

Mémoire présenté devant l'ENSAE Paris
pour l'obtention du diplôme de la filière Actuariat
et l'admission à l'Institut des Actuaires
le 17/03/2022

Par : **Alexandre Morisse**

Titre : **Application des réseaux de neurones récurrents
à l'estimation des calculs réglementaires**

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

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

Membres présents du jury de la filière

*Entreprise : **KPMG***

Nom : Christian-Yann Robert

Signature : 

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

Directeur du mémoire en entreprise :

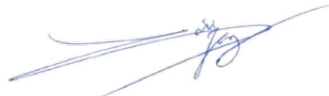
Nom : Loïc Seitz

Signature : 

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


Secrétariat :

Signature du responsable entreprise



Bibliothèque :

Signature du candidat



Résumé

Les entreprises d'assurance-vie utilisent des modèles ALM pour estimer leurs engagements vis-à-vis des assurés. Ils sont cependant très chronophages et freinent la mise en place de l'ORSA. Ce mémoire traite de la mise en place d'un outil de valorisation des passifs des entreprises d'assurance, destiné à des fins réglementaires ou de pilotage. Cet outil est un modèle proxy répliquant les résultats des modèles ALM. Plus précisément, il propose une nouvelle architecture permettant la prédiction des flux financiers en lieu et place des indicateurs économiques (BE, VIF, SCR). Cette architecture a pour but la prise en compte du vieillissement du portefeuille lors des prédictions pluri-annuelles. Elle se base sur l'utilisation de réseaux de neurones récurrents, permettant la transmission de l'information avec le temps. Ce mémoire présente dans un premier temps le modèle ALM dupliqué. Il détaille ensuite les différentes étapes ayant mené à l'architecture finalement retenue et analyse les résultats du modèle en fonction des modifications des paramètres en entrée. Enfin, le modèle est utilisé pour calculer un BSCR et effectuer une projection pluriannuelle.

Mots-clés : Solvabilité II, Best Estimate, Proxy, Modèle ALM, Réseaux de neurones récurrents

Summary

Life-insurance companies use ALM models to estimate their liabilities. These models are very time-expensive and are not well-suited to ORSA. In this actuarial thesis, a tool that estimates the liabilities of a life-insurance company is implemented, on regulatory or management purposes. This tool is a proxy model that replicates the results of ALM models. More precisely, this thesis introduces a new architecture that forecasts financial cash flows, whereas existing proxy models directly compute financial indicators (BE, VIF, SCR). This architecture aims at better taking into account the ageing of the portfolio when forecasting on several years (when doing ORSA for example). This architecture is based on recurrent neural networks. It allows transmission of the information through time. First, this master thesis describes the ALM model that is duplicated. Then, it explains the different steps that led to the final architecture and analyzes the robustness of the model against several inputs. Finally, the model is used to compute the BSCR of the company and conduct an ORSA on several years.

Keywords : Solvency II, Best Estimate, Proxy, ALM model, Recurrent neural networks

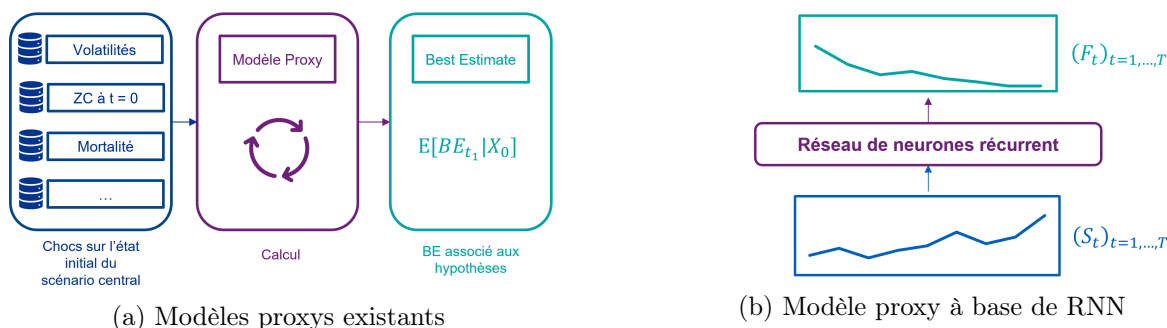
Note de synthèse

Contexte de l'étude

Solvabilité 2 impose aux assureurs-vie différentes contraintes afin de garantir leur solvabilité. La norme leur impose notamment de disposer d'un montant minimal de fonds propres, appelé SCR (*Solvency Capital Requirement*). Ce montant correspond à la valeur théorique minimale de fonds propres à détenir pour garantir leur solvabilité à 1 an avec une probabilité de 99,5%. Pour vérifier que l'entreprise détient bien ce montant, elle doit évaluer son bilan en valeur de marché. Pour ce faire, elle doit en particulier comptabiliser ses engagements vis-à-vis de ses assurés, regroupés sous la dénomination de Best Estimate (BE). Le BE se calcule comme la somme des flux futurs actualisés sous probabilité risque-neutre. En notant BE sa valeur, F_t le flux financier associé au pas de temps t et r_t le taux sans risque de maturité t , on a :

$$BE = \mathbb{E}^{\mathbb{Q}} \left[\sum_{t \geq 0} \frac{F_t}{(1 + r_t)^t} \right]$$

Pour calculer les cash flows, les assureurs ont recourt à des modèles ALM très coûteux en temps de calcul. Ce temps de calcul est un frein à l'application de l'ORSA et à la gestion des risques d'une manière générale. Des modèles proxys ont alors été proposés, permettant de répliquer les calculs de BE issus du modèle ALM de manière instantanée. Ces modèles sont des régressions ou des interpolations (Perceptrons, Krigeages, régressions polynomiales...) permettant de prédire un BE en fonction de différents indicateurs comme la volatilité des marchés, le niveau du CAC40, les taux d'intérêts, la duration des obligations en portefeuille, etc. Cependant, ces modèles proxys sont calibrés pour prédire un BE à une échéance donnée (BE à l'instant présent ou à un an par exemple), qui est immuable. Cette architecture est source d'erreur pour les prédictions pluriannuelles puisqu'elle oblige à négliger l'effet du vieillissement des model points.



Schémas du fonctionnement des différents modèles proxys

Pour corriger ce problème, une nouvelle architecture de modèle proxy est introduite. Cette architecture, basée sur des réseaux de neurones récurrents, prédit les flux financiers passifs en fonction de scénarios économiques donnés. Le BE étant calculé comme la somme des flux futurs actualisés, elle permet des prédictions de BE à des échéances quelconques.

Création de la base de données

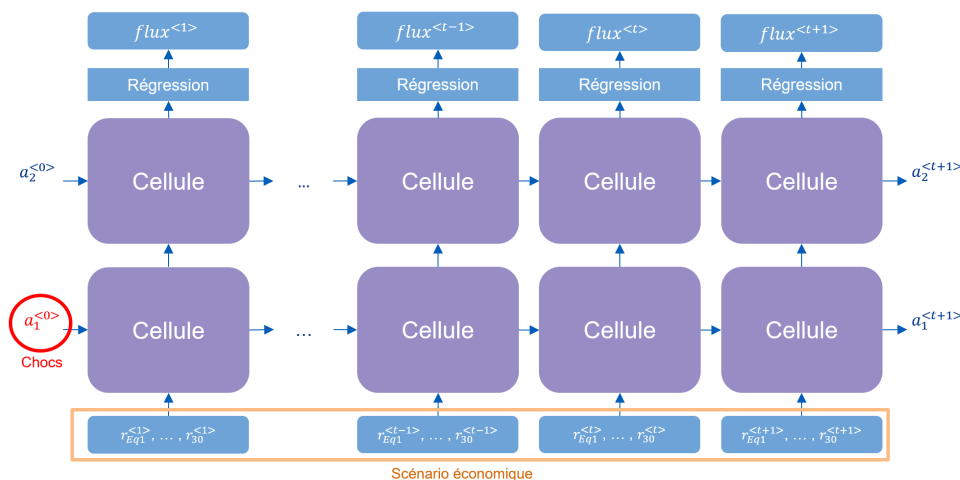
Pour générer la base de données, le modèle ALM est lancé avec une multitude d'inputs différents. Ces inputs sont regroupés dans trois types de fichiers :

- Les model points, qui sont inchangés au fil des projections. Les model points représentent les actifs et passifs de l'entreprise. Les model points passifs ne comportent que des contrats d'épargne d'assurance-vie avec supports en euros.
- Les hypothèses techniques et financières, dont seuls les chocs sur les facteurs de risque varient. Les hypothèses techniques et financières précisent les lois de rachats, les frais, chargements et autres paramètres utilisés lors de la projection du modèle ALM. Les chocs sur les facteurs de risque sont les chocs applicables pour le calcul du SCR à partir de la formule standard. Ils permettent de tester différents changements sur les hypothèses de projection et de faire des calculs de SCR.
- Les scénarios économiques, qui varient à chaque projection. Ils définissent les rendements des actions et de l'immobilier ainsi que les courbes des taux spots applicables sur chaque pas de temps, pour chaque projection.

Les chocs sur les facteurs de risque sont générés selon les matrices de corrélation fournies par Solvabilité II pour agréger les modules et sous-modules. Les courbes des taux initiales sont générées de manière synthétique, via des lois uniformes sur différents points de la courbe, ainsi que l'outil d'interpolation de Smith-Wilson fourni par l'EIOPA. Enfin, les rendements des actions et de l'immobilier sont choisis selon des volatilités aléatoires.

Construction du modèle proxy

Le modèle retenu est basé sur un réseau de neurones récurrents, composé de deux couches, et utilisant des cellules GRU avec des dimensions d'états cachés égales à 64. Ces deux couches sont surplombées d'une couche de régression permettant la prédiction des cash flows actualisés.



Architecture du modèle

Les inputs sont des vecteurs $x^{<t>}$:

$$x^{<t>} = (r_{Eq1}^{<t>, r_{Eq2}^{<t>, r_{Immo}^{<t>, r_1^{<t>, r_5^{<t>, r_{10}^{<t>, r_{20}^{<t>, r_{30}^{<t>})$$

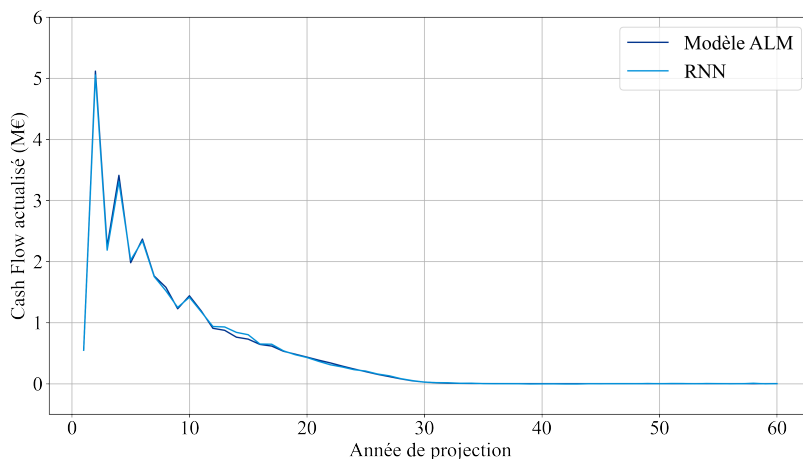
Les deux premières coordonnées du vecteur contiennent les rendements des actions modélisées lors de l'année de projection t . La troisième coordonnée contient les rendements de l'immobilier lors de cette même année. Enfin, les 5 dernières coordonnées du vecteur contiennent les taux spots aux échéances 1, 5, 10, 20 et 30 ans en vigueur au pas de temps t .

La fonction de coût utilisée pour entraîner le modèle est une fonction de coût personnalisée. Elle est composée d'un premier terme mesurant la distance quadratique entre vecteurs prédits et vecteurs à prédire, et d'un second terme pénalisant le biais de prédiction sur le Best Estimate :

$$L(y_i, \hat{y}_i) = \sum_{t=1}^T (\hat{y}_i^{<t>} - y_i^{<t>})^2 + C \left(\sum_{t=1}^T (\hat{y}_i^{<t>} - y_i^{<t>}) \right)^2$$

Analyse des résultats

Le modèle créé permet de prédire des cash flows actualisés. Les cash flows prédits permettent



Exemple de prédiction sur un scénario stochastique

de calculer un BE sur scénario central avec une erreur de 0.02%. Sur la base de données de test, le RMSE est de 0.49%, soit une racine d'erreur quadratique moyenne de 160 k€ sur un bilan de 32 968 k€.

La robustesse du modèle vis-à-vis des différents inputs est ensuite étudiée. L'erreur maximale observée, tout input confondu, est de 0,41% (tous les inputs sont étudiés de manière séparée). Seule l'erreur en fonction des taux dépasse ce seuil : elle est de l'ordre de 1% pour les chocs à la hausse supérieurs à 30BP.

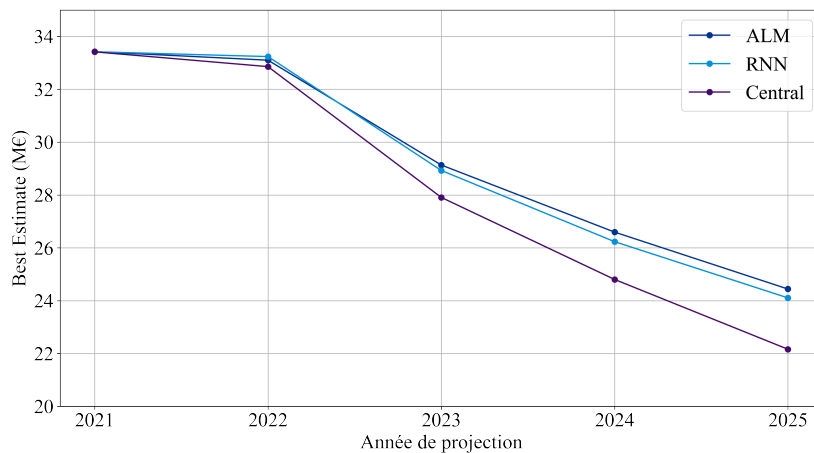
Applications

Le modèle créé permet divers applications comme le calcul de SCR à l'instant présent ou la valorisation du BE à des échéances pluriannuelles.

Le SCR de base est calculé sous la formule standard avec le modèle ALM et le modèle proxy. Il est obtenu comme le montant agrégé des modules risque de marché et risque de souscription vie, les autres modules étant nuls ou négligeables. Les résultats obtenus sont de 935 k€ pour le modèle proxy et de 932 k€ pour le modèle ALM.

Pour vérifier la capacité de prise en compte du vieillissement des model points, un scénario de hausse des taux couplée à une chute de l'immobilier est testé. Les valeurs de BE calculées

via le modèle ALM et prédites via le modèle proxy pour chaque pas de temps sont représentées dans la figure ci-contre. Les BE calculés et prédits y sont mis en perspective par rapport aux BE calculés sur le scénario central. Les résultats du modèle ALM et ceux prédits par le modèle



BE en fonction de l'horizon de projection

proxy sont très proches : l'erreur maximale est de 2% pour une prédiction à un horizon de 4 ans.

Toutefois, plusieurs limites subsistent. Le modèle proxy développé ne permet pas la prise en compte d'affaires nouvelles. Enfin, il est impossible de calculer un SCR à un instant ultérieur à l'instant initial.

Synthesis Note

Context

Solvency II imposes various constraints on life-insurance companies to enhance confidence in their solvency. In particular, it introduces the Solvency Capital Requirement (SCR). It is the theoretic amount of available capital that an insurance company must own to guarantee its solvency with a probability of 99,5%. To ensure it detains this amount of capital, it must compute its balance-sheet in a market-consistent way. To do so, it must compute all its policy insurance liabilities, aggregated together in the Best Estimate (BE). The Best Estimate is computed as the sum of dicounted future cash flows under the risk-neutral probability. Noting BE the best estimate value, F_t the cash flow at time t and r_t the risk free rate with maturity t , the BE is given by :

$$BE = \mathbb{E}^{\mathbb{Q}} \left[\sum_{t \geq 0} \frac{F_t}{(1 + r_t)^t} \right]$$

Life-insurers use ALM models to compute cash flows. These models involve heavy computations that require a lot of time to execute. It prevents companies from fully implementing *Own Risk and Solvency Assessment* and using it to manage risk. Therefore, proxy models have been developped, replicating instantaneously the Best Estimate computation of the ALM model. These models consist in regressions or interpolations (Perceptrons, Krigings, polynomial regressors...). They allow the computation of the BE based on several factors like the volatility, the value of the CAC40, interest rates, portfolio's bond duractions, etc. However, these proxy models only allow to compute the BE at a given date, that can't be changed. This architecture is a source of error for pluriannual predictions and forces to ignore the impact of model points ageing.

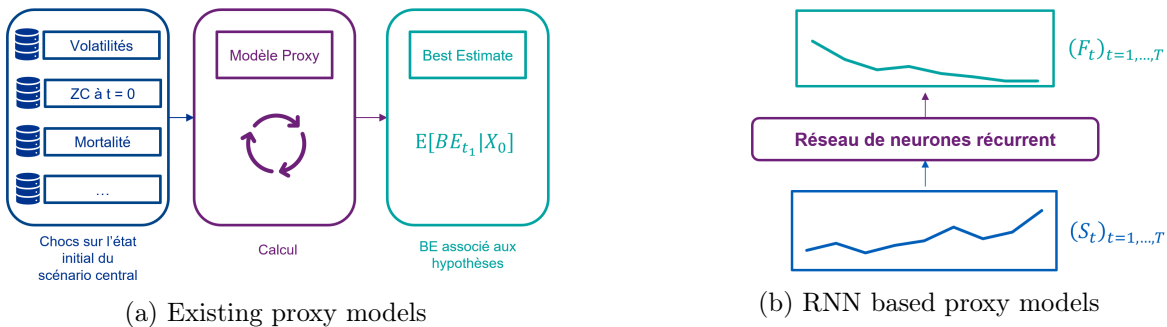


Diagram of the different proxy models

This actuarial thesis introduces a new architecture to deal with this problem. This recurrent neural network based architecture predicts cash flows depending on economical scenarios given as inputs. As the BE is the sum of the discounted cash flows, starting at a specific date, it can be computed to any time horizon.

Database generation

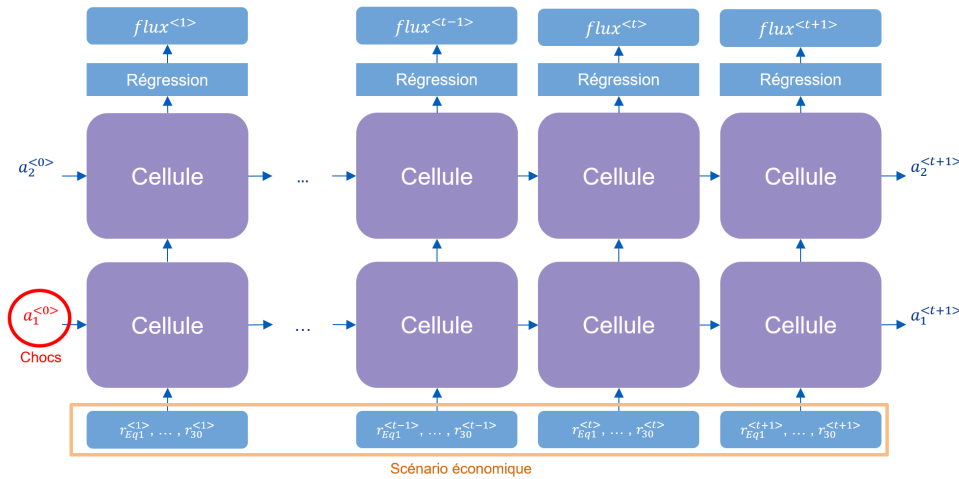
To create the database, the ALM model is launched with several different inputs. These inputs are of three types :

- The model points that don't change during the different projections. The model points represent the company's assets and liabilities. The liabilities only include with profit contracts.
- The technical and financial assumptions that don't change during the different projections. The only assumptions that vary are the shocks on the risk factors. The technical and financial assumptions specify the lapse rates, fees, loading rates and other parameters that are used by the ALM model during the projection. Shocks on risk factors are those used to compute the SCR with the standard formula. They allow to test different assumption changes and compute the SCR.
- The economic scenarios, which vary with each projection. They define the equity and real estate returns as well as the interest rate curves for each time step and each projection.

The shocks on the risk factors are generated according to the correlation matrices provided by Solvency II to aggregate the modules and sub-modules. The initial yield curves are generated synthetically with uniform laws on different points of the curve and with the interpolation tool provided by EIOPA. Finally, equity and real estate yields are chosen according to random volatilities.

Construction of the proxy model

The model chosen is based on two layers of recurrent neural networks using GRU cells with hidden state dimensions equal to 64. These two layers are topped by a regression layer allowing the prediction of discounted cash flows.



Architecture of the proxy model

The inputs are vectors $x^{<t>}$:

$$x^{<t>} = (r_{Eq1}^{<t>}, r_{Eq2}^{<t>}, r_{Immo}^{<t>}, r_1^{<t>}, r_5^{<t>}, r_{10}^{<t>}, r_{20}^{<t>}, r_{30}^{<t>})$$

The first two coordinates of the vector contain the returns of the modeled stocks in the projection year t . The third coordinate contains the real estate yields for the same year. Finally, the

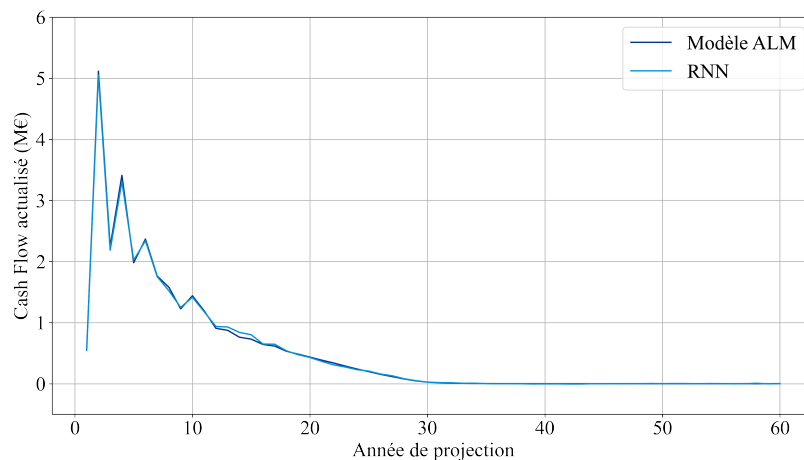
last 5 coordinates of the vector contain the spot rates at maturities of 1, 5, 10, 20 and 30 years at time step t .

The model is trained with a custom loss function. This loss function is defined as the sum of a quadratic loss between predicted and targeted cash flows, and a penalization term on bias of BE prediction :

$$L(y_i, \hat{y}_i) = \sum_{t=1}^T (\hat{y}_i^{<t>} - y_i^{<t>})^2 + C \left(\sum_{t=1}^T (\hat{y}_i^{<t>} - y_i^{<t>}) \right)^2$$

Analysis of results

The created model allows prediction of discounted cash flows. The predicted cash flows



Example of prediction on a stochastic scenario

allows to compute a BE on the central scenario with an error of 0.02%. On the testing set, the RMSE is equal to 0.49%, which represents an error of 160 k€ on a 32 968 k€ balance sheet.

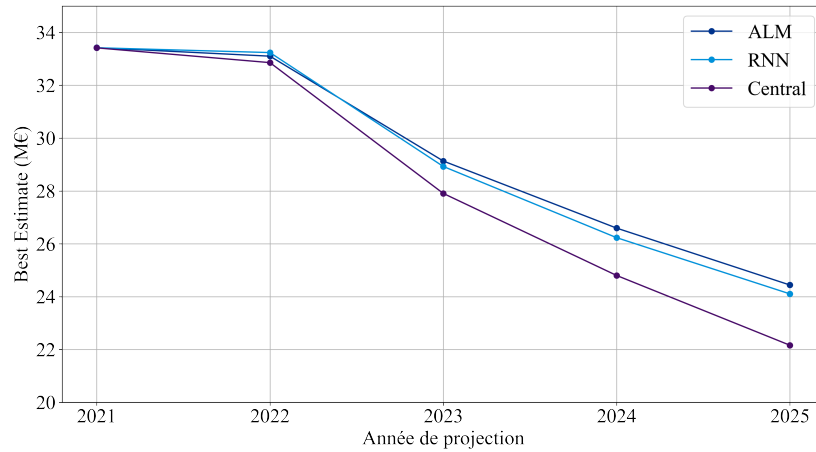
Then, this actuarial thesis studies the model robustness against the different inputs. Each input was tested separately. The maximal error found is of 0,41%. The only error that goes beyond this error is the one related to the interest rate curve. It amounted to 1% for chocs that were superior to 30 BP.

Applications

The final proxy model is suited to immediate SCR computation or computation of Best Estimate on different time horizon (several years).

First, the proxy model and the ALM model are used to compute the Basic SCR under the standard formula assumptions. Its is computed as the aggregated amounts of the market risk module and the life risk module. The proxy model forecasts a 935k€ BSCR whereas the ALM model computes a 932 k€ BSCR.

To verify the ability of the model to take into account the model points ageing through time, a scenario based on an increase of interest rates and a fall of real estate value is tested. The BE computed thanks to the ALM model and predicted through the proxy model are represented in the figure below. They are put in perspective against the BE computed on the central scenario.



BE computed and estimated through time

The results of the ALM model and those of the proxy model are very close : the maximal error is equal to 2% and is obtained for the maximum maturity.

However, some limits remain. The proxy model doesn't allow new businesses to be taken into account. Moreover, it only allows SCR computation at initial time.

Remerciements

Je tiens tout d'abord à remercier Loïc Seitz, directeur au sein du cabinet, pour m'avoir fait confiance et sans qui ce mémoire n'aurait pas pu voir le jour. Je tiens aussi à remercier Ahmed Taïtaï, pour son implication, son suivi et ses nombreuses relectures. Enfin, je les remercie tous les deux pour le temps qu'ils m'ont accordé et les conseils qu'ils m'ont prodigués tout au long de l'écriture de ce mémoire.

Je remercie aussi vivement Christian-Yann Robert, pour son suivi, sa disponibilité et ses conseils tout au long de l'élaboration de ce mémoire.

Je remercie l'ensemble des membres de l'équipe Actuariat et Analytics de KPMG pour leur accueil, leur bonne humeur et toute l'aide qu'ils ont pu m'apporter.

Enfin, je remercie mes parents, mes grand-parents, Guillaume et Modana pour m'avoir aidé durant toutes ces années.

Table des matières

Introduction	15
1 Contexte de l'étude	16
1.1 Assurance-Vie	16
1.1.1 Types de contrat & support	16
1.1.2 Caractéristiques des contrats en euros	16
1.2 Présentation de la réglementation Solvabilité II	18
1.3 Bilan sous Solvabilité II	19
1.3.1 Structure du bilan	19
1.3.2 Best Estimate	20
1.3.3 Risk margin	20
1.4 Calcul du <i>Best Estimate</i>	21
1.4.1 Formulation mathématique du <i>Best Estimate</i>	21
1.4.2 Simulations dans les Simulations	23
1.5 Réduction du temps de calcul	25
1.5.1 Least Squares Monte Carlo	25
1.5.2 Curve Fitting	26
1.5.3 Modèles proxys	27
2 Création de la base de données	31
2.1 Outils de simulation	31
2.1.1 Modèle ALM	31
2.1.2 Générateur de Scénarios Économiques	31
2.2 Actifs	32
2.2.1 Obligations	33
2.2.2 Actions et Immobilier	34
2.2.3 Cash	34
2.3 Passifs	35
2.4 Configuration du GSE	37
2.4.1 Matrice de corrélation	37
2.4.2 Taux d'inflation	37
2.4.3 Modèle de Black-Scholes	38
2.4.4 Volatilités des actions et de l'immobilier	38
2.4.5 Caps	39
2.4.6 Modèle de taux	39
2.4.7 Courbes de taux	41
2.5 Modèle ALM	43
2.5.1 Risque-neutralisation des obligations	43
2.5.2 Allocation d'actifs	44
2.5.3 Algorithme de participation aux bénéfices	44

2.5.4	Rachats	46
2.5.5	Frais et chargements	48
2.6	Chocs sur les facteurs de risque	48
2.6.1	Loi	49
2.6.2	Simulations	50
2.6.3	Implémentation des chocs	51
3	Construction du modèle proxy	53
3.1	Réseaux de neurones récurrents	53
3.1.1	Présentation	53
3.1.2	Fonctionnement	54
3.1.3	Calibration	56
3.2	Architecture initiale du modèle proxy	58
3.2.1	Inputs	58
3.2.2	Outputs	58
3.3	Conditions d'entraînement	59
3.3.1	Loss function et sur-apprentissage	59
3.3.2	Root Mean Square Error	60
3.3.3	Données d'entraînement	61
3.3.4	Données de validation	61
3.4	Choix des cellules	61
3.5	Nombre de units et epochs	62
3.6	Prise en compte des taux d'intérêts	63
3.6.1	Identification du problème	63
3.6.2	Ajout de taux zéro-coupons	65
3.7	Suppression du biais	66
3.7.1	Identification du problème	66
3.7.2	Modification de la loss function	67
3.8	Comparaison des modèles	68
4	Analyse des résultats	69
4.1	Précision sur les données de validation	69
4.1.1	Présentation des données	69
4.1.2	Résultats	69
4.2	Scénario central	70
4.2.1	Description	70
4.2.2	Scénario déterministe	70
4.2.3	Scénario stochastique	71
4.3	Courbe des taux	73
4.4	Volatilités	74
4.5	Chocs sur les facteurs de risque	76
4.5.1	Choc sur les rachats exceptionnels	76
4.5.2	Choc sur les rachats	77
4.5.3	Choc sur le taux de mortalité	77
4.5.4	Choc sur les frais	78
4.5.5	Choc sur les valeurs de marché action	79
4.5.6	Choc sur l'immobilier	79

5 Applications	81
5.1 Calcul du SCR	81
5.1.1 Module "risque de marché"	82
5.1.2 Module "risque de souscription en vie"	83
5.1.3 BSCR	84
5.2 ORSA	84
5.2.1 Présentation	84
5.2.2 Description du scénario	84
5.2.3 Résultats	85
5.2.4 Limites	86
Conclusion	86
A Libor Market Model	92

Introduction

Ce mémoire traite de la mise en place d'un outil de valorisation des passifs des entreprises d'assurance, destiné à des fins réglementaires ou de pilotage. Cet outil est un méta-modèle basé sur des réseaux de neurones récurrents, visant à reproduire les résultats des modèles ALM.

Solvabilité 2 impose aux assureurs-vie d'évaluer leurs provisions techniques en valeur de marché, dont la principale composante est appelée Best Estimate. Pour ce faire, les assureurs ont recours à des modèles ALM, qui permettent de projeter les flux financiers selon un scénario économique donné. L'évaluation du Best Estimate repose alors souvent sur une méthode de Monte-Carlo très coûteuse en temps de calcul.

Pour résoudre ce problème, des méta-modèles (ou modèles proxys) ont été développés. Ces modèles ont pour but de répliquer les résultats des modèles ALM des assureurs, de manière quasi-instantanée. Plusieurs mémoires ont déjà traité la mise en place de ces méta-modèles à des fins de calculs du Best Estimate ou des ratios de solvabilité. Ces mémoires se sont principalement basés sur des méthodes d'interpolation ou d'estimation comme le *Krigeage* [2, 13], les régressions linéaires ou les réseaux de neurones à propagation avant [1]. Cependant, comme le souligne V. Cescutti [2], ils peinent à reproduire les résultats issus du modèle ALM lorsque l'horizon de calcul du Best Estimate s'éloigne. En effet, ces modèles ne peuvent pas prendre en compte le vieillissement des model points. Ceci incite à l'utilisation des modèles proxys pour valoriser des bilans à des horizons proches de celui utilisé pour la calibration, limitant notamment leurs utilisations dans le cadre d'ORSA sur plusieurs années.

Ce mémoire propose de résoudre ce problème via l'utilisation de réseaux de neurones récurrents. Au lieu de prédire le BE à un instant t en fonction d'un choc et de paramètres économiques, il est proposé de construire un modèle prédisant les flux financiers associés à un scénario économique donné, comme le fait le modèle ALM. Cette architecture se rapproche de celle des modèles ALM : elle permet donc toujours le calcul du Best Estimate (comme somme des flux actualisés) et permet de résoudre le problème de vieillissement des model points. Le modèle créé est donc adapté à des calculs de Best Estimate à des échéances pluri-annuelles.

Chapitre 1

Contexte de l'étude

1.1 Assurance-Vie

Nous nous intéressons à la modélisation du bilan d'une entreprise d'assurance-vie. Une entreprise d'assurance-vie est une entreprise proposant des produits d'assurance liés à la vie de l'assuré. Cette branche se distingue de l'assurance non-vie par le fait qu'elle assure des événements certains (décès de l'assuré) mais de dates inconnues. Les produits proposés par les entreprises d'assurance-vie se regroupent en produits d'assurance-vie, produits d'assurance-décès et produits mixtes.

Les produits d'assurance-vie permettent l'obtention de droits, exprimés sous forme de capital ou de rente, en cas de survie de l'assuré. Les contrats d'assurance-vie représentent le moyen d'épargne préféré des Français : selon les chiffres clés de la Fédération française de l'assurance, les encours des contrats d'assurance-vie atteignaient 1 786 milliards d'euros en janvier 2021. Dans cette étude, nous nous intéressons à une entreprise d'assurance-vie disposant uniquement de contrats d'assurance-vie.

1.1.1 Types de contrat & support

Les contrats d'assurance-vie sont principalement composés de deux types de supports :

- les supports en euros,
- les supports en unités de comptes (UC).

Dans les supports en unités de compte, l'assureur garantit à l'assuré un nombre de part dans un ou plusieurs fonds ou valeurs mobilières. Les valorisations de ces fonds ou valeurs mobilières sont variables, et les variations sont totalement imputées à l'assuré : l'assureur ne porte presque aucun risque.

Au contraire, dans les supports en fonds euros, l'assureur garantit une revalorisation donnée du capital en euros placé sur le contrat, quel que soit le rendement de ses investissements. Cette garantie fait porter un risque à l'assureur : en cas de mauvais résultats financiers, il se voit quand même dans l'obligation de revaloriser ou maintenir les provisions mathématiques des assurés.

Les contrats d'assurance vie peuvent être mono-supports ou multi-supports, ce qui revient à disposer dans un même contrat d'un support en euros et d'un support en unités de comptes.

1.1.2 Caractéristiques des contrats en euros

Dans notre étude, nous nous limitons à un portefeuille composé de contrats en euros.

Taux Minimum Garanti

Le Taux Minimum Garanti (TMG) est le taux de revalorisation minimum que l'assureur s'engage à appliquer sur l'épargne de l'assuré. Ce taux est défini contractuellement entre l'assureur et l'assuré. Ce taux est plafonné par l'article A 132-3 du codes des assurances. Il doit être inférieur au minimum entre

- 150% du taux d'intérêt technique maximal,
- le maximum entre 110% de la moyenne des taux servis sur les deux dernières années et 120% du taux d'intérêt technique maximal.

Le taux d'intérêt technique maximal est défini comme $60\% \times TME$, le TME étant le Taux Moyen d'emprunt d'État. Sa méthode de calcul est donnée dans l'article A 132-1. On a donc :

$$TMG_{max} = \min(150\% \times 75\% \times TME, \max(110\% \times \frac{TS_{n-1} + TS_{n-2}}{2}, 120\% \times 75\% \times TME))$$

L'assureur a pour obligation de revaloriser tous les contrats à un taux supérieur à leur TMG. Cette obligation tient quel que soit le taux de rendement moyen de ses actifs.

Participation aux bénéfices

La participation aux bénéfices est la part des bénéfices reversée à un assuré détenant un contrat en euros. La participation aux bénéfices peut être réglementaire, contractuelle ou discrétionnaire.

La participation aux bénéfices réglementaire est la participation aux bénéfices minimale que l'assureur a pour obligation de verser aux assurés. Elle est fixée par la réglementation à travers le taux de participation aux bénéfices réglementaire :

$$PB_{min} = \max(PR_{min} - TMG, 0)$$

avec PR_{min} la participation aux résultats définie comme :

$$PR_{min} = 85\% \times \text{résultat financier} + 90\% \times \text{résultat technique}$$

On remarque que la formule de PB_{min} est cohérente avec l'obligation de l'assureur de verser le TMG à ses assurés. La participation aux bénéfices vient s'ajouter au TMG : elle vient augmenter le taux servi aux assurés si les résultats de l'entreprise le permettent. Le taux finalement versé correspond donc au TMG si les résultats de l'assureur sont insuffisants, et à $85\% \times \text{résultat financier} + 90\% \times \text{résultat technique}$ sinon. Si les résultats techniques de l'assureur sont négatifs, alors l'assureur peut prendre en compte 100% du résultat technique dans le calcul de la participation aux résultats.

La participation aux bénéfices peut aussi être contractuelle. Il s'agit d'une participation aux bénéfices définie dans le contrat, et non définie par la réglementation. Elle est alors forcément supérieure à la participation aux bénéfices réglementaire.

Enfin, l'assureur peut verser des participations aux bénéfices discrétionnaires. Il s'agit de participations aux bénéfices supérieures à celles prévues par la réglementation et le contrat. Le montant de participation aux bénéfices discrétionnaire versé dépend essentiellement de l'environnement économique et de la stratégie commerciale de l'entreprise. Par exemple, en cas de bons rendements des actifs, verser une participation aux bénéfices trop faible expose l'entreprise à des risques de rachats. En effet, cela pourrait entraîner le mécontentement des assurés ayant connaissance de taux servis plus élevés chez la concurrence.

Provision pour participation aux bénéfices

Les montants de participation aux bénéfices peuvent être directement crédités aux assurés, ou alors mis de côté dans les provisions pour participations aux bénéfices (PPB). La réglementation impose le versement des participations aux bénéfices sous 8 ans. Les assureurs mettent en général une part des participations aux bénéfices dans la PPB afin de lisser dans le temps les taux servis aux assurés.

1.2 Présentation de la réglementation Solvabilité II

Entré en application au 1^{er} janvier 2016, Solvabilité II est un ensemble de règles, représenté par trois piliers, qui vise à garantir la solvabilité des entreprises d'assurance et de réassurance exerçant dans l'Union Européenne. Le cadre Solvabilité II est défini par différents textes, rassemblés en trois niveaux [11].

Les textes de niveau 1 sont les deux directives décrivant les principes généraux de Solvabilité II. Ces directives sont la directive 2009/138/CE ("directive Solvabilité II") qui prévoit les grands principes du régime prudentiel Solvabilité II; et la directive 2014/51/UE ("directive Omnibus II") qui amende la première. La directive Omnibus II intègre des prérogatives de l'EIOPA et modifie les règles du pilier 1 concernant la courbe des taux.

Les textes de niveau 2 regroupent le règlement délégué (UE) 2015/35 du 10 octobre 2014 et des notes pédagogiques. Les règlements s'appliquent totalement et directement, contrairement aux directives qui donnent les objectifs à atteindre. Il est à noter que le règlement 2015/35 a fait l'objet d'un amendement par le règlement 2016/467. Ce dernier a pour but de favoriser les investissements des compagnies d'assurance en Europe, conformément au Plan d'investissement pour l'Europe adopté en Novembre 2014 par la Commission Européenne.

Enfin, les textes de niveau 3 regroupent les normes techniques d'exécution (*ITS - Implementation Technical Standards*) ainsi que les orientations (*Guidelines*) proposées par l'EIOPA. Les ITS sont adoptées par la Commission Européenne et directement applicables, alors que les orientations s'appliquent au niveau de chaque autorité nationale selon le schéma "*comply or explain*".

Le règlement délégué 2015/35 regroupe les exigences que doivent satisfaire les entreprises d'assurance en trois piliers :

- le pilier I qui regroupe les exigences quantitatives relatives à la valorisation des actifs et passifs de l'entreprise, les exigences de capital minimal requis (ou *MCR - Minimum Capital Requirement*), de capital de solvabilité requis (ou *SCR - Solvency Capital Requirement*), et le mode de calcul de ce dernier (choix entre formule standard ou modèle interne). Il précise aussi les conditions d'utilisation des modèles internes.
- le pilier II, regroupant d'une part les exigences en termes de gouvernance et de gestion des risques, d'autre part les exigences quant à l'évaluation interne des risques et de la solvabilité (ou *ORSA - Own Risk and Solvency Assessment*).
- le pilier III, regroupant les exigences de communication de l'information à destination du public et des superviseurs. Celui-ci vise notamment à harmoniser la communication des entreprises d'assurance au niveau européen. Il précise aussi les fréquences de remise ou de publication des différents rapports.

Le contenu de chacun de ces trois piliers est résumé dans la figure 1.1.



FIGURE 1.1 – Piliers de Solvabilité II

1.3 Bilan sous Solvabilité II

1.3.1 Structure du bilan

Le pilier I impose aux entreprises d'assurance de valoriser leurs actifs et leurs passifs avec des méthodes de valorisation au prix du marché. Cette valorisation doit couvrir les engagements se rapportant à l'activité existante uniquement, sous l'hypothèse de continuité d'exploitation de l'entreprise. Les actifs sont donc, sauf exceptions, comptabilisés en valeur de marché.

Les passifs des entreprises d'assurance et de réassurance sont composés des fonds propres et provisions techniques valorisées en valeur de marché. Les fonds propres sont calculés comme la différence entre les actifs et les provisions techniques.

Le pilier I de Solvabilité II introduit deux nouvelles exigences relatives aux fonds propres : le MCR et le SCR. Le MCR correspond à la quantité minimale de fonds propres dont doit disposer toute entreprise exerçant une activité d'assurance ou de réassurance. Le non-respect de cette règle déclenche automatiquement l'intervention des autorités de contrôle, et peut se solder par le retrait de l'agrément. Le SCR correspond quant à lui à la quantité de fonds propres requis pour garantir la solvabilité de l'entreprise d'assurance ou de réassurance à un an avec une probabilité de 99,5%. Toute entreprise ne disposant pas des fonds propres requis pour respecter l'exigence du SCR s'expose à des sanctions des autorités de contrôle, comme une exigence de capital requis supplémentaire.

Pour calculer la quantité de fonds propres disponibles, il est nécessaire de calculer les provisions techniques. Comme les actifs, les passifs doivent être valorisés à leur valeur de marché. Etant donné l'impossibilité d'observer directement les valeurs de provisions techniques sur un marché (inexistence d'un marché liquide des passifs), les régulateurs ont défini les provisions techniques comme la somme entre la meilleure estimation (plus connue dans sa version anglaise *Best Estimate* et abrégée BE ou BEL) et la marge de risque (*Risk Margin* abrégée RM). Une représentation schématique du bilan d'une entreprise d'assurance ou de réassurance sous Solvabilité II est donnée figure 1.2.

La marge de risque est calculée de manière à ce que la valeur totale des provisions techniques corresponde "au montant qu'une autre entreprise d'assurance ou de réassurance demanderait pour reprendre et honorer les engagements d'assurance et de réassurance" de l'entreprise. En

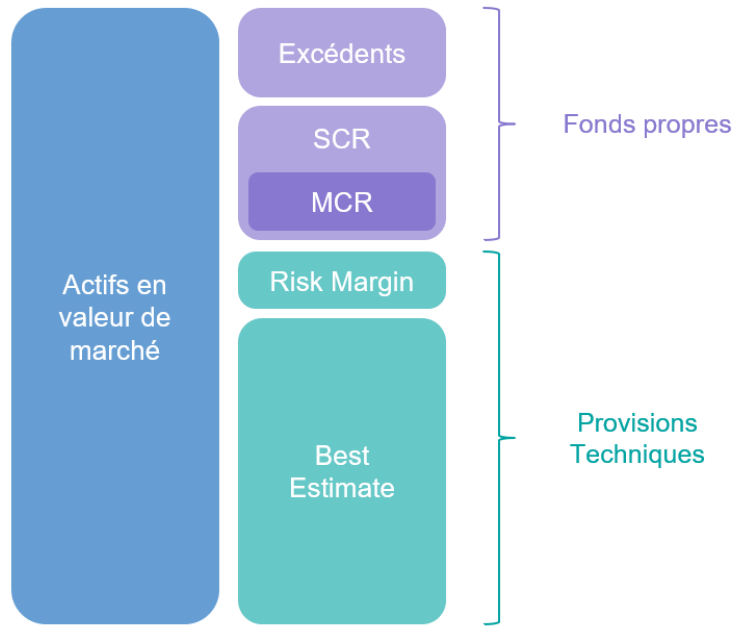


FIGURE 1.2 – Bilan économique sous Solvabilité II

notant PT les provisions techniques, on a donc :

$$PT = BE + RM$$

.

1.3.2 Best Estimate

Le Best Estimate est défini par la réglementation comme "la moyenne des flux de trésorerie futurs, pondérés par leur probabilité, et compte tenu de la valeur temporelle de l'argent" estimée sur la base de la courbe des taux sans risques pertinents. Notons que le calcul s'effectue brut de réassurance. En notant F_t le flux de trésorerie de l'année t et r_t le taux sans risque de maturité t , on en déduit que :

$$BE = \mathbb{E}^{\mathbb{Q}} \left[\sum_{t \geq 0} \frac{F_t}{(1+r_t)^t} \right] \quad (1.1)$$

En introduisant les déflateurs $\delta_t = \frac{1}{(1+r_t)^t}$ on obtient :

$$BE = \mathbb{E}^{\mathbb{Q}} \left[\sum_{t \geq 0} F_t \delta_t \right]$$

1.3.3 Risk margin

Elle correspond au coût du capital requis pour couvrir l'exigence de capital additionnel associée au portefeuille, jusqu'à son extinction. Notons CoC le coût du capital. Si la SCR_t désigne l'exigence en capital additionnel associée au portefeuille, CoC le coût du capital et r_t le

taux sans risque de maturité t , la marge de risque du portefeuille est calculée par la formule :

$$RM = CoC \sum_{t \geq 0} \frac{SCR_t}{(1 + r_{t+1})^{t+1}}$$

Le coût du capital CoC est fixé à 6% par la réglementation.

1.4 Calcul du *Best Estimate*

Par convention, on parle de BE déterministe lorsque le BE est calculé comme la simple somme des flux futurs actualisés sur un scénario central déterministe. Un scénario central déterministe est un scénario dans lequel tous les actifs évoluent selon la "moyenne". Dans ce scénario, les actions et l'immobilier ont par exemple des rendements constants égaux à leurs rendements moyens. Le bilan étant en valeur de marché, le rendement moyen des deux actifs précédents correspond au taux sans risque. On parle au contraire de BE stochastique lorsque le BE est calculé comme la moyenne des flux futurs actualisés sur un "grand" nombre de scénarios aléatoires.

La réglementation suggère aux entreprises d'assurance vie d'utiliser des méthodes de simulation pour produire des calculs appropriés de la meilleure estimation [9]. En effet, les calculs déterministes de la meilleure estimation ne permettent pas de correctement prendre en compte les effets des options et garanties des contrats en portefeuille. Sur un scénario central déterministe, un certain nombre des options des contrats ne sont pas déclenchées, par exemple parce que les rendements des actifs sont suffisant pour servir des taux supérieurs aux TMG. Cependant, dans un scénario plus pessimiste où les rendements des actifs sont faibles, certaines options auparavant moins contraignantes peuvent le devenir.

C'est pourquoi seul le calcul d'un BE stochastique permet de satisfaire les exigences de la réglementation. Dénotons la valeur temporelle des options et garanties contenues dans le portefeuille passif par l'acronyme *TVOG* (Time Value of Options and Guarantees). Comme le BE déterministe ne permet pas de prendre en compte la valeur des options et garanties en portefeuille, on a :

$$TVOG = BE_{sto} - BE_{det}$$

Les entreprises d'assurance se sont tournées vers les méthodes de Monte-Carlo et ont logiquement recours à des modèles stochastiques de projection des flux financiers. Ces modèles sont basés sur la génération de scénarios économiques et l'utilisation d'un modèle ALM. Le modèle ALM permet la simulation des interactions entre les variations de l'environnement économique et financier (définies dans les scénarios), les comptes de l'assureur, les taux servis aux assurés, les taux de rachats (structurels et conjoncturels), les décès et la politique de gestion actif-passif de l'entreprise d'assurance. Ces simulations se déroulent de manière itérative : variation des actifs, calcul du taux cible, réajustement du portefeuille d'actif selon la stratégie de gestion actifs-passifs et calcul du taux servi et des taux de rachats et décès.

La méthode classique utilisée pour calculer le BE stochastique est la méthode des Simulations dans les Simulations (abrégée *SdS*). Elle permet de valoriser un bilan économique à un instant t quelconque en procédant à un grand nombre de simulations, et en utilisant la méthode de Monte-Carlo.

A partir de maintenant et sauf mention contraire, l'acronyme BE (ou BEL) fera référence au BE stochastique.

1.4.1 Formulation mathématique du *Best Estimate*

La formule (1.1) du Best Estimate donnée n'est valable que pour le calcul d'un BE immédiat, *i.e.* la valeur actuelle (au sens temporel du terme) des engagements vis-à-vis des assurés.

Cependant, on peut vouloir calculer la valeur la plus probable de nos engagements à un horizon temporel éloigné. Supposons par exemple que l'on se trouve au 1er janvier de l'année t_0 , et que l'on souhaite donner la meilleure estimation du BE de clôture. Celui-ci correspond au BE le plus probable du 31 décembre de cette même année, qui par continuité du BE est égal au BE au 1^{er} janvier de l'année suivante. On cherche donc à calculer le BE à la date $t_0 + 1$. De manière plus générale, on peut vouloir estimer le BE stochastique à une échéance quelconque notée $t_1 \geq t_0$.

Avant d'introduire la formulation mathématique du problème, il est nécessaire de bien comprendre l'intuition derrière les différentes probabilités qui vont être introduites.

Lorsque l'on souhaite valoriser un produit quelconque de manière "market-consistent" (qu'il s'agisse d'un élément de bilan ou d'un produit financier), on calcule l'espérance sous la probabilité risque-neutre des flux futurs actualisés adossés à ce produit. Dans un marché complet sans opportunité d'arbitrage, la probabilité risque-neutre est unique et garantit l'unicité des prix. La valeur de l'espérance des flux futurs actualisés obtenue correspond alors au coût minimal du portefeuille répliquant les flux de ce "produit" (et donc permettant de couvrir ce risque). Le portefeuille répliquant est créé à partir des actifs **disponibles sur les marchés financiers**, sous les hypothèses classiques utilisées en mathématiques financières (absence d'opportunité d'arbitrage, possibilité de vente à découvert, existence de taux sans risque et divisibilité parfaite des actifs et sous-jacents). La valeur d'une option d'achat ou de vente d'un actif financier calculée sous probabilité risque-neutre diffère bien de la valeur calculée sous probabilité historique, puisqu'il est possible de trouver une couverture à moindre frais en achetant ou vendant les actifs sous-jacents du contrat. Prenons maintenant l'exemple d'une police d'assurance automobile. Les risques couverts par les polices d'assurance automobile ne sont pas reproductibles par un portefeuille d'instruments financiers. Par conséquent, le prix de la couverture minimale de ces risques correspond bien à la simple espérance de la somme des flux futurs actualisés adossés aux contrats en portefeuille.

Lorsque l'on souhaite donner la valorisation "la plus probable" du bilan à une date quelconque t_1 ultérieure à l'instant initial, on cherche l'espérance sous la probabilité historique (aussi appelée probabilité "monde réel") de cette valorisation en t_1 . Cette valorisation dépend alors de l'évolution de l'environnement entre l'instant initial et l'instant ultérieur, ce dernier étant le point de départ de la valorisation sous probabilité risque-neutre.

Notons :

- F_t les flux de l'année t ,
- $r_{t_2}^{t_1}$ le taux sans risque à la date t_1 et de maturité t_2 ,
- \mathbb{P} la probabilité historique ou "monde réel",
- \mathbb{Q} la probabilité risque-neutre,
- $(\mathcal{F}_t)_t$ la filtration naturelle associée à l'ensemble des processus permettant de définir le contexte dans lequel évolue notre entreprise.

Alors le BE "le plus probable" à la date t_1 est :

$$\mathbb{E}^{\mathbb{P}} \left[\mathbb{E}^{\mathbb{Q}} \left[\sum_{t=t_1}^{\infty} \frac{F_t}{(1 + r_t^{t_1})^{t-t_1}} \middle| \mathcal{F}_{t_1} \right] \right] \quad (1.2)$$

Précédemment, nous avons introduit la formulation du Best Estimate d'une entreprise comme $BE = \mathbb{E}^{\mathbb{Q}} \left[\sum_{t=0}^{\infty} \frac{F_t}{(1+r_t)^t} \right]$. Il s'agissait donc d'une valeur déterministe. Cependant, le

Best Estimate à un instant futur $t_1 > 0$ est une variable aléatoire dont l'expression est :

$$\text{BE}_{t_1} = \mathbb{E}^{\mathbb{Q}} \left[\sum_{t=t_1}^{\infty} \frac{F_t}{(1+r_t^{t_1})^{t-t_1}} \middle| \mathcal{F}_{t_1} \right] \quad (1.3)$$

On remarque que cette formule est bien cohérente avec la formule précédente puisque

$$\text{BE}_0 = \mathbb{E}^{\mathbb{Q}} \left[\sum_{t=0}^{\infty} \frac{F_t}{(1+r_t^0)^t} \middle| \mathcal{F}_0 \right] = \mathbb{E}^{\mathbb{Q}} \left[\sum_{t=0}^{\infty} \frac{F_t}{(1+r_t^0)^t} \right] = BE$$

On peut donc réécrire la quantité d'intérêt (1.2) comme :

$$\mathbb{E}^{\mathbb{P}} [\text{BE}_{t_1}] \quad (1.4)$$

Ce qui est cohérent puisque cette quantité correspond à l'espérance d'un Best Estimate futur.

1.4.2 Simulations dans les Simulations

Principe de la méthode SdS

La méthode de Monte-Carlo consiste en l'approximation d'une espérance de variable aléatoire par la somme de ses réalisations. Dit autrement, si la suite $(X_i)_{i=1, \dots, N}$ est indépendante et identiquement distribuée (i.i.d.), nous avons pour N suffisamment grand :

$$\mathbb{E} [X_1] \approx \frac{1}{N} \sum_{i=1}^N X_i$$

Pour estimer (1.4), nous utilisons deux Monte-Carlo imbriqués (d'où le nom de Simulations dans les Simulations). Nous simulons d'abord N fois sous \mathbb{P} l'évolution de l'ensemble des processus définissant \mathcal{F}_t , entre t_0 et t_1 . Nous obtenons alors un ensemble de N scénarios possibles définissant N états du monde probables en t_1 représentés par les N filtrations $(\mathcal{F}_{t_1}^n)_{n=1, \dots, N}$.

Nous appliquons ensuite une première méthode de Monte-Carlo à chaque filtration pour n allant de 1 à N , en simulant K scénarios de flux sous \mathbb{Q} pour les années $t = t_1, \dots, T$, en partant de l'état du monde $\mathcal{F}_{t_1}^n$ (T suffisamment grand pour que les flux actualisés pour $t > T$ soient négligeables). Nous avons alors pour tout n allant de 1 à N :

$$\text{BE}_{t_1}^n = \mathbb{E}^{\mathbb{Q}} \left[\sum_{t=t_1}^{\infty} \frac{F_t}{(1+r_t^{t_1})^{t-t_1}} \middle| \mathcal{F}_{t_1}^n \right] \approx \frac{1}{K} \sum_{k=1}^K \sum_{t=t_1}^T \frac{F_t^k}{(1+r_t^{t_1})^{t-t_1}}$$

Par soucis de simplification, nous n'avons pas indicé les flux sommés et les taux zéro-coupons par n , bien qu'ils soient à chaque fois propres au scénario n .

Il suffit alors d'appliquer la deuxième méthode de Monte-Carlo en agrégeant les résultats obtenus pour n allant de 1 à N :

$$\mathbb{E}^{\mathbb{P}} [\text{BE}_{t_1}] \approx \frac{1}{N} \sum_{n=1}^N \widehat{\text{BE}}_{t_1}^n \quad (1.5)$$

avec

$$\forall n \quad \widehat{\text{BE}}_{t_1}^n = \frac{1}{K} \sum_{k=1}^K \sum_{t=t_1}^T \frac{F_t^k}{(1+r_t^{t_1})^{t-t_1}}$$

Soit en développant (1.5) :

$$\mathbb{E}^{\mathbb{P}} [\text{BE}_{t_1}] \approx \frac{1}{N} \frac{1}{K} \sum_{n=1}^N \sum_{k=1}^K \sum_{t=t_1}^T \frac{F_t^k}{(1+r_t^{t_1})^{t-t_1}}$$

BE probable à échéance donnée

Plus concrètement, supposons que nous nous situons à l'instant $t = 0$ et que nous souhaitons connaître le bilan le plus probable en $t = t_1$. Comme vu dans la section précédente, la méthode des simulations dans les simulations s'effectue en deux phases :

1. Simulations "monde réel" : nous simulons N trajectoires sous probabilité historique \mathbb{P} , produisant N séries de cash flows de $t = 0$ à $t = t_1$.
2. Simulations "risque-neutre" : nous simulons K trajectoires sous probabilité risque-neutre \mathbb{Q} , produisant, pour chacun des N scénarios monde réel, K séries de cash flows de $t = t_1$ à $T \gg t_1$

Il suffit ensuite d'utiliser la formule 1.5 pour obtenir le résultat.

Ces N et K trajectoires respectives sont représentées figure 1.3. Bien que pour simplifier, nous n'ayons représenté que $N = 4$ et $K = 3$ trajectoires, ces deux nombres sont en général de l'ordre du millier.

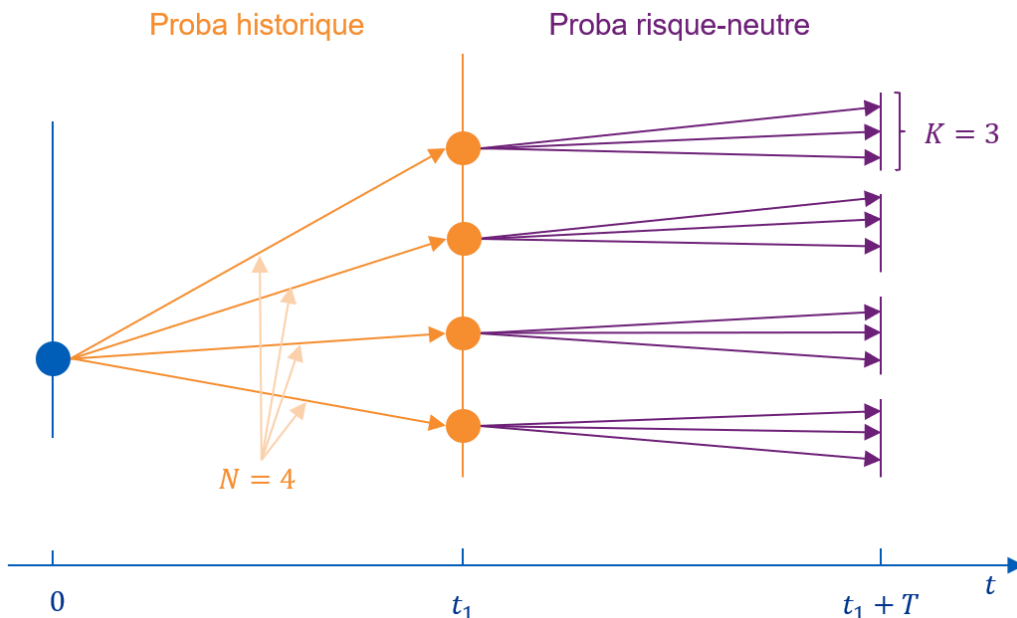


FIGURE 1.3 – Schéma de la méthode SdS

Scénario historique prédéfini

Jusqu'ici, nous avons cherché à évaluer l'espérance du BE stochastique à échéance donnée. Cependant, nous voulons aussi pouvoir connaître le BE stochastique à échéance donnée, selon un scénario monde réel bien précis. C'est le cas par exemple lorsqu'on effectue des projections ORSA sur plusieurs années. Pour faire écho à la section précédente, ceci revient à estimer

$$\mathbb{E}^{\mathbb{Q}} \left[\sum_{t=t_1}^{\infty} \frac{F_t}{(1 + r_t^{t_1})^{t-t_1}} \middle| \mathcal{F}_{t_1} \right]$$

avec \mathcal{F}_{t_1} fixé au lancement de l'algorithme.

Un scénario monde réel prédéfini est par exemple un scénario dans lequel les marchés actions chutent de 10% chaque année pendant 2 ans, les taux restent les mêmes, l'immobilier augmente de 2% par an, etc.

Il s'agit alors d'un problème plus simple que celui énoncé précédemment. Le calcul s'effectue toujours en deux phases :

1. Évolution du bilan en "monde réel" : nous calculons une unique série de cash flows de $t = 0$ à $t = 1$
2. Simulations "risque-neutre" : nous simulons K trajectoires sous probabilité risque-neutre \mathbb{Q} , produisant K séries de cash flows de $t = t_1$ à T

Il suffit alors d'appliquer la méthode de Monte-Carlo pour estimer l'espérance conditionnelle. Cette configuration est représentée figure 1.4.

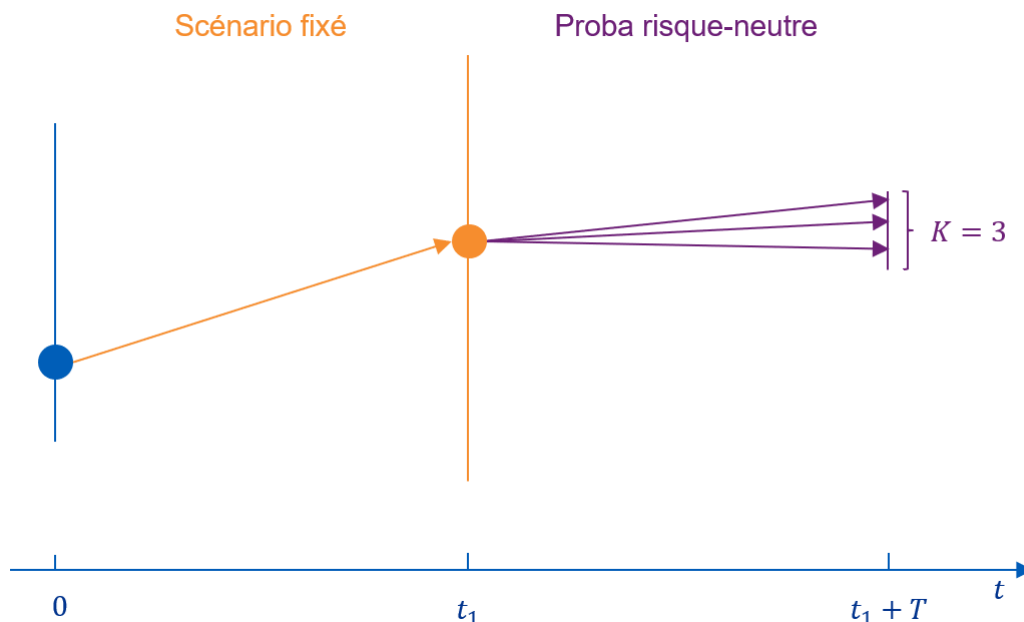


FIGURE 1.4 – Schéma de la méthode pour des projections ORSA sur plusieurs années

1.5 Réduction du temps de calcul

Le temps de calcul important de la méthode des Simulations dans les Simulations vient principalement de la composition des deux méthodes de Monte-Carlo. Il induit de fortes contraintes opérationnelles comme lors de la réalisation d'études ORSA. Pour se dégager de ces contraintes, les entreprises peuvent avoir recours à différentes techniques. Elles peuvent mettre en place des méthodes de calcul basées sur une réduction du temps de calcul pour chaque résultat, comme dans les méthodes Least Squares Monte Carlo et Curve-Fitting. Elles peuvent aussi effectuer un certain nombre de simulations en amont et déduire de ces simulations des résultats complémentaires, de manière quasi-immédiate (évaluation d'une fonction d'interpolation ou de régression en un point).

Deux des méthodes les plus connues permettant d'obtenir des résultats avec moins de temps de calcul sont les méthodes Least Squares Monte Carlo (LSMC) et Curve Fitting. Ces deux méthodes sont assez similaires : elles se basent toutes les deux sur une réduction du nombre de simulations effectuées.

1.5.1 Least Squares Monte Carlo

La méthode LSMC se base sur une réduction du nombre de simulations risque-neutre utilisées pour effectuer la méthode de Monte-Carlo risque-neutre, puis une estimation par régression

(par exemple polynomiale). En effet, en réduisant le nombre de simulations risque-neutres utilisées pour chacune des N trajectoires monde réel, *i.e.* en prenant $K' \ll K$, nous diminuons le temps de calcul des N BE stochastiques, au prix d'une imprécision accrue. Pour K' très réduit (par exemple $K' = 2$), il devient nécessaire d'estimer les vraies valeurs des BE stochastiques à partir des valeurs calculées, trop imprécises. Pour ce faire, on cherche à estimer la vraie valeur de ces BE stochastiques en effectuant une régression (suffisamment "régulière"), que l'on ajuste par la méthode des moindres carrés. Par cet ajustement, on cherche à éliminer l'incertitude ou bruit autour des valeurs des BE calculés, incertitude que l'on suppose centrée. La figure 1.5 schématise le fonctionnement de cette méthode.

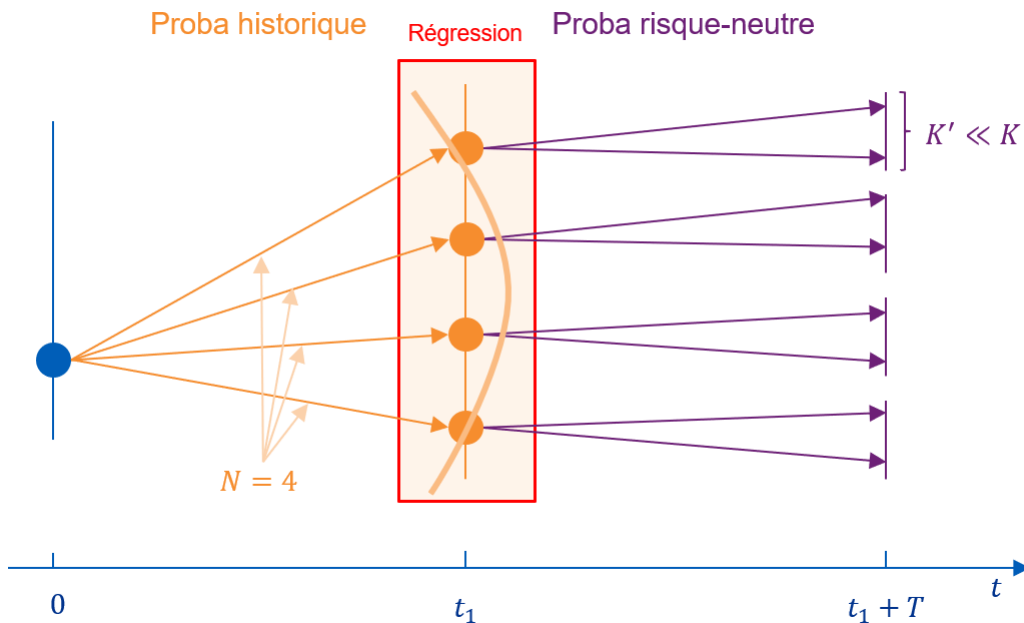


FIGURE 1.5 – Schéma de la méthode Least Squares Monte-carlo

1.5.2 Curve Fitting

La méthode de Curve Fitting se base sur la réduction du nombre de simulations monde réel. Nous effectuons $N' \ll N$ simulations sous probabilité historique. Nous calculons ensuite les N' BE stochastiques associés de manière exacte en prenant un nombre K similaire à celui utilisé dans la méthode SdS. Rappelons à titre de comparaison que dans la méthode LSMC, les N BE calculés sont très imprécis, car $K' \ll K$. Nous effectuons alors une interpolation entre les N' BE obtenus. Il est important de noter qu'ici, une méthode d'interpolation et non de régression s'impose. En effet, il ne s'agit plus d'estimer la vraie valeur d'un paramètre que l'on suppose incertain ou bruité (à cause d'un Monte-Carlo risque-neutre imprécis), mais plutôt de calculer de nouvelles valeurs inconnues à partir de notre connaissance de quelques points exacts (K suffisamment grand pour garantir une bonne précision des $N' \ll N$ BE stochastiques calculés).

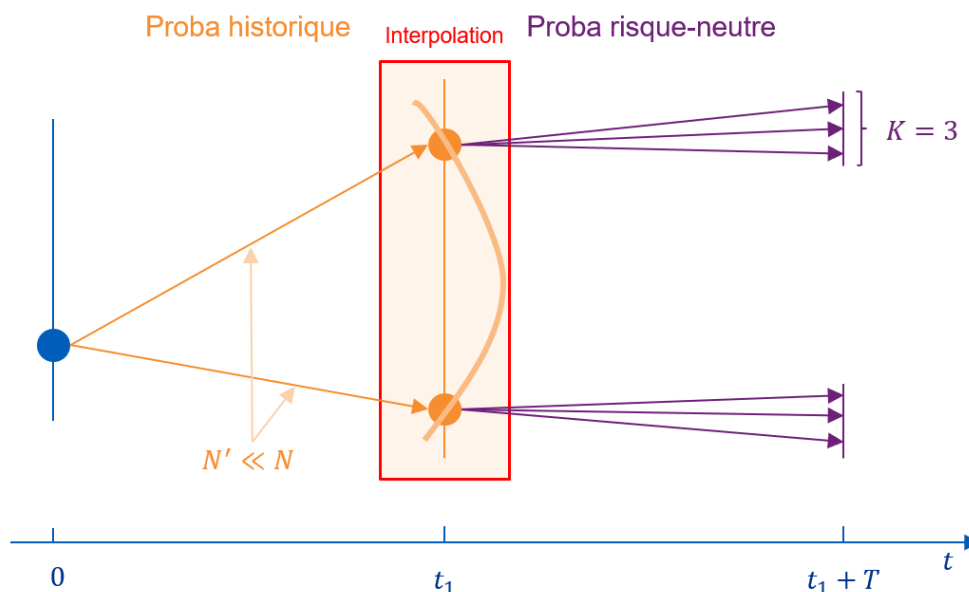


FIGURE 1.6 – Schéma de la méthode Curve Fitting

1.5.3 Modèles proxys

Modèles existants

L'autre approche revient à utiliser des modèles proxys. Ces modèles ont pour but de répliquer au plus proche les résultats des modèles de projection des flux des assureurs, en un temps très restreint ou quasi-instantané. Ces modèles sont en général développés à partir d'une base de données contenant les variations de paramètres d'entrée du modèle comme variables explicatives, et les indicateurs économiques d'intérêts calculés en sortie du modèle comme variable à expliquer. Cette base de donnée est obtenue en lançant un grand nombre de simulations sur le modèle complet de projection des flux, avec des inputs différents. Les modèles proxys consistent alors en l'ajustement d'interpolations ou de régressions sur cette base de données. Dès lors, il suffit de donner les paramètres de simulation souhaités exprimés en entrée de la fonction pour obtenir une évaluation de ce qu'aurait été la valeur calculée par le modèle complet.

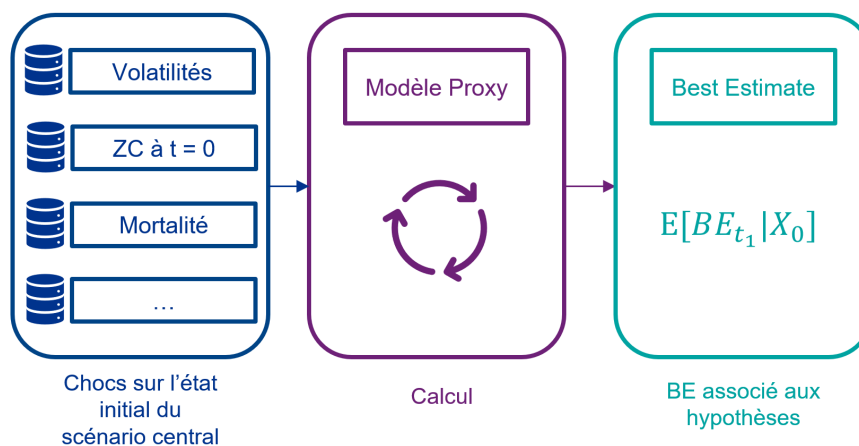


FIGURE 1.7 – Fonctionnement des modèles proxys existants

On cherche par exemple une fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$ telle $\forall i, y_i \approx f(x_i)$. f peut par exemple être une régression linéaire, une fonction polynomiale, un krigeage, ou un réseau de neurones. $X = (x_1, \dots, x_n)^T$ contient les chocs sur les paramètres d'entrée par rapport à un scénario

central, comme par exemple les différentes configurations de volatilité des actions et de l'immobilier, de taux, de mortalité et de frais. $y = (y_1, \dots, y_n)^T$ contient quant à lui les *BE* calculés avec chacune de ces configurations. Il peut aussi contenir les *SCR* ou *VIF*, selon le résultat souhaité. Plusieurs mémoires ont déjà traité la mise en place de ces méta-modèles à des fins de calculs réglementaires, basés sur des méthodes d'interpolation comme le Krigeage [2, 13] ou d'estimation comme les régressions linéaires ou les réseaux de neurones à propagation avant [1].

Limites

Ces modèles proxys produisent de très bons résultats pour la valorisation du bilan (BE stochastique) et le calcul des indicateurs financiers comme le ratio SCR. Cependant, comme le souligne V. Cescutti [2], ils peinent à reproduire les résultats issus du modèle ALM lorsque la date de valorisation du bilan est éloignée de la date de calibration (t_1) utilisée dans la base de données.

Ces modèles sont calibrés sur des bases de données $(x_1, \dots, x_n)^T$ et $(BE_1^{t_1}, \dots, BE_n^{t_1})^T$, avec $(x_i)_{i=1, \dots, n}$ les différentes configurations des chocs sur les facteurs de risque à $t = 0$. Cette structure de données force l'utilisation du modèle pour la prédiction des résultats à échéance t_1 . En effet, pour produire des résultats à horizon t_2 supérieur à t_1 , il devient nécessaire de supposer que le BE en t_2 sachant la valeurs des facteurs de risque en $t_2 - t_1$ est égal au BE en t_1 sachant ces mêmes facteurs de risque en 0. Cette hypothèse peut se réécrire sous la forme :

$$\forall x \in \mathbb{R}^d, t_2 > 0 \quad \mathbb{E}^{\mathbb{Q}} [BE_{t_2} | X_{t_2-t_1} = x] = \mathbb{E}^{\mathbb{Q}} [BE_{t_1} | X_0 = x] \quad (1.6)$$

Cette hypothèse traduit bien l'indépendance de la valeur du Best Estimate par rapport :

1. à l'horizon de valorisation : elle ne permet donc pas de prendre en compte le passage du temps entre t_1 et t_2 . La mortalité, le vieillissement des contrats et des assurés et les rachats structurels ayant eu lieu entre t_1 et t_2 sont donc ignorés.
2. à la trajectoire économique de l'environnement entre l'instant initial et l'instant de valorisation, qui peut avoir une influence long terme sur la projection.

Elle empêche l'utilisation de ces modèles pour effectuer des ORSA sur plusieurs années car elle engendre des erreurs importantes pour les dates de projection éloignées. En effet, elle sous-entend que la valeur du BE ne dépend pas de l'horizon de projection. Elle ne permet donc pas de prendre en compte le passage du temps entre t_1 et t_2 . La mortalité, le vieillissement des contrats et des assurés et les rachats structurels ayant eu lieu entre t_1 et t_2 sont donc ignorés. Cette simplification entraîne logiquement des erreurs croissantes avec le temps, *i.e.* avec t_2 .

De plus, l'équation précédente implique que le BE est indépendant de la "trajectoire de l'environnement" dans lequel évolue l'entreprise entre l'instant initial et l'instant de valorisation. L'environnement correspond à tous les paramètres pouvant avoir une influence sur la valorisation de l'entreprise : les marchés financiers, la mortalité, les taux servis par la concurrence...

Supposons que nous devons valoriser le bilan d'une entreprise d'assurance à un horizon de 4 ans et que $t_1 = 1$ an. Par exemple, si les marchés actions et immobiliers chutent fortement les deux premières années et remontent rapidement à un niveau supérieur au niveau initial l'année suivante, alors le BE prédit par les modèles existants devrait être élevé (en négligeant le vieillissement du portefeuille). Cependant, la baisse soutenue des marchés financiers les deux premières années a un effet important sur le BE puisque, via les rachats conjoncturels, la taille du portefeuille (et donc le BE) est très réduite. De même, nous pouvons imaginer des scénarios de mortalité plus élevée sur une période courte. Nous voyons bien dans cet exemple que le BE

(via les flux) est non seulement dépendant de l'état du monde à l'instant de valorisation (ou précédent l'instant de valorisation), mais aussi de tous les états par lesquels le monde est passé entre le début de projection et le moment de valorisation. Mathématiquement, nous pouvons dire que les modèles proxys prédisent $\mathbb{E}^{\mathbb{Q}}[\text{BE}_{t_2}|X_{t_2-t_1}]$ au lieu de prédire $\mathbb{E}^{\mathbb{Q}}[\text{BE}_{t_2}|\mathcal{F}_{t_2-t_1}]$ (ou $\mathbb{E}^{\mathbb{Q}}[\text{BE}_{t_2}|X_{t_2}]$ au lieu de $\mathbb{E}^{\mathbb{Q}}[\text{BE}_{t_2}|\mathcal{F}_{t_2}]$ si la valorisation du bilan s'effectue à horizon immédiat).

Le fait que le modèle ne prenne ni en compte le vieillissement du portefeuille entre t_1 et t_2 , ni la trajectoire économique du monde entre t_1 et t_2 rend son utilisation difficile pour une projection ORSA sur plusieurs années.

Solution : Réseaux de neurones récurrents

La solution que nous proposons s'inspire d'un papier de recherche publié par Moody's Analytics [4]. L'article propose de prédire les flux de trésorerie en lieu et place des BE ou autres indicateurs d'intérêts, grâce à des réseaux de neurones récurrents (RNN). Les modèles proxys présentés précédemment prédisent directement un BE, comme représenté figure 1.7. Nous proposons donc au contraire de prédire une série de flux de trésorerie pour chaque scénario économique donné en entrée, comme dans la figure 1.8 où $(S_t)_{t=1,\dots,T}$ et $(F_t)_{t=1,\dots,T}$ représentent respectivement le cours d'un marché action et les cash flows passifs correspondants, calculés par le réseau de neurones récurrent. Pour obtenir un BE grâce à notre réseau de neurones récurrent, il suffit de lui fournir en entrée K scénarios économiques pour obtenir K séries de flux. Nous pouvons alors calculer le BE comme la moyenne des sommes des flux actualisés sur ces K scénarios.

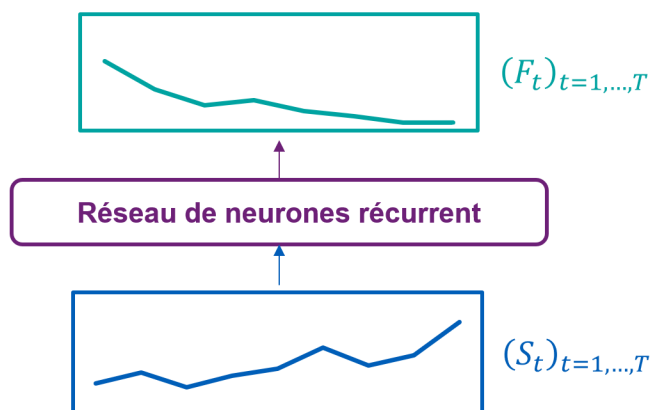


FIGURE 1.8 – Fonctionnement des modèles basés sur des réseaux de neurones récurrents

Prédire les flux grâce à un réseau de neurones récurrent présente plusieurs avantages. Les bénéfices principaux de la prédiction des flux à la place du BE sont d'abord la possibilité de calculer un BE à n'importe quel horizon, tout en prenant en compte le vieillissement de l'environnement et des portefeuilles actifs et passifs de l'entreprise. En effet, étant donné que les flux calculés prennent en compte le passage du temps, la somme actualisée à partir d'une certaine maturité des flux futurs prend bien en compte le passage du temps.

De plus, négliger la trajectoire passée empêche la prise en compte d'information impactant la taille du portefeuille et donc la valeur du BE. Les modèles existants utilisent uniquement l'information à l'instant initial ou l'instant précédent l'instant de valorisation pour prédire le BE. La structure très particulière du réseau de neurone récurrent permet de prendre en compte toute l'information acquise en t et avant t pour chaque prédiction de flux. Le sens de transmission de l'information vers la sortie du réseau mais aussi à sens unique entre les neurones d'une même couche (détaillé dans le chapitre dédié) permet la transmission de l'information

avec le temps. Cette transmission joue un rôle de stockage croissant de l'information. Cette croissance fait penser à la propriété de croissance des filtrations mathématiques $(\mathcal{F}_t)_{t=1,\dots,T}$ représentant toute l'information disponible à chaque instant.

Enfin, la génération de la base de données est beaucoup plus rapide. Les modèles proxys présentés précédemment requièrent le calcul de K trajectoires (voir KN dans le cas d'un BE de cloture) pour obtenir un résultat de BE. Ce résultat de BE ne permet d'ajouter qu'une ligne à la base d'entraînement. Dans le cas de l'architecture avec réseaux de neurones récurrents, une ligne de base de données correspond à une trajectoire. Le nombre de lignes obtenues pour notre base de données est donc K , voir KN fois plus important pour un même temps de calcul du modèle ALM.

Chapitre 2

Création de la base de données

La calibration de notre modèle proxy nécessite la création d'une base de données très large. Cette base de données doit couvrir l'ensemble des entrées possibles en termes de taux ou de trajectoires économiques. La qualité des prédictions dépendra en grande partie de la construction de la base de données. Pour créer cette base de données, nous décidons de jouer sur plusieurs paramètres : les courbes des taux, les valeurs de marché des actions et de l'immobilier, les volatilités des marchés, les taux de rachats, les taux de décès, les frais et l'inflation. Nous présentons dans ce chapitre les outils utilisés, le portefeuille et la procédure de création de la base de données.

2.1 Outils de simulation

2.1.1 Modèle ALM

Pour créer la base de données nécessaire à l'élaboration de notre modèle proxy, il est nécessaire d'utiliser un logiciel de projection des flux financiers, contenant le modèle ALM. Ce logiciel a pour objectif de simuler les interactions entre actifs et passifs au bilan de l'entreprise d'assurance, dans un environnement économique et social paramétrable. Pour l'élaboration de ce mémoire, nous utilisons le logiciel appelé DEvent. DEvent est une plateforme de modélisation, permettant l'implémentation complète de modèles actuariels et leur utilisation pour effectuer des projections. Ce logiciel reçoit trois groupes de fichiers en entrée :

- les model points, décrivant l'ensemble des actifs et passifs en portefeuille ;
- les hypothèses, décrivant à la fois les paramètres de l'environnement et les paramètres de fonctionnement et de gestion de l'entreprise dans le modèle, comme les frais, l'allocation cible du portefeuille d'actifs, la mortalité, les courbes de rachats conjoncturels et structurels ;
- les scénarios économiques, décrivant l'évolution des courbes des taux, des marchés actions et immobilier, des taux de défauts.

La figure 2.1 donne un exemple non exhaustif des fichiers utilisés par le logiciel.

2.1.2 Générateur de Scénarios Économiques

Les fichiers des scénarios économiques en entrée de DEvent sont créés à partir d'un Générateur de Scénarios Economiques (GSE). Le GSE permet de simuler différents scénarios économiques aléatoires pour les actions, l'immobilier, les taux, les spreads, les taux de défauts

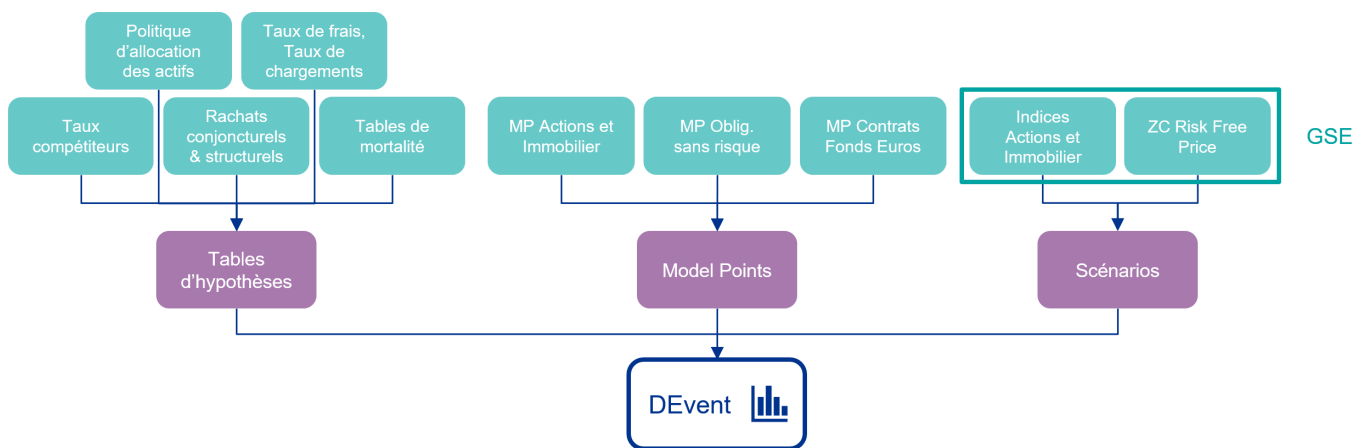


FIGURE 2.1 – Fichiers fournis en entrée du logiciel de projection

et l'inflation. C'est la diversité des scénarios économiques utilisés qui permet de déclencher les options asymétriques des contrats en portefeuille et donc la bonne évaluation des engagements vis-à-vis des assurés. Les fichiers à fournir en entrée du GSE sont représentés figure 2.2.

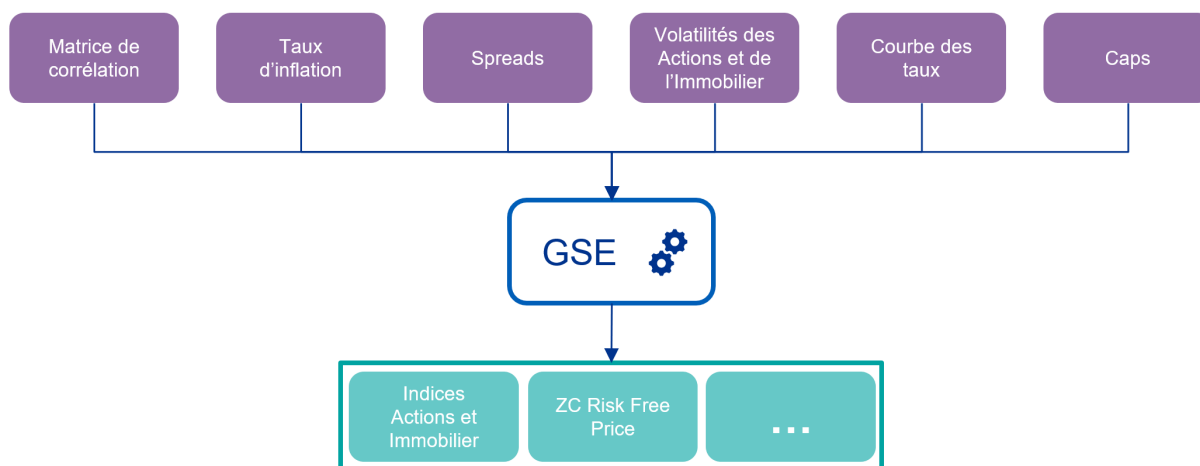


FIGURE 2.2 – Entrées et sorties du GSE

2.2 Actifs

Les actifs de l'entreprise que nous nous proposons d'étudier sont composés d'obligations sans risques, d'actions, d'immobilier et de liquidités. Les obligations sans risques représentent plus de 80% du portefeuille. Les proportions en valeurs comptables et valeurs de marché sont données dans le tableau 2.1.

Actif	Valeur Comptable (€)	Valeur de Marché (€)	%VM
Obligations sans risque	23 579 047	26 575 350	81%
Actions	3 050 026	3 437 607	10%
Immobilier	2 023 574	2 280 720	7%
Cash	674 525	674 525	2%
Total	29 327 173	32 968 203	100%

TABLE 2.1 – Actifs de l'entreprise

2.2.1 Obligations

Une obligation est une valeur mobilière représentant une créance, émise dans une devise donnée, et dont la maturité est fixée. Les maturités moyennes de ces créances se situent entre 1 et 50 ans. Les obligations possèdent plusieurs caractéristiques :

- le montant initialement prêté appelé nominal,
- les intérêts versés au titre de la créance et appelés coupons,
- leur méthode de calcul,
- leur fréquence de versement,
- les notations données par les agences de notation,

Nominal

Le nominal, noté N , peut faire l'objet de différents profils de remboursement (appelés également amortissements) :

- le profil *in fine* dans lequel le remboursement du nominal s'effectue en une fois à la maturité,
- le profil linéaire dans lequel le remboursement s'effectue par fractions égales du nominal, jusqu'à la maturité,
- le profil à annuités constantes dans lequel la somme des remboursements du nominal et des intérêts est constante au cours du temps.

La totalité des obligations en portefeuille sont à profil de remboursement *in fine*.

Coupons

On appelle coupons les intérêts versés au détenteur d'une obligation. Il existe différents types d'obligations, déterminant les méthodes de calcul des coupons :

- les obligations à taux fixes. Le taux et la méthode de calcul sont indiqués sur le contrat.
- les obligations à taux révisables. Le taux d'intérêt est indexé sur une référence qui varie dans le temps, comme par exemple le Libor 3 mois. L'emprunteur ne connaît donc pas à l'avance le montant des coupons qu'il devra verser.
- les obligations à taux variables : les taux sont calculés à partir d'une formule, dont le résultat est connu uniquement en fin de période.
- les obligations à taux structurés : ces taux dépendent de performances d'actions, de spreads de taux... Leur valorisation est complexe et fait appel à la théorie des options.

Les deux types d'obligations les plus communément rencontrés sont les obligations à taux fixes et à taux révisables, que l'on appelle à taux variables par abus de langage.

Les obligations à taux fixe disposent d'un taux d'intérêt facial, qui sert de base de calcul pour les coupons. Les coupons étant calculés sur le capital restant à rembourser, ils dépendent à la fois de la fréquence des versements et du profil d'amortissement de l'obligation. Dans le cadre des obligations à profil de remboursement *in fine*, la base de calcul est constante égale à N . Si les coupons sont versés annuellement, alors ils sont tous égaux à $C \times N$.

Dans le cas des obligations à taux révisables, le taux appliqué est celui constaté à un moment défini par le contrat. Si les dates de versements des coupons sont les $(t_i)_{i \in \{1, \dots, n\}}$, le taux appliqué pour le coupon versé en t_i est celui constaté en t_{i-1} . La durée $t_i - t_{i-1}$ doit être cohérente avec la durée sur laquelle porte le taux d'intérêt. Par exemple, si le taux d'intérêt est le Libor 3 mois, alors $t_i - t_{i-1} = 3$ mois.

Le portefeuille est uniquement composé d'obligations à taux fixe, dont la fréquence de versement est annuelle.

Notes

La notation d'une obligation, appelée en anglais *rating*, est un ensemble de lettres ou symboles représentant le risque de crédit d'un emprunteur. Les notations sont données par les agences de notations, dont les trois plus célèbres sont Moody's, Standard & Poor's et Fitch. Le logiciel de modélisation ALM utilise un système de notation simplifié utilisant 8 symboles classés dans l'ordre croissant des risques de défaut : AAA, AA, A, BBB, BB, B, CCC, NR. Ces notations sont proches de celles utilisées par les agences de notations. La répartition des notes des obligations sans risques détenues par l'entreprise, en fonction des valeurs de marché, est donnée figure 2.3.

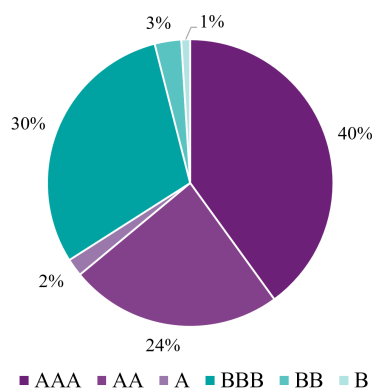


FIGURE 2.3 – Notes des obligations sans risque

2.2.2 Actions et Immobilier

Dans ce mémoire, il n'est question que d'actions cotées sur un marché réglementé. Une action est une part du capital d'une entreprise, qui donne un droit de propriété à son détenteur. Sa valeur fluctue au cours du temps, en fonction de différents facteurs plus ou moins subjectifs comme la confiance des investisseurs dans l'entreprise, contexte économique et financier, secteur d'activité. L'actionnariat donne le droit de toucher des dividendes, qui correspondent à une part des bénéfices de l'entreprise.

De même l'immobilier représente la possession de biens immobiliers, dont la valeur fluctue. Le détenteur de bien immobilier peut percevoir des loyers.

Le modèle ALM ne prend pas en compte les dividendes et les loyers. Les revenus proviennent uniquement des reventes des actions et actifs immobilier. Ceci revient à supposer que les loyers et dividendes sont directement réinvestis.

2.2.3 Cash

Le cash est le montant d'actif stocké dans la devise d'exercice de l'entreprise. Il s'agit ici de l'Euro.

2.3 Passifs

Les passifs associés aux contrats sont modélisés par des model points. Il s'agit d'agrégats de contrats ayant des caractéristiques similaires. L'agrégation des contrats en model points est obligatoire, sans quoi il serait impossible d'effectuer les projections en un temps raisonnable. La réglementation autorise l'agrégation des contrats, tant qu'elle n'altère pas la précision des résultats des projections. Les contrats sont agrégés en 100 model points. Le portefeuille de contrat est uniquement composé de contrats en euros. Les provisions mathématiques totales sont de 29.3 Mds €. Chaque model point dispose des caractéristiques suivantes :

- type de contrat et support,
- genre de l'assuré,
- année de naissance,
- année de souscription,
- réserve mathématique,
- TMG,
- taux de participation aux bénéfices contractuel,
- taux de chargement contractuel.

Type de contrat & support

L'outil DEvent permet de modéliser deux types de contrats : les contrats d'épargne et les contrats de rente.

Parmi les contrats d'épargne, DEvent permet de modéliser deux types de supports : euros et UC. Les contrats d'épargne peuvent être mono ou multi-supports. Les contrats multi-supports sont des contrats dont une partie est placée en unités de compte et l'autre partie sur un fond en Euros ; ou des contrats placés sur plusieurs supports UC différents. La hiérarchie des différents types de contrats et supports est représentée figure 2.4

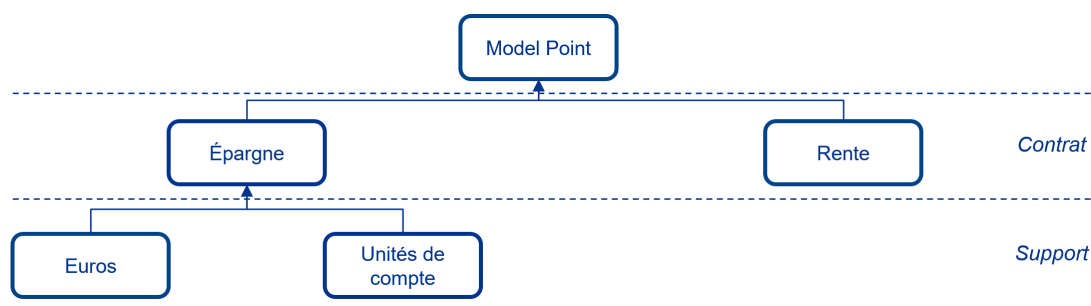


FIGURE 2.4 – Types de contrats dans l'outil de modélisation

Année de naissance & souscription

Les assurés en portefeuille sont nés entre 1936 et 1995. L'âge moyen des assurés est de 55 ans. L'ensemble des contrats en portefeuille ont été souscrits entre 1996 et 2020 et ont en moyenne 13 ans d'ancienneté. Les répartitions des années de naissance des assurés et de souscription des

contrats sont représentées dans les figures 2.5 et 2.6. Ces représentations tiennent compte des provisions mathématiques associées à chaque model point.

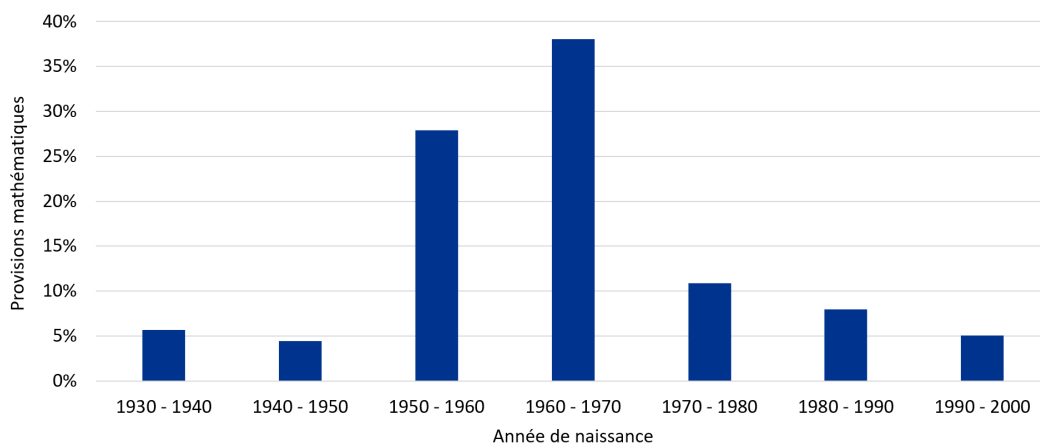


FIGURE 2.5 – Répartition des années de naissance des assurés

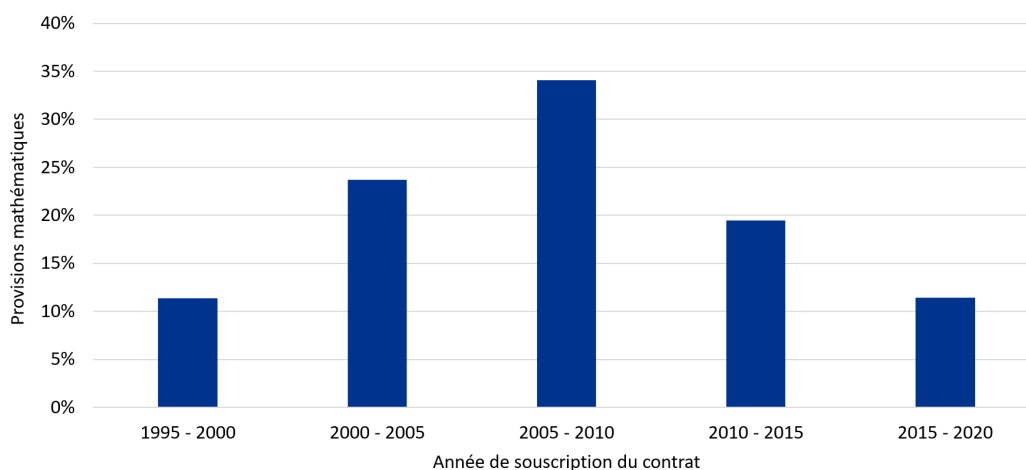


FIGURE 2.6 – Répartition des années de souscription des contrats

TMG

Les contrats souscrits peuvent avoir un TMG égal à 0, 0.5, 0.75 et 1%. La majorité des contrats souscrits après 2010 ont un TMG égal à 0%. Les contrats dont le TMG est égal à 1% ont été souscrits avant 2009. Seulement quelques contrats souscrits entre 2004 et 2009 ont des TMG égaux à 0.5 et 0.75%. La répartition des TMG en fonction des provisions mathématiques est donnée figure 2.7.

Participation aux bénéfices et chargement

Aucun contrat en portefeuille ne dispose d'une participation aux bénéfices contractuelle. Les seules participations aux bénéfices versées sont les participations aux bénéfices réglementaires et discrétionnaires. Les participations aux bénéfices discrétionnaires sont définies à partir de l'algorithme de participations aux bénéfices présenté dans la section 2.5.

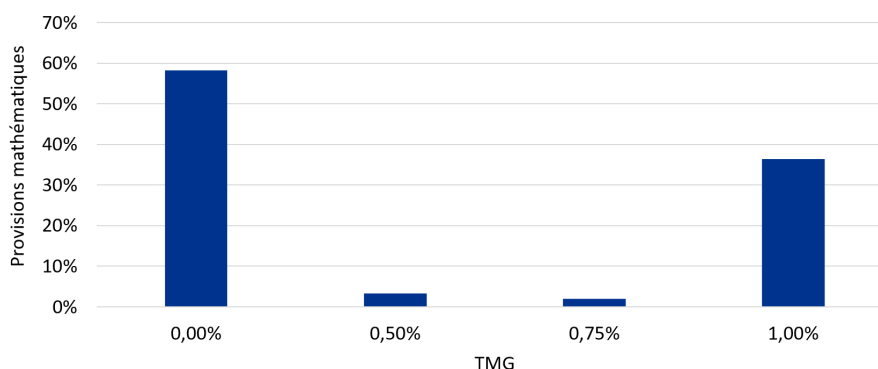


FIGURE 2.7 – Répartition des TMG

2.4 Configuration du GSE

2.4.1 Matrice de corrélation

La matrice de corrélation utilisée (représentée dans la table 2.2) est inspirée de la matrice utilisée pour Solvabilité 2 dans le cadre de l'agrégation des sous-modules du SCR de marché. Elle sera constante pour toutes les simulations. L'outil permet de simuler deux cours d'actions différents, le cours de l'immobilier et les taux d'intérêts.

	Equity 1	Equity 2	Immobilier	Taux courts	Taux longs
Equity 1	1	0.75	0.75	0.5	0
Equity 2	0.75	1	0.75	0.5	0
Immobilier	0.75	0.75	1	0.5	0
Taux courts	0.5	0.5	0.5	1	0
Taux longs	0	0	0	0	1

TABLE 2.2 – Matrice de corrélation fournie au GSE

Cette matrice diffère de la matrice d'agrégation de Solvabilité 2 par l'ajout d'une variable taux longs, et par la corrélation entre les cours Equity / Immobilier et les taux courts, constante égale à 0.5. Dans Solvabilité 2, elle dépend du type de choc (haussier ou baissier) sur les taux d'intérêts.

2.4.2 Taux d'inflation

Les taux d'inflation simulés sont définis à partir d'une courbe des taux spots d'inflation. Cette courbe est calibrée à partir des taux d'inflation pour les maturités à 1 mois, 3 mois, 6 mois, un an et toutes les années ultérieures, par pas de temps annuel. La courbe utilisée est donnée figure 2.8.

L'inflation dans le scénario $i \in 1, \dots, N$ est calculée comme :

$$I_i(t_n) = \frac{B^{infl}(0, t_n)}{\mathcal{D}_i(0, t_n)^\xi}$$

avec $B^{infl}(0, t)$ la courbe des taux spot d'inflation, $\mathcal{D}_i(0, t)$ le déflateur à l'instant t du scénario i , et ξ le coefficient d'élasticité égal à 0.8.

Si la courbe d'inflation est constante entre les différentes projections, nous disposons d'un paramètre permettant de la modifier. L'inflation est ajustée d'une constante ΔI permettant de choquer l'inflation : l'inflation réelle appliquée est alors $I_i(t_n) + \Delta I$. L'utilisation détaillée de ΔI est donnée dans la section 2.6.

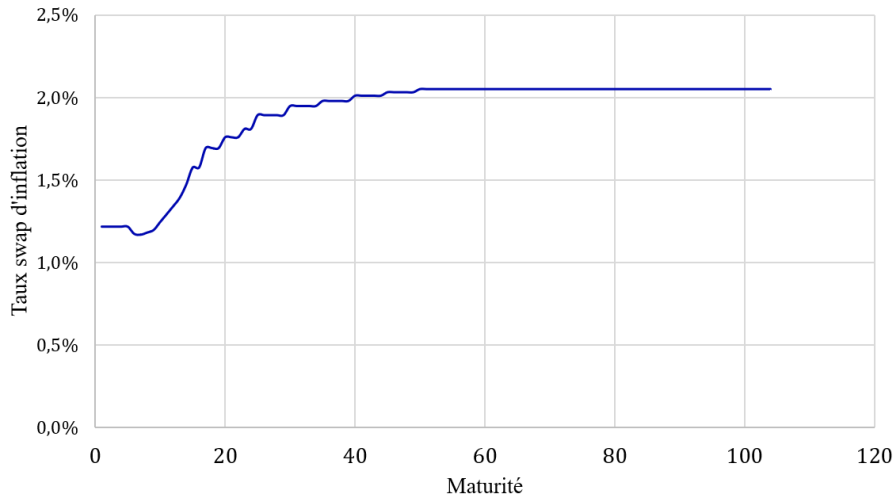


FIGURE 2.8 – Courbe des taux swaps d'inflation

2.4.3 Modèle de Black-Scholes

Le modèle utilisé pour la simulations des cours des actions et de l'immobilier est le modèle de Black-Scholes. Bien que ce modèle ait été historiquement introduit pour la valorisation des produits dérivés, il constitue une bonne solution pour la modélisation des prix des actifs. On note :

- S_t le prix de l'actif d'intérêt à l'instant t
- σ_t la volatilité à l'instant t
- $W_t^{\mathbb{Q}}$ un mouvement brownien sous la probabilité risque-neutre \mathbb{Q}

r_t peut être vu comme le taux d'intérêt que l'on peut espérer toucher à l'instant t et pour une maturité $t + dt$. Alors, le prix de l'actif est régi sous probabilité risque-neutre par l'équation :

$$dS_t = S_t(r_t dt + \sigma_t dW_t^{\mathbb{Q}})$$

Le prix de l'actif à l'instant t est alors :

$$S_t = S_0 \exp\left(\int_0^t (r_s - \frac{\sigma_s^2}{2}) ds + \int_0^t \sigma_s dW_s^{\mathbb{Q}}\right)$$

En cas de corrélations entre actifs, on utilise le cas plus général d'un mouvement brownien multidimensionnel.

2.4.4 Volatilités des actions et de l'immobilier

Pour diversifier les scénarios dans la base de données, les volatilités utilisées sont choisies aléatoirement. Elles sont supposées constantes dans le temps et sont tirées selon des lois uniformes :

$$\begin{aligned}\sigma^{Immo} &\sim \mathcal{U}[1\%; 21\%] \\ \sigma^{Eq1} &\sim \mathcal{U}[7\%; 42\%] \\ \sigma^{Eq2} &= \sigma^{Eq1}\end{aligned}$$

Il est important de signaler que l'hypothèse de constance de la volatilité n'est pas réaliste. Cependant, compte-tenu du nombre d'hypothèses simplificatrices utilisées dans ce mémoire, dans le GSE et dans le modèle ALM, l'apport d'un modèle de volatilité plus complexe n'a pas été jugé nécessaire.

2.4.5 Caps

Caplet

Un caplet est l'équivalent d'une option d'achat d'action appliqué à un taux : il permet de se couvrir contre une hausse de taux à échéance donnée. Notons $L(t, T, T + \theta)$ le taux forward proposé à l'instant t pour la date T de durée θ . Alors le caplet garantit à l'acheteur emprunteur l'application d'un taux maximum ne pouvant pas être dépassé, grâce à un pay-off $(L(T, T, T + \theta) - K)^+$. Le terme pay-off fait ici référence au taux appliqué par le contrat et non au montant perçu. Celui-ci s'obtient en calculant le montant des intérêts avec la convention de calcul en vigueur sur le taux sous-jacent, pour une durée θ et sur un notionnel Q . Par exemple, dans le cas d'une convention d'intérêts proportionnels, sur un notionnel Q , avec θ exprimé dans la bonne convention, le pay-off est :

$$(L(T, T, T + \theta) - K)^+ \times \theta \times Q$$

Caps

Un cap est une combinaison linéaire de caplets de maturités différentes. Ainsi, le pay-off d'un cap de maturité T_α , de strike K et de durée T_M est [12] :

$$\sum_{i=\alpha+1}^M (L(T_{i-1}, T_{i-1}, T_i) - K)^+ \times (T_i - T_{i-1}) \times Q$$

Les volatilités des taux d'intérêts sont calibrées à partir des prix des caps observés sur les marchés. En effet, de la même manière que les prix des options sur actions permettent de calculer des volatilités implicites des sous-jacents, les prix des caps permettent de calculer les volatilités implicites des taux. Les prix des caps à renseigner sont les prix des caps ATM (*At The Money*) sur le taux Libor 6 mois. Nous décidons de fixer les prix des caps et donc les volatilités des taux.

Maturité	Prix du Cap
3	0.00298
4	0.00598
5	0.01035
7	0.02354
10	0.05187
12	0.07366
15	0.10497
20	0.14674

TABLE 2.3 – Prix des Caps renseignés au GSE

2.4.6 Modèle de taux

Le modèle utilisé pour simuler l'évolution des taux est le modèle LMM+, calibré sur les prix des caps. Le modèle LMM+ étant dérivé du modèle LMM, nous présentons d'abord ce dernier.

Notons en plus des notations précédemment introduites :

- $L(t, T, T + \theta)$ le taux forward proposé à l'instant t pour la date T de durée θ

- $B(t, T) = \mathbb{E} \left[e^{\int_t^T -r_s ds} \middle| \mathcal{F}_t \right]$ la valeur du zéro-coupon de maturité T vu à la date t
- $\mathbb{Q}^{T+\theta}$ la probabilité associée au numéraire $(B(t, T + \theta))_{t \geq 0}$

Pour les taux courts comme le taux Libor, la convention de taux proportionnels s'impose, par opposition aux taux longs où les taux actuariels sont utilisés. Nous avons donc :

$$\frac{1}{B(t, T)}(1 + \theta L(t, T, T + \theta)) = \frac{1}{B(t, T + \theta)}$$

D'où :

$$L(t, T, T + \theta) = \frac{1}{\theta} \left(\frac{B(t, T)}{B(t, T + \theta)} - 1 \right)$$

LMM

Le modèle LMM est un modèle basé sur les taux forwards. Il suppose que sous chaque mesure forward-neutre $\mathbb{Q}^{T_{i+1}}$, le taux forward associé a une dynamique log-normale. En reprenant les différentes notations des caplets, on a :

$$\frac{dL(t, T_i, T_{i+1})}{L(t, T_i, T_{i+1})} = \sigma_i(t) dW_{i,t}^{\mathbb{Q}^{T_{i+1}}} \quad (2.1)$$

Notons $V_{caplet_i}(t)$ la valeur du caplet i dont le pay-off en T_{i+1} est $(L(T_i, T_i, T_{i+1}) - K)^+(T_{i+1} - T_i)$. Alors, en l'absence d'opportunité d'arbitrage, la valeur de ce caplet est :

$$V_{caplet_i}(t) = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{T_{i+1}} r_s ds} (L(T_i, T_i, T_{i+1}) - K)^+(T_{i+1} - T_i) \middle| \mathcal{F}_t \right]$$

En calculant explicitement la valeur de l'espérance, on effectue un passage sous le numéraire $(B(t, T_{i+1}))_t$. C'est lors de ce passage qu'apparaît naturellement la nouvelle probabilité forward-neutre $\mathbb{Q}^{T_{i+1}}$ sous laquelle on suppose 2.1. En utilisant les mêmes formules que pour le calcul du prix des options d'achats européennes sur actions, et sachant que $\mathbb{E}^{\mathbb{Q}^{T_{i+1}}} [L(T_i, T_i, T_{i+1}) | \mathcal{F}_t] = L(t, T_i, T_{i+1})$ on obtient :

$$V_{caplet_i}(t) = B(t, T)(T_{i+1} - T_i)(L(t, T_i, T_{i+1})\mathcal{N}(d_1) - K\mathcal{N}(d_2)), \quad \forall t \leq T_i$$

Avec :

$$d_1 = \frac{\ln\left(\frac{L(t, T_i, T_{i+1})}{K}\right) + \frac{1}{2}\gamma_i^2(T_i - t)}{\gamma_i\sqrt{T_i - t}}$$

$$d_2 = d_1 - \gamma_i\sqrt{T_i - t}$$

$$\gamma_i = \sqrt{\int_t^{T_i} \frac{1}{T_i - t} \sigma_i(u)^2 du}$$

Pour calibrer les valeurs des γ_i qui sont inconnues, on cherche à minimiser :

$$\sum_i |V_{caplet_i}^{BS} - V_{caplet_i}^{VM}|$$

Avec $V_{caplet_i}^{BS}$ la valeur théorique d'un caplet calculée à partir de la formule précédente et $V_{caplet_i}^{VM}$ la valeur de marché du caplet.

Enfin, pour simuler les taux, qui sont corrélés entre eux, on réécrit l'équation 2.1 par rapport au brownien $W_{i,t}^{\mathbb{Q}^d}$, le i^e composant du brownien multidimensionnel $W_t^{\mathbb{Q}^d}$. Introduisons

$m(t) = \inf\{j : T_j \geq t\}$. Alors ce brownien est celui associé à la mesure de probabilité issue du numéraire :

$$B_d(t) = \frac{B(t, T_{m(t)})}{\prod_{j=0}^{m(t)-1} B(T_j, T_{j+1})}$$

Après réécriture on a :

$$\frac{dL(t, T_i, T_{i+1})}{L(t, T_i, T_{i+1})} = \sigma_i(t) \sum_{j=m(t)}^N \frac{(T_{i+1} - T_i)\rho_{i,j}\sigma_j(t)L(t, T_j, T_{j+1})}{1 + (T_{j+1} - T_j)L(t, T_j, T_{j+1})} dt + \sigma_i(t)dW_{i,t}^{\mathbb{Q}^d}$$

où : $\rho_{i,j} = d \langle W_{t,i}^{\mathbb{Q}^d}, W_{t,j}^{\mathbb{Q}^d} \rangle$

Ce modèle présente le défaut de ne pas pouvoir reproduire les courbes des taux initiales en cas de taux négatifs. C'est la principale raison pour laquelle le modèle LMM+ a été introduit.

LMM+

Pour permettre la prise en compte des taux négatifs, le modèle LMM+ introduit des coefficients de *shift* b_i et la volatilité du taux *shifté* $\sigma_i^b(t)$.

Les dynamiques des taux sous leurs probabilités forward-neutre respectives s'écrivent :

$$dL(t, T_i, T_{i+1}) = \sigma_i^b(t)(L(t, T_i, T_{i+1}) + b_i)dW_{i,t}^{\mathbb{Q}^{T_{i+1}}}$$

LMM+ suppose donc que les taux forwards suivent la dynamique :

$$\frac{dL(t, T_i, T_{i+1})}{L(t, T_i, T_{i+1}) + b_i} = \sigma_i^b(t) \sum_{j=m(t)}^i \frac{(T_{i+1} - T_i)\rho_{i,j}\sigma_j^b(t)(L(t, T_j, T_{j+1}) + b_j)}{1 + (T_{j+1} - T_j)(L(t, T_j, T_{j+1}) + b_j)} dt + \sigma_i^b(t)dW_{i,t}^d$$

En suivant le même raisonnement que précédemment, on obtient :

$$V_{caplet_i}^{LMM+}(t) = B(t, T)(T_{i+1} - T_i)((L(t, T_i, T_{i+1}) + b_i)\mathcal{N}(d_1) - (K + b_i)\mathcal{N}(d_2)), \quad \forall t \leq T_i$$

Avec :

$$d_1 = \frac{\ln\left(\frac{L(t, T_i, T_{i+1}) + b_i}{K}\right) + \frac{1}{2}\gamma_i^2(T_i - t)}{\gamma_i\sqrt{T_i - t}}$$

$$d_2 = d_1 - \gamma_i\sqrt{T_i - t}$$

$$\gamma_i = \sqrt{\int_t^{T_i} \frac{1}{T_i - t}\sigma_i(u)^2 du}$$

2.4.7 Courbes de taux

Outil

Le GSE permet de créer une multitude de trajectoires de taux, grâce aux prix des caps et d'une courbe de taux initiale. Pour diversifier au maximum les trajectoires de taux de la base d'entraînement, il est donc nécessaire de disposer de plusieurs courbes de taux initiales. Deux méthodes permettent de remplir la base de données de courbes de taux : générer des courbes synthétiques ou sélectionner des courbes existantes ou ayant existé. Dans ce mémoire, les courbes des taux sont créées de manière synthétique. Ces courbes doivent cependant avoir des formes usuelles, être suffisamment régulières et avoir des valeurs cohérentes. L'outil d'interpolation avec la méthode de Smith-Wilson, fourni par l'EIOPA, permet de répondre à ces critères. L'outil permet de créer une courbe de taux en fournissant un ensemble de points de la courbe et un taux "limite". Les entrées de l'outil sont :

- r_1 , le taux spot à 1 an,
- r_{30} , le taux spot à 30 ans,
- UFR, l'*Ultimate Forward Rate* vers lequel la courbe de taux doit converger.

Les courbes de taux générées seront utilisées pour créer les scénarios économiques fournis en entrée du modèle proxy, grâce au GSE.

UFR

Les compagnies d'assurance couvrent des risques dont la durée moyenne dépasse parfois très largement le dernier point de liquidité (LLP) des instruments de taux disponibles sur les marchés. Ne pouvant directement observer les taux en vigueur pour ces maturités tout en respectant le critère DLT (*Deep Liquid Transparent*), il devient nécessaire d'extrapoler la courbe des taux. Pour ce faire, l'EIOPA a introduit l'*Ultimate Forward Rate*. Il s'agit du taux vers lequel converge la courbe des taux pour les grandes maturités. Ce taux est calculé par l'EIOPA comme la somme des taux d'intérêts réels attendus et du taux d'inflation attendu, qui ne doivent pas dépendre du contexte économique et financier à court et moyen terme. Cependant, l'UFR ne peut varier de plus de 15 points de base par an. Cette mesure permet de limiter l'impact du changement d'UFR sur les provisions des assureurs. Ainsi, l'UFR applicable dépend des taux d'intérêts et de l'inflation, mais aussi de la valeur de l'UFR l'année précédente. Pour 2021, l'UFR est égal à 3,6%, alors qu'il était égal à 3,75% en 2020. Les UFR appliqués lors des années précédentes sont donnés dans la table 2.4.

Année	UFR
2018	4,05%
2019	3.90%
2020	3.75%
2021	3.60%

TABLE 2.4 – Historique des valeurs de l'UFR

Génération des courbes de taux

Les courbes de taux générées doivent être suffisamment diverses pour permettre une utilisation prospective du modèle final. Cependant, elles doivent aussi être suffisamment vraisemblables. En effet, un entraînement sur un ensemble de courbes de taux plus large se fait au détriment de la précision sur le scénario central. Les taux spots et UFR utilisés doivent donc être à la fois suffisamment éloignés des taux en vigueur pour permettre une bonne adaptabilité du modèle, et suffisamment proches afin de garantir une grande précision des résultats.

Les taux r_1 et r_{30} sont donc générés à partir de déformations de la courbe de taux actuelle. Les taux spots en vigueur au 31 août 2021 sont $r_1 = -0.615\%$ et $r_{30} = 0.905\%$. Les chocs appliqués pour le calcul du SCR associé au module de taux amènent les utilisateurs à modifier la courbe utilisée en entrée du modèle ALM. Les valeurs des taux choqués à la hausse et à la

Maturité	r^{down}	r	r^{up}
1	-0.615%	-0.615%	0.385%
30	0.654%	0.905%	1.905%

TABLE 2.5 – Taux utilisés pour le calcul du SCR du module de taux

baisse sont données dans la table 2.5. On remarque que $r_1^{down} = r_1$. En effet, l'article 167-2 du

règlement 2015/35 précise que les chocs à la baisse sont nuls en cas de taux d'intérêts négatifs. Pour permettre l'utilisation de l'outil pour les calculs de SCR, les taux sont choisis selon les lois uniformes :

- $r_1 = \mathcal{U}[-1\%, 0.5\%]$
- $r_{30} = \mathcal{U}[0.5\%, 2\%]$

La borne inférieure de la loi utilisée pour r_1 est plus basse que la valeur utilisée pour le choc baissier. Cette borne est abaissée afin de permettre l'utilisation du modèle dans le cadre d'études prospectives, comme un scénario de baisse des taux.

L'UFR applicable pour 2021 est quant à lui égal à 3.6%. Le fait que les variations soient limitées à 15 points de base par an incite à sélectionner des UFR proches de 3.6%. Les UFR fournis pour la création des courbes sont donc générés selon la loi :

- $\text{UFR} = \mathcal{U}[3\%, 4.2\%]$

La courbe des taux actuelle et les déformations des différents points de la courbe sont représentées figure 2.9.

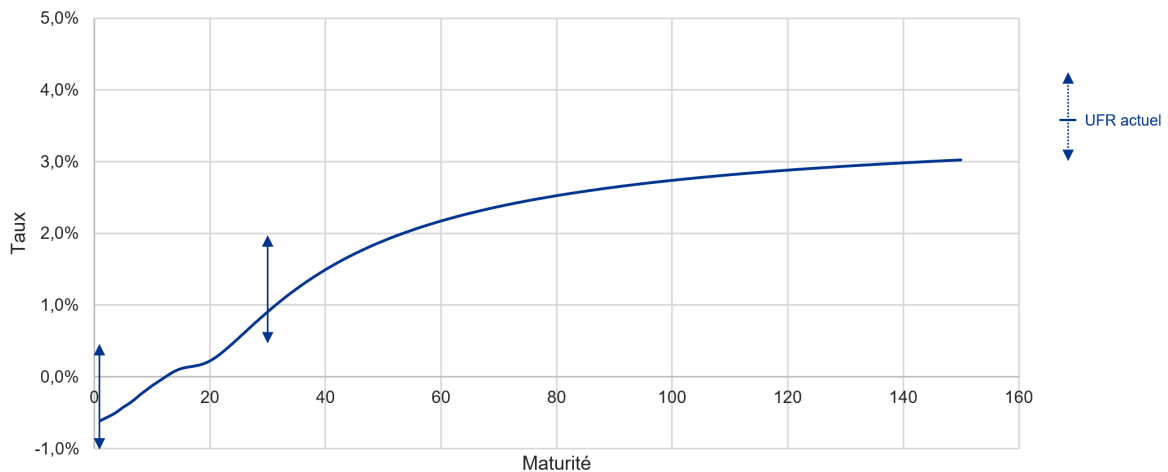


FIGURE 2.9 – Courbe de taux sans risque au 31/08/2021 et plages des déformations

Un ensemble de 20 courbes est généré, chacune associée à des volatilités pour les actions et l'immobilier générées conformément à 2.4.4.

2.5 Modèle ALM

2.5.1 Risque-neutralisation des obligations

La projection étant effectuée en univers risque-neutre, il est nécessaire d'ajuster les prix des obligations à $t = 0$. En effet, les prix des obligations observés sur les marchés prennent en compte l'aversion au risque des investisseurs. Ces risques peuvent être de divers types : risque de défaut, de liquidité, etc. Ainsi, si on calcule le prix théorique de l'obligation de coupons annuels C_t , de nominal N et de maturité T en actualisant avec la courbe des taux sans risques on a :

$$VT = \sum_{t=1}^T \frac{N \times C_t}{(1 + r_t)^t} + \frac{N}{(1 + r_T)^T}$$

La valeur est différente de la valeur de marché observée (en général $VT \geq VM$ par aversion au risque) et renseignée dans les Model Points. Pour être "market-consistent", on doit modifier les propriétés de ces obligations pour obtenir $VT = VM$. Pour ce faire, on peut jouer sur deux facteurs : les taux de coupons ou le nominal. Notre algorithme ALM risque-neutralise les obligations en modifiant le nominal. Nous modifions le nominal en prenant $N' = \alpha N$. Dès lors, la valeur théorique de l'obligation est :

$$VT' = \sum_{t=1}^T \frac{\alpha N \times C_t}{(1+r_t)^t} + \frac{\alpha N}{(1+r_T)^T} = \alpha VT$$

Il suffit de prendre $\alpha = \frac{VM}{VT}$ pour obtenir $VT' = VM$. D'un point de vue pratique, la multiplication du nominal par α permet de diminuer tous les cash flows de l'obligation de manière proportionnelle, comme représenté figure 2.10

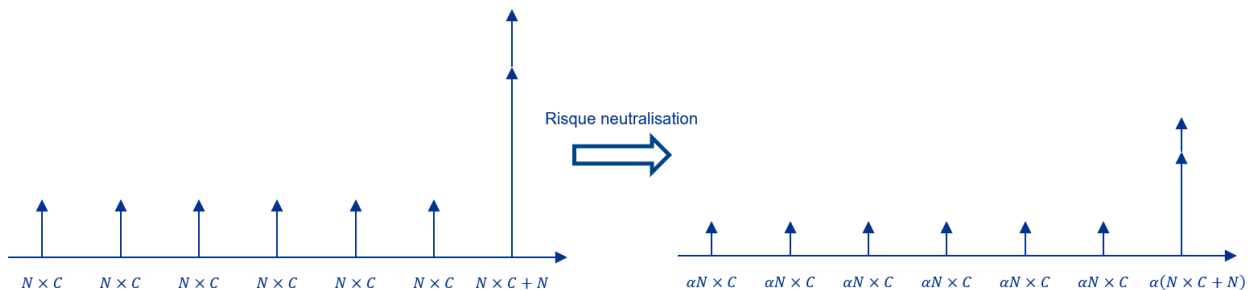


FIGURE 2.10 – Risque neutralisation des obligations sur le nominal

2.5.2 Allocation d'actifs

La stratégie d'allocation d'actifs définit des règles de gestion du portefeuille d'actifs. Notre modèle fixe les proportions d'actifs en valeur de marché comme égales aux proportions initiales. En cas de mouvement de marché modifiant la composition du portefeuille, l'entreprise achète ou vend les actifs dont elle dispose pour retrouver la composition d'origine. La composition est donc fixée comme dans le tableau 2.6.

Actif	Part
Obligations sans risque	81%
Actions	10%
Immobilier	7%
Cash	2%

TABLE 2.6 – Allocation des actifs

2.5.3 Algorithme de participation aux bénéfices

Les participations aux bénéfices versées sont calculées à partir d'un taux cible, des provisions pour participation aux bénéfices et d'un taux concurrent.

Taux concurrent

Le taux concurrent est défini comme le taux moyen servi par les concurrents à leurs assurés. Ce taux est calculé à partir des rendements des marchés financiers sur l'année, d'un portefeuille type et des taux de chargement et de partage de la concurrence. Le taux de partage correspond à la part moyenne de résultats, principalement financiers, que les concurrents reversent à leurs assurés. Ce taux définit donc le taux servi moyen par la concurrence à leurs assurés comme une fonction affine des rendements du portefeuille concurrent :

$$\text{Taux concurrent}_n = \max(\text{PS} \times \text{Rendement portefeuille}_n - \text{Taux chargement}, 0)$$

avec PS le taux de partage qui est fixé à 70%. Le taux de chargement concurrent est fixé à 0%. La composition du portefeuille d'actifs des concurrents est présentée dans le tableau 2.7. La durée moyenne des obligations du portefeuille concurrent est de 6 ans.

Actif	Part
Obligations	70%
Actions	15%
Immobilier	15%

TABLE 2.7 – Composition du portefeuille concurrent

Taux versé

Pour chaque année n , un taux cible est calculé comme :

$$\text{Taux cible}_n = \text{Taux concurrent}_n + 0.2\%$$

La participation aux bénéfices cible est alors :

$$\text{PB cible}_n = \sum_{i=1}^N \max(\text{Taux cible}_n - TMG^i, 0) * PM_n^i$$

Avec PM_n^i la provision mathématique du contrat i pendant la n^{ieme} année de projection. Cette participation aux bénéfices correspond à la somme totale idéalement reversée aux assurés, si les bénéfices le permettent. Le montant de participation aux bénéfices réglementaire à verser aux assurés, noté PB_n^{\min} , est calculé conformément à la section 1.1.2. Les provisions pour participations aux bénéfices sont notées PPB_n . La participation aux bénéfices versée est alors calculée grâce à la disjonction suivante :

- Si $PB_n^{\min} + PPB_n > \text{PB cible}_n$:
 - $PB_n = \text{PB cible}_n$
- Sinon :
 - $PB_n = PB_n^{\min} + PPB_n$

Le taux de participation aux bénéfices versé est donc :

$$\text{Taux PB versé}_n = \frac{PB_n}{\sum_{i=1}^N PM_n^i}$$

Et le taux versé pour un contrat i est :

$$\text{Taux versé}_n^i = \text{Taux PB versé}_n + TMG^i$$

Enfin, le taux moyen versé est :

$$\text{Taux moyen versé}_n = \frac{\sum_{i=1}^N \text{Taux versé}_n^i \times PM_n^i}{\sum_{i=1}^N PM_n^i}$$

2.5.4 Rachats

On distingue deux formes de rachats : les rachats structurels et les rachats conjoncturels. Si les définitions de ces deux types de rachats sont données dans les sous-sections suivantes, nous pouvons d'ores et déjà dire que :

$$\text{Rachats} = \text{Rachats structurels} + \text{Rachats conjoncturels}$$

Les méthodes de modélisation utilisées sont conformes aux Orientations Nationales Complémentaires (ONC) publiées par l'ACPR pour la préparation de l'exercice 2013 sous le régime Solvabilité II [8].

Rachats structurels

Les rachats structurels sont les rachats qui interviennent chaque année, indépendamment du contexte économique. Ils sont modélisés comme une fonction de l'ancienneté du contrat, même s'ils peuvent parfois dépendre d'autres paramètres. Conformément aux recommandations contenues dans l'ONC, la courbe utilisée (figure 2.11) est une courbe d'expérience.

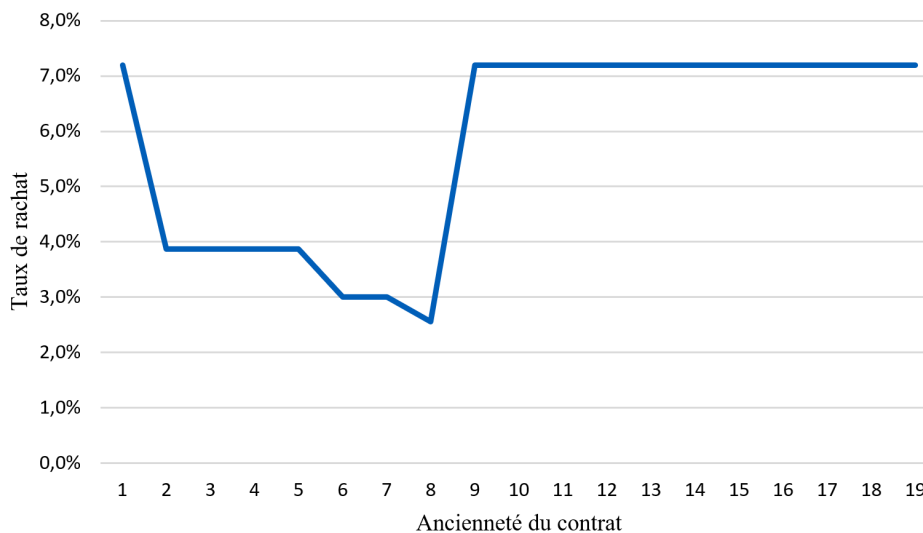


FIGURE 2.11 – Courbe des rachats structurels

La fiscalité avantageuse des assurances vie d'ancienneté supérieure à 8 ans se traduit par un pic de rachats à partir de cette ancienneté, puis par un plateau pour les anciennetés supérieures.

Rachats conjoncturels

Les rachats conjoncturels correspondent aux rachats liés à la conjoncture économique. Ils sont généralement modélisés par une fonction dépendant de la différence entre le taux servi aux assurés et un taux attendu par ces mêmes assurés.

L'ONC introduit une fonction générique de rachats dépendant de 6 paramètres. Notons σ_n l'écart entre le taux servi lors de l'année $n - 1$ et le taux attendu. Alors les rachats sont calculés comme suit :

$$Rachat_n = \begin{cases} RC_{max} & \sigma_n < \alpha \\ RC_{max} \frac{\sigma_n - \beta}{\alpha - \beta} & \alpha < \sigma_n < \beta \\ 0 & \beta < \sigma_n < \gamma \\ RC_{min} \frac{\sigma_n - \gamma}{\delta - \gamma} & \gamma < \sigma_n < \delta \\ RC_{min} & \delta < \sigma_n \end{cases}$$

L'ONC impose aux entreprises d'utiliser une courbe de rachats dont le niveau se situe entre les courbes "plancher" et "plafond". Ces courbes correspondent aux courbes obtenues à partir de la fonction de rachat précédente avec les paramètres donnés dans le tableau 2.8.

Paramètre	Plancher	Plafond	Choix
α	-6%	-4%	-5%
β	-2%	0%	-1%
γ	1%	1%	1%
δ	2%	4%	3%
RC_{min}	20%	40%	30%
RC_{max}	-6%	-4%	-5%

TABLE 2.8 – Planchers et plafonds des paramètres sous QIS5

Les paramètres utilisés sont ceux égaux aux moyennes des paramètres planchers et plafonds (confère la colonne "Choix" de la table 2.8). Cette courbe se situe donc bien entre les courbes "plancher" et "plafond", comme représenté figure 2.12.

L'écart entre taux servi et taux attendu σ_n est calculé comme :

$$\sigma_n = \max(\text{TMG moyen}, \text{Taux moyen versé}_{n-1}) - \text{Taux concurrent}_{n-1}$$

Avec le taux concurrent défini dans la section 2.5.3.

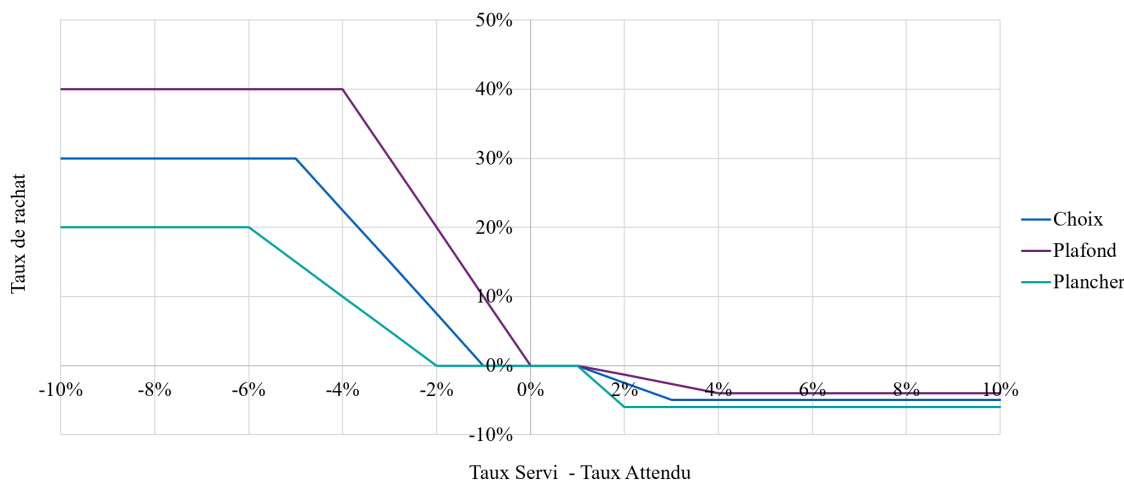


FIGURE 2.12 – Courbe des rachats conjoncturels

2.5.5 Frais et chargements

L'outil ALM permet la prise en compte de frais et chargements.

Frais

Les frais sont exprimés en pourcentage de provisions mathématiques. Ils peuvent être de 5 sortes :

- frais liés aux décès,
- frais de rachats,
- frais administratifs,
- frais d'acquisition,
- frais de gestion des investissements.

Notre portefeuille fonctionnant en run-off, les frais d'acquisition sont fixés à 0%. Les frais de gestion des investissements sont aussi à 0%. On suppose que les frais de rachats et de décès sont les mêmes et valent 0,6%. Enfin, les frais administratifs sont fixés à 0,3%.

Chargements

Les chargements peuvent être de différentes sortes : chargements sur rachats partiels, rachats totaux, arbitrages, provisions mathématiques, primes versées. N'ayant pas de primes versées, et ne modélisant ni les rachats partiels, ni les arbitrages, les taux de chargements associés sont fixés à 0%. Enfin, le taux de chargement sur provisions mathématiques est de 0,3%.

2.6 Chocs sur les facteurs de risque

Pour chaque projection, le modèle ALM se voit fournir un ensemble de paramètres permettant de définir un choc sur la projection. Un choc est une modification de certains paramètres de projection, de manière temporaire ou durable. Il est possible de choquer les valeurs de marché des actifs, la mortalité, les frais, les rachats, etc. La méthode d'application des chocs dépend du type de choc. Par exemple, les chocs sur la mortalité sont proportionnels. Si en l'absence de choc, le taux de mortalité appliqué est q_x , alors pour un choc de mortalité à la hausse ΔDeath , le taux appliqué est $q_x \times (1 + \Delta\text{Death})$. Les chocs implémentés sont directement inspirés de ceux utilisés pour le calcul des SCR dans le cadre de Solvabilité II. L'objectif de l'étude est de rendre le modèle utilisable à la fois pour les calculs de SCR et pour des études prospectives. 7 types de chocs sont disponibles :

- $\Delta\text{LapseY1}$, le choc sur les rachats lors de la première année de projection,
- ΔLapse , le choc sur les rachats pendant toute la durée de projection,
- ΔDeath , le choc sur les taux de mortalité pendant toute la projection,
- ΔInfl , le choc sur l'inflation pendant toute la projection,
- ΔExp , le choc sur les dépenses initiales,
- ΔEq , le choc sur les valeurs de marché action,
- ΔProp , le choc sur les valeurs de marché de l'immobilier,

2.6.1 Loi

Le choix des combinaisons de chocs appliquées est un problème important et dépend des capacités de calcul disponibles. Les combinaisons doivent à la fois permettre de couvrir l'ensemble des cas d'utilisation tout en limitant au maximum le nombre de combinaisons. Par exemple, effectuer toutes les combinaisons possibles sur une grille de 3 valeurs pour chaque choc reviendrait à effectuer $3^7 = 2187$ chocs. Ces chocs devant eux-mêmes être testés sur une multitude de scénarios de taux et de volatilités d'actions différents, le temps de calcul serait trop important.

La grille définissant les chocs appliqués est de dimension 7 chocs \times 200. Le nombre 200 est choisi de manière arbitraire : il s'agit du nombre maximum de chocs applicables, permettant de conserver un temps de calcul raisonnable compte-tenu des capacités de calcul disponibles au moment de l'étude et du nombre de scénarios différents souhaité. Avec 20 scénarios, l'exécution du calcul a duré 24 heures.

Chacun des 200 chocs appliqués sont tirés aléatoirement. Il est supposé que les 7 types de chocs suivent des lois normales centrées :

$$\forall i \in \{\text{LapseY1}, \dots, \text{Prop}\}, \quad \Delta_i \sim \mathcal{N}(0, \sigma_i^2)$$

Les écart-types σ_i sont choisis à partir des quantiles à 99,5% suggérés par Solvabilité 2. Plus précisément, la volatilité σ_i est choisie de telle sorte que le quantile à 99,5% de la loi normale $\mathcal{N}(0, \sigma_i^2)$ corresponde au quantile $q_i^{99,5\%}$ donné par Solvabilité 2 (39% pour les actions par exemple). Ainsi on cherche σ_i tel que :

$$\mathbb{P}(\Delta_i \leq q_i^{99,5\%}) = 99,5\%$$

Sachant que $\sigma\mathcal{N}(0, 1) \sim \mathcal{N}(0, \sigma^2)$, et en notant N une variable aléatoire de loi normale centrée réduite :

$$\mathbb{P}(\Delta_i \leq q_i^{99,5\%}) = \mathbb{P}(N \leq \frac{q_i^{99,5\%}}{\sigma_i})$$

En notant Φ la fonction de répartition de la loi normale centrée réduite, on a donc :

$$\Phi^{-1}(99,5\%) = \frac{q_i^{99,5\%}}{\sigma_i} \iff \sigma_i = \frac{q_i^{99,5\%}}{\Phi^{-1}(99,5\%)}$$

En notant $\Phi^{-1}(99,5\%) = q_{\text{centrée réduite}}^{99,5\%}$, on retient finalement :

$$\forall i, \sigma_i = \frac{q_i^{99,5\%}}{q_{\text{centrée réduite}}^{99,5\%}}$$

Les quantiles utilisés et les écart-types associés sont donnés dans la table 2.9. Le choc Δ_{Death} permet à la fois de simuler les risques de longévité lorsqu'il est positif et de mortalité lorsqu'il est négatif. Pour permettre le calcul du SCR dans les cas de longévité et mortalité, il est défini comme :

$$q_{\Delta_{\text{Death}}}^{99,5\%} = \max(|q_{\text{Longevite}}^{99,5\%}|, |q_{\text{Mortalite}}^{99,5\%}|) = \max(15\%, 20\%) = 20\%$$

Les valeurs utilisées pour les chocs sur l'inflation et les frais sont elles issues des valeurs utilisées pour le calcul du SCR des dépenses en vie.

Les lois normales utilisées sont corrélées entre elle. Les corrélations utilisées sont données dans la table 2.10. Ces corrélations sont issues des matrices de corrélation indiquées par la formule standard pour les calculs de SCR.

i	$\Delta\text{LapseY1}$	ΔLapse	ΔDeath	ΔInfl	ΔExp	ΔEq	ΔProp
q_i	40%	50%	20%	1%	10%	39%	25%
σ_i	15,5%	19,4%	7,8%	0,4%	3,9%	15,1%	9,7%

TABLE 2.9 – Valeurs des quantiles à 99,5% et écart-types associés

	$\Delta\text{LapseY1}$	ΔLapse	ΔDeath	ΔInfl	ΔExp	ΔEq	ΔProp
$\Delta\text{LapseY1}$	100%	50%	0%	0%	0%	25%	25%
ΔLapse	50%	100%	0%	50%	50%	25%	25%
ΔDeath	0%	0%	100%	25%	25%	25%	25%
ΔInfl	0%	50%	25%	100%	75%	0%	25%
ΔExp	0%	50%	25%	75%	100%	0%	25%
ΔEq	25%	25%	25%	0%	0%	100%	75%
ΔProp	25%	25%	25%	25%	25%	75%	100%

TABLE 2.10 – Matrice de corrélation des chocs

Sachant que les risques de longévité et de mortalité sont simulés via la même variable ΔDeath , il est impossible d'appliquer des corrélations différentes à ces deux risques. Les corrélations entre chocs sur les rachats $\Delta\text{LapseY1}$ et ΔLapse , et le choc sur la mortalité ΔDeath sont donc fixées à 0%.

2.6.2 Simulations

La simulation du vecteur gaussien $(\Delta\text{LapseY1}, \dots, \Delta\text{Prop})$ requiert l'utilisation de certaines propriétés des vecteurs gaussiens.

Diagonalisation et loi des vecteurs gaussiens

Soit Σ une matrice de covariance. Σ est symétrique, définie positive à coefficients réels. Elle peut donc s'écrire sous la forme :

$$\Sigma = USU^T \quad (2.2)$$

avec U orthogonale, U^T la transposée de U et $S = \text{diag}(s_1, \dots, s_n)$ diagonale à coefficients strictement positifs. En posant $S^{1/2} = \text{diag}(\sqrt{s_1}, \dots, \sqrt{s_n})$ et $A = US^{1/2}$, Σ se réécrit comme :

$$\Sigma = AA^T \quad (2.3)$$

Dès lors, si X est un vecteur gaussien centré de matrice de covariance Z :

$$AX \sim \mathcal{N}(0, AZA^T)$$

Dans le cas particulier où $Z = I_n$ la matrice identité de dimension n :

$$AX \sim \mathcal{N}(0, \Sigma) \quad (2.4)$$

Méthode

Pour générer les vecteurs gaussiens de volatilités et corrélations indiquées dans les tables 2.9 et 2.10, il est nécessaire d'écrire la matrice de covariance finale sous la même forme que dans l'équation 2.3.

Notons C la matrice de corrélation 2.10. Elle s'écrit sous la forme :

$$C = \begin{pmatrix} 1 & \rho_{1,2} & \dots & \rho_{1,n-1} & \rho_{1,n} \\ \rho_{1,2} & \ddots & \ddots & \ddots & \rho_{2,n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \rho_{1,n-1} & \ddots & \ddots & \ddots & \rho_{n-1,n} \\ \rho_{1,n} & \rho_{2,n} & \dots & \rho_{n-1,n} & 1 \end{pmatrix}$$

Notons $S = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$ la matrice diagonale contenant les variances des facteurs de risque, données dans la table 2.9. Alors :

$$S^{1/2}CS^{1/2} = \begin{pmatrix} \sigma_1^2 & \dots & \sigma_1\rho_{1,n}\sigma_n \\ \vdots & \ddots & \vdots \\ \sigma_n\rho_{1,n}\sigma_1 & \dots & \sigma_n^2 \end{pmatrix} = \Sigma \quad (2.5)$$

où Σ correspond à la matrice de covariance recherchée. En diagonalisant la matrice de corrélation C comme dans l'équation 2.2 puis en la réécrivant comme dans 2.3, C s'écrit : $C = BB^T$. Grâce à l'équation 2.5, Σ s'écrit alors comme :

$$\Sigma = S^{1/2}BB^TS^{1/2} = (S^{1/2}B)(S^{1/2}B)^T$$

Si X est un vecteur gaussien de dimension n dont les coordonnées sont des lois normales centrées réduites indépendantes, alors d'après l'équation 2.4 :

$$S^{1/2}BX \sim \mathcal{N}(0, \Sigma)$$

Les combinaisons de chocs utilisées sont donc générées à partir des volatilités des facteurs de risque et de la matrice B , elle-même issue de la diagonalisation de C .

2.6.3 Implémentation des chocs

Ci-dessous sont présentées les implémentations des différents facteurs de risque au sein du modèle ALM.

Choc sur les rachats exceptionnels

Le choc sur les rachats permettant de définir des rachats plus élevés sur la première année de projection est $\Delta\text{LapseY1}$. Il est appliqué de telle sorte que les rachats lors de la première année soient définis comme :

$$\text{Taux Rachats Structurels} = \Delta\text{LapseY1}$$

$$\text{Taux Rachats Conjoncturels} = 0$$

D'où :

$$\text{Taux Rachats} = \Delta\text{LapseY1}$$

La valeur donnée à $\Delta\text{LapseY1}$ est donc le taux de rachats appliqué lors de la première année de projection.

Choc sur les rachats

Le choc sur les rachats permettant de définir des rachats plus élevés sur toute la durée de la projection est ΔLapse . Il est appliqué de telle sorte que pour toutes les années de projection, les rachats soient définis comme :

$$\text{Taux Rachats} = \text{Taux Rachats Conjoncturels} + \text{Taux Rachats Structurels} \times (1 + \Delta\text{Lapse})$$

La valeur donnée à ΔLapse est donc le taux d'augmentation des rachats structurels appliqué sur toute la durée de la projection.

Choc sur le taux de mortalité

Le choc sur le taux de mortalité permet de modifier la mortalité sur toute la durée de la projection. Il est noté ΔDeath . Notons l_x le nombre de survivant d'âge x . Sans choc, le taux de mortalité q_x est défini comme $q_x = 1 - \frac{l_{x+1}}{l_x}$. En cas de choc, le taux de mortalité est défini comme :

$$q_x = \left(1 - \frac{l_{x+1}}{l_x}\right) \times \Delta\text{Death}$$

La valeur donnée à ΔDeath est donc le taux d'augmentation des décès, quel que soit la tranche d'âge, et ce appliqué sur toute la durée de la projection.

Choc sur l'inflation

Le choc sur l'inflation noté ΔInfl modifie le taux d'inflation annuel sur toute la durée de la projection. Si l'inflation annuelle en absence de choc vaut Infl , alors :

$$\text{Infl} := \text{Infl} + \Delta\text{Infl}$$

La valeur donnée à ΔInfl est donc l'augmentation du taux d'inflation en points de pourcentage.

Choc sur les frais

Le choc sur les frais noté ΔExp modifie les taux des différents frais sur toute la durée de la projection. Si le taux de frais en absence de choc valent Exp , alors :

$$\text{Exp} := \text{Exp} \times (1 + \Delta\text{Exp})$$

La valeur donnée à ΔExp est donc le pourcentage d'augmentation du taux de frais.

Choc sur les valeurs de marché action

Le choc sur les actions noté ΔEq modifie la valeur de marché des actions utilisée pour la projection. Ainsi, on a : Valeur de marché actions = Valeur de marché actions $\times (1 + \Delta\text{Eq})$

La valeur de ΔEq est donc un pourcentage d'augmentation ou de diminution de la valeur de marché des actions en portefeuille, par rapport au scénario central.

Choc sur l'immobilier

Le choc sur l'immobilier noté ΔProp modifie la valeur de marché des actions utilisée pour la projection. Ainsi, on a : Valeur de marché immobilier = Valeur de marché immobilier $\times (1 + \Delta\text{Prop})$

La valeur de ΔProp est donc un pourcentage d'augmentation ou de diminution de la valeur de marché des biens immobiliers en portefeuille, par rapport au scénario central.

Chapitre 3

Construction du modèle proxy

La construction du modèle proxy est détaillée dans ce chapitre. L'architecture finale du modèle est le fruit de plusieurs modifications successives. Par conséquent, ce chapitre présente :

1. Le fonctionnement des réseaux de neurones récurrents,
2. L'architecture de base du modèle,
3. Les modifications effectuées pour parvenir à l'architecture finale.

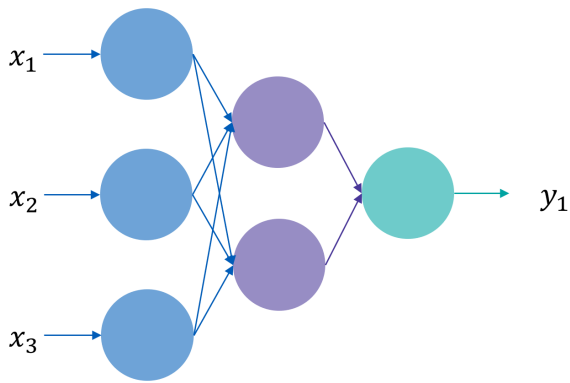
3.1 Réseaux de neurones récurrents

3.1.1 Présentation

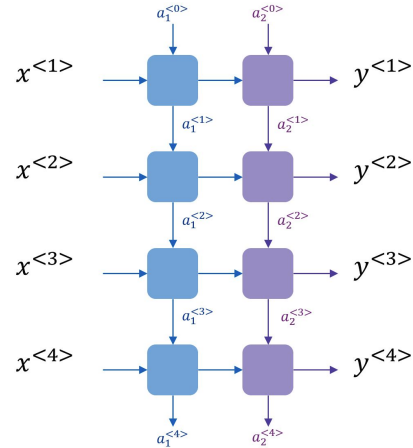
Les réseaux de neurones à propagation avant, appelés en anglais *Feedforward Neural Networks* et abrégés FNN sont des réseaux de neurones dans lesquels l'information est uniquement transmise aux "couches" suivantes, *i.e.* des noeuds d'entrée vers les noeuds de sortie. Ce sont les réseaux de neurones les plus connus et les plus utilisés, que ce soit pour les applications classiques de régression ou pour des problèmes plus complexes. L'exemple le plus connu de FNN est le perceptron multicouches, dont un exemple est présenté figure 3.1.

Ces réseaux de neurones ne permettent cependant pas de traiter des suites ordonnées d'inputs. Ils ne sont donc pas adaptés pour la prédiction de séries temporelles.

Les réseaux de neurones récurrents, appelés en anglais *Recurrent Neural Networks* et abrégés RNN sont des réseaux de neurones dans lesquels l'information transite à la fois de couche en couche, mais aussi au sein d'une même couche. Dans les réseaux de neurones récurrents, les neurones sont appelés cellules. Ces réseaux de neurones ont naturellement trouvé des applications dans les problèmes de traduction de texte ou de prédiction de séries temporelles. Au sein d'une même couche, l'information est en général transmise dans un sens unique (voir figure 3.1). Cependant, il existe des réseaux de neurones récurrents à double sens de circulation de l'information au sein d'une même couche.



(a) Réseau de neurones à propagation avant



(b) Réseau de neurones récurrents

FIGURE 3.1 – Sens de propagation de l'information en fonction du type de réseau

3.1.2 Fonctionnement

Couche

Les cellules des RNN sont regroupées en couches successives. Toutes les cellules d'une même couche sont identiques. Le réseau reçoit en entrée une suite ordonnée de vecteurs $(x^{<t>})_{t \in \{1, \dots, T\}} \in \mathbb{R}^{T \times p}$. À titre de comparaison, les réseaux de neurones à propagation avant ne reçoivent que de simples vecteurs $x \in \mathbb{R}^p$. Chacun des T vecteurs est donné en entrée d'une des T cellules de la première couche. Chaque cellule reçoit aussi un autre vecteur appelé "état caché" et noté $a^{<t-1>}$, fournit par la cellule adjacente précédente. Chaque cellule produit alors deux sorties : un vecteur $y^{<t>}$ et un nouvel état caché $a^{<t>}$. Ces vecteurs sont représentés dans la figure 3.2.

Les vecteurs $y^{<t>}$ sont les sorties calculées par les cellules en fonction des entrées $x^{<t>}$ et des états cachés $a^{<t-1>}$. Les états cachés permettent la circulation de l'information au sein d'une même couche. Cette circulation de l'information est à l'origine de l'effet "mémoire" propre aux réseaux de neurones récurrents. Il est facilement mis en avant lors de l'utilisation des RNN pour la détection d'émotions dans des textes. On parle en général de "units" pour désigner la dimension de l'état caché, *i.e.* la dimension du vecteur $a^{<t>}$.

Cette structure permettant la transmission d'information à travers les états cachés est parti-

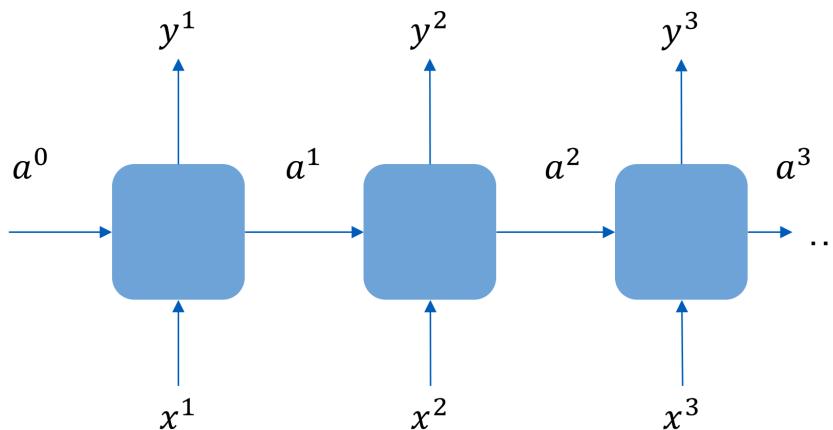


FIGURE 3.2 – Structure de réseau de neurones récurrent

culièrement adéquate pour la prédiction des flux financiers. En effet, circulant en sens unique,

l'information transmise via les états cachés agit comme une filtration, au sens mathématique du terme. Plus précisément, soit $\mathcal{F}_t = \sigma(X_0, X_1, \dots, X_t)$ la filtration naturelle associée à un processus discret $(X_t)_{t \in \mathbb{N}}$. Supposons qu'au pas de temps t , $a^{<t-1>}$ contient l'information contenue dans F_{t-1} . Alors si le RNN est suffisamment bien calibré, en fournissant l'entrée $x^t = X_t$, il est capable de calculer $a^{<t>}$ de telle sorte qu'il contienne la même information que celle contenue dans F_t . Ainsi, les entrées $(x^{<t>})_{t \in [1, T]}$ permettent de mettre à jour l'information qu'a le modèle sur l'environnement économique, via le calcul de $a^{<t>}$, et permettent d'ajuster la valeur du cash flow $y^{<t>}$ en conséquence.

Cellule

Chaque cellule contient une suite d'opérations arithmétiques et des passages dans des fonctions d'activation. Une cellule de RNN est représentée figure 3.3. Cette sous-section présente le fonctionnement des cellules RNN simples. En effet, il existe plusieurs types de cellules. Les plus connues sont les cellules RNN simples, les cellules LSTM (Long Short-Term Memory) initialement introduites en 1997 [5], et les cellules GRU (Gated Recurrent Unit), introduites en 2014 par Cho Kyunghyun et al [3]. Leurs différences principales vis-à-vis du modèle présenté ci-dessous résident dans le fait qu'elles disposent de structures légèrement plus complexes permettant de résoudre le problème de "vanishing gradient", évoqué à la fin de la sous-section 3.1.3.

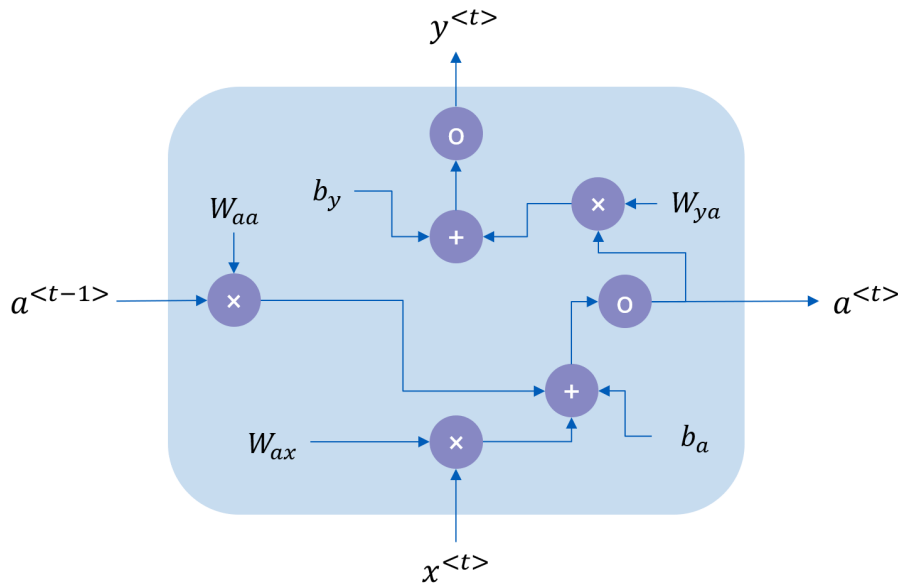


FIGURE 3.3 – Structure d'un neurone dans une couche RNN simple

Les matrices W_{aa}, W_{ax}, W_{ya} et les vecteurs b_a, b_y sont les paramètres caractérisant le fonctionnement du réseau. Ils sont communs à toutes les cellules d'une même couche. Les symboles "o" de la figure 3.2 représentent des fonctions d'activation alors que les symboles "x" représentent des produits matriciels. Les fonctions d'activation sont des fonctions par lesquelles sont composés les résultats. Si σ est une fonction d'activation, et x une valeur fournie en entrée, la sortie correspond simplement à $\sigma(x)$. Les fonctions d'activation les plus connues et utilisées sont la sigmoïde, softmax, ReLu et tanh.

Si σ_1 et σ_2 désignent les deux fonctions d'activation de la cellule, alors nous disposons (pour la structure figure 3.3) des équations :

$$a^{<t>} = \sigma_1(W_{aa}a^{<t-1>} + W_{ax}x^{<t>} + b_a) \quad (3.1)$$

$$y^{<t>} = \sigma_2(W_{ya}a^{<t>} + b_y) \quad (3.2)$$

Dans la suite, W désigne l'ensemble des coefficients contenus dans les matrices W_{aa}, W_{ax}, W_{ya} et les vecteurs b_a, b_y .

3.1.3 Calibration

Fonction de coût

Comme dans tout problème de régression (dont font partie les calibration des réseaux de neurones classiques), on dispose d'un ensemble de données $\mathcal{D}_N = (x_i, y_i)_{i=1, \dots, N}$ et d'une fonction de coût $L(y, y')$ (ou *loss function*). Cette fonction calcule une distance entre deux valeurs dans l'espace de sortie du modèle. Cette fonction coût permet de quantifier l'erreur du modèle et de l'entraîner. En effet, si $f_W(x_i)$ est la valeur prédite par le modèle pour l'input x_i , et y_i la valeur à prédire, alors $L(f_W(x_i), y_i)$ correspond bien à une distance entre la prédiction du modèle et la valeur à prédire. Les paramètres du modèle sont donc choisis de telle sorte à minimiser cette fonction coût. Les fonctions de coût usuelles pour les problèmes de régressions unidimensionnelles sont $L(y, y') = (y - y')^2$ ou $L(y, y') = |y - y'|$. Dans le cas multidimensionnel (de dimension d), ces dernières sont remplacées par leurs équivalents vectoriels, à savoir $L(y, y') = \|y - y'\|_2^2 = \sum_{i=1}^d (y_i - y'_i)^2$ ou $L(y, y') = \|y - y'\|_1 = \sum_{i=1}^d |y_i - y'_i|$.

Parfois sont introduits des termes de pénalisation, permettant de limiter le sur-apprentissage. Ces termes provoquent des coûts dans la fonction de coût plus importants si les poids de la matrice W sont important. C'est le cas dans les régressions Ridge et LASSO.

Dans le cas des réseaux de neurones classiques (FNN), les fonctions de coût utilisées correspondent généralement à une distance quadratique $L(y, y') = \|y - y'\|_2^2$. Cependant, comme représenté figure 3.2, le réseau dispose d'une suite de sorties vectorielles. Si le réseau dispose de T sorties, il est possible de calculer T fonctions de coût $L^{<t>}(y^{<t>}, y'^{<t>})$. On définit donc la fonction de coût à minimiser comme : $L(f_W(x_i), y_i) = \sum_t L^{<t>}(f_W^{<t>}(x_i), y_i^{<t>})$, comme représenté figure 3.4.

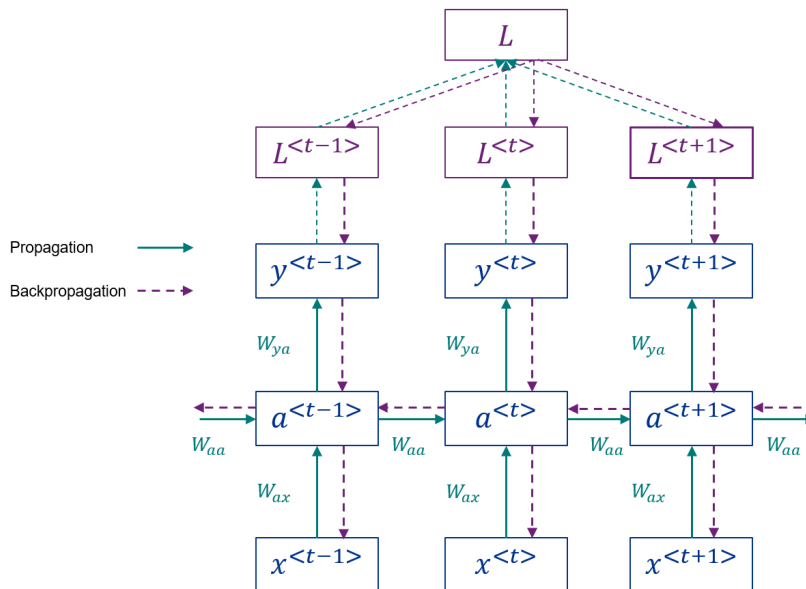


FIGURE 3.4 – Représentation de la propagation et backpropagation dans un RNN

Descente de gradient

La mise à jour des coefficients s'effectue via un algorithme de descente de gradient. L'algorithme de descente de gradient le plus simple consiste en l'équation suivante :

$$W^{t+1} = W^t - \eta \nabla_{W^t} L(f_{W^t} W(x_i), y_i)$$

où η est appelé learning rate. η définit la taille des "sauts" de la matrice W^t à chaque pas de temps. Si dans cet exemple basique, le learning rate est constant, il est variable en fonction de t pour la quasi-totalité des algorithmes de descente de gradient plus sophistiqués. La valeur du gradient précédent permet de définir la direction de la descente.

Bien qu'aujourd'hui, l'algorithme de descente de gradient précédent ne soit plus utilisé, il est à la base de toutes les méthodes développées. On peut par exemple citer les algorithmes RMSProp et Adam [6]. Ce dernier, paru en 2014, est sélectionné pour ce mémoire.

Chain rule et backpropagation

Pour calculer les dérivées par rapport à chaque coefficient de la matrice W , il est nécessaire d'utiliser la "chain rule". Soit f et g deux fonctions dérivables. Alors si on note $x \in \mathbb{R}^n$, $y = f(x) \in \mathbb{R}^m$ et $z = g(y) \in \mathbb{R}$, on a :

$$\frac{\partial z}{\partial x_i} = \sum_j \frac{\partial z}{\partial y_j} \frac{\partial y_j}{\partial x_i}$$

Cette loi permet de calculer la dérivée de notre erreur en sortie du réseau par rapport aux coefficients de W , en calculant les dérivées successives depuis la sortie jusqu'au coefficient d'intérêt. C'est le fait de propager les corrections des valeurs des coefficients en calculant des dérivées successives dans le réseau depuis la sortie qui a inspiré le nom de backpropagation.

Notons L l'erreur totale. Pour effectuer la descente de gradient, il est nécessaire de calculer $\frac{\partial L}{\partial W_{ya}}$, $\frac{\partial L}{\partial W_{aa}}$ et $\frac{\partial L}{\partial W_{ax}}$. La première dérivée partielle est simple à calculer puisque la chain rule s'applique directement :

$$\frac{\partial L}{\partial W_{ya}} = \sum_t \frac{\partial L^{<t>}}{\partial W_{ya}} = \sum_t \frac{\partial L^{<t>}}{\partial y^{<t>}} \frac{\partial y^{<t>}}{\partial W_{ya}}$$

$\frac{\partial y^{<t>}}{\partial W_{ya}}$ est facilement calculable grâce à l'équation 3.2.

Pendant, les calculs des deux autres dérivées sont plus complexes. Calculons $\frac{\partial L}{\partial W_{aa}}$. Il est nécessaire de déterminer $\frac{\partial L}{\partial W_{aa}} = \sum_t \frac{\partial L^{<t>}}{\partial W_{aa}}$. Mais la chain rule ($\frac{\partial L^{<t>}}{\partial W_{aa}} = \frac{\partial L^{<t>}}{\partial y^{<t>}} \frac{\partial y^{<t>}}{\partial a^{<t>}} \frac{\partial a^{<t>}}{\partial W_{aa}}$) ne s'applique pas puisque $a^{<t>}$ dépend aussi de W_{aa} via $a^{<t-1>}$, $a^{<t-2>}$, etc. Il faut donc utiliser la formule de la dérivée totale qui donne :

$$\frac{\partial L^{<t>}}{\partial W_{aa}} = \frac{\partial L^{<t>}}{\partial y^{<t>}} \frac{\partial y^{<t>}}{\partial a^{<t>}} \left(\frac{\partial a^{<t>}}{\partial W_{aa}} + \frac{\partial a^{<t>}}{\partial a^{<t-1>}} \frac{\partial a^{<t-1>}}{\partial W_{aa}} + \dots + \frac{\partial a^{<t>}}{\partial a^{<t-1>}} \dots \frac{\partial a^{<0>}}{\partial W_{aa}} \right)$$

Le calcul est le même pour obtenir $\frac{\partial L^{<t>}}{\partial W_{ax}}$.

Vanishing gradient

La somme $\sum_i \frac{\partial a^{<t>}}{\partial a^{<t-1>}} \dots \frac{\partial a^{<t-i>}}{\partial W_{aa}}$ a des propriétés particulières. Si les normes des dérivées successives $\frac{\partial a^{<t>}}{\partial a^{<t-1>}}$ sont inférieures à 1, alors le produit tend exponentiellement vite vers 0. Au contraire, si cette norme est supérieure à 1, ce produit explose. Le cas où les normes sont

inférieures à 1 correspond au problème du vanishing gradient. Il empêche la prise en compte des dépendances long terme lors de l'apprentissage et lors des prédictions, puisque le produit tend vers 0. Les cellules GRU et LSTM évoquées précédemment résolvent ce problème. Pour en savoir plus, nous vous renvoyons aux ressources [3, 5].

3.2 Architecture initiale du modèle proxy

Le modèle proxy proposé prend en entrée une trajectoire issue d'un scénario économique et des valeurs de chocs et prédit en sortie les flux de passif actualisés associés au scénario et au choc.

3.2.1 Inputs

Un scénario économique est décrit par les rendements des actions et de l'immobilier et les courbes des taux sur 60 ans. Les scénarios économiques sont initialement pris en compte dans le modèle via la variable

$$x^{<t>} = (r_{Eq1}^{<t>}, r_{Eq2}^{<t>}, r_{Immo}^{<t>}, r_{30}^{<t>})$$

avec $r_{Eq1}^{<t>}$, $r_{Eq2}^{<t>}$ et $r_{Immo}^{<t>}$ les rentabilités sur l'année de projection des marchés actions et de l'immobilier, et $r_{30}^{<t>}$ le taux zéro-coupon de maturité 30 ans sur l'année de projection. Les taux zéro-coupons sont obtenus à partir des prix des zéro-coupons grâce à la formule :

$$r_t = \left(\frac{1}{ZC_t}\right)^{\frac{1}{t}} - 1$$

Cette formule s'obtient directement en inversant la formule du prix des zéro-coupons en convention de taux actuariels.

Les chocs sur les facteurs ne peuvent être pris en compte dans les vecteurs $x^{<t>}$ car ils ne dépendent pas du temps. Ils sont donc pris en compte via l'état caché de la première couche. Le nombre de chocs (7) étant inférieur à la dimension de l'état caché (supérieure à 8), l'état caché est initialisé avec la valeur :

$$a_1^0 = (\Delta\text{LapseY1}, \dots, \Delta\text{Prop}, 0, \dots, 0)$$

L'état caché de la seconde couche est initialisé à $(0, \dots, 0)$. La prise en compte en compte des chocs et des scénarios économiques est présentée dans la figure 3.5.

3.2.2 Outputs

Le modèle est calibré de telle sorte qu'il prédise les flux actualisés des passifs liés aux contrats.

La fonction d'activation en sortie des cellules des réseaux de neurones récurrents est une tangente hyperbolique (\tanh), prenant ses valeurs dans $] -1; 1[$. De plus, la sortie $y^{<t>}$ d'une cellule d'une couche de RNN est un vecteur. Une couche de régression est donc placée en sortie du réseau de neurones afin de permettre les prédictions sur un intervalle suffisamment large. Il s'agit d'une régression linéaire appliquée à chaque sortie de cellule. Les coefficients de la régression sont communs à toutes les sorties de la dernière couche. Cette couche de régression est représentée dans la figure 3.5.

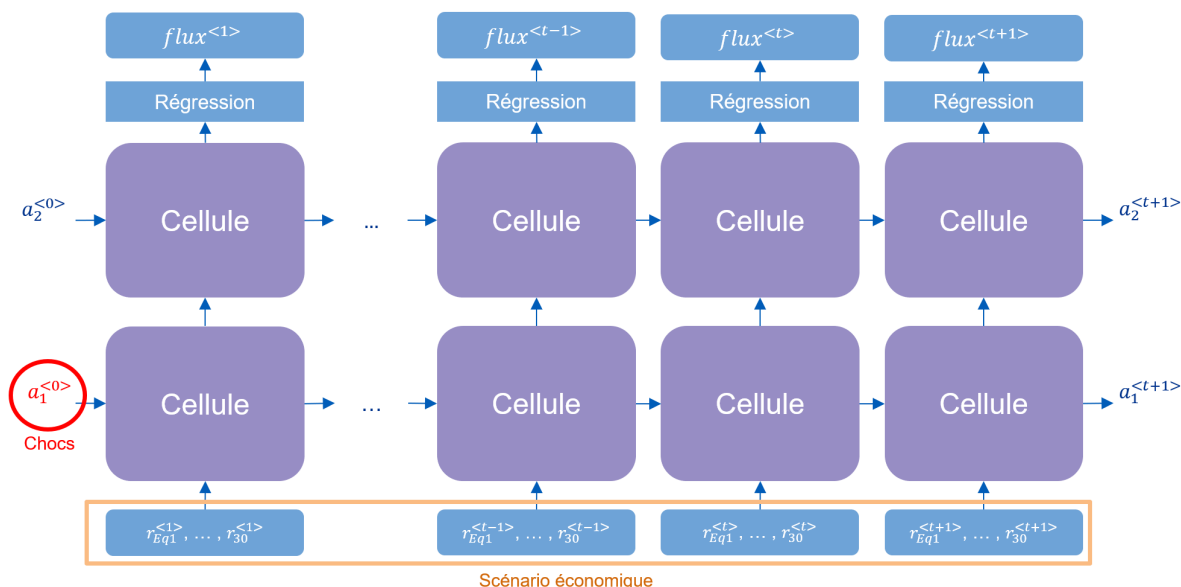


FIGURE 3.5 – Architecture de notre réseau

3.3 Conditions d'entraînement

3.3.1 Loss function et sur-apprentissage

Un premier critère de performance permettant de déterminer quels sont les modèles les plus performants est la fonction de coût introduite dans la section 3.1.3. Soit un ensemble de n scénarios aléatoires créés à partir d'un seul couple de courbe de taux et de volatilités, lui-même associé à un unique choc. Soient $y_i^{<t>$ le cash flow à prédire pour le scénario i au pas de temps t , et $\hat{y}_i^{<t>$ le cash flow prédit, alors la fonction de coût est :

$$L(y_i, \hat{y}_i) = \sum_{t=1}^T (\hat{y}_i^{<t> - y_i^{<t>})^2$$

Elle correspond à la somme des erreurs quadratiques sur chaque pas de temps, pour un unique scénario.

Le calcul de la fonction de coût sur les données utilisées pour l'entraînement donne l'erreur de prédiction des cash flows. Cependant, pour connaître les performances du modèle, il est nécessaire de calculer l'erreur sur des données indépendantes des données d'entraînement. En effet, l'entraînement de l'algorithme consiste en la minimisation de cette fonction sur le jeu d'entraînement. Le modèle peut donc apprendre des poids permettant de minimiser l'erreur spécifiquement sur les données d'entraînement, tout en donnant des résultats médiocres sur de nouvelles données : le modèle est alors sujet à de l'*overfitting* ou sur-apprentissage. Pour éviter ce cas, des techniques d'apprentissage spécifiques peuvent être introduites comme le dropout ou l'utilisation de termes de pénalisation dans la fonction de coût. Cependant, le seul moyen de détecter l'*overfitting* est de calculer la fonction de coût sur des données absentes des données d'entraînement : ce sont les données de validation ou *test set*.

Si la fonction de coût diminue à la fois sur les données d'entraînement et de validation, le modèle apprend le lien ou "mécanisme" reliant les entrées aux sorties. Si la fonction de coût diminue sur les données d'entraînement mais augmente sur les données de validation, alors le modèle apprend des poids permettant de minimiser spécifiquement la fonction sur les données d'entraînement et n'apprend plus "le mécanisme". La figure 3.6 illustre ce phénomène.

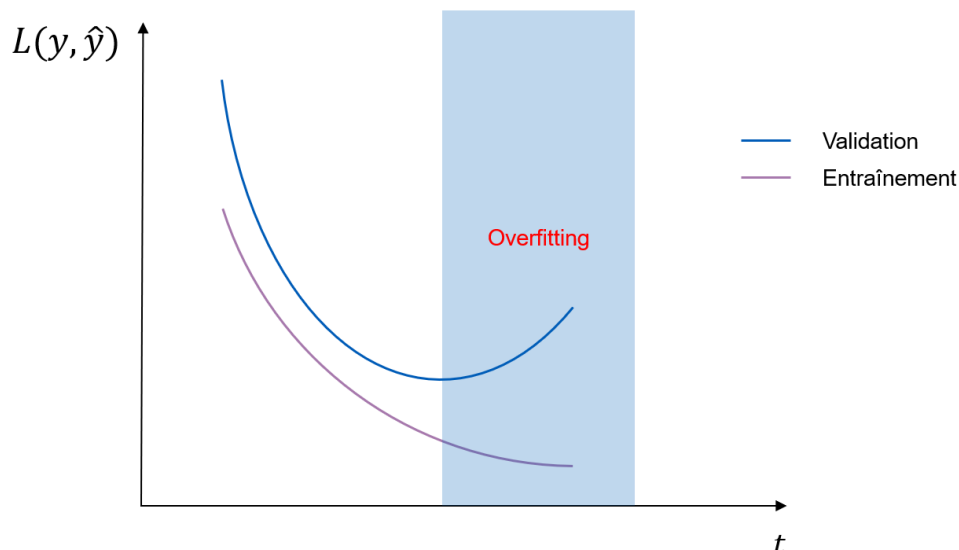


FIGURE 3.6 – Détection de l'overfitting sur les courbes d'apprentissage

3.3.2 Root Mean Square Error

La fonction de coût permet de définir une distance dans l'espace de prédiction du modèle. C'est cette distance qui est utilisée pour calibrer le modèle. Cependant, cette distance ne permet parfois pas de quantifier la précision du modèle pour une utilisation donnée. Par exemple, la fonction de coût utilisée ici est la distance entre la série de cash flows actualisés prédits et la série à prédire. Cependant, bien que liée, la fonction de coût ne permet pas de quantifier l'erreur de prédiction du modèle sur les Best Estimate. Une mesure de l'erreur sur les prédictions de BE est donc nécessaire.

Si le BE est calculé sur n scénarios aléatoires, on a :

$$\widehat{BE} = \frac{1}{n} \sum_{i=1}^n \sum_{t=1}^T \hat{y}_i^{<t>}$$

$$BE = \frac{1}{n} \sum_{i=1}^n \sum_{t=1}^T y_i^{<t>}$$

Si on dispose de N ensembles de n scénarios permettant de calculer N BE, la racine de l'erreur quadratique moyenne (notée RMSE) est définie comme :

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{j=1}^N (\widehat{BE}_j - BE_j)^2}$$

Cette valeur correspond intuitivement à un écart-type des erreurs de prédictions, si on suppose qu'elles sont centrées. Le RMSE est homogène au BE : il est exprimé en euros. On peut aussi définir un RMSE normalisé comme :

$$\text{RMSE}_{\%} = \sqrt{\frac{1}{N} \sum_{j=1}^N \left(\frac{\widehat{BE}_j - BE_j}{BE_j} \right)^2}$$

Il s'agit simplement de la moyenne quadratique des erreurs relatives. Par abus de langage, le terme RMSE pourra par la suite désigner le RMSE en euros ou le RMSE normalisé.

3.3.3 Données d'entraînement

La base de données est construite à partir de 20 couples de courbes de taux et volatilités des marchés actions et immobilier. Chacun de ces couple a permis de générer 100 scénarios économiques aléatoires. Un scénario économique correspond à un ensemble de rendements et de taux zéros coupons aléatoires, décrivant l'évolution de l'environnement économique sur 60 ans. Autrement dit, il s'agit d'une projection aléatoire de l'environnement économique sur 60 ans. Ces couples (*courbe de taux ; volatilités*) ont permis de générer un total de $100 \times 20 = 2000$ scénarios économiques aléatoires. La construction des courbes de taux, la sélection des volatilités et la génération des scénarios sont décrits dans la section 2.4.

Ces 2000 scénarios aléatoires ont été croisés avec 200 combinaisons aléatoires de chocs, dont la construction est présentée dans la section 2.6. Le modèle dispose donc d'une base d'entraînement comprenant 400 000 lignes.

3.3.4 Données de validation

Pour le calcul de la loss pendant l'entraînement

Les données de test créées pour évaluer la fonction de coût doivent être indépendantes des données d'entraînement. Pour cela sont générés 50 vecteurs de chocs, conformément à 2.6. Ces chocs sont croisés à 4 couples de courbes de taux et volatilités générés conformément à 2.4. Chacun des 200 couples de scénarios et chocs génère 100 trajectoires aléatoires.

Cette base de données permet de tester les performances du modèle sur des données indépendantes et représentatives de l'utilisation finale de l'outil.

Pour le calcul du RMSE sur le BE

Comme expliqué dans la sous-section 3.3.2, la mesure d'intérêt est le RMSE sur les BE. Pour le calculer, il est nécessaire de disposer d'ensembles de scénarios aléatoires homogènes avec N assez grand, pour garantir la convergence de la méthode de Monte Carlo vers le BE stochastique. Les données de validation utilisées pour le calcul de la fonction de coût ne disposent que de $N = 100$ trajectoires aléatoires par couple de scénario et choc, ce qui est insuffisant. Un nouvel ensemble de données est donc créé avec $N = 1000$, de la même manière que les données utilisées pour la loss, au détail près que le nombre de chocs générés est réduit (pour réduire le temps de calcul, déjà important à cause de la projection sur 1000 scénarios). Sont donc générés 10 vecteurs de chocs, conformément à 2.6. Ces chocs sont encore croisés à 4 couples de courbes de taux et volatilités générés conformément à 2.4.

Ces 10×4 scénarios choqués génèrent alors chacun $N = 1000$ trajectoires sous probabilité risque-neutre permettant le calcul de 40 BE.

3.4 Choix des cellules

Comme évoqué dans la section 3.1.2, il existe trois types de cellules différentes :

- les cellules RNN "simples",
- les cellules de type Gated Recurrent Units (GRU),
- les cellules de type Long Short-Term Memory (LSTM).

Ces trois types de cellules diffèrent principalement par leur traitement des états cachés.

L'hyper-paramètre principal des réseaux de neurones récurrents est la dimension de l'état caché. Ce paramètre est appelé *units* dans TensorFlow. Une grande dimension implique un plus

grand transport de l'information via l'état caché (l'état caché disposant de plus de coefficients pour stocker l'information). Cependant, l'état caché est obtenu par produit matriciel : une grande dimension implique donc aussi un plus grand nombre de paramètres dans la matrice W . Bien que l'augmentation de cette dimension facilite l'apprentissage des mécanismes complexes, il favorise aussi le sur-apprentissage et augmente le temps d'entraînement. Afin de choisir les cellules permettant d'obtenir le meilleur modèle proxy possible, 12 modèles ont été calibrés sur l'ensemble de la base de données d'entraînement. Ces 12 modèles sont issus de l'entraînement des réseaux de neurones avec les trois types de cellules sur 4 nombres de *units* différents. Les valeurs des *loss* sur l'ensemble d'entraînement et l'ensemble de validation sont données dans la table 3.1. Les données étant normalisées pour l'apprentissage, les valeurs des *loss* n'ont pas d'unités.

Units	RNN		LSTM		GRU	
	Entraînement	Validation	Entraînement	Validation	Entraînement	Validation
8	2,64%	2,11%	1,15%	1,15%	0,99%	1,04%
16	0,95%	1,38%	0,48%	0,71%	0,53%	0,74%
32	0,74%	0,81%	0,32%	0,67%	0,37%	0,39%
64	0,71%	0,77%	0,27%	0,44%	0,28%	0,37%

TABLE 3.1 – Comparaison des erreurs des différentes cellules

La table 3.1 montre clairement que les cellules RNN simples donnent des résultats en deçà de ceux obtenus avec les cellules GRU et LSTM, quelque soit le nombre de *units*. De même, plus le nombre de units est important et plus l'erreur est faible, que ce soit sur la base d'entraînement ou la base de validation.

Les résultats des cellules GRU et LSTM sont proches. Cependant, on observe que les *loss* sont plus faibles sur les données d'entraînement pour les LSTM pour $Units \geq 16$. Au contraire, elles sont plus faibles sur les données de validation pour les GRU. Ceci peut s'expliquer par le nombre de paramètres internes (poids à calibrer dans les matrices W) des cellules LSTM et GRU : les cellules LSTM disposent de plus de paramètres que les cellules GRU pour une dimension d'état caché donnée. À titre d'exemple, pour l'état caché de dimension 64, le réseau à base de LSTM dispose de 50 779 paramètres alors que le GRU n'en a que 38 491. Le temps d'entraînement sur l'ensemble des données, pour 5 itérations (appelées "epochs"), est de 2 250 secondes pour les cellules GRU et de 2 800 secondes pour les cellules LSTM . Le temps d'entraînement pour les cellules LSTM est donc environ 25% plus long que pour les GRU. Les cellules GRU ont donc un temps d'entraînement, une *loss* sur les données de validation et un risque de sur-apprentissage plus faibles. Elles sont sélectionnées pour créer le modèle proxy.

3.5 Nombre de units et epochs

Un epoch est une itération sur le jeu d'entraînement. A chaque epoch, le modèle parcourt l'ensemble du jeu d'entraînement, et modifie ses poids au fur à mesure. Si le nombre d'epochs est égal à 5, le modèle parcourt 5 fois l'ensemble des données d'entraînement au cours de son entraînement. TensorFlow permet l'affichage des *loss* à la fin de chaque epoch, comme dans la figure 3.7. Elle montre comme attendu une erreur décroissante avec le nombre d'epochs, puisque les poids sont ajustés au fur et à mesure pour minimiser la *loss* (ou résultat de la fonction de coût sur les données d'entraînement).

Bien que la table 3.1 indique que la *loss* est plus faible pour un nombre de units plus important, ce résultat est donné pour un nombre d'epochs égal à 5. Le nombre de units modifie le temps d'entraînement pour un epoch, ainsi que la vitesse de convergence. Pour savoir s'il est

```

Epoch 1/5
11250/11250 [=====] - 449s 40ms/step - loss: 0.0154 - mse: 0.0154 - val_loss: 0.0073 - val_mse: 0.0073
Epoch 2/5
11250/11250 [=====] - 621s 55ms/step - loss: 0.0032 - mse: 0.0032 - val_loss: 0.0064 - val_mse: 0.0064
Epoch 3/5
11250/11250 [=====] - 656s 58ms/step - loss: 0.0025 - mse: 0.0025 - val_loss: 0.0061 - val_mse: 0.0061
Epoch 4/5
11250/11250 [=====] - 682s 61ms/step - loss: 0.0022 - mse: 0.0022 - val_loss: 0.0064 - val_mse: 0.0064
Epoch 5/5
11250/11250 [=====] - 680s 60ms/step - loss: 0.0020 - mse: 0.0020 - val_loss: 0.0057 - val_mse: 0.0057

```

FIGURE 3.7 – Exemple d’affichage des informations d’entraînement pour 5 epochs

optimal d’utiliser 32 ou 64 units, il est nécessaire de tracer les courbes de loss en fonction des epochs.

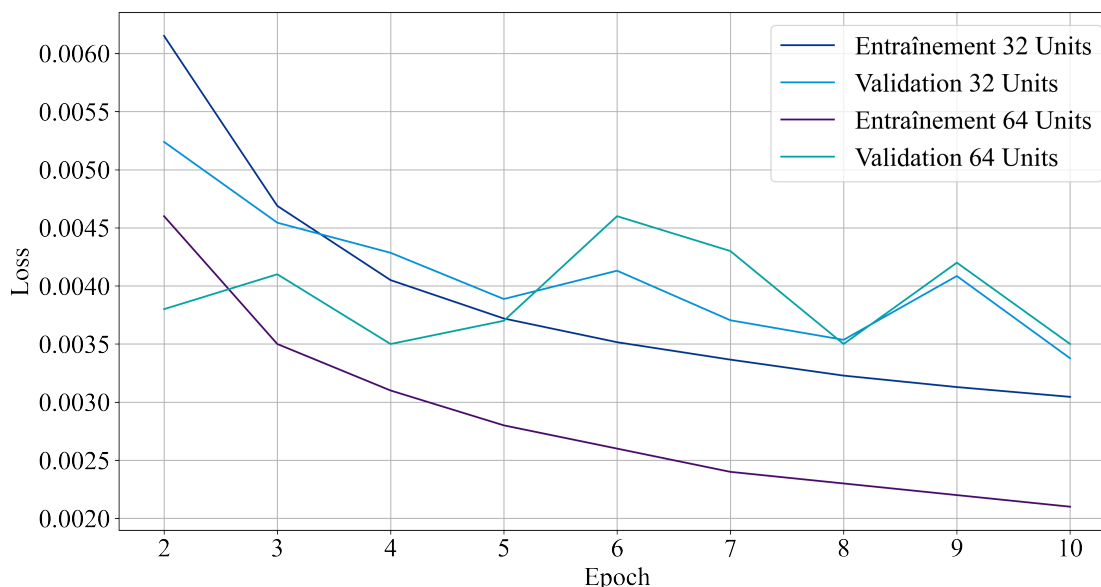


FIGURE 3.8 – Courbes d’apprentissage des modèles GRU

La figure 3.8 montre que les modèles optimaux à 32 et 64 units par rapport à l’erreur sur les données de validation sont respectivement obtenus pour 10 et 8 epochs. Les erreurs minimales atteintes par les modèle à 32 et 64 units sont de 0.35%, pour respectivement 2 900 et 3 120 secondes d’entraînement.

Les **racines des erreurs quadratiques moyennes sur les BE**, calculées grâce aux données de validation pour les calculs de BE, sont respectivement de **1.77%** et **1.51%** pour les modèles à 32 et 64 units. **Le modèle à 64 units est donc sélectionné.**

3.6 Prise en compte des taux d’intérêts

3.6.1 Identification du problème

Le modèle présenté précédemment démontre des difficultés dans la prise en compte des courbes des taux. Une analyse similaire à celle détaillée dans la section 4.3 a permis d’identifier cette faiblesse. En effet, bien que ce chapitre présente une construction simple et linéaire du modèle final, les modifications ont été effectuées par phases successives d’implémentations et de tests. C’est pendant l’élaboration des tests du chapitre 4 qu’ont été décelées les difficultés de prédiction du modèle en fonction des taux, et que les améliorations présentées dans cette

section et la section suivante ont pu être apportées. Le problème a été décelé lors de la mise en place des chocs de taux, comme dans la section 4.3.

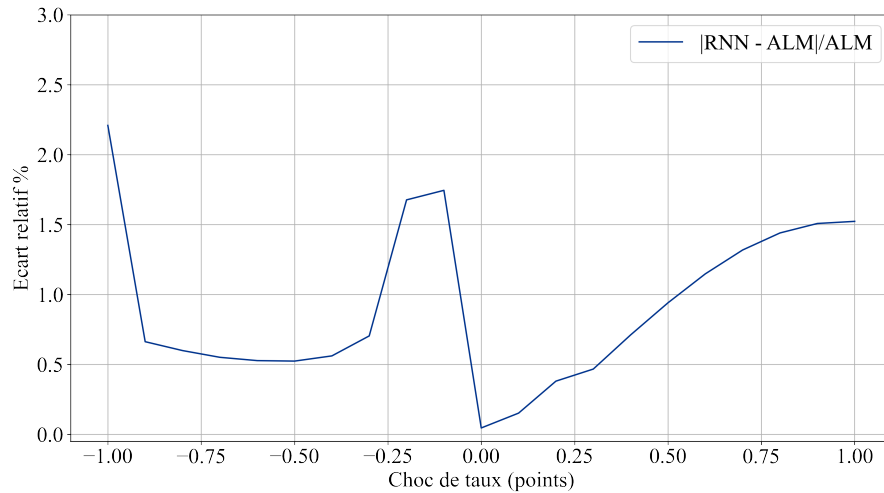
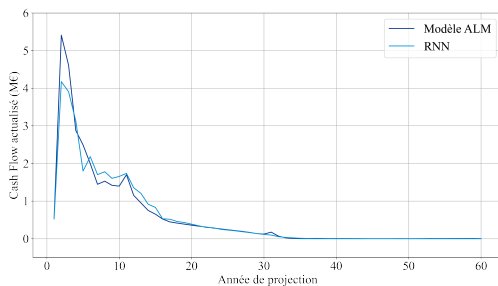
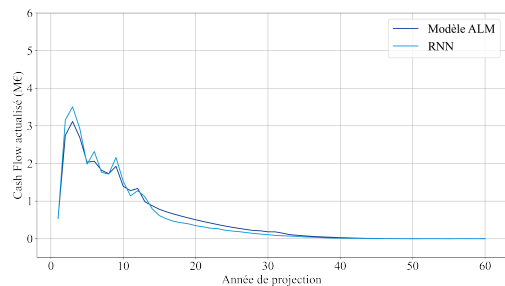


FIGURE 3.9 – Ecart relatif de prédiction du BE en fonction du choc de taux

La figure 3.9 présente l’erreur de prédiction en fonction de l’éloignement des taux spots à 1 an et 30 ans par rapport au scénario central. L’axe des abscisses donne la translation appliquée en points de pourcentage. Le scénario central est présenté dans la section 4.3. Il est remarquable que l’erreur croît grandement avec l’éloignement des taux par rapport au scénario central.



(a) Choc de taux de 1 point à la baisse



(b) Choc de taux de 1 point à la hausse

FIGURE 3.10 – Cash flows obtenus pour des scénarios stochastiques avec chocs de taux

La figure 3.10 présente deux exemples de projection des cash flows sur des scénarios stochastiques. La figure (a) montre les cash flows prédits par le RNN et ceux calculés par le modèle ALM pour un choc de 1 point de pourcentage à la baisse. La figure (b) présente ce même résultat pour un choc de 1 point à la hausse. Les prédictions étant faites sur les cash flows actualisés, les taux d’intérêts ont une influence sur leurs valeurs. Une hausse de la courbe des taux a tendance, par effet d’actualisation, à diminuer la valeur des cash flows. À l’inverse, une baisse de la courbe a tendance à faire augmenter la valeur des cash flows. Les deux figures montrent clairement que le RNN n’est pas capable d’ajuster l’actualisation des cash flows. En cas de baisse des taux, le modèle proxy prédit des cash flows trop faibles. Au contraire, en cas de hausse des taux, les cash flows prédits sont trop élevés.

3.6.2 Ajout de taux zéro-coupons

Le choix avait été fait de ne conserver qu'un unique taux d'intérêt en entrée du modèle : le taux d'intérêt sans risque associé au zéro-coupon de maturité 30 ans. Ainsi, la courbe des taux n'était représentée que par un point. Ceci s'est traduit par l'incapacité du modèle à correctement actualiser les flux de maturités courtes, comme montré dans la figure 3.10. Le choix est donc fait d'ajouter un certain nombre de points de la courbe des taux. Le choix retenu après plusieurs phases de tests est d'utiliser 5 points : **les taux d'intérêts à 1, 5, 10, 20 et 30 ans**. Les vecteurs donnés en entrée du modèle s'écrivent donc dorénavant sous la forme :

$$x^{<t>} = (r_{Eq1}^{<t>}, r_{Eq2}^{<t>}, r_{Immo}^{<t>}, r_1^{<t>}, r_5^{<t>}, r_{10}^{<t>}, r_{20}^{<t>}, r_{30}^{<t>})$$

Ce choix résulte d'un arbitrage entre temps d'entraînement, et précision des résultats sur les données de validation. La courbe est représentée par 5 points, permettant au modèle d'appréhender des creux, bosses ou tout autre type de déformation pour les maturités inférieures à 30 ans. Les nouveaux écarts relatifs obtenus sont donnés dans la figure 3.11.

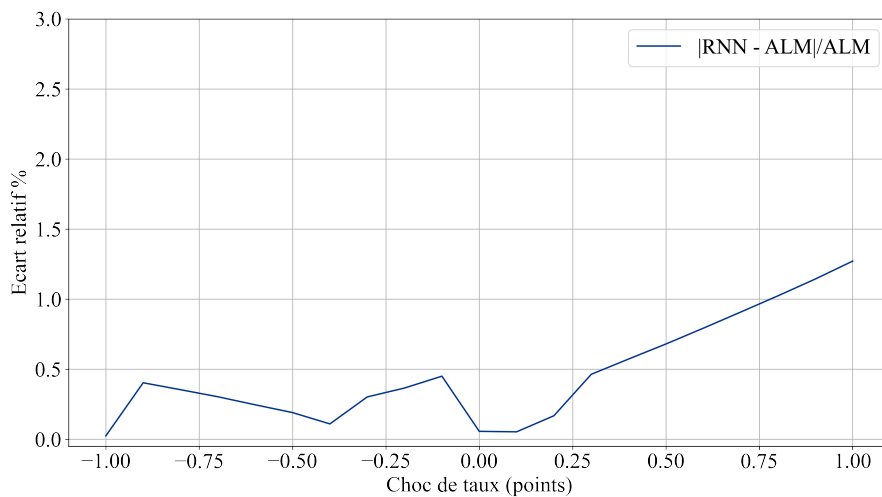
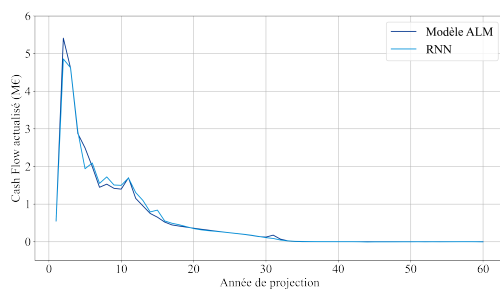
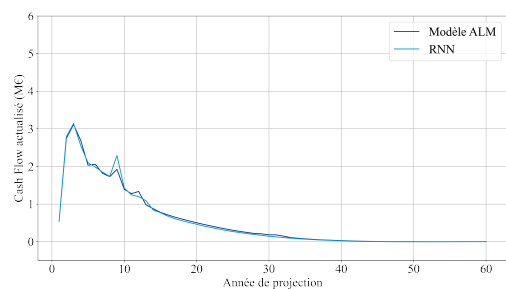


FIGURE 3.11 – Ecart relatif de prédiction avec 5 taux en entrées

On observe que les écarts relatifs ont grandement diminué pour les chocs proches de 1% et -1%. Ces observations se traduisent par des cash flows mieux actualisés, comme le montre la figure 3.12.



(a) Choc de taux de 1 point à la baisse



(b) Choc de taux de 1 point à la hausse

FIGURE 3.12 – Cash flows obtenus pour des scénarios stochastiques avec chocs de taux

3.7 Suppression du biais

3.7.1 Identification du problème

Le modèle obtenu précédemment est optimal en terme de minimisation de la loss function (fonction de coût). Alors que le modèle prenant en compte un taux d'intérêts ne permet de réduire la loss que jusqu'à 0.23% sur les données d'entraînement et à 0.29% sur les données de test, le modèle à 5 taux permet de les réduire respectivement à 0.08% et 0.1%. Pour rappel, la loss est calculée comme :

$$L(y_i, \hat{y}_i) = \sum_{t=1}^T (\hat{y}_i^{<t>} - y_i^{<t>})^2$$

La loss est ici exactement égale à une mesure de l'erreur quadratique moyenne dans l'espace des vecteurs de dimension T contenant les T cash flows (prédits ou à prédire). La minimisation de la loss consiste donc en la minimisation des distances quadratiques entre les cash flows prédits et les cash flows à prédire. Cependant, il est impossible de connaître les capacités de prédiction du modèle pour les BE. Une erreur faible sur la prédiction des cash flows implique évidemment une erreur faible pour les prédictions de BE. Cependant, l'optimalité du modèle pour les prédictions de cash flows n'implique pas l'optimalité pour la prédiction des BE : un biais pourrait par exemple être introduit. C'est l'hypothèse que suggère la figure 3.11, : tous les écarts relatifs sont négatifs. Pour confirmer qu'il existe bien un biais de prédiction des BE, il est nécessaire de tester le modèle sur un jeu de données contenant plus de scénarios et de chocs.

Les données de validation présentées précédemment permettent de calculer la loss sur un grand ensemble de couples de chocs et scénarios économiques différents. Cependant, chacun n'ayant généré que 100 trajectoires, il est impossible de mesurer l'efficacité du modèle dans le cadre de prédiction de BE. Pour cela, il est nécessaire de disposer de 1000 trajectoires pour chaque couple de scénario et choc, afin de garantir une convergence suffisante de la méthode de Monte Carlo associée. Ainsi, même si un premier jeu de données de validation est à notre disposition pour calculer la loss pendant l'entraînement, un deuxième jeu encore indépendant des données d'entraînement est créé. Il est toujours constitué de 4 scénarios économiques générés aléatoirement. Le temps de calcul étant approximativement multiplié par 10 à cause du passage de 100 à 1000 trajectoires aléatoires, il est nécessaire de réduire le nombre de couples (*scénario économique, choc sur les facteurs de risques*). Ainsi, les scénarios économiques ne sont croisés qu'à 10 chocs différents sur les facteurs de risque. Cette base de données est présentée plus en détails dans la section 4.1.1.

Choc	Scénario 1	Scénario 2	Scénario 3	Scénario 4
1	-1,23	-1,37	-1,24	-1,40
2	-0,25	-0,23	-0,77	-1,37
3	-0,69	-0,35	-1,08	-2,11
4	-0,53	-0,44	-0,87	-1,31
5	-0,35	-0,14	-0,81	-1,55
6	-0,41	-0,33	-0,71	-1,21
7	-0,48	-0,15	-0,92	-1,66
8	-0,61	-0,12	-0,63	-0,89
9	-0,63	-0,58	-0,8	-1,00
10	-0,44	-0,32	-0,81	-1,34

TABLE 3.2 – Écarts relatifs de prédiction des BE en %

Les résultats de prédiction de BE du modèle à 5 taux sur les 40 couples de scénarios et chocs sont présentés dans la table 3.2.

Les courbes de taux et volatilités utilisées ne présentent pas de caractéristiques communes. La table 3.2 indique donc un biais à la baisse dans les prédictions de BE.

3.7.2 Modification de la loss function

Pour corriger ce biais, on propose d'introduire un nouveau terme dans la fonction de coût.

Cette idée s'inspire d'abord des termes de pénalisation que l'on peut rencontrer dans les méthodes Ridge et LASSO. Ils permettent d'introduire une contrainte supplémentaire dans la convergence du modèle. Dans le cas de Ridge et LASSO, ce terme permet de limiter l'overfitting, en ajoutant la norme L^1 ou L^2 des paramètres de régression (matrice W) dans la fonction de coût. C'est en s'inspirant de ces termes de pénalisation que Loisel, Piette et Tsai sont parvenus à optimiser leur algorithme de prédiction des rachats en assurance-vie [7]. Le papier cité propose de calibrer le modèle de telle sorte à minimiser l'erreur de prédiction des pertes économiques engendrées par les rachats, au lieu des erreurs sur les prédictions des rachats.

La proposition qui est faite dans ce mémoire est donc d'ajouter un terme de pénalisation du biais de prédiction sur les BE. Le terme s'apparentant à la MSE des cash flows prédits est bien conservé. Cependant, le terme représentant le biais de prédiction des BE est introduit. Ce terme est calculé comme l'erreur quadratique de prédiction de BE que l'on obtiendrait sur un unique scénario stochastique :

$$\left(\sum_{t=1}^T (\hat{y}_i^{<t>} - y_i^{<t>}) \right)^2$$

Ainsi, la nouvelle fonction de coût utilisée est :

$$L(y_i, \hat{y}_i) = \sum_{t=1}^T (\hat{y}_i^{<t>} - y_i^{<t>})^2 + C \left(\sum_{t=1}^T (\hat{y}_i^{<t>} - y_i^{<t>}) \right)^2$$

avec C un hyper-paramètre à calibrer permettant de régler l'importance de la pénalisation du biais dans la loss function. Après plusieurs entraînements successifs pour différentes valeurs de C , la valeur $C = 2$ est retenue.

Choc	Scénario 1	Scénario 2	Scénario 3	Scénario 4
1	-0.22	0.37	0.03	1.09
2	-0.74	0.20	-0.24	0.54
3	-0.69	0.26	-0.24	0.45
4	-0.71	0.15	-0.30	0.58
5	-0.72	0.20	-0.30	0.36
6	-0.63	0.18	-0.28	0.59
7	-0.76	0.09	-0.37	0.43
8	-0.19	0.56	0.09	1.02
9	-0.41	0.36	-0.15	0.84
10	-0.65	0.18	-0.30	0.58

TABLE 3.3 – Écarts relatifs de prédiction des BE en %

La table 3.3 montre des écarts relatifs dont le signe est principalement corrélé au scénario. La sous-estimation du BE (biais de prédiction à la baisse) par le modèle a bien été supprimée. C'est ce modèle, comportant 5 taux en entrée et entraîné avec une fonction de coût personnalisée qui est conservé dans toute la suite de ce mémoire.

3.8 Comparaison des modèles

Ainsi, 3 versions du modèle ont été proposées après sélection du type de cellule et du nombre de units. Elles ont toutes la même architecture de base, à savoir :

- deux couches de réseaux de neurones récurrents,
- utilisant des cellules GRU,
- dont l'état caché est de dimension 64 ;
- et dont les vecteurs donnés en inputs contiennent les rendements des actions et de l'immobilier à chaque pas de temps.

Cependant les trois versions ont des caractéristiques différentes :

1. Utilisation en input d'un unique taux zéro-coupon de maturité 30 ans,
2. Utilisation des taux zéro-coupons de maturités 1, 5, 10, 20 et 30 ans,
3. Utilisation des taux zéro-coupons de maturités 1, 5, 10, 20, et 30 ans ; optimisation avec une fonction de coût pénalisant le biais de prédiction du BE.

Différents résultats des trois algorithmes présentés précédemment sont donnés dans la table 3.4. Le RMSE CF correspond aux premières loss utilisées, *i.e.* la racine de l'erreur quadratique moyenne sur les prédictions des cash flows. Elle est calculée sur les données d'entraînement et le premier jeu de données de validation, utilisé pour les cash flows. Le RMSE BE correspond à la racine de l'erreur quadratique moyenne sur les prédictions de Best Estimate. Cette mesure est calculée sur le deuxième jeu de données de validation, permettant de calculer les erreurs sur les BE.

Modèle	1 Taux ZC	5 Taux ZC	Avec Pénalisation du biais
RMSE CF - Entraînement	0.21%	0.08%	0.13%
RMSE CF - Validation	0.35%	0.10%	0.15%
RMSE relatif BE - Validation	1.51%	0.93%	0.49%
RMSE BE (k€) - Validation	499	304	160

TABLE 3.4 – Statistiques de prédiction

Le dernier modèle, à 5 maturités de taux différentes en input, et optimisé avec une fonction de coût pénalisant le biais a des erreurs quadratiques de prédiction des cash flows plus élevées. Pour autant, il obtient bien les meilleures performances de prédiction du BE, avec un RMSE de 0.49%, puisque le biais à la prédiction des BE a été supprimé.

Chapitre 4

Analyse des résultats

4.1 Précision sur les données de validation

4.1.1 Présentation des données

Les données permettant d’analyser la précision du modèle pour la prédiction du BE sont indépendantes des données d’entraînement. Elles sont composées de 4 scénarios économiques, ayant chacun généré 1000 trajectoires, croisés à 10 chocs sur les facteurs de risque. Ces scénarios et chocs sont créés conformément à 2.4. Les données utilisées pour la validation dans le chapitre 3 étaient composées de 50 chocs croisés à 4 scénarios économiques. Cependant, ceux-ci ne génèrent que 100 trajectoires chacun, afin de limiter le temps de calcul. Ici, il est nécessaire de créer des ensembles de 1000 trajectoires afin de pouvoir calculer des Best Estimate suffisamment précis. La table 4.1 présente certaines statistiques sur les chocs sur les facteurs de risque utilisés dans l’ensemble de validation, ainsi que sur les Best Estimate résultants.

Variable	Moyenne	Min	$q_{25\%}$	$q_{50\%}$	$q_{75\%}$	Max
$\Delta\text{LapseY1}$ (%)	7.43	-22.53	-5.23	0.14	28.62	45.06
ΔLapse (%)	4.76	-24.72	-11.19	1.6	25.44	29.7
ΔDeath (%)	-0.6	-13.76	-9.14	0.34	4.12	10.66
ΔInfl (%)	-0.01	-0.66	-0.29	-0.15	0.38	0.73
ΔExp (%)	-0.09	-8.63	-3.6	-1.26	3.76	8.6
ΔEq (%)	-7.58	-24.09	-17.78	-8.91	-5.03	14.35
ΔProp (%)	-4.31	-21.88	-16.04	-1.0	7.15	9.32
BE (k€)	32 771	30 933	32 118	32 712	33 342	34 937

TABLE 4.1 – Statistiques des données de validation

Cet ensemble de données permet donc de calculer 40 BE différents, utilisés pour calculer la précision de notre modèle.

4.1.2 Résultats

Les RMSE, MAE (Mean Absolute Error) et erreurs minimales et maximales sont calculées sur les 40 BE de la base de données.

Il est important de garder en tête que la précision du modèle dépend directement des données utilisées et de l’adaptabilité souhaitée du modèle. Si l’utilisateur final souhaite pouvoir faire varier les taux sur une plage moins large que celle présentée, alors la calibration peut être effectuée sur des données plus proches du scénario central. Ceci augmente la précision du modèle. A l’inverse, augmenter la largeur de la plage d’entraînement augmente l’adaptabilité

Variable	RMSE	MAE	Min	Max
BE (k€)	160	130	-240	350
BE (%)	0.49	0.42	-0.76	1.09

TABLE 4.2 – Statistiques de prédiction sur les données de validation

du modèle au détriment de sa précision. Il faut donc garder à l’esprit que les chiffres donnés ne sont valables que toutes choses étant égales par ailleurs, et que de simples modifications de la plage d’utilisation du modèle pourraient radicalement changer les résultats.

Choc	Scénario 1	Scénario 2	Scénario 3	Scénario 4
1	-0.22	0.37	0.03	1.09
2	-0.74	0.20	-0.24	0.54
3	-0.69	0.26	-0.24	0.45
4	-0.71	0.15	-0.30	0.58
5	-0.72	0.20	-0.30	0.36
6	-0.63	0.18	-0.28	0.59
7	-0.76	0.09	-0.37	0.43
8	-0.19	0.56	0.09	1.02
9	-0.41	0.36	-0.15	0.84
10	-0.65	0.18	-0.30	0.58

TABLE 4.3 – Écarts relatifs de prédiction des BE (%)

La table 4.3 montre les écarts relatifs de prédiction des BE. Il est remarquable que des tendances de prédictions à la baisse ou à la hausse apparaissent sur les scénarios, indépendamment des chocs appiqués.

4.2 Scénario central

4.2.1 Description

Le scénario central est le scénario qui est déformé pour générer la base de données d’entraînement. Ce scénario est décrit par une courbe des taux, et des volatilités actions et immobilier. La courbe des taux du scénario central est fixée comme égale à la courbe des taux fournie par l’EIOPA au 31/08/2021. Cette courbe des taux n’est pas exactement la courbe des taux moyenne obtenue en prenant les moyennes des lois uniformes utilisées dans la section 2.4.7. En effet, les plages des lois uniformes utilisées pour générer les taux ont été sélectionnées en fonction des chocs Solvabilité 2 applicables, alors que les méthodes de calcul des chocs de taux à la hausse et à la baisse ne sont pas symétriques.

Les volatilités utilisées correspondent presque exactement aux moyennes des lois utilisées pour générer les scénarios :

- $\sigma^{Immo} = 10\%$
- $\sigma^{Eq} = 23\%$

4.2.2 Scénario déterministe

Le modèle proxy est capable de calculer un BE déterministe. En effet, il suffit de fournir un scénario risque-neutre déterministe en entrée pour obtenir une prédiction des cash flows déterministes en sortie.

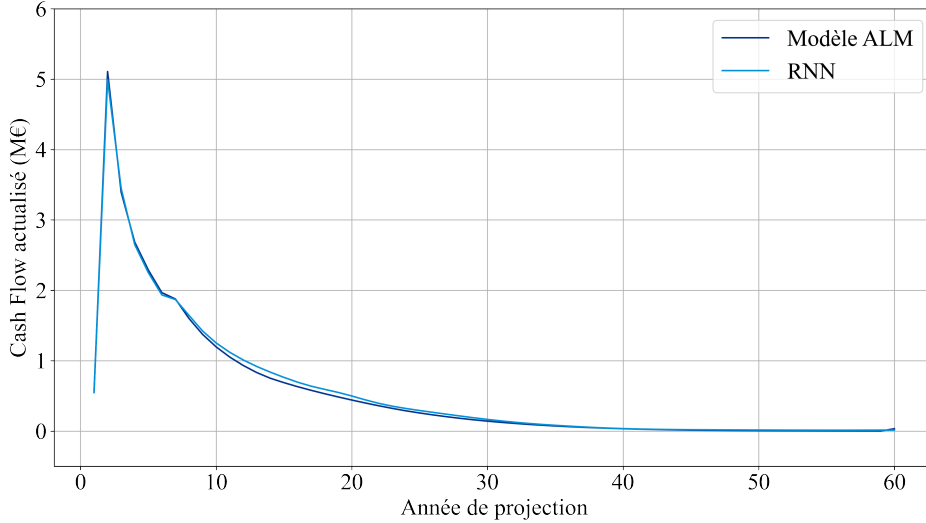


FIGURE 4.1 – Trajectoire déterministe du scénario central

On remarque que le RNN sur-estime légèrement les cash flows prédits à partir de 7 ans de projection. Cette sur-estimation mène à un écart de prédiction :

$$\widehat{BE} - BE = 1241\text{k€}$$

Cet écart correspond à 3.9% d'erreur relative de prédiction. **Il est important de noter que si le modèle créé est capable de donner une estimation du BE déterministe, il ne s'agit pas de sa vocation première.** En effet, si le modèle est sans biais, la prédiction d'un BE stochastique est plus précise que celle d'un BE déterministe. Plus précisément, si \hat{y}_t désigne le cash flow de passif prédit pour l'année de projection t , le BE déterministe est calculé comme :

$$BE^{det} = \sum_{t=1}^T \hat{y}_t$$

Alors que dans le cas d'un BE stochastique :

$$BE^{sto} = \frac{1}{N} \sum_{n=1}^N \sum_{t=1}^T \hat{y}_t^n$$

Si les $\sum_{t=1}^T \hat{y}_t^n$ sont des lois normales *i.i.d.* (chacun des N scénarios stochastiques générés sont créés à partir de mouvements browniens multidimensionnels indépendants et identiquement distribués) :

$$\text{Var}(BE^{sto}) = \frac{1}{N} \text{Var}(BE^{det})$$

Les erreurs des prédictions stochastiques sont donc de l'ordre de $\frac{1}{\sqrt{N}} \times$ Erreur déterministe. Sachant qu'ici $N = 1000$, il est possible d'attendre une erreur de l'ordre de $\frac{3.8\%}{\sqrt{1000}} \approx 0.1\%$.

4.2.3 Scénario stochastique

Un exemple de cash flows prédits sur un scénario stochastique est donné figure 4.2.

Le modèle proxy prédit un BE stochastique de 33 416 k€ alors que le BE stochastique calculé par le modèle ALM est de 33 424 k€. On a donc :

$$\widehat{BE} - BE = -8\text{k€}$$

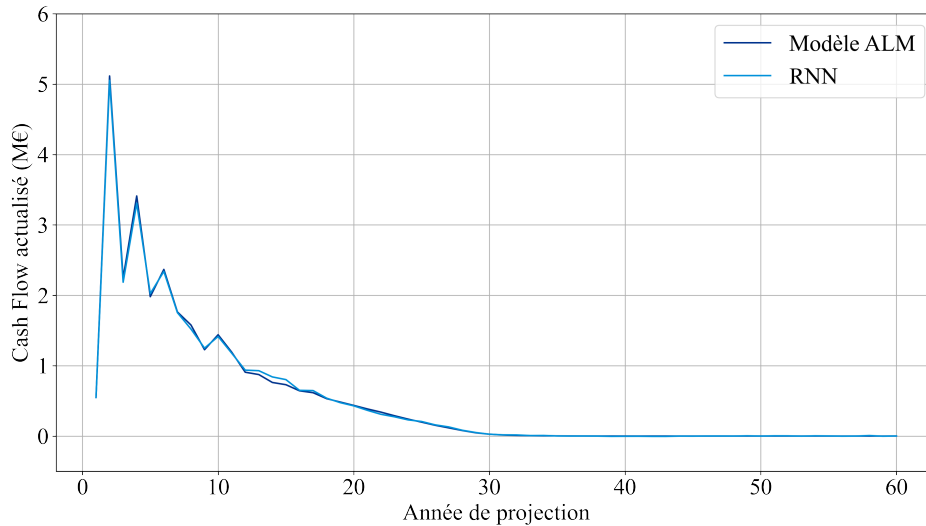


FIGURE 4.2 – Exemple de prédiction sur une trajectoire stochastique issue du scénario central

Cette erreur correspond à une erreur relative de -0.02%. Elle est obtenue alors que les courbes des taux utilisées pour générer la base de données d’entraînement ne présentent pas de creux ni de bosses. En effet, elles sont générées en n’utilisant que 3 points : r_1 , r_{30} et l’UFR. Ainsi même si une courbe de la base d’entraînement était générée avec $r_1 = -0.615\%$, $r_{30} = 0.905\%$ et UFR = 3.6%, *ie* les paramètres de la courbe de taux au 31/08/2021 (ce qui n’est pas le cas), elle serait différente de la courbe de taux interpolée par l’EIOPA.

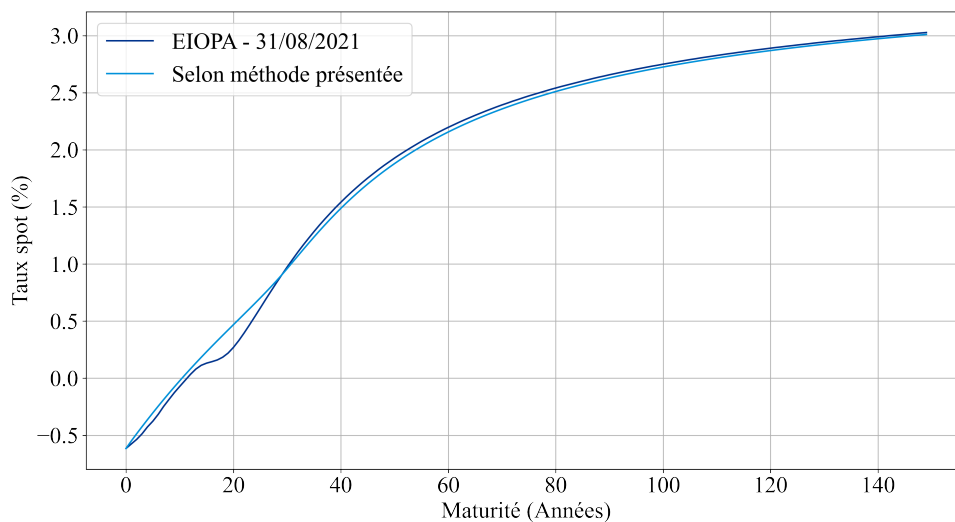


FIGURE 4.3 – Courbe de taux réelle de l’EIOPA et interpolée avec 3 points

Les résultats démontrent une certaine capacité du modèle à prendre en compte des déformations relativement complexes de la courbe des taux.

4.3 Courbe des taux

Afin de vérifier la robustesse du modèle vis-à-vis des courbes de taux, un ensemble de 21 courbes de taux translattées par rapport au scénario central est créé. Les scénarios générés sont proches du scénario central, à savoir :

- Volatilités :
 - Action : 23%
 - Immobilier : 10%
- Choc nul sur les facteurs de risques : $a_0^1 = (0, \dots, 0)$

Les courbes de taux utilisées sont générées en translattant les trois paramètres r_1 , r_{30} et l'UFR. Les 21 triplets définissant ces courbes sont obtenus pour $i = 0, \dots, 10$:

- $r_1 = -0.615\% \pm \frac{1\% \times i}{10}$
- $r_{30} = 0.905\% \pm \frac{1\% \times i}{10}$
- $\text{UFR} = 3.6\% \pm \frac{0.6\% \times i}{10}$

Il est à noter que les intervalles parcourus par r_1 et r_{30} sont de largeurs égales à 2 points de pourcentages (200 BP) alors que les intervalles utilisés pour créer la base de données n'ont qu'une largeur égale à 1.5 points de pourcentages. Cette largeur plus importante permet de tester les performances du modèle hors de sa zone d'entraînement, et donc de mesurer les capacités d'extrapolation du modèle.

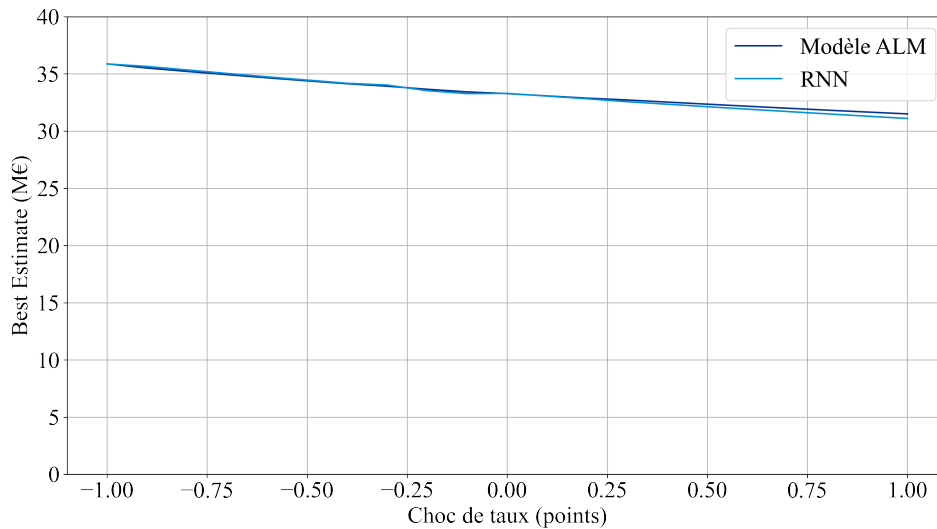


FIGURE 4.4 – Best Estimate en fonction de la translation appliquée sur r_1

La figure 4.4 représente les BE stochastiques prédits et calculés par le modèle ALM. L'abscisse correspond au terme $\pm \frac{1\% \times i}{10}$: c'est la translation appliquée aux taux r_1 ou r_{30} par rapport au scénario central. On observe d'abord que plus les taux augmentent, plus le Best Estimate diminue. En effet, les taux d'actualisation sont plus élevés, ce qui diminue la valeur du BE.

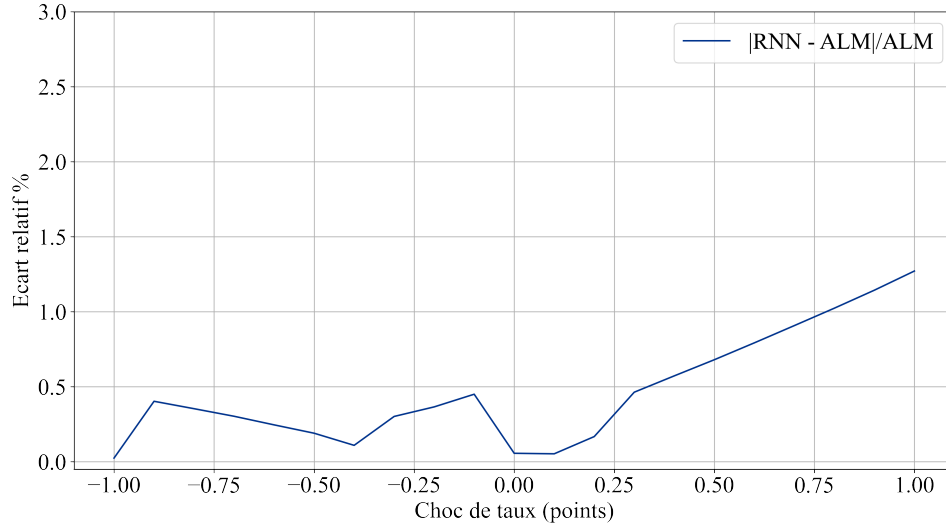


FIGURE 4.5 – Écarts relatifs en fonction de la translation appliquée sur r_1

On observe sur la figure 4.4 que, hors de l'intervalle $[-0, 20\%, -0, 05\%]$, les résultats prédits par le RNN sont proches de ceux donnés par le modèle ALM. La figure 4.5 permet de le confirmer : l'écart relatif $\frac{\widehat{BE} - BE}{BE}$ absolu est inférieur à 0.5%, hormis pour les chocs à la hausse supérieurs à 0.5 points de pourcentages. De manière générale, on remarque que l'erreur de prédiction augmente avec l'importance du choc. De plus, un choc de taux positif entraîne un écart de prédiction $\widehat{BE} - BE$ négatif, alors qu'un choc à la baisse entraîne un écart de prédiction positif. Si le modèle ne prenait pas en compte l'impact des taux d'intérêts sur les cash flows prédits, \widehat{BE} serait constant alors que BE est négativement corrélé au choc sur les taux (par effet d'actualisation). Les erreurs seraient croissantes avec la valeur du choc de taux. Puisque c'est l'effet inverse qui est observé, il est possible que le modèle sur-interprète les effets des courbes de taux sur le Best Estimate.

4.4 Volatilités

Afin de vérifier la robustesse du modèle vis-à-vis de la volatilité des scénarios économiques fournis en entrée, un ensemble de 11 scénarios initiaux est créé. Ces 11 scénarios utilisent la courbe des taux réelle présentée dans la figure 4.3. Aucun choc sur les facteurs de risque n'est appliqué, *i.e.* $a_1^0 = (0, \dots, 0)$. Les scénarios sont obtenus pour $i = 0, \dots, 11$:

- $\sigma^{Eq1} = 7\% + 3.5\% \times i$
- $\sigma^{Immo} = 1\% + 2\% \times i$
- $\sigma^{Eq2} = \sigma^{Eq1}$

Les résultats sont présentés dans la figure 4.6. Le BE augmente avec la volatilité des marchés actions et immobilier. Ceci est dû à la hausse des rachats conjoncturels. En effet, une plus grande volatilité des marchés implique une plus grande volatilité des résultats financiers de l'entreprise. Celle-ci entraîne à son tour une plus grande différence entre taux servis (dépendant des résultats financiers) et taux attendus (dont le plancher est le TMG moyen du portefeuille).

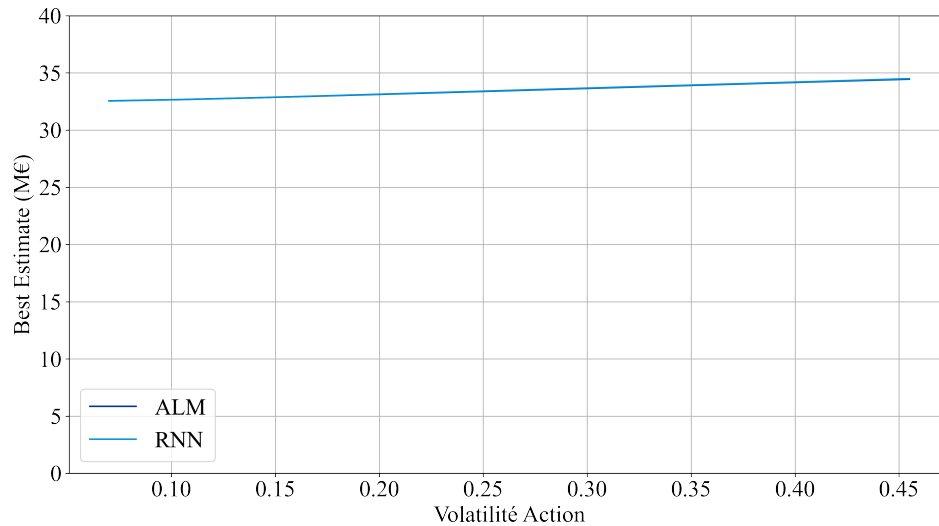


FIGURE 4.6 – Best Estimate en fonction de la volatilité action

Les volatilités testées sur les actions et l’immobilier sont plus étendues que celles utilisées pour calibrer le modèle. En effet, les volatilités maximales testées sont respectivement de 45.5% et 23% pour les actions et l’immobilier, alors que les volatilités maximales des lois uniformes utilisées pour l’entraînement étaient respectivement de 42% et 21 %. Contrairement aux chocs sur les taux, les volatilités n’ont pas d’impact sur la précision des résultats du modèle, comme le montre la figure 4.7.

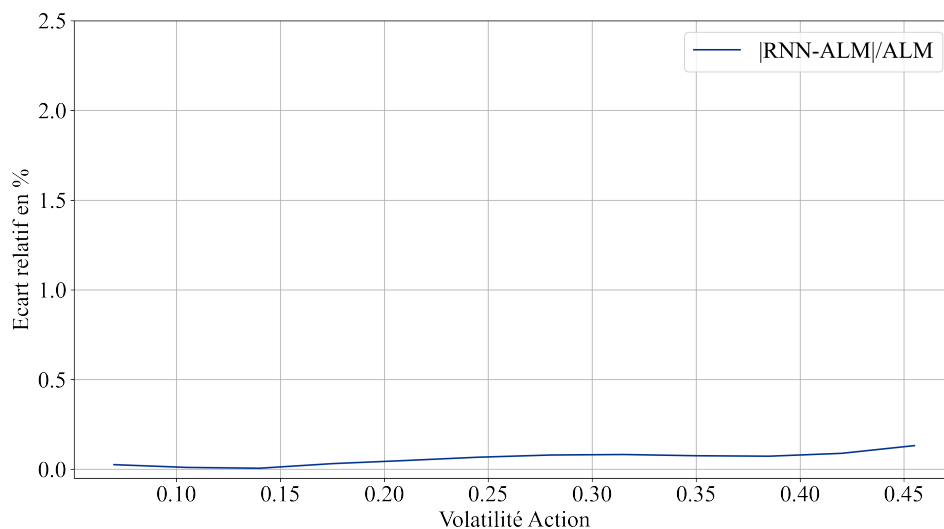


FIGURE 4.7 – Écart relatif en fonction de la volatilité action

L’erreur relative maximale est de **0.13%**, ce qui représente un écart maximal de 45k€. La MAE est de **0.06%**, ce qui représente un écart absolu moyen de **20 k€** sur un bilan de 33 000 k€ environ.

4.5 Chocs sur les facteurs de risque

Cette section s'intéresse à la robustesse du modèle vis-à-vis des chocs sur les facteurs de risque. Tous les tests sont effectués sur le scénario central, c'est-à-dire avec :

- des volatilités actions et immobilier respectivement égales à 23% et 10%,
- une courbe de taux initiale égale à la courbe des taux extrapolée par l'EIOPA au 31/08/2021.

Par soucis de lisibilité, les graphiques suivant "zooment" sur les BE calculés par le modèle ALM et le RNN. Ces zooms permettent de mieux entrevoir les différences de prédiction et les tendances à la hausse ou à la baisse des BE en fonction des chocs sur les facteurs de risque. Par conséquent, l'espacement apparent entre les courbes ALM et RNN semble plus important. Pour autant, la précision des prédictions reste la même.

4.5.1 Choc sur les rachats exceptionnels

La figure 4.8 montre que le BE diminue avec l'augmentation des rachats lors de la première année. En effet, même si des rachats exceptionnels positifs se traduisent par un cash flow de passif important lors de la première année, la taille du portefeuille s'en retrouve diminuée. Les cash flows étant proportionnels aux provisions mathématiques, tous les cash flows des années ultérieures sont réduits.

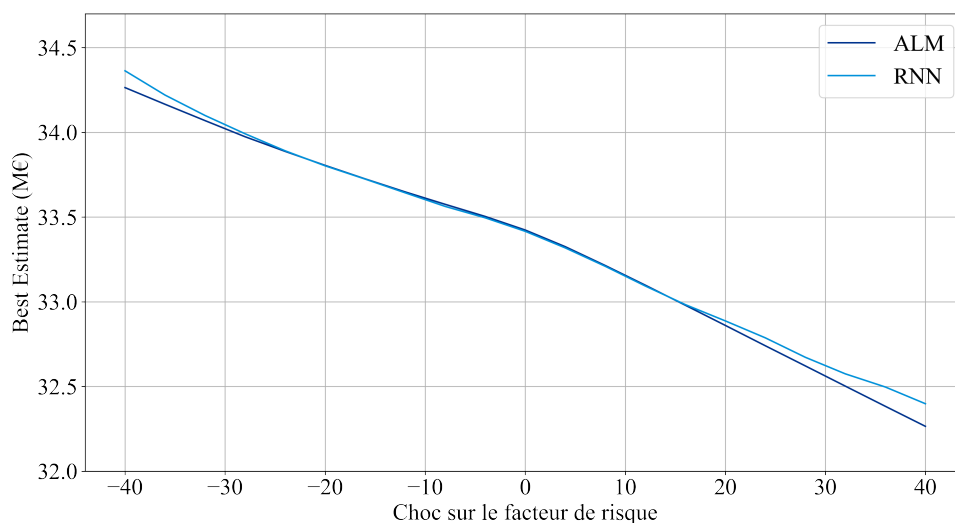


FIGURE 4.8 – Best Estimate en fonction des rachats exceptionnels

On remarque que l'erreur de prédiction est croissante lorsque le choc sur les rachats exceptionnels est important, qu'il soit à la hausse ou à la baisse. Les erreurs sont des surestimations du RNN, ce qui est prudent. L'**erreur maximale** est de **133 k€**, ce qui correspond à une **erreur relative** de **0.41%**.

Une première explication de cette erreur peut être que les chocs sont sélectionnés selon des lois normales. Ces lois étant des distributions à queues fines, les chocs extrêmes sont peu fréquents dans la base d'entraînement, ce qui justifie les difficultés de prédiction du modèle pour les chocs extrêmes. Une autre raison peut être la corrélation des lois normales utilisées pour générer la base de données. La base d'entraînement a été générée afin de maximiser la précision du modèle sur des chocs probables, *i.e.* dont les facteurs de risques ont pour matrice

de corrélation la matrice 2.10. Par conséquent, le vecteur de choc $(\delta, \epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_5, \epsilon_6)$ peut ne pas se trouver dans une zone à forte densité de points connus. Cette raison est valable pour l'ensemble des chocs étudiés dans cette section.

4.5.2 Choc sur les rachats

La figure 4.9 montre que le BE diminue aussi avec l'augmentation durable des rachats. Cette diminution est cependant moins marquée que pour les rachats exceptionnels. En effet, le choc sur les rachats exceptionnels est un taux de rachats directement appliqué alors que le choc sur les rachats (qui est appliqué à toutes les années de projection) est un taux d'augmentation ou de diminution du taux applicable sans choc (confère 2.6.3). Par conséquent, l'effet de l'augmentation ou de la diminution des rachats ne se fait sentir sur les cash flows qu'au bout de plusieurs années de projections. Les premiers cash flows sont calculés sur une provision mathématique proche de celle qui serait en vigueur sans choc : l'impact sur le BE est plus faible que pour les rachats exceptionnels.

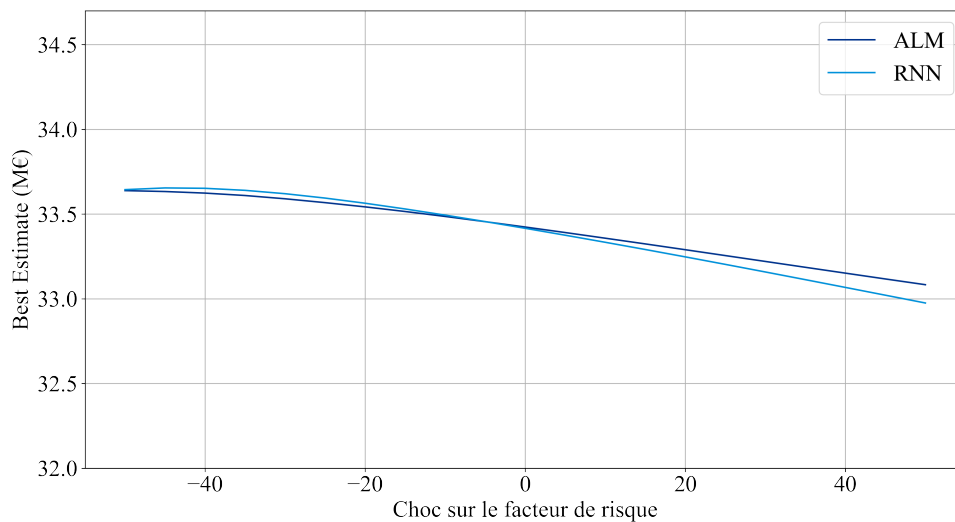


FIGURE 4.9 – Best Estimate en fonction du choc sur les rachats

L'erreur de prédiction sur les rachats augmente avec la hausse des rachats. L'erreur pour les chocs à la baisse sur les rachats est plus faible. Cette fois-ci, le modèle proxy sous-estime les montants de Best Estimate pour les chocs à la hausse. L'erreur maximale est de **108 k€**, soit une erreur relative de **0.41%**.

4.5.3 Choc sur le taux de mortalité

Au quantile de 99.5%, les risques de longévité et de mortalité ont moins d'impact sur le BE que les risques de rachats. Le Best Estimate est légèrement décroissant en fonction du choc sur la mortalité. Cependant, la variation de BE en fonction du choc est très dépendante de la composition du portefeuille passif de l'assureur : un portefeuille plus âgé sera plus impacté par les hausses ou baisses de taux de mortalité qu'un portefeuille jeune (pour lequel les probabilités de décès sont très faibles).

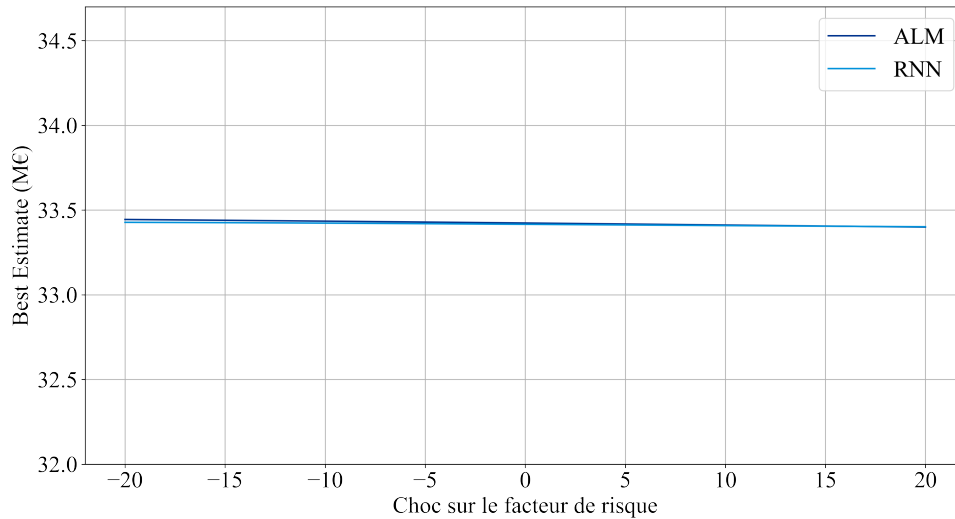


FIGURE 4.10 – Best Estimate en fonction du choc sur la mortalité

L'erreur de prédiction maximale est de 17 k€, ce qui représente un écart de prédiction maximal de 0.05%.

4.5.4 Choc sur les frais

L'augmentation des frais est naturellement liée à une augmentation du Best Estimate, comme le montre la figure 4.11.

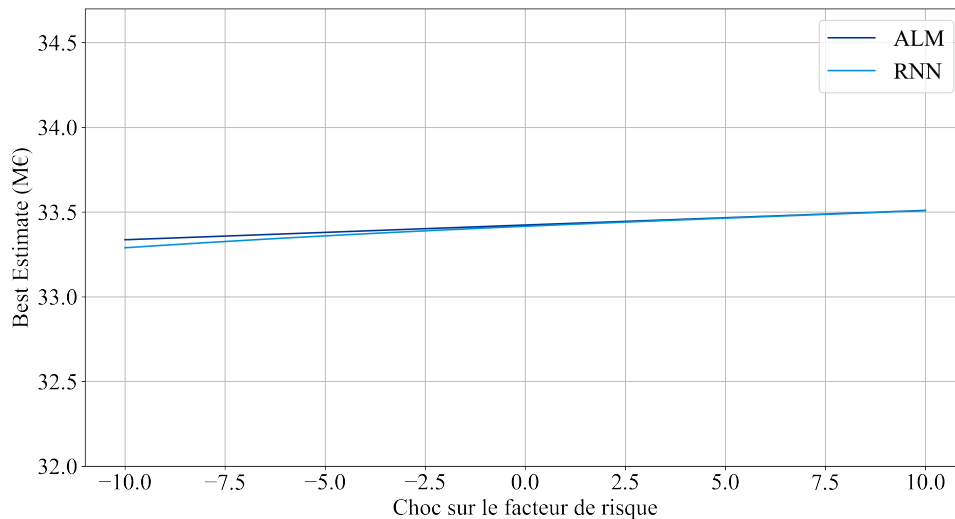


FIGURE 4.11 – Best Estimate en fonction du choc sur les frais

L'erreur de prédiction est croissante avec la baisse des frais. L'erreur maximale est de 47 k€, ce qui correspond à une erreur relative de 0.14%.

4.5.5 Choc sur les valeurs de marché action

La figure 4.12 montre clairement une forte augmentation du BE avec l'augmentation de la valeur de marché des actions en portefeuille. La taille du BE diminue donc lorsque les valeurs des actions sont choquées négativement. Cette diminution ne se traduit pas par une augmentation des fonds propres économiques de l'assureur, puisque un choc sur les valeurs de marché entraîne aussi une chute de la taille du bilan.

Par exemple, une chute de 20% des valeurs de marché des actions représente une réduction de 690 k€ de la taille du bilan. Le BE diminue quant à lui d'environ 250 k€. Par conséquent, en supposant la variation de Risk Margin nulle, la chute de 20% des valeurs de marché des actions se traduit bien par une réduction de 440 k€ des fonds propres.

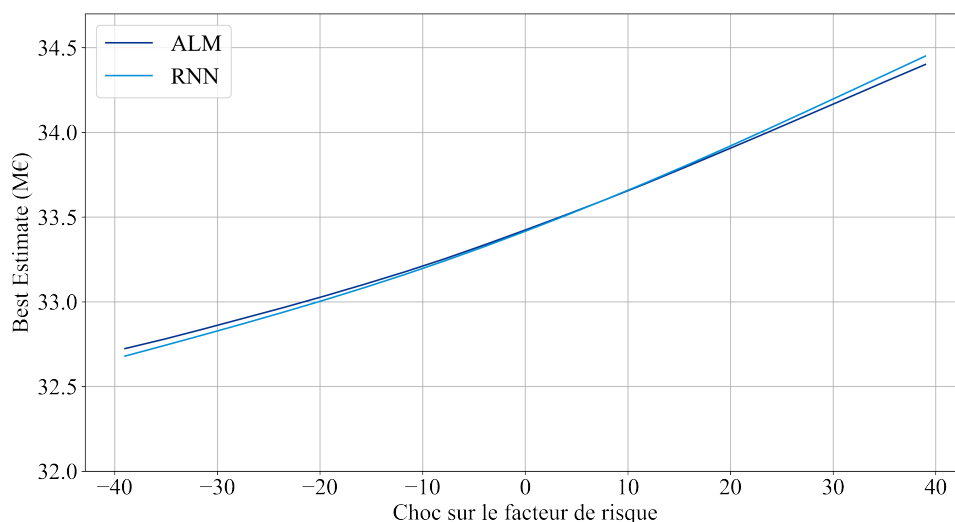


FIGURE 4.12 – Best Estimate en fonction du choc sur les actions

L'erreur maximale de prédiction du modèle RNN vis-à-vis des résultats du modèle ALM est de **26 k€**, soit **0.08%** d'erreur relative.

4.5.6 Choc sur l'immobilier

Le phénomène observé est similaire à celui des actions. La seule différence semble être la pente de la courbe, moins importante pour l'immobilier. Cependant, la courbe pour l'immobilier est tracée sur la plage $[-25\%, 25\%]$ alors que la courbe action est tracée sur $[-39\%, 39\%]$.

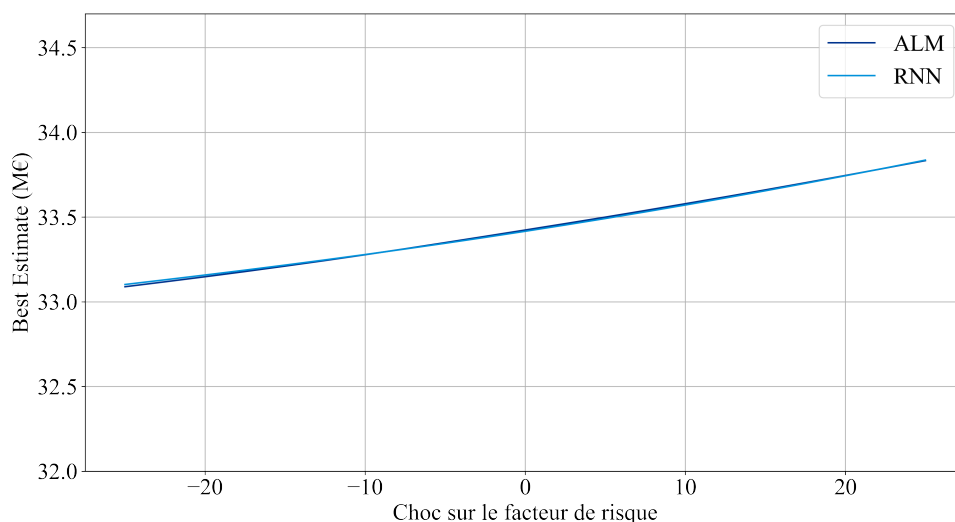


FIGURE 4.13 – Best Estimate en fonction du choc sur l’immobilier

La sensibilité moyenne du BE par rapport aux valeurs de marché actions est approximativement de 26.5k€ par point de pourcentages. La sensibilité moyenne du BE par rapport aux valeurs de marché de l’immobilier est de 16.8k€ par point de pourcentages. Cette différence s’explique par les parts plus ou moins importantes qu’occupent les actions et l’immobilier dans le portefeuille. À parts égales dans le portefeuille, le BE aurait des sensibilités égales aux actions et à l’immobilier.

Sachant que la valeur de marché des actions est de 3 438 k€ et que la valeur de marché des actifs immobiliers est de 2 281 k€, on trouve :

$$\frac{16.8}{2281} = 0.007 = \frac{26.5}{3438}$$

Cette simple vérification confirme bien qu’à parts égales dans le portefeuille, le BE aurait les mêmes sensibilités par rapport aux actions et à l’immobilier.

L’**erreur maximale** de prédiction du modèle RNN vis-à-vis des résultats du modèle ALM est de **13 k€**, soit **0.04%** d’erreur relative.

Chapitre 5

Applications

5.1 Calcul du SCR

Cette section se focalise sur le calcul du SCR via la formule standard. Celle-ci privilégie une approche modulaire. La compagnie d'assurance ou de réassurance calcule les pertes en capital associées à une trentaine de facteurs de risque différents, chacun associé à un sous-module. Ces sous-modules sont regroupés en 6 modules : Marché, Santé, Défaut, Vie, Non-vie, Intangible. Ces 6 modules permettent de calculer le SCR de base (BSCR), auquel on ajoute les montants en capitaux associés au risque opérationnel, ainsi que les effets d'atténuation des pertes brutes. Les atténuations correspondent principalement à la capacité qu'a l'assureur à transférer une partie de ses pertes aux assurés, ainsi qu'à l'économie d'impôts effectuée en cas de baisse du résultat fiscal.

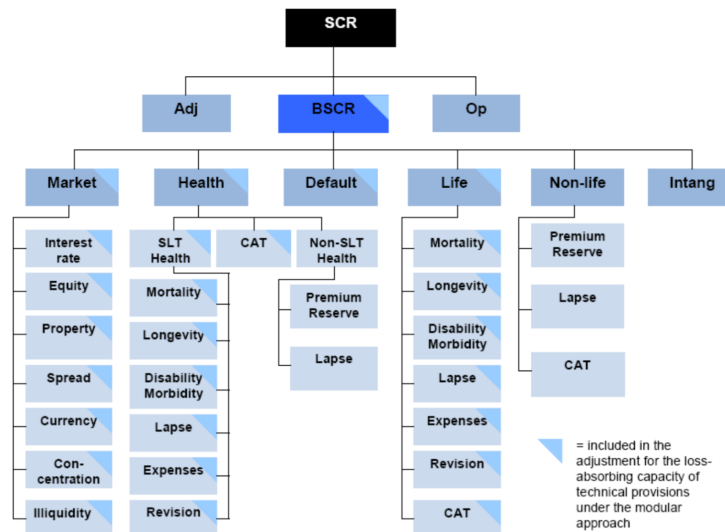


FIGURE 5.1 – Calcul du SCR sous la formule standard

Les agrégations des sous-modules en modules et des modules en BSCR sont faites via des matrices de corrélation. Ces matrices de corrélation permettent la prise en compte de la diversification des risques, qui réduit *in fine* le montant SCR éligible.

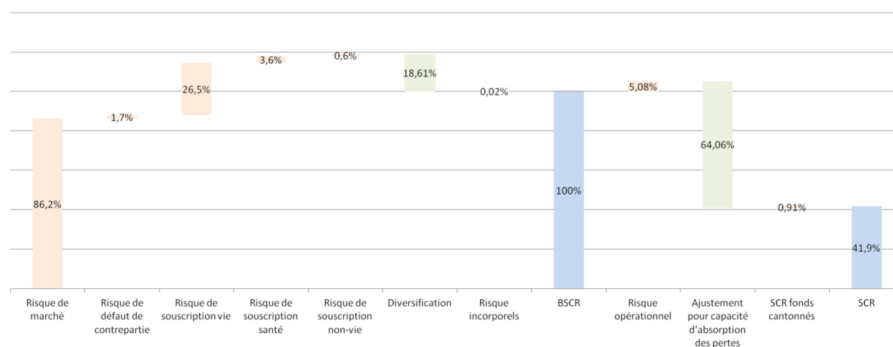


FIGURE 5.2 – Poids des différents sous-modules et modules dans le SCR final d’une compagnie d’assurance vie

La figure 5.2 présente l’impact des différents modules et les effets de la diversification sur le SCR, dans le cas d’une compagnie assurance-vie mixte "type". Ces résultats sont ceux présentés par l’ACPR dans la synthèse des résultats de l’étude d’impact QIS5 [10]. Les modules "risque de marché" et "risque de souscription-vie" (noms utilisés dans la réglementation de 2015/35 pour les modules "Marché" et "Vie") sont les deux principaux modules permettant de calculer le montant du BSCR. Ils représentent respectivement 86,2% et 26,5% (avant diversification) du BSCR, alors que la somme des 4 autres modules ne représente que 5,9%.

Dans notre étude, les montants associés aux risques de souscription non-vie et santé sont nuls, puisque le portefeuille est uniquement composé de fonds euros. Ce chapitre se focalise donc sur les calculs des modules "Marché" et "Vie". Les calculs des montants SCR sont effectués sur le scénario central, correspondant à la courbe des taux et les volatilités en vigueur au moment de l’étude.

5.1.1 Module "risque de marché"

Le module «risque de marché» est constitué des sous-modules «risque de taux d’intérêt», «risque sur actions», «risque sur actifs immobiliers», «risque lié à la marge» (*spread*), «risque de change» et «concentration du risque de marché».

L’entreprise d’assurance étudiée n’est ni exposée au risque de change, ni au risque de spread (car elle ne possède que des obligations sans risque). De plus, on suppose que le risque de concentration est nul (actifs suffisamment diversifiés). Ainsi, seuls les SCR actions, immobilier et taux d’intérêts sont pris en compte dans cette étude. Cependant, le module de taux d’intérêts n’a pas été implémenté dans le modèle. Le module de risque de marché se résume donc aux calculs des SCR actions et immobilier.

Risque sur actions

L’exigence de capital pour le sous-module "risque sur actions" correspond à la perte en fonds propres associée à une baisse soudaine des valeurs de marché des actions en portefeuille. Le pourcentage de baisse appliqué dépend du type d’action, de la durée d’investissement et de l’organisme financé. Il est ici considéré que toutes les actions en portefeuille entrent dans la catégorie choquée à 39%.

Les montants calculés pour le risque sur actions sont de **640 k€** pour le modèle ALM et de **604 k€** pour le modèle proxy.

Risque sur actifs immobiliers

L'exigence de capital pour le sous-module "risque sur actifs immobiliers" correspond à la perte en fonds propres associée à une baisse soudaine de 25% des valeurs actifs immobiliers.

Les montants calculés pour le risque sur les actifs immobiliers sont de **235 k€** pour le modèle ALM et de **256 k€** pour le modèle proxy.

5.1.2 Module "risque de souscription en vie"

Le module "Vie" ou "risque de souscription en vie" regroupe tous les risques liés aux activités de souscription d'assurance-vie. Ce module regroupe les sous-modules "risque de mortalité", «risque de longévité», «risque d'invalidité — de morbidité», «risque de dépenses en vie», «risque de révision», "risque de cessation», et «risque de catastrophe en vie».

Risque de cessation

L'exigence de capital pour risque de cessation (que l'on appelle aussi risque de rachat) est égale à la plus élevée des exigences de capital en cas de :

1. hausse permanente de 50% des rachats,
2. baisse permanente de 50% des rachats,
3. rachat massif de 40%.

On a donc :

$$SCR_{lapse} = \max(SCR_{up}, SCR_{down}, SCR_{Massive})$$

Les montants calculés sont donnés dans la table 5.1

	Modèle ALM	Proxy RNN
Hausse	-341.0	-441.0
Baisse	214.0	228.0
Massif	-1159.0	-1017.0

TABLE 5.1 – Montants calculés et prédits de pertes en fond propres (k€)

Les montants retenus correspondent donc aux choc à la baisse des rachats, pour un SCR calculé par le modèle ALM de **214 k€**, et estimé par le modèle ALM à **228€**.

Risque de Longévité

L'exigence de capital pour le sous-module "risque de longévité" correspond à la perte en fonds propres associée à une baisse soudaine et durable de 20% des taux de mortalité utilisés pour calculer les provisions techniques.

Les montants calculés pour le risque de longévité sont de -25 k€ pour le modèle ALM et de -14 k€ pour le modèle proxy. Ils sont donc **tous les deux nuls**, puisque les montants des SCR sont positifs.

Risque de Mortalité

L'exigence de capital pour le sous-module "risque de mortalité" correspond à la perte en fonds propres associée à une hausse soudaine et durable de 15% des taux de mortalité utilisés pour calculer les provisions techniques.

Les montants calculés pour le risque de mortalité sont de **16 k€** pour le modèle ALM et de **10 k€** pour le modèle proxy.

Risque de dépenses en vie

L'exigence de capital pour le sous-module "risque de dépenses en vie" correspond à la perte de fonds propres associée à la hausse soudaine de 10% des montant de dépenses conjuguée à la hausse d'un point de pourcentages du taux d'inflation des dépenses.

Les montants calculés pour le risque de dépenses sont de **86 k€** pour le modèle ALM et de **103 k €** pour le modèle proxy.

5.1.3 BSCR

Le montant SCR associé au module de marché après agrégation avec la matrice de corrélation est de 831 k€ pour le modèle proxy et de 814 k€ pour le modèle ALM. De même, le montant SRC associé au module de souscription vie après agrégation des différents sous-modules est de 269 k€ pour le modèle ALM et de 294 k€ pour le modèle proxy.

Après agrégation des modules marché et souscription vie, le modèle ALM calcule un montant BSCR de **935 k€** alors que le modèle proxy prédit un montant de **932 k€**, soit une **erreur de 3k€** correspondant à un **écart de 0,3%**.

5.2 ORSA

5.2.1 Présentation

L'évaluation interne des risques et de la solvabilité (*Own Risk and Solvency Assesment* - *ORSA*) est une notion introduite par le deuxième pilier de la directive Solvabilité II. Il a pour objectif la mise en place d'un processus de suivi régulier des risques auxquels est soumise l'entreprise d'assurance ou de réassurance. Ce suivi doit permettre une conformité continue avec les exigences de capital (SCR). Pour ce faire, le processus ORSA peut être subdivisé en deux composantes :

- l'ORSA permanent,
- l'ORSA prospectif.

L'ORSA permanent est le processus permettant de vérifier le respect constant des exigences réglementaires. L'ORSA prospectif est le processus permettant de s'assurer que l'entreprise sera toujours capable de respecter ces exigences sur le moyen ou long terme. Elle inclut donc des projections pluri-annuelles.

Les entreprises sont invitées à élaborer leurs "propres processus [ORSA] avec des techniques appropriées et adéquates". L'utilisation de modèles proxys s'avère particulièrement intéressante, étant donné les gains temporels et opérationnels qu'ils permettent.

Avec un processus calibré, il suffit de 24h pour obtenir un modèle capable d'effectuer toutes sortes d'études prospectives de manière instantanée. Le modèle permet alors d'effectuer des recherches de scénarios mettant à mal l'entreprise.

Le nouveau type de modèle proxy introduit dans ce mémoire a pour but de permettre des études prospectives plus précises que les modèles proxys précédemment utilisés, puisque prenant en compte le vieillissement des portefeuilles et la trajectoire des marchés entre l'instant initial et l'instant de valorisation du Best Estimate.

5.2.2 Description du scénario

Il est proposé d'étudier un scénario ORSA prospectif de hausse des taux couplée à une chute de l'immobilier. Les rendements et taux zéro-coupons par année de projection sont donnés

dans la table 5.2. Ce scénario est un des scénarios envisageables, étant donné les pressions inflationnistes au sortir de la crise de la Covid. En effet, une hausse des taux est censée permettre la réduction de la masse monétaire en circulation via la réduction des crédits souscrits par les ménages. Cette hausse des taux d'intérêts pourrait provoquer une réduction des souscriptions de crédits immobiliers (et/ou un regain d'intérêt pour les bons du trésor et obligations privées). Cette réduction entraînerait une baisse de la demande pour les biens immobiliers.

Année	2021	2022	2023	2024	2025
Actions 1	NA	-2,2%	1,1%	2,1%	1,1%
Actions 2	NA	0,1%	1,0%	2,4%	1,0%
Immobilier	NA	-1,2%	-1,1%	0,1%	-2,0%
Taux 1 an	-0,62%	-0,54%	-0,46%	-0,34%	-0,19%
Taux 30 ans	0,91%	1,03%	1,15%	1,27%	1,39%

TABLE 5.2 – Rendements et taux du scénario *High Yield - Real Estate*

5.2.3 Résultats

La figure 5.3 montre les résultats obtenus pour ce scénario. La courbe "Central" correspond au BE prédit pour un scénario déterministe. Les résultats du RNN restent proches des résultats du modèle ALM, même après plusieurs années de projection, et ce sur un scénario assez éloigné du scénario central.

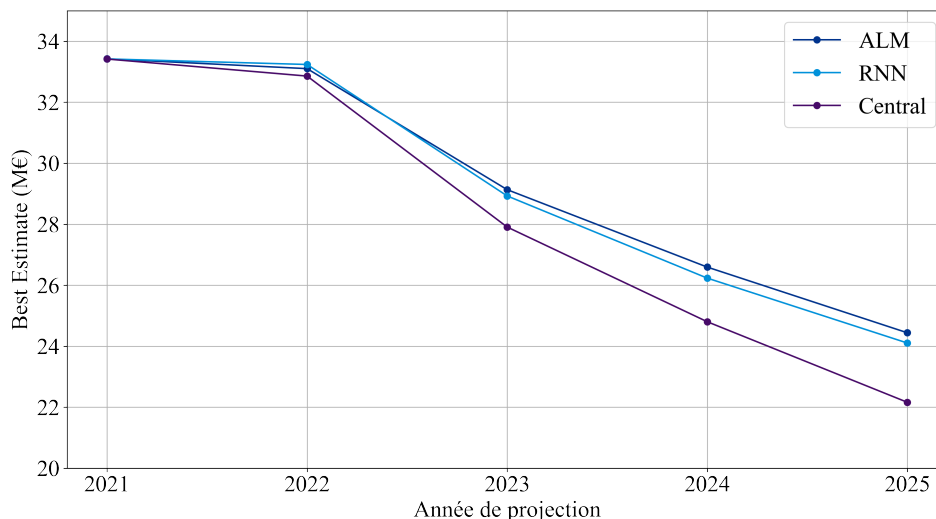


FIGURE 5.3 – BE en fonction du temps.

L'erreur maximale est de 2%, et est obtenue en fin de projection. Compte-tenu de l'incertitude vis-à-vis des paramètres de la projection (déroulement du scénario sur 5 ans, rentabilité, taux, actions, portefeuille de contrats, environnement juridique...), cette erreur n'est plus un frein à l'utilisation du modèle. L'objectif de prise en compte du vieillissement et de la trajectoire économique via l'état caché est donc *a priori* rempli.

5.2.4 Limites

Toutefois, plusieurs limites subsistent. Si l'erreur de prédiction de 2% au bout de 4 ans d'ORSA ne représente pas un frein à l'utilisation du modèle, elle pourrait être corrigée. Cette erreur est probablement due à la risque-neutralisation des obligations, qui est systématiquement appliquée par le modèle proxy. Celle-ci fausse les rendements des obligations en portefeuille lors des premières projections de flux, effectuées sous probabilité historique (monde réel). Le rendement plus faible des obligations en portefeuille provoque des rachats plus élevés qui se traduisent naturellement par une sous-estimation du Best Estimate sur les projections à horizon lointain.

De plus, s'il permet la prise en compte du vieillissement des model points, le modèle n'est calibré que pour effectuer des prédictions sans affaires nouvelles. Hors, l'ORSA prospectif inclut des affaires nouvelles (uniquement pour les premiers pas de temps sous probabilité historique, *ie* avant la valorisation du BE sous probabilité risque-neutre).

Enfin, de part sa structure, il est impossible de calculer un SCR à un instant ultérieur à l'instant initial, sans effectuer d'approximation. En effet, les chocs à appliquer pour calculer le SCR sont fournis dans la variable a_1^0 , qui par définition, modifie la projection à partir du premier cash flow.

Il pourrait être intéressant d'essayer de corriger ces problèmes, en entraînant un modèle similaire sur des scénario contenant des projections historiques puis risque-neutres de longueurs variables. Le modèle pourrait recevoir pour chaque pas de temps t un indicateur permettant de lui notifier si le cash flow de la période est à calculer :

- avec ou sans risque-neutralisation,
- avec ou sans affaires nouvelles.

Ces paramètres permettraient alors de créer un modèle plus adapté à l'ORSA prospectif.

Conclusion

Le contexte économique et financier incertain dans lequel évoluent désormais les compagnies d'assurance vie menace grandement leur solvabilité. En introduisant l'ORSA, la réglementation Solvabilité II définit un processus permettant de garantir le respect permanent des exigences en capital qu'elle introduit. Cependant, l'utilisation de méthodes stochastiques pour l'ORSA est très chronophage. L'ORSA prospectif, qui consiste en l'évaluation de la solvabilité de l'entreprise dans le futur, selon un scénario économique, démographique et financier donné, est difficile à mettre en pratique : il requiert le lancement d'une multitude de simulations, prenant chacune plusieurs heures.

L'utilisation de modèles proxys permet de contourner le problème du temps de calcul des méthodes stochastiques. Leurs résultats pour la prédiction des indicateurs financiers à court terme permettent leur mise en pratique. Cependant, leur difficulté à prendre en compte le vieillissement des model points limite leurs applications dans le cadre de l'ORSA prospectif.

Ce mémoire propose donc l'utilisation d'une nouvelle architecture de modèle proxy, basée sur les réseaux de neurones récurrents. Au lieu de directement prédire les indicateurs économiques du bilan, il propose de prédire les flux de passifs du bilan.

Les résultats obtenus sont très encourageants. L'outil permet un calcul très précis des cash flows et donc du Best Estimate final. Si seule la sensibilité aux taux d'intérêts présente des imprécisions, le modèle permet toutefois un calcul précis du SCR de base de la compagnie d'assurance-vie. De plus, il remplit sa promesse en délivrant de bons résultats pour l'ORSA prospectif.

Toutefois le modèle développé présente des limites : il ne permet ni la désactivation de la risque-neutralisation, ni l'ajout d'affaires nouvelles. Il pourrait être intéressant de développer ces fonctionnalités dans un modèle similaire, afin de le rendre pleinement fonctionnel.

Enfin, les complications techniques induites par la calibration et l'utilisation des modèles proxys en limitent pour l'instant l'utilisation. La mise en place de modèles proxys nécessite la création de processus adéquats pour la génération de la base de données, la calibration du modèle et l'obtention des résultats. Sans eux, les gains opérationnels obtenus grâce à l'instantanéité des résultats sont perdus dans le temps passé à générer les modèles. Si des solutions existent, seulement quelques assureurs y ont recourt. Finalement, leur utilisation dépendra de la démocratisation des techniques de mise en production des modèles de Machine Learning.

Bibliographie

- [1] Ian BERDAH. *Calibration de modèles d'apprentissage sur des modèles épargne*. 2020.
- [2] Vincent CESCUTTI. *Estimés de solvabilité par méta-modélisation*. 2016.
- [3] Kyunghyun CHO et al. *Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation*. 2014.
- [4] Aubrey CLAYTON. *Deep-learning the Cash Flow Model*. Juin 2020.
- [5] Sepp HOCHREITER et Jürgen SCHMIDHUBER. *Long Short-Term Memory*. 1997.
- [6] Diederick P. KINGMA et Jimmy L. BA. *Adam : A Method for Stochastic Optimization*. 2014.
- [7] Stéphane LOISEL, Pierrick PIETTE et Jason TSAI. *Applying economic measures to lapse risk management with machine learning approaches*. 2021.
- [8] *Orientations Nationales Complémentaires aux Spécifications Techniques pour l'exercice 2013 de préparation à Solvabilité II*. ACPR, 2012. URL : https://acpr.banque-france.fr/sites/default/files/20130527-onc-2013_0.pdf.
- [9] *Règlement délégué (UE) 2015/35*. Commission Européenne. Jan. 2015. URL : <https://acpr.banque-france.fr/europe-et-international/assurances/reglementation-europeenne/solvabilite-ii>.
- [10] *Solvabilité 2 : principaux enseignements de la cinquième étude quantitative d'impact (QIS5)*. ACPR, mar. 2011.
- [11] *Solvabilité II*. ACPR. URL : <https://acpr.banque-france.fr/europe-et-international/assurances/reglementation-europeenne/solvabilite-ii>. (Accès : 06/07/2021).
- [12] Igor TODER. *Produits de taux d'intérêts : présentation et principes de valorisation*. Nov. 2020.
- [13] Eve ZURFLUH. *Utilisation du Machine Learning dans l'estimation du ratio de solvabilité d'un assureur vie et application au Reverse Stress Tests*. 2019.

Table des figures

1.1	Piliers de Solvabilité II	19
1.2	Bilan économique sous Solvabilité II	20
1.3	Schéma de la méthode SdS	24
1.4	Schéma de la méthode pour des projections ORSA sur plusieurs années	25
1.5	Schéma de la méthode Least Squares Monte-carlo	26
1.6	Schéma de la méthode Curve Fitting	27
1.7	Fonctionnement des modèles proxys existants	27
1.8	Fonctionnement des modèles basés sur des réseaux de neurones récurrents	29
2.1	Fichiers fournis en entrée du logiciel de projection	32
2.2	Entrées et sorties du GSE	32
2.3	Notes des obligations sans risque	34
2.4	Types de contrats dans l’outil de modélisation	35
2.5	Répartition des années de naissance des assurés	36
2.6	Répartition des années de souscription des contrats	36
2.7	Répartition des TMG	37
2.8	Courbe des taux swaps d’inflation	38
2.9	Courbe de taux sans risque au 31/08/2021 et plages des déformations	43
2.10	Risque neutralisation des obligations sur le nominal	44
2.11	Courbe des rachats structurels	46
2.12	Courbe des rachats conjoncturels	47
3.1	Sens de propagation de l’information en fonction du type de réseau	54
3.2	Structure de réseau de neurones récurrent	54
3.3	Structure d’un neurone dans une couche RNN simple	55
3.4	Représentation de la propagation et backpropagation dans un RNN	56
3.5	Architecture de notre réseau	59
3.6	Détection de l’overfitting sur les courbes d’apprentissage	60
3.7	Exemple d’affichage des informations d’entraînement pour 5 epochs	63
3.8	Courbes d’apprentissage des modèles GRU	63
3.9	Écarts relatifs de prédiction du BE en fonction du choc de taux	64
3.10	Cash flows obtenus pour des scénarios stochastiques avec chocs de taux	64
3.11	Écarts relatifs de prédiction avec 5 taux en entrées	65
3.12	Cash flows obtenus pour des scénarios stochastiques avec chocs de taux	65
4.1	Trajectoire déterministe du scénario central	71
4.2	Exemple de prédiction sur une trajectoire stochastique issue du scénario central	72
4.3	Courbe de taux réelle de l’EIOPA et interpolée avec 3 points	72
4.4	Best Estimate en fonction de la translation appliquée sur r_1	73
4.5	Écarts relatifs en fonction de la translation appliquée sur r_1	74
4.6	Best Estimate en fonction de la volatilité action	75

4.7	Écart relatif en fonction de la volatilité action	75
4.8	Best Estimate en fonction des rachats exceptionnels	76
4.9	Best Estimate en fonction du choc sur les rachats	77
4.10	Best Estimate en fonction du choc sur la mortalité	78
4.11	Best Estimate en fonction du choc sur les frais	78
4.12	Best Estimate en fonction du choc sur les actions	79
4.13	Best Estimate en fonction du choc sur l'immobilier	80
5.1	Calcul du SCR sous la formule standard	81
5.2	Poids des différents sous-modules et modules dans le SCR final d'une compagnie d'assurance vie	82
5.3	BE en fonction du temps.	85

Liste des tableaux

2.1	Actifs de l'entreprise	32
2.2	Matrice de corrélation fournie au GSE	37
2.3	Prix des Caps renseignés au GSE	39
2.4	Historique des valeurs de l'UFR	42
2.5	Taux utilisés pour le calcul du SCR du module de taux	42
2.6	Allocation des actifs	44
2.7	Composition du portefeuille concurrent	45
2.8	Planchers et plafonds des paramètres sous QIS5	47
2.9	Valeurs des quantiles à 99,5% et écart-types associés	50
2.10	Matrice de corrélation des chocs	50
3.1	Comparaison des erreurs des différentes cellules	62
3.2	Écarts relatifs de prédiction des BE en %	66
3.3	Écarts relatifs de prédiction des BE en %	67
3.4	Statistiques de prédiction	68
4.1	Statistiques des données de validation	69
4.2	Statistiques de prédiction sur les données de validation	70
4.3	Écarts relatifs de prédiction des BE (%)	70
5.1	Montants calculés et prédits de pertes en fond propres (k€)	83
5.2	Rendements et taux du scénario <i>High Yield - Real Estate</i>	85

Annexe A

Libor Market Model

Rappelons les notations :

- $L(t, T, T + \theta)$ le taux forward proposé à l'instant t pour la date T de durée θ
- $B(t, T) = \mathbb{E}^{\mathbb{Q}} \left[e^{\int_t^T -r_s ds} \middle| \mathcal{F}_t \right]$ la valeur du zéro-coupon de maturité T vu à la date t

Notons $V_{caplet_i}(t)$ la valeur du caplet i dont le pay-off en T_{i+1} est $(L(T_i, T_i, T_{i+1}) - K)^+(T_{i+1} - T_i)$. Alors, en l'absence d'opportunité d'arbitrage, la valeur de ce caplet est :

$$\begin{aligned}
 V_{caplet_i}(t) &= \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{T_{i+1}} r_s ds} (L(T_i, T_i, T_{i+1}) - K)^+(T_{i+1} - T_i) \middle| \mathcal{F}_t \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[\mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{T_{i+1}} r_s ds} (L(T_i, T_i, T_{i+1}) - K)^+(T_{i+1} - T_i) \middle| \mathcal{F}_{T_i} \right] \middle| \mathcal{F}_t \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[\mathbb{E}^{\mathbb{Q}} \left[e^{-\int_{T_i}^{T_{i+1}} r_s ds} (L(T_i, T_i, T_{i+1}) - K)^+(T_{i+1} - T_i) \middle| \mathcal{F}_{T_i} \right] e^{-\int_t^{T_i} r_s ds} \middle| \mathcal{F}_t \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[\mathbb{E}^{\mathbb{Q}} \left[e^{-\int_{T_i}^{T_{i+1}} r_s ds} \middle| \mathcal{F}_{T_i} \right] e^{-\int_t^{T_i} r_s ds} (L(T_i, T_i, T_{i+1}) - K)^+(T_{i+1} - T_i) \middle| \mathcal{F}_t \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[B(T_i, T_{i+1}) e^{-\int_t^{T_i} r_s ds} (L(T_i, T_i, T_{i+1}) - K)^+ \middle| \mathcal{F}_t \right] (T_{i+1} - T_i)
 \end{aligned}$$

Pour continuer le calcul, il est nécessaire d'effectuer un changement de numéraire. Rappelons la formule de Bayes généralisée :

Théorème 1. Soit \mathbb{P} et \mathbb{Q} deux probabilités équivalentes. Soit $\frac{d\mathbb{Q}}{d\mathbb{P}} \big|_{\mathcal{F}_t} = Z_t$, processus martingale sous \mathbb{P} . $Z_t > 0$, $\mathbb{E}[Z_T] = 1$ puisque \mathbb{P} et \mathbb{Q} sont équivalentes (Radon Nikodyn). Alors, si l'on considère un processus X \mathcal{F} -adapté, la formule de Bayes généralisée s'écrit :

$$\mathbb{E}^{\mathbb{Q}} [X_T | \mathcal{F}_t] = \frac{1}{Z_t} \mathbb{E}^{\mathbb{P}} [Z_T X_T | \mathcal{F}_t]$$

qui peut se réécrire :

$$\mathbb{E}^{\mathbb{Q}} [X_T | \mathcal{F}_t] Z_t = \mathbb{E}^{\mathbb{P}} [Z_T X_T | \mathcal{F}_t]$$

On peut appliquer la formule de Bayes et changer de numéraire ($T_i = T$) :

$$\begin{aligned}
 V_{caplet_i}(t) &= \mathbb{E}^{\mathbb{Q}} \left[B(T_i, T_{i+1}) e^{-\int_0^{T_i} r_s ds} (L(T_i, T_i, T_{i+1}) - K)^+ \middle| \mathcal{F}_t \right] e^{\int_0^t r_s ds} (T_{i+1} - T_i) \\
 &= \mathbb{E}^{\mathbb{Q}^{T_{i+1}}} \left[(L(T_i, T_i, T_{i+1}) - K)^+ \middle| \mathcal{F}_t \right] B(t, T_{i+1}) (T_{i+1} - T_i)
 \end{aligned}$$

La nouvelle probabilité $\mathbb{Q}^{T_{i+1}}$ est appelée probabilité forward-neutre associée au numéraire $(B(t, T_{i+1}))_t$. Sous cette probabilité, $L(t, T_i, T_{i+1})$ est une martingale :

$$dL(t, T_i, T_{i+1}) = L(t, T_i, T_{i+1})\sigma_i(t)dW_{i,t}^{\mathbb{Q}^{T_{i+1}}}$$

En utilisant les mêmes formules que pour le calcul du prix des options d'achats européennes sur actions, et sachant que $\mathbb{E}^{\mathbb{Q}^{T_{i+1}}} [L(T_i, T_i, T_{i+1})|\mathcal{F}_t] = L(t, T_i, T_{i+1})$ on obtient :

$$V_{caplet_i}(t) = B(t, T)(T_{i+1} - T_i)(L(t, T_i, T_{i+1})\mathcal{N}(d_1) - K\mathcal{N}(d_2)), \quad \forall t \leq T_i$$

Avec :

$$d_1 = \frac{\ln\left(\frac{L(t, T_i, T_{i+1})}{K}\right) + \frac{1}{2}\gamma_i^2(T_i - t)}{\gamma_i\sqrt{T_i - t}}$$

$$d_2 = d_1 - \gamma_i\sqrt{T_i - t}$$

$$\gamma_i = \sqrt{\int_t^{T_i} \frac{1}{T_i - t}\sigma_i(u)^2 du}$$