98%
Immo	152					152.1	152.0	-87	-0.06%	7 596	7 807	211	2.78%
Spread	153					151.7	151.5	-152	-0.10%	7 327	7 405	78	1.06%
Up	154					144.4	145.3	941	0.65%	9 917	9 917	0	0.00%
Down	155					168.5	166.9	-1 531	-0.91%	5 179	5 097	-82	-1.59%
Central	160	-25 bps	+10%	+5%	+25 bps	152.9	152.8	-105	-0.07%	7 890	7 888	-2	-0.03%
Action	161					145.6	145.6	10	0.01%	3 208	3 099	-109	-3.38%
Immo	162					151.0	150.8	-176	-0.12%	7 219	7 198	-21	-0.29%
Spread	163					150.3	150.1	-291	-0.19%	7 052	7 077	25	0.35%
Up	164					143.1	143.8	620	0.43%	9 417	9 405	-12	-0.13%
Down	165					167.0	166.1	-909	-0.54%	5 376	4 747	-629	-11.69%
Central	170	-25 bps	+5%	-	-	156.1	156.2	38	0.02%	9 167	9 358	191	2.08%
Action	171					147.5	147.6	137	0.09%	5 179	4 878	-301	-5.81%
Immo	172					154.2	154.2	-45	-0.03%	8 375	8 557	182	2.18%
Spread	173					153.5	153.5	-65	-0.04%	8 244	8 432	187	2.27%
Up	174					146.1	147.1	941	0.64%	10 714	10 642	-71	-0.67%
Down	175					169.6	168.2	-1 439	-0.85%	7 519	7 675	156	2.08%
Central	180	-50 bps	-	-	-	159.9	160.0	134	0.08%	9 434	9 464	29	0.31%
Action	181					151.0	151.2	205	0.14%	5 008	4 610	-398	-7.94%
Immo	182					158.1	158.1	-10	-0.01%	8 534	8 717	183	2.14%
Spread	183					157.5	157.2	-222	-0.14%	8 268	8 370	102	1.23%
Up	184					147.3	148.3	927	0.63%	11 194	11 061	-133	-1.19%
Down	185					170.6	168.7	-1 925	-1.13%	8 096	8 543	448	5.53%
Central	190	-50 bps	-	-	+25 bps	158.3	158.1	-257	-0.16%	8 632	8 596	-36	-0.42%
Action	191					149.8	149.9	83	0.06%	3 790	3 554	-236	-6.24%
Immo	192					156.4	156.1	-289	-0.19%	7 811	7 890	79	1.01%
Spread	193					155.9	155.6	-266	-0.17%	7 566	7 576	10	0.14%
Up	194					145.9	146.7	795	0.55%	10 529	10 382	-147	-1.39%
Down	195					169.4	167.5	-1 877	-1.11%	6 689	6 799	109	1.64%

Central	200	-50 bps	+10%	+5%	-	162.6	162.8	207	0.13%	10 516	10 476	-39	-0.37%
Action	201					152.3	152.5	188	0.12%	6 179	6 151	-28	-0.46%
Immo	202					160.7	160.8	63	0.04%	9 539	9 592	54	0.56%
Spread	203					160.4	160.4	-76	-0.05%	9 089	9 159	70	0.77%
Up	204					150.3	151.0	701	0.47%	12 038	12 004	-35	-0.29%
Down	205					172.9	171.3	-1 688	-0.98%	9 538	10 096	558	5.85%
Central	210	-50 bps	+10%	+5%	+25 bps	161.2	161.1	-140	-0.09%	9 537	9 570	33	0.35%
Action	211					150.9	150.9	-3	0.00%	5 151	5 062	-89	-1.72%
Immo	212					159.2	159.1	-94	-0.06%	8 740	8 692	-48	-0.55%
Spread	213					158.9	159.0	75	0.05%	8 312	8 405	93	1.12%
Up	214					148.9	149.6	690	0.46%	11 333	11 278	-55	-0.49%
Down	215					172.0	170.3	-1 729	-1.01%	7 877	8 251	374	4.75%