

Mémoire présenté le :

pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA
et l'admission à l'Institut des Actuaires

Par : Margot SCHMUTZ

Titre : Rationnalisation du passage du BE déterministe au BE stochastique par des modèles de machine learning

Confidentialité : NON (Durée : 1 an 2 ans)

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

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

*Entreprise : FORSIDES
Nom :*

Signature :

Directeur de mémoire en entreprise :

Nom : Daniel ZERBIB

Signature :

Invité :

Nom :

Signature :

*Membres présents du jury de
l'ISFA*

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

Signature du responsable entreprise



Signature du candidat



Table des matières

Résumé	4
Abstract	5
Remerciements	6
Synthèse	7
Synthesis	8
Introduction	1
1 Cadre théorique et réglementaire du régime Solvabilité II	3
1.1 Solvabilité II : principes et objectifs	4
1.1.1 Historique de la directive européenne	4
1.1.2 Les trois piliers de S2 : exigences de capital, gouvernance, et transparence	4
1.2 Définition et calcul du ratio de solvabilité S2	4
1.2.1 Une mesure synthétique de la solidité financière d'un assureur	5
1.2.2 Le Solvency Capital Requirement (SCR) : estimation des risques extrêmes	5
1.2.3 Les fonds propres éligibles : capacité à couvrir les pertes	5
1.2.4 Un indicateur sensible aux chocs économiques et aux arbitrages de gestion	5
1.3 Enjeux de pilotage de la solvabilité	5
1.3.1 Rôle de l'anticipation du ratio dans la gestion stratégique	6
2 Modèle ALM interne et construction de la base de données	7
2.1 Le modèle ALM interne	8
2.1.1 Générateur de scénarios économiques	8
2.1.2 Fonctionnement de SALLTO	8
2.2 Description du portefeuille étudié	8
2.2.1	9
2.3 Méthodologie de la construction de la base de données	9
2.3.1 Définition de la variable cible et des variables explicatives retenues	10
2.3.2 Construction de la base de données	10
2.3.3 Justification des variables retenues	12

3 Apports du machine learning pour la rationalisation des écarts	13
3.1 Problématique de régression dans le cadre du BE	14
3.1.1 Formulation du problème	14
3.1.2 Pourquoi le machine learning est adapté à ce contexte	14
3.1.3 Positionnement par rapport aux approches traditionnelles	14
3.2 Présentation des modèles de machine learning utilisés	14
3.2.1 Modèles de référence	15
3.2.2 Modèles non linéaires	15
3.2.3 Avantages et limites de chaque modèles dans le cadre d'étude	15
3.3 Méthodologie d'apprentissage	15
3.3.1 Séparation des données	16
3.3.2 Choix des hyperparamètres	16
3.3.3 Prévention de surapprentissage	16
3.4 Critère de performance	16
3.4.1 Indices retenus	17
3.4.2 Comparaison des modèles	17
3.4.3 Lecture des performances	17
4 Analyse des résultats et rationalisation des écarts	18
4.1 Qualité de la modélisation du BE stochastique	19
4.1.1 Capacité des modèles à reproduire le BE stochastique	19
4.1.2 Comparaison avec le BE déterministe seul	19
4.1.3 Apports du machine learning	19
4.2 Analyse des variables explicatives majeures	19
4.2.1 Variables dominantes dans l'explication du BE stochastique	20
4.2.2 Rôle du BE déterministe	20
4.2.3 Impact des caractéristiques du passif	20
4.3 Lecture des écarts BE déterministe/BE stochastique	20
4.3.1 Identification des principaux facteurs d'écarts	21
4.3.2 Analyse par classification	21
4.3.3 Mise en évidence des effets non linéaires	21
4.4 Apports pour la traçabilité et la compréhension des résultats	21
4.4.1 Amélioration de la lisibilité des écarts	22
4.4.2 Contribution à la documentation des modèles	22
4.4.3 Intérêts pour la gouvernance	22
5 Limites, apports et perspectives	23
5.1 Limites de l'approche	24
5.1.1 Limites liées aux données	24
5.1.2 Limites des modèles de machine learning	24
5.1.3 Dépendance au modèle ALM utilisé	24
5.2 Apports opérationnels et réglementaires	24
5.2.1 Aide à l'analyse des résultats	25

5.2.2 Outils de compréhension complémentaire	25
5.2.3 Contribution à la transparence	25
5.3 Perspectives d'amélioration	25
5.3.1 Extension à d'autres portefeuilles	26
5.3.2 Intégration d'autres variables	26
5.3.3 Approfondissement des méthodes d'explicabilité	26
6 Conclusion	27
Annexes	28
Bibliographie	29

Résumé

Petit résumé en français de mon mémoire ?

Abstract

It's a brief sum up in english !

Remerciements

Synthèse

Un long résumé en français de mon mémoire

Synthesis

Un long résumé en anglais de mon mémoire

Introduction

Dans un contexte réglementaire exigeant et en constante évolution, les organismes d'assurance sont amenés à renforcer la robustesse, la traçabilité et la transparence de leurs méthodes de valorisation des engagements. Depuis l'entrée en vigueur de la directive Solvabilité II, le calcul des provisions techniques repose sur une approche économique cohérente avec les conditions de marché, dont le Best Estimate (BE) constitue un élément central. Il représente l'espérance actualisée des flux futurs liés aux engagements d'assurance et joue un rôle déterminant dans l'évaluation de la situation financière des assureurs.

Dans le cadre de Solvabilité II, le Best Estimate stochastique est considéré comme une référence pour la valorisation des engagements, dans la mesure où il permet d'intégrer l'incertitude liée aux évolutions futures des variables économiques et financières. Néanmoins, en pratique, les assureurs ont également recours à des approches déterministes, notamment pour des besoins d'analyse, de comparaison ou de compréhension des résultats. Le passage du BE déterministe au BE stochastique soulève alors des enjeux importants de traçabilité, de justification et de maîtrise des écarts, en particulier dans un contexte où les résultats produits par les modèles doivent être compris, expliqués et documentés.

Ainsi, le passage du BE déterministe au BE stochastique constitue un enjeu à la fois technique, opérationnel et réglementaire. S'il est admis que le BE stochastique offre une vision plus complète du risque, la compréhension des facteurs expliquant les différences entre ces deux approches reste un sujet clé, notamment en termes de gouvernance des modèles, de transparence vis-à-vis des parties prenantes et de maîtrise des résultats.

Dans ce contexte, les avancées récentes en machine learning ouvrent de nouvelles perspectives. Ces méthodes permettent d'exploiter des bases de données complexes afin de modéliser des relations non linéaires entre variables explicatives et variables cibles, tout en offrant des outils d'analyse de l'importance des facteurs explicatifs. Appliquées au cadre actuariel, elles peuvent constituer un outil complémentaire aux modèles traditionnels, non pas dans une logique de substitution, mais dans une démarche de compréhension et de rationalisation des résultats.

Ce mémoire s'inscrit dans cette dynamique et a pour objectif de rationnaliser le passage du BE stochastique au BE déterministe à l'aide de modèles de machine learning, en mettant l'accent sur la transparence et l'interprétabilité des résultats. Plus précisément, il vise à identifier et analyser les variables expliquant les écarts entre ces deux approches, en s'appuyant sur des données issues d'un modèle ALM interne utilisé en cabinet de conseil. Le BE stochastique constitue la variable cible des modèles, tandis que le BE déterministe et d'autres variables majoritairement liées au passif sont intégrées comme variables explicatives.

Ce mémoire s'articule autour des parties suivantes :

La première partie rappelle les fondements réglementaires de Solvabilité II relatifs au calcul du Best Estimate, en présentant les principes du BE déterministe et du BE stochastique ainsi que leurs rôles respectifs dans le cadre prudentiel.

La deuxième partie décrit le modèle ALM interne utilisé pour cette étude et la méthodologie de construction de la base de données. Elle précise les hypothèses retenues, les variables explicatives considérées, principalement au passif, ainsi que les choix de paramétrage effectués.

La troisième partie présente les principes théoriques des modèles de machine learning utilisés pour les problèmes de régression, les critères de performance retenus, ainsi que les enjeux d'interprétabilité et d'explicabilité, essentiels dans un cadre actuariel et réglementaire.

La quatrième partie est consacrée à l'analyse des résultats obtenus. Elle met en évidence les variables les plus explicatives du BE stochastique, propose une lecture actuarielle des écarts observés entre BE déterministe et BE stochastique, et en dégage les grandes tendances.

Enfin, la dernière partie discute les limites de l'approche proposée ainsi que les perspectives d'amélioration et d'extension du cadre d'analyse.

Ainsi, ce mémoire se situe à l'interface entre les exigences prudentielles, les contraintes opérationnelles des modèles ALM et les nouvelles méthodes d'analyse issues du machine learning. Il s'inscrit dans une démarche visant à renforcer la transparence, la compréhension et la maîtrise du calcul du Best Estimate en assurance vie.

Chapitre 1

Cadre théorique et réglementaire du régime Solvabilité II

1.1 Solvabilité II : principes et objectifs

1.1.1 Historique de la directive européenne

1.1.2 Les trois piliers de S2 : exigences de capital, gouvernance, et transparence

1.2 Définition et calcul du ratio de solvabilité S2

1.2.1 Une mesure synthétique de la solidité financière d'un assureur

Définition du ratio S2 : Fonds propres éligibles / SCR

Interprétation économique du ratio (seuil réglementaire de 100)

Rôle central du ratio dans la communication avec les régulateurs et les investisseurs

1.2.2 Le Solvency Capital Requirement (SCR) : estimation des risques extrêmes

Logique de calcul : capital requis pour absorber un choc extrême avec un niveau de confiance de 99,5

Méthodes de calcul : formule standard vs modèle interne partiel ou complet

Décomposition du SCR : risque de marché, risque de souscription, risque de crédit, risque opérationnel

1.2.3 Les fonds propres éligibles : capacité à couvrir les pertes

Classification des fonds propres : Tier 1, Tier 2, Tier 3

Critères d'éligibilité et ajustements réglementaires

Impact des réévaluations d'actifs, des plus-values latentes et des instruments hybrides

1.2.4 Un indicateur sensible aux chocs économiques et aux arbitrages de gestion

Sensibilité du ratio aux taux d'intérêt, à la volatilité, aux spreads de crédit

Mécanismes de gestion du ratio : couverture du SCR, allocation d'actifs, revalorisation du passif

Enjeux stratégiques : pilotage dynamique de la solvabilité, gestion de la marge de manœuvre

1.3 Enjeux de pilotage de la solvabilité

1.3.1 Rôle de l'anticipation du ratio dans la gestion stratégique

1.3.1.1 Limites opérationnelles des modèles actuels

Chapitre 2

Modèle ALM interne et construction de la base de données

2.1 Le modèle ALM interne

Au sein de Forsides, le modèle d'ALM s'appelle SALLTO (Solvency Asset Liability Life TOols). Il s'agit d'un outil développé et codé en interne en C#. Ce modèle permet de représenter l'évolution temporelle d'une compagnie d'assurance vie en intégrant explicitement les interactions entre l'actif et le passif, à partir de scénarios économiques.

Pour fonctionner, le modèle nécessite un fichier d'input dans lequel sont renseignées différentes hypothèses. Ces hypothèses concernent à la fois l'actif et le passif, ainsi que la politique de gestion de la compagnie d'assurance étudiée. Le fichier d'entrée inclut également les scénarios économiques, générés indépendamment par le générateur de scénarios économiques (GSE), qui servent de base aux projections réalisées par le modèle.

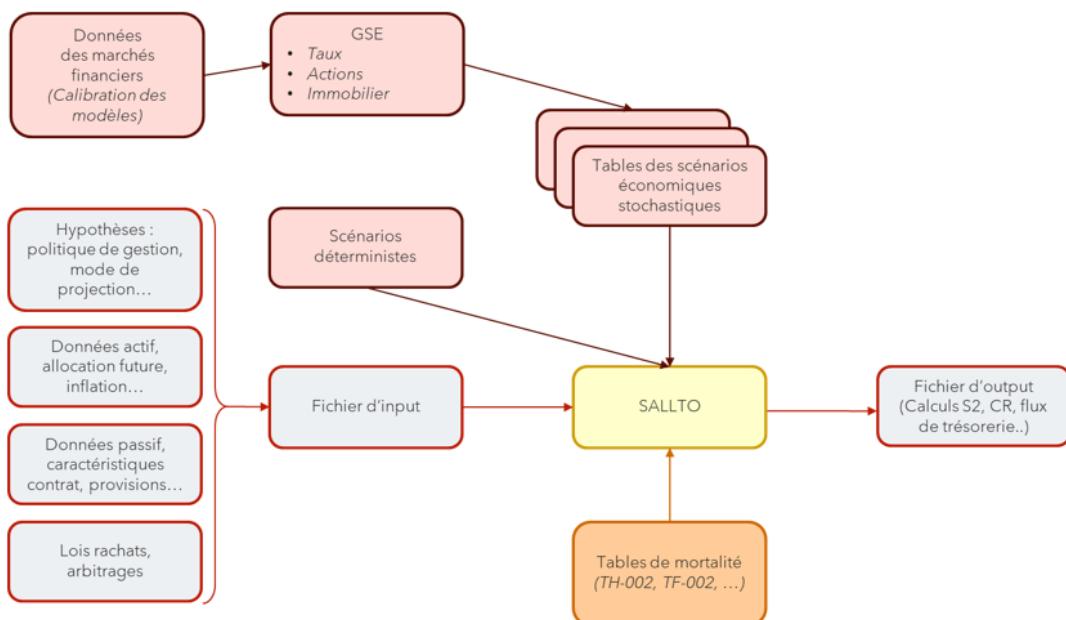


FIGURE 2.1 – Schéma de fonctionnement de SALLTO

2.1.1 Générateur de scénarios économiques

Théorie mathématique derrière

2.1.2 Fonctionnement de SALLTO

Etapes de fonctionnement

2.2 Description du portefeuille étudié

2.2.1

2.3 Méthodologie de la construction de la base de données

2.3.1 Définition de la variable cible et des variables explicatives retenues

L'objectif principal de ce mémoire est de comprendre et d'expliquer les mécanismes conduisant aux différences de valorisation entre un BE déterministe et un BE stochastique. L'enjeu n'est pas de comparer ces deux approches d'un point de vue normatif, mais d'analyser les facteurs qui contribuent aux écarts observés, afin d'en améliorer la traçabilité, la justification et la compréhension dans un cadre réglementaire. Afin d'aborder la problématique sous des angles complémentaires, deux bases de données distinctes ont été construites. Elles reposent sur un même socle de données explicatives, mais se distinguent par le choix de la variable cible.

La première base de données consiste à considérer le BE stochastique comme variable à expliquer, à partir d'un ensemble de variables explicatives majoritairement liées au passif, incluant notamment des éléments issus du cadre déterministe. Cette approche permet d'évaluer dans quelle mesure les caractéristiques du portefeuille et les hypothèses déterministes contribuent à la formation du niveau du BE stochastique.

La seconde base de données adopte une approche centrée sur l'analyse des écarts, en définissant comme variable cible la différence entre le Best Estimate stochastique et le Best Estimate déterministe. Cette formulation vise à isoler et à analyser plus directement les facteurs expliquant les divergences entre les deux valorisations, en mettant en évidence les effets liés à la prise en compte de l'incertitude, des non-linéarités et des options financières.

L'utilisation conjointe de ces deux approches permet ainsi de proposer une lecture actuarielle approfondie des résultats, en apportant des éléments de compréhension sur le passage du BE déterministe au BE stochastique. Elle s'inscrit dans une démarche de transparence et de gouvernance des modèles, en cohérence avec les exigences de Solvabilité II.

Les variables explicatives retenues sont majoritairement liées aux caractéristiques du passif, conformément à l'objectif du mémoire, qui vise à analyser les écarts de valorisation du point de vue des engagements d'assurance.

2.3.2 Construction de la base de données

Le jeu de données utilisé dans ce mémoire est construit à partir d'une démarche de sensibilités appliquée à SALLTO. Cette approche consiste à faire varier de manière contrôlée certaines hypothèses du modèle, principalement liées au passif, afin d'observer l'impact de ces variations sur la valorisation des engagements d'assurance.

L'objectif de cette démarche est de générer un ensemble d'observations structurées permettant d'analyser, dans un cadre maîtrisé, les relations entre les variables explicatives issues des sensibilités et les résultats de Best Estimate. Chaque observation du jeu de données correspond à une configuration donnée des hypothèses du modèle, définie par un ensemble de valeurs prises par les variables soumises à sensibilité.

Les sensibilités sont réalisées à une date de valorisation donnée, sur un périmètre de passif constant. En dehors des variables explicitement soumises à sensibilité, l'ensemble des hypothèses est maintenu identique sur l'ensemble des calculs. Cette hypothèse garantit que les variations observées sur les résultats de valorisation proviennent exclusivement des paramètres étudiés.

Pour chaque configuration testée, le SALLTO permet de calculer :

un Best Estimate déterministe ;

un Best Estimate stochastique.

Le Best Estimate déterministe constitue à la fois un résultat de valorisation à part entière et une information explicative essentielle pour l'analyse des résultats stochastiques.

Le jeu de données repose sur un ensemble de variables explicatives comprenant à la fois des variables issues des sensibilités et le Best Estimate déterministe.

Les sensibilités réalisées portent principalement sur des variables du passif, complétées par une variable relative à l'actif. Les variables explicatives soumises à sensibilité sont les suivantes :

taux servi ;

rachats conjoncturels ;

rachats structurels ;

âge des assurés ;

taux minimal garanti ;

ancienneté des contrats ;

courbe de taux utilisée pour la valorisation de l'actif.

En complément des variables soumises à sensibilité, le Best Estimate déterministe est intégré au jeu de données en tant que variable explicative. Il constitue une synthèse des caractéristiques du portefeuille et des hypothèses retenues dans le cadre déterministe, et permet de capturer une part importante de l'information expliquant le niveau du Best Estimate stochastique.

L'inclusion du BE déterministe parmi les variables explicatives vise à analyser dans quelle mesure celui-ci permet d'expliquer le BE stochastique, ainsi que les limites de cette explication lorsque des effets non linéaires ou liés à l'incertitude sont présents.

Pour chacune des variables soumises à sensibilité, cinq sensibilités sont définies :

une configuration centrale, correspondant aux hypothèses de référence ;

deux sensibilités à la hausse ;

deux sensibilités à la baisse.

Ces variations sont choisies de manière symétrique autour du scénario central et permettent de couvrir une plage de valeurs suffisamment large pour faire apparaître des effets non linéaires, tout en restant cohérentes.

2.3.3 Justification des variables retenues

Chapitre 3

Apports du machine learning pour la rationalisation des écrats

3.1 Problématique de régression dans le cadre du BE

3.1.1 Formulation du problème

3.1.2 Pourquoi le machine learning est adapté à ce contexte

3.1.3 Positionnement par rapport aux approches traditionnelles

3.2 Présentation des modèles de machine learning utilisés

3.2.1 Modèles de référence

3.2.2 Modèles non linéaires

3.2.3 Avantages et limites de chaque modèles dans le cadre d'étude

3.3 Méthodologie d'apprentissage

3.3.1 Séparation des données

3.3.2 Choix des hyperparamètres

3.3.3 Prévention de surapprentissage

3.4 Critère de performance

3.4.1 Indices retenus

3.4.2 Comparaison des modèles

3.4.3 Lecture des performances

Chapitre 4

Analyse des résultats et rationnalisation des écarts

4.1 Qualité de la podélisation du BE stochastique

4.1.1 Capacité des modèles à reproduire le BE stochastique

4.1.2 Comparaison avec le BE déterministe seul

4.1.3 Apports du machine learning

4.2 Analyse des variables explicatives majeurs

4.2.1 Variables dominantes dans l'explication du BE stochastique

4.2.2 Rôle du BE déterministe

4.2.3 Impact des caractéristiques du passif

4.3 Lecture des écarts BE déterministe/BE stochastique

4.3.1 Identification des principaux facteurs d'écarts

4.3.2 Analyse par classification

4.3.3 Mise en évidence des effets non linéaires

4.4 Apports pour le traçabilité et la compréhension des résultats

- 4.4.1 Amélioration de la lisibilité des écarts
- 4.4.2 Contribution à la documentation des modèles
- 4.4.3 Intérêts pour la gouvernance

Chapitre 5

Limites, approts et perspectives

5.1 Limites de l'approche

- 5.1.1 Limites liées aux données
- 5.1.2 Limites des modèles de machine learning
- 5.1.3 Dépendance au modèle ALM utilisé

5.2 Apports opérationnels et réglementaires

- 5.2.1 Aide à l'analyse des résultats
- 5.2.2 Outils de compréhension complémentaire
- 5.2.3 Contribution à la transparence

5.3 Perspectives d'amélioration

- 5.3.1 Extension à d'autres portefeuilles
- 5.3.2 Intégration d'autres variables
- 5.3.3 Approfondissement des méthodes d'explicabilité

Chapitre 6

Conclusion

Annexes

Bibliographie

- [1] DORNAIKA Fadi, HAMAD Denis, CONSTANTIN Joseph, TRONG HOANG Vinh. *Advances in Data Clustering*. Springer, 2024.
- [2] GOFFARD Pierre-Olivier, GUERRAULT Xavier. « Is it optimal to group policyholders by age, gender, and seniority for BEL computations based on model points ? ». *European Actuarial Journal*, volume 5, 17 Avril 2015, p. 165-180.
- [3] BEN FADHEL Amine. « Accélération de l'évaluation de la solvabilité prospective d'un assureur épargne ». *Mémoire pour l'Institut des Actuaires*, 2022.
- [4] FRANCE ASSUREURS. *L'assurance vie en 2023*. (20 septembre 2024). Consulté le 1er Septembre 2025, sur <https://www.franceassureurs.fr/nos-chiffres-cles/assurance-vie/etude-statistique-assurance-vie-2023/>