

Métamodélisation ALM : étude de performance sur différents cas d'usage en épargne-retraite

Nexialog Consulting, Département R&D, Paris, France
Medhi Jacqueline, Valentin Erades, Areski Cousin

22 octobre 2024

Résumé

Cet article présente les principaux résultats d'une étude réalisée dans le cadre d'une collaboration entre l'équipe R&D de Nexialog Consulting et le département ALM d'un grand groupe d'assurance. En réponse aux défis posés par la complexité croissante des modèles ALM, une solution consiste à construire une version *proxy* du modèle ALM (ou métamodèle ALM), capable de reproduire rapidement et avec précision les principaux indicateurs du bilan, améliorant ainsi la disponibilité des ressources de simulation. Dans cette étude, nous avons testé et comparé plusieurs algorithmes d'apprentissage statistique pour prédire le BE et le SCR dans le cadre de Solvabilité 2. La base d'entraînement est constituée des chocs et sensibilités réalisés pour le calcul de la formule standard et enrichie de l'information collectée au fil des exercices réglementaires. Les résultats obtenus sur le périmètre de l'épargne-retraite montrent que cette approche est particulièrement prometteuse pour plusieurs cas d'usage spécifiques. Parmi ceux-ci, la prédiction du niveau de BE sur l'exercice réglementaire suivant, l'évaluation des sensibilités avec la connaissance des chocs S2, ainsi que la prédiction du SCR. Ces cas d'usage démontrent l'efficacité du méta-modèle dans des contextes de forte charge opérationnelle, notamment durant les périodes de production des indicateurs ALM.

Keywords — ALM, PERP, Best Estimate, Solvabilité II, SCR, machine learning, approche proxy, métamodélisation

Table des matières

Introduction	3
1 Base de données	4
2 Implémentation	6
2.1 Approches de prédiction des variables cibles	6
2.2 Sélection des variables explicatives pour les différents objectifs	6
2.3 Méthode de validation	8
2.3.1 Découpage base d’entraînement / base test	8
2.3.2 Cross validation temporelle	8
2.4 Validation des résultats	9
2.5 Modèles retenus	10
3 Cas d’usages du proxy	12
3.1 Prédiction du niveau de BE	13
3.2 Prédiction des sensibilités connaissant les chocs S2	15
3.3 Prédiction du SCR	17
Conclusion	20
Références	21
A Annexes	22
A.1 Définition des chocs S2 considérés dans l’étude	22
A.2 Définition des sensibilités	22
B Glossaire	24
Contacts	25

Introduction

L'introduction de nouvelles réglementations a relevé le niveau d'exigence du pilotage de l'activité qui doit désormais se faire en vision multinorme (Normes françaises, IFRS, Solvabilité 2, etc...). Le modèle ALM doit ainsi intégrer un grand nombre de facteurs de risque et d'hypothèses pour calculer les indicateurs nécessaires au suivi de l'activité dans chaque référentiel. Cette complexité croissante se traduit par un allongement du temps d'exécution des calculs. Face à cette situation, il devient essentiel de trouver des solutions permettant de réduire ces temps de calcul tout en maintenant un niveau de précision adéquat. L'une des approches les plus prometteuses pour relever ce défi est la méta-modélisation.

L'étude vise à prédire des indicateurs clés de gestion d'actifs et de passifs (ALM), tels que le Best Estimate (BE), la Net Asset Value (NAV) et le Solvency Capital Requirement (SCR), à l'aide d'un modèle d'apprentissage statistique. Celui-ci est entraîné sur une base de données constituée d'un ensemble de points (entrées, sortie) générés par le modèle d'origine, l'objectif étant de reproduire certaines variables cibles du modèle ALM. L'avantage d'un tel proxy est d'obtenir une estimation rapide et précise du BE (Best Estimate) ou de la NAV (Net Asset Value) et par conséquent du SCR en formule standard sans consommer les ressources de la plateforme de simulation interne, améliorant ainsi le temps de production des indicateurs ALM et la disponibilité des ressources. Cette approche par méta-modélisation vient compléter d'autres approches de régression ou de simplification utilisées notamment dans le cadre des calculs de solvabilité en modèles internes ou de l'ORSA : simulations dans les simulations, Curve Fitting, LSMC (voir, par exemple, Gauville (2017)), Scenario Picking (voir par exemple Choquer (2019) ou Arrouy et al. (2023)).

Les études récentes sur le sujet se concentrent sur deux axes principaux : l'amélioration des proxies du Best Estimate (BE) et la méta-modélisation du SCR. Par exemple, dans le contexte des contrats d'épargne en Euros, Benyoussef (2023) propose un proxy du BEL en intégrant la sensibilité de l'indicateur aux variations du marché, comme un choc sur la courbe des taux et sur les PMVL/PPE. De même, Ear (2023) explore l'application de modèles de machine learning, tels que Random Forest et eXtreme Gradient Boost, pour accélérer le calcul du SCR, en particulier pour les portefeuilles de fonds en Euros. D'autres recherches, comme celles de Messoussi (2017), se concentrent sur la prédiction des ratios de couverture en utilisant des algorithmes de machine learning pour simplifier le processus de simulation de l'évolution du capital requis. Les travaux de Hounkonnou and Behi (2022), appliqués aux passifs d'épargne en Euros, ainsi que ceux de Cescutti (2016), concernant les contrats d'épargne italiens, montrent comment la méta-modélisation peut offrir une solution efficace pour le calcul dynamique du SCR, en utilisant des techniques d'apprentissage automatique pour améliorer la réactivité et la précision des estimations. Ces recherches soulignent une tendance croissante vers l'intégration du machine learning pour surmonter les défis des calculs réglementaires complexes. Notre étude utilise des techniques de machine learning, telles que Random Forest et la régression par Processus Gaussiens, pour prédire simultanément le BE, la NAV et le SCR. Une approche de cross-validation temporelle a été retenue pour renforcer la robustesse des prédictions sur différentes périodes. Le modèle est spécialement adapté aux portefeuilles type PERP, ce qui permet des prédictions précises sans recalibrations fréquentes, même en période de stress de marché.

Cette note se concentre sur un portefeuille de contrats retraite, formant un canton au sein du modèle ALM. Plus précisément, le PERP, devenu PER en 2019, est un portefeuille relativement simple d'utilisation qui ne prévoit pas de rachats (sauf dans des cas exceptionnels) et offre une sortie en rente au départ à la retraite. Cela présente l'avantage de limiter les variables explicatives liées aux assurés (et au passif du portefeuille). De plus, le PERP a connu peu d'évolution commerciale, avec peu de nouveaux entrants (portefeuille en run-off), et a été peu affecté par les évolutions de modèles sur la période étudiée, ce qui en fait un candidat idéal pour une première implémentation.

Le modèle ALM permet d'estimer le niveau des engagements et les différents postes du bilan et du compte de résultat, à partir de projections des cash-flows et provisions futurs, dans différents scénarios d'évolution des facteurs de risque. Ces projections prennent en compte les interactions actif-passif ainsi que certaines règles de gestion de l'assureur. Ces différentes composantes sont illustrées en Figure 1. À gauche de cette figure, on trouve les données connues concernant la composition du bilan, tant pour l'actif que pour le passif, ainsi que les hypothèses économiques retenues. En bas sont représentées les hypothèses relatives au comportement du passif, telles que la mortalité et les rachats. En haut sont présentées les différentes trajectoires économiques

fournies par le Générateur de Scénarios Économiques. Le modèle ALM se compose donc d'un grand nombre d'entrées et inclut même plusieurs « sous-modèles », ce qui nécessite un temps d'exécution conséquent pour chaque calcul réglementaire.

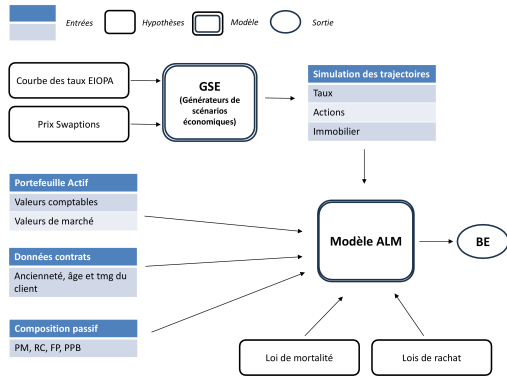


Figure 1 Description du modèle ALM

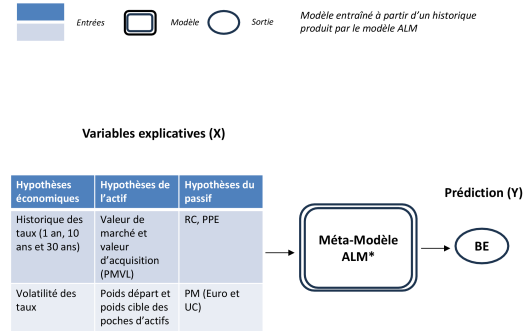


Figure 2 Proxy au modèle ALM

L'objectif de notre étude est de construire un proxy du modèle en entraînant un modèle d'apprentissage statistique sur des données reconstituées à partir d'historiques de « run » produits par le modèle ALM. Le schéma décrit en Figure 2 illustre ce processus.

Dans Messoussi (2017), les variables ont été sélectionnées et testées en raison de leur potentiel à influencer significativement le capital économique. La base d'apprentissage utilisée dans cette étude provient d'une sélection rigoureuse fondée sur l'expertise technique, visant à capturer l'ensemble des effets intrinsèques associés au canton retraite. Elle regroupe majoritairement des variables liées à l'actif, parmi lesquelles : la courbe des taux, la valeur de marché et le poids des actifs dans le portefeuille. Au passif, nous retrouvons principalement les réserves et provisions suivantes RC, PPE, PM Euros et PM UC, la durée du passif, ainsi que les BE Euros et UC, qui constituent nos variables cibles.

La note est articulée autour de trois parties. Dans une première partie, nous aborderons la construction de la base de données utilisée pour les différentes simulations et prédictions, en détaillant la sélection des variables explicatives. Ensuite, nous décrirons les méthodes d'implémentation des modèles, mettant en avant les approches statistiques et de machine learning retenues pour optimiser la prédiction des indicateurs ALM. Enfin, nous explorerons les cas d'usages du modèle proxy développé, en illustrant son application concrète à travers des exemples de prédiction du Best Estimate, du SCR et des sensibilités aux chocs réglementaires.

1 Base de données

Les variables explicatives ont été sélectionnées sur la base d'une représentation « simplifiée » du bilan de l'assureur. Elles peuvent être regroupées en trois familles de variables :

- Composition du portefeuille d'actifs : la répartition initiale et son évolution au cours du temps.
- Variables caractéristiques des facteurs de risques financiers : courbe des taux, volatilité du marché actions et immobilier, etc.
- Variables de passif : provision mathématique du fonds Euros et des UC, durations, etc.

Pour décrire la composition du portefeuille d'actifs, nous avons considéré une agrégation en 3 poches : monétaires, crédit et diversifié (action, immobilier).

La composition du portefeuille d'actifs inclut des variables qui caractérisent la manière dont notre portefeuille sera rebalancé au fil du temps. Nous considérons un portefeuille d'actifs agrégé selon trois poches : monétaires, crédit et diversifié. La composition de chaque poche sera définie à partir des variables suivantes :

- Répartition cible en constant-mix.

- Valeur d'acquisition des actifs.
- Valeur de marché des actifs à la date d'évaluation.
- Duration de l'actif : la duration de l'actif fait référence à la durée moyenne pondérée des flux de trésorerie attendus de l'actif de l'entreprise. Cela inclut notamment les intérêts, les remboursements de capital. Elle ne concerne que les flux obligataires.

Les facteurs de risque financier sont des variables considérées comme exogènes dans le modèle, c'est-à-dire qu'elles ne dépendent pas directement de la composition du portefeuille, mais influencent les risques associés à celui-ci. Ces variables incluent les risques de marché, les spreads, et le risque de crédit sur différents segments de marché. La courbe des taux et son évolution auront un impact significatif à la fois sur l'actif et le passif. Elles influencent plus généralement toutes les poches d'actifs, notamment les poches monétaires et de crédit. Le niveau des taux 1 an, 10 ans et 30 ans ont donc été choisis. Nous pouvons aussi citer la volatilité de la poche diversifiée (mélange d'actions et d'immobilier) qui est représentée par la moyenne des volatilités implicites des options EX50 à date d'évaluation.

Concernant la typologie de nos variables de passif, nous avons en premier lieu les variables caractérisant le stock de richesse à partager entre l'assureur et les assurés (épargnants) et en second lieu les variables de composition du passif.

Les variables caractérisant le stock de richesse sont :

- La PPE (Provision pour Participation aux Excédents) : c'est une provision constituée par les organismes de retraite pour redistribuer les éventuels excédents financiers aux épargnants.
- La RC (Réserve de capitalisation) : la réserve de capitalisation permet aux assureurs de lisser leurs résultats en mettant de côté une partie des plus-values réalisées sur les ventes d'obligations pour compenser d'éventuelles moins-values futures ou passées, assurant ainsi une stabilité financière face aux fluctuations des marchés obligataires.

Les variables qui décrivent la composition du passif sont :

- La PM (Provision mathématique) du fonds Euro et du fonds UC : estimation financière des engagements futurs d'une compagnie d'assurance envers ses assurés. Elle est basée sur des calculs mathématiques qui prennent en compte les probabilités de réalisation des sinistres et les montants des indemnités à verser.
- Duration du passif : la duration du passif mesure la durée moyenne pondérée des flux de trésorerie relatifs aux engagements financiers de l'assureur, tel que les rentes et les autres engagements envers les assurés dans le cadre du canton retraite.

La variable cible à prédire diffère selon les besoins et les cas d'usage de l'assureur. Dans cette étude, les variables cibles considérées sont :

- BE Euros
- BE UC
- BE Agrégé (BE Euros + BE UC)
- Delta NAV sur BE Agrégé

Les deux dernière variables n'étaient pas présentes à l'origine dans la base de données, mais leurs valeurs peuvent être déduites directement des données.

La base de données utilisée pour cette étude est composée de 870 observations réparties sur 25 variables, couvrant la période allant du premier trimestre 2022 au quatrième trimestre 2023. Chaque trimestre comporte six scénarios de sensibilité distincts, représentant des environnements économiques variés ainsi que des comportements du passif différents, dont un scénario sans sensibilité particulière (voir l'Annexe A.2). Pour chacun de ces scénarios et pour chaque trimestre, 15 chocs intervenant dans la construction du SCR en formule standard S2 ont été calculés avec le modèle ALM, soit 15 chocs S2 par scénario de sensibilité et par trimestre. Ainsi, chaque choc de la formule standard correspond à un point de données dans la base, ce point représentant l'effet conjoint du choc sur les variables d'entrée et sur les variables cibles du modèle, dans le trimestre et la sensibilité considérées (voir l'Annexe A.1 pour un descriptif détaillé de ces chocs).

Cela donne une structure initiale de 6 scénarios par trimestre, combinés avec 15 chocs, soit une base de 720 lignes. À partir du premier trimestre 2023, des sensibilités supplémentaires ont été introduites avec des variations de taux de plus ou moins 200 points de base, ajoutant 120 lignes

supplémentaires à la base. Au total, on compte 90 lignes par arrêté (trimestre) en 2022 et 120 lignes par arrêté en 2023. Enfin, certaines sensibilités distinctes sont également testées dans la base de test pour évaluer comment le modèle réagit face à des chocs qu'il n'a pas observés précédemment.

2 Implémentation

Dans cette section, nous allons aborder les étapes, méthodes et modèles utilisés pour mettre en place le proxy.

2.1 Approches de prédiction des variables cibles

La prédiction de chaque variable, une sous-sélection de variables explicatives a été effectuée sur la base de l'expertise, l'objectif étant une fois de plus de capter les mécanismes de la poche concernée et de ne retenir que les variables les plus influentes.

Les variables cibles BE Agrégé et UC ont été relativement simples à prédire. En effet, pour ces cibles, l'approche que nous nommerons « Approche directe », qui consiste à prédire sans modifier la base de données, donc uniquement en considérant une sélection de variables appropriées, a fourni des résultats convaincants. Le BE Euros a été la cible la plus difficile à prédire. En effet, celui-ci présente des relations non linéaires avec les variables explicatives retenues. Afin d'augmenter le pouvoir prédictif des algorithmes pour le BE Euros, nous avons considéré plusieurs approches :

- Approche directe : prédiction via la base de données sans modification.
- Approche avec ajout du BE Euros (t-1) : ajout du BE Euros de l'arrêté précédent en tant que variable explicative. En introduisant une autocorrélation dans la base de données, nous espérons gagner en pouvoir prédictif.
- Approche par agrégation : prédiction avec les meilleurs modèles du BE Agrégé (Euros + UC) et du BE UC, puis soustraction des prédictions pour obtenir le BE Euros.
- Approche par agrégation avec ajout du BE Euros (t-1) : combinaison des deux méthodes précédentes.

Dans un premier temps, ces méthodes ont été implémentées pour obtenir de meilleures prédictions du BE Euros. Par la suite, nous les avons adaptées pour la prédiction du BE Agrégé. En effet, dans l'approche par agrégation, une meilleure prédiction du BE Agrégé conduit à une meilleure prédiction du BE Euros. Ces méthodes ont permis d'obtenir de meilleurs résultats, notamment une division par deux du RMSE global et de la MRE.

2.2 Sélection des variables explicatives pour les différents objectifs

Cette section explore la méthodologie de sélection des variables parmi les variables présentes dans la base d'origine, visant à identifier les facteurs les plus pertinents pour chaque cible spécifique, afin d'optimiser la performance des modèles prédictifs.

Chaque cible étudiée (BE, BE UC, BE Euros, Delta NAV) possède des spécificités et des relations différentes avec les variables explicatives. Nous avons donc utilisé plusieurs méthodes pour identifier une sélection de variables adaptées :

- Recursive Feature Elimination (RFE)
- Stepwise regression (backward et forward)
- Élimination sur la base des corrélations
- Sélection sur la base de l'expertise.

Les méthodes de sélection RFE et de régression stepwise ne se sont pas avérées efficaces. En effet, ces méthodes de sélection sont basées sur la maximisation d'un critère d'information (tel que l'AIC), mais elles ne tiennent pas compte de l'interprétabilité des variables ni de leur importance actuarielle. Pour cette étape de sous-sélection, nous nous sommes donc basés sur une sélection fondée sur l'avis d'experts, complétée par l'élimination de certaines variables en fonction des corrélations. De plus, ces méthodes présentent un autre inconvénient, notamment en ce qui concerne le temps de calcul requis, comme souligné dans Kópczyk (2018).

La Figure 3 illustre un exemple de corrélogramme pour la cible BE avant l'étape de suppression des variables. Sur ce graphique, nous pouvons voir que les variables caractérisant le niveau des taux et de volatilité sont très corrélées entre elles. Cela est néfaste pour les modèles, car une forte corrélation entre les variables peut entraîner des problèmes de multicollinéarité. La multicollinéarité peut rendre les coefficients des modèles instables et difficiles à interpréter, et peut également affecter la précision des prédictions en augmentant la variance des estimations. En conséquence, il peut être nécessaire de réduire la redondance en éliminant ou en combinant les variables fortement corrélées pour améliorer la performance et la robustesse du modèle. Par conséquent, nous optons, entre autres, pour :

- Suppression du taux à 30 ans au profit du taux à 10 ans¹ : compte tenu des durées actives et passives dans notre base de données, il a été plus judicieux de privilégier le taux à 10 ans afin de maintenir une cohérence entre les variables explicatives.
- Suppression de la volatilité du taux à 1 an au profit de la volatilité du taux agrégé. La volatilité du taux à 1 an est une variable largement liée à la poche monétaire du portefeuille, qui représente 3,5 % du portefeuille. En revanche, la volatilité du taux agrégé est liée à la poche crédit (obligations), qui représente 80 % du portefeuille.
- Suppression du poids de départ de la poche monétaire au profit de la poche crédit, qui représente la plus grande partie du portefeuille (80%).

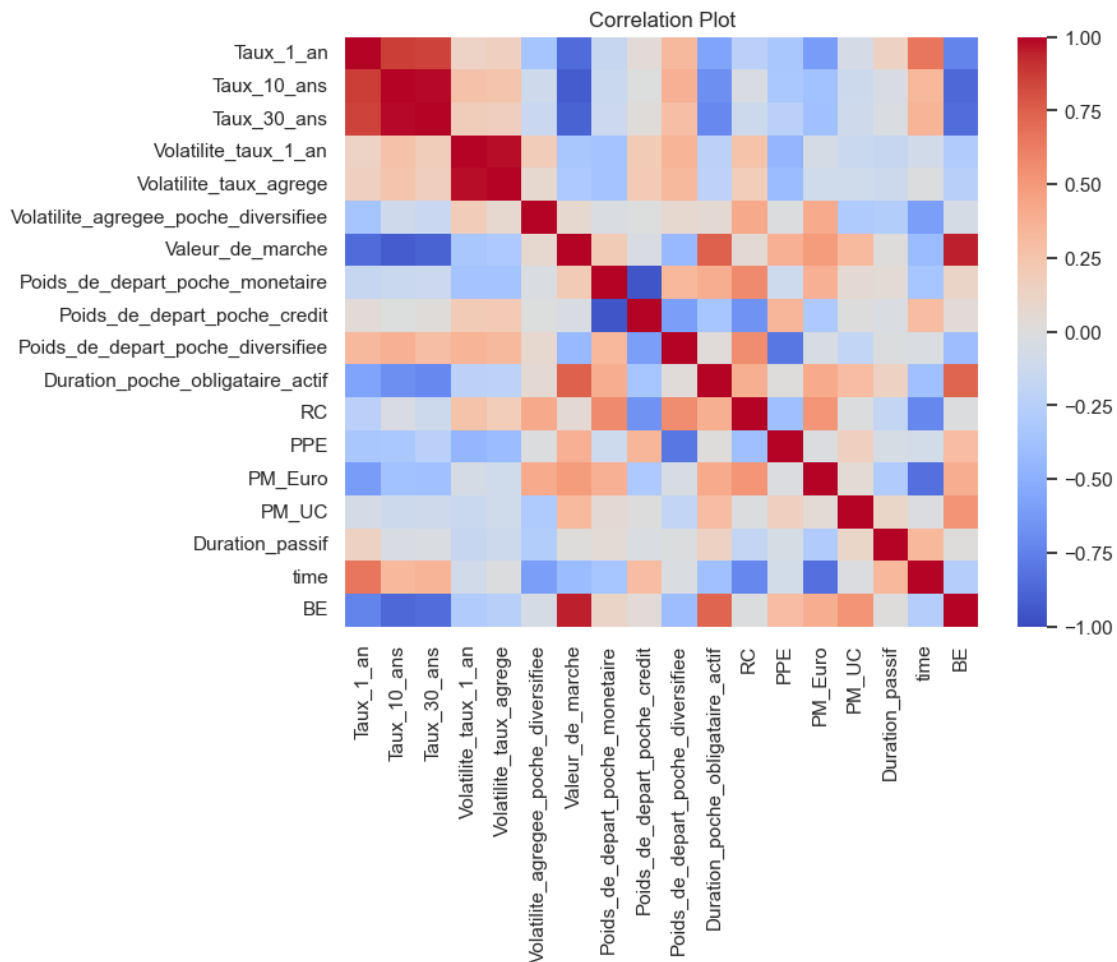


Figure 3 Corrélogramme pour la cible BE

Le Tableau 1 récapitule les variables explicatives sélectionnées pour la prédiction de chaque variable cible.

1. Indépendamment du portefeuille, les assureurs ont tendance à investir dans des obligations de maturité de 10 ans, et la durée des passifs en assurance vie est en moyenne comprise entre 8 et 15 ans. Par conséquent, le taux à 10 ans aurait été privilégié même sans information spécifique sur le portefeuille.

BE	BE UC	BE Euros
Taux 1 an	Taux 1 an	Taux 1 an
Taux 10 ans	Taux 10 ans	Taux 10 ans
Volatilité taux agrégée	Volatilité taux agrégée	Volatilité taux agrégée
Volatilité agrégée poche diversifiée	Volatilité agrégée poche diversifiée	Volatilité agrégée poche diversifiée
Valeur de marché des actifs	PM UC	Valeur de marché des actifs
Duration poche obligataire actif		Duration poche obligataire actif
RC		Poids départ poche crédit
PPE		RC
PM Euros		PPE
PM UC		PM Euros
Duration passif		Duration passif

Table 1 Variables explicatives pour les différentes variables cibles prédites

Nous pouvons observer que, pour prédire le BE UC, un sous-ensemble de 5 variables a été sélectionné. En effet, le BE UC est principalement décrit par la PM UC en raison du mode de fonctionnement de l'épargne en unités de compte : la corrélation entre le BE UC et la PM UC est de 99 % et une MRE de 0,6 % en utilisant uniquement la PM UC comme variable explicative dans une régression linéaire. Néanmoins, cela n'est pas aussi simple avec le BE Euros car celui-ci admet de nombreuses relations non linéaires avec les variables explicatives

2.3 Méthode de validation

Pour garantir la fiabilité et la robustesse des modèles prédictifs développés dans cette étude, il est essentiel de mettre en place des méthodes de validation rigoureuses. La validation permet de s'assurer que les modèles ne sont pas simplement ajustés aux données historiques, mais qu'ils sont également capables de généraliser à de nouvelles données non vues. Dans ce cadre, nous avons exploré différentes stratégies de découpage des données, prenant en compte la dimension temporelle inhérente à notre base de données, afin de reproduire au mieux les conditions d'utilisation pratiques de l'outil. Nous présentons ici deux approches principales : la méthode de division train/test et la cross-validation temporelle, chacune adaptée aux particularités de notre ensemble de données et de notre problématique.

2.3.1 Découpage base d'entraînement / base test

Une dimension temporelle apparaît dans les données. Effectivement, la base est alimentée de manière trimestrielle à chaque arrêté, ce qui permet de suivre l'évolution des variables au fil du temps et d'incorporer ces dynamiques dans nos modèles prédictifs.

Dans ce contexte, le découpage des données ne peut se faire aléatoirement car cela reviendrait à entraîner les modèles avec des données « futures », ce qui, en pratique, ne se produira jamais.

Concernant la proportion de données à incorporer dans chaque ensemble, nous avons décidé de faire une répartition semblable à la manière dont l'outil serait utilisé en pratique : nous accumulons toute l'information disponible en T (donc tous les arrêts précédents disponibles) pour prédire l'arrêt T+1.

2.3.2 Cross validation temporelle

Du fait d'un nombre limité d'observations par arrêté (en comptant les sensibilités et les chocs), soit 75 observations par arrêté sur 8 arrêts (voir Annexes A.1 et A.2), la cross-validation temporelle s'est avérée utile pour évaluer la robustesse de nos modèles et détecter de potentielles

erreurs liées à la qualité ou la modélisation de nos données. Nous avons retenu deux méthodes de cross-validation temporelle.

- **Cross-validation temporelle avec une fenêtre étendue.** Les données sont divisées en ensembles d'entraînement et de validation en utilisant des intervalles temporels non consécutifs. Nous augmentons donc la quantité de données d'entraînement au fur et à mesure (voir illustration 4).

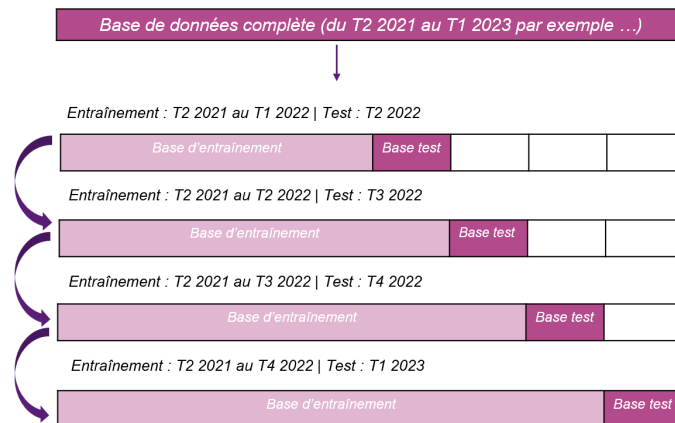


Figure 4 Cross Validation temporelle avec une fenêtre étendue

- **Cross-validation temporelle avec une fenêtre glissante.** Dans cette méthode, nous utilisons une fenêtre de validation qui se déplace à travers les données avec un pas fixe, maintenant ainsi une quantité de données d'entraînement constante. Cette approche est particulièrement utile lorsque la performance du modèle peut varier au fil du temps. En utilisant une fenêtre glissante, nous nous assurons que le modèle est évalué sur des périodes temporelles variées, permettant ainsi d'obtenir une évaluation plus robuste de sa capacité à prédire dans des conditions temporelles changeantes (voir illustration 5).

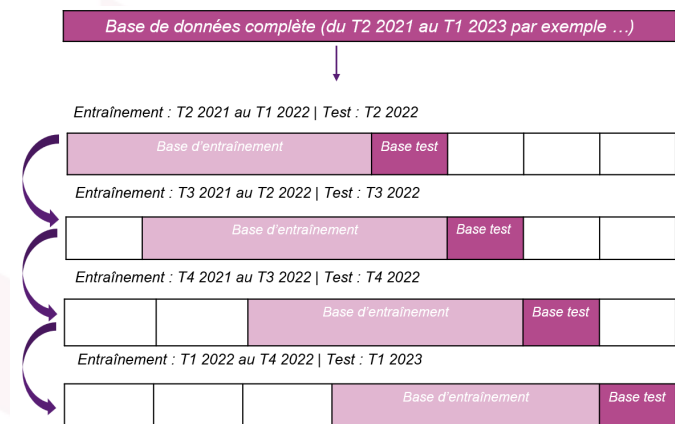


Figure 5 Cross Validation temporelle avec une fenêtre glissante

A la suite de notre étude, la méthode de cross validation temporelle avec fenêtre étendue s'avérerait être la plus appropriée en raison d'un nombre limité d'arrêts. En effet, une fenêtre glissante représenterait une trop grande perte d'information.

2.4 Validation des résultats

Pour évaluer la performance de nos modèles prédictifs, nous avons choisi d'utiliser plusieurs indicateurs : RMSE (Root Mean Square Error), MRE (Mean Relative Error), R^2 (Coefficient de Détermination), et R^2 ajusté. Chacun de ces indicateurs présente des avantages spécifiques qui les rendent particulièrement adaptés à la validation des résultats en machine learning.

Chaque indicateur a été sélectionné pour la raison suivante :

- Le **RMSE** est largement utilisé en machine learning car il est facile à interpréter et permet de comparer les performances de différents modèles. Une valeur de RMSE plus faible indique une meilleure précision prédictive.
- Le **MRE** est particulièrement utile dans les contextes où l'importance de l'erreur est proportionnelle à la taille de la valeur mesurée. Cela permet d'évaluer l'efficacité du modèle indépendamment de l'échelle des données.
- Le R^2 est un indicateur classique et intuitif de la performance des modèles en régression. Il est couramment utilisé pour comparer l'efficacité de différents modèles et comprendre leur capacité explicative.
- Le R^2 **ajusté** est crucial dans les modèles de machine learning où il est possible d'inclure un grand nombre de variables explicatives. Il permet d'évaluer la qualité du modèle tout en tenant compte de la complexité, garantissant que l'ajout de nouvelles variables améliore réellement la performance du modèle et n'ajoute pas simplement du bruit.

S'agissant des modèles, la validation des résultats et notamment l'étude de la stabilité temporelle de ceux-ci a été faite par Cross Validation Temporelle (cf Section 2.3).

En accumulant l'information, nous nous attendons à une diminution du RMSE (Root Mean Squared Error) et du MRE (Mean Relative Error), ainsi qu'à une augmentation du R^2 et du R^2 ajusté. C'est ce que l'on observe sur la Figure 6. En effet, la RMSE diminue de plus en plus, et ce phénomène est particulièrement marqué pour le modèle GP. Ces métriques ont été calculées au cours de la validation croisée temporelle du modèle Random Forest pour la prédiction du BE Euros. L'axe des abscisses est commun à tous les graphiques et indique l'arrêté que nous prédisons. De plus, la figure montre que le R^2 ajusté converge vers une valeur très proche de 1 pour les 3 modèles.

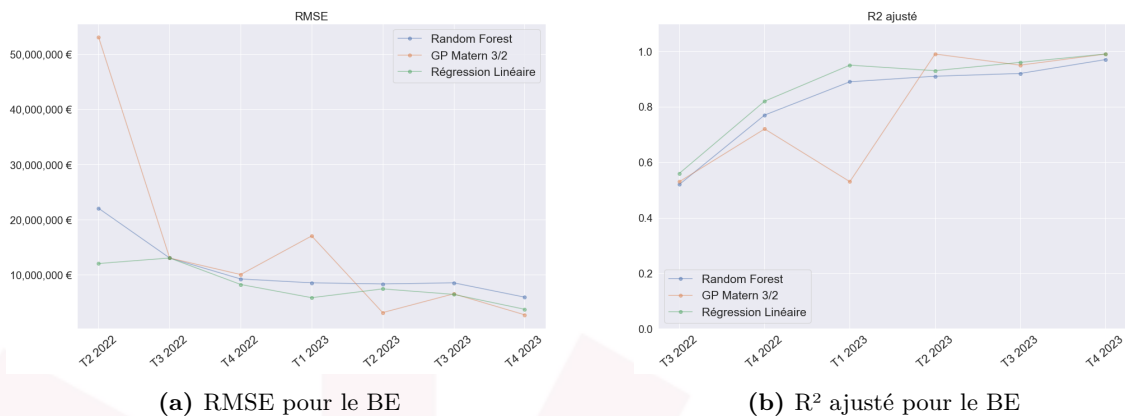


Figure 6 Résultats de Cross Validation temporelle avec fenêtre étendue

Une métrique supplémentaire a été ajoutée car elle est utilisée par le département ALM pour contrôler l'erreur liée à la fuite de modèle. Dans ce contexte, le seuil de tolérance est fixé à 0.02%.

$$\frac{\text{RMSE}}{\text{BE Moyen sur l'arrêté}} \tag{1}$$

2.5 Modèles retenus

Compte tenu de la taille de l'ensemble de données et en s'appuyant sur les conclusions d'autres travaux, plusieurs algorithmes ont été considérés afin de comparer leurs performances et de choisir le plus adéquat. Les modèles qui ont été retenus sont explicités dans le Tableau 2.

Famille de Modèles	Modèles Retenus
Modèles Linéaires	Régression Linéaire Ridge Lasso Elastic Net GLM Gamma
Arbres de Décision	XGBoost Random Forest
Processus Gaussiens	Gaussian Process (Plusieurs noyaux)

Table 2 Modèles retenus, regroupés par famille

Des algorithmes de type réseaux de neurones ont également été implémentés, mais ils n'ont pas permis d'obtenir des résultats satisfaisants. Une des causes probables est la taille de la base de données. En effet, nous ne disposons que d'un millier de lignes, ce qui s'avère insuffisant pour ce type de techniques de régression.

La plupart des modèles étant assez communs ne seront pas présentés mais des éléments peuvent être retrouvés dans Duchesnay et al. (2021). Nous détaillons, en revanche, le fonctionnement de la Régression par Processus Gaussiens (ou GP).

La régression par processus gaussien (GP) est une technique puissante d'apprentissage automatique utilisée pour les tâches de régression. C'est une approche non paramétrique et probabiliste qui peut capturer des relations complexes dans les données. Voici un bref aperçu de son fonctionnement :

1. **Idée Fondamentale** : Dans la régression par processus gaussien, au lieu d'ajuster une fonction spécifique aux données, nous modélisons l'ensemble de la distribution des fonctions possibles. Cette distribution est représentée par un processus gaussien qui définit une préférence sur les fonctions. Le GP est entièrement spécifié par une fonction moyenne et une fonction de covariance (noyau).
2. **Prior** : Au début, le GP suppose une distribution préalable sur l'espace des fonctions. La fonction moyenne représente la valeur attendue de la fonction en chaque point, et la fonction de covariance décrit la similarité entre les valeurs de la fonction à différents points. À mesure que nous recevons des points de données, le GP est mis à jour vers une distribution postérieure sur les fonctions cohérentes avec les données observées. La distribution postérieure est calculée à l'aide du théorème de Bayes.
3. **Prédictions** : Étant donné un ensemble de données d'entraînement et le GP postérieur, nous pouvons faire des prédictions pour de nouvelles données non observées en échantillonnant à partir de la distribution postérieure sur les fonctions. La moyenne de cette distribution représente la valeur prédite, et l'incertitude est donnée par la variance.
4. **Hyperparamètres** : Les Processus Gaussiens ont des hyperparamètres qui contrôlent le comportement du modèle, comme l'échelle de longueur de la fonction de noyau pour chaque dimension du domaine d'entrée. Ces hyperparamètres sont généralement appris à partir des données en maximisant la vraisemblance des données observées.

La Figure 7 présente, en abscisse, les valeurs observées (la cible de test) et, en ordonnée, les valeurs prédites dans ce modèle. Si les prédictions sont précises, nous devrions observer deux droites alignées. Ici, c'est presque le cas : les valeurs sont quasiment alignées à l'exception de quelques-unes. Le point le plus significatif est que toutes les valeurs prédites se situent à l'intérieur de l'intervalle de confiance à 95%. Cet intervalle est la force du GP et garantit la robustesse des prédictions. D'autres éléments sur ce type de modèle peuvent être retrouvés dans Rasmussen and Williams (2006).

Le Tableau 3 récapitule les avantages et inconvénients de la régression par processus Gaussien.

Analyse comparative des valeurs prédites et observées avec GP Matern 5/2

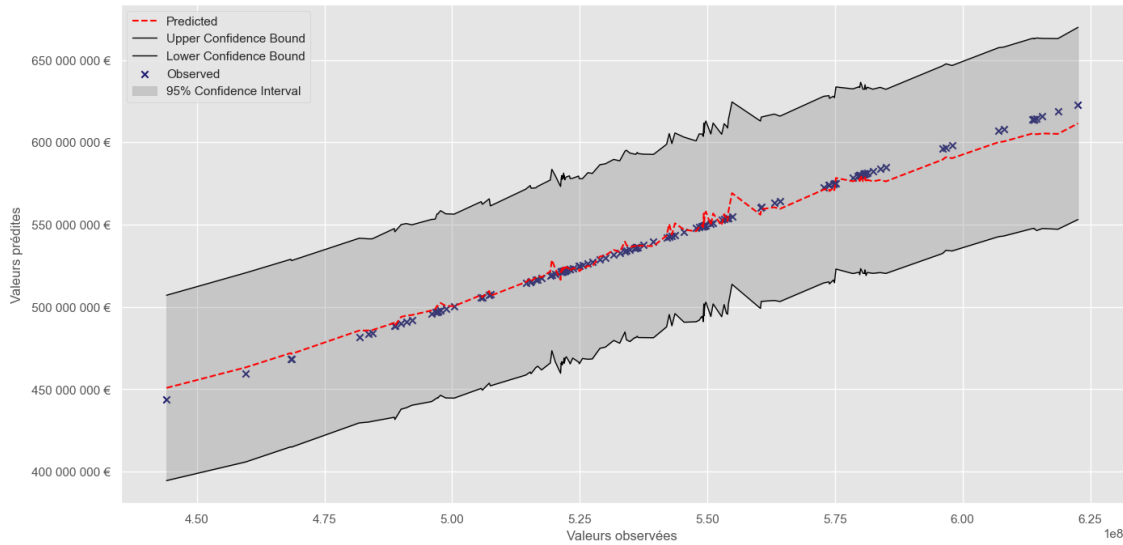


Figure 7 Exemple d'utilisation du GP

Avantages	Inconvénients
<p>Estimation de l'incertitude : Les GP fournissent non seulement des prédictions ponctuelles, mais aussi une mesure d'incertitude pour chaque prédiction. Cela est crucial dans les cas où les données sont bruitées ou limitées.</p> <p>Flexibilité : Les GP peuvent modéliser des fonctions complexes sans spécifier de structure de modèle fixe. Ils peuvent s'adapter aux données.</p> <p>Interpolation et extrapolation : Les GP peuvent interpoler entre les points de données et fournir des prédictions raisonnables même dans les régions sans données d'entraînement.</p>	<p>Complexité informatique : La complexité informatique des GP augmente rapidement à mesure que le nombre de points de données augmente. L'inversion de la matrice de covariance peut être difficile pour les grands ensembles de données.</p> <p>Ajustement des hyperparamètres : La sélection des hyperparamètres du noyau peut être délicate et nécessiter un ajustement manuel.</p> <p>Manque d'interprétabilité : Les GP offrent une interprétabilité pour les modèles plus simples, ils peuvent devenir moins interprétables lorsqu'ils utilisent des noyaux plus complexes ou des données de haute dimension.</p>

Table 3 Avantages et inconvénients des processus gaussiens (GP)

3 Cas d'usages du proxy

Dans cette section, nous présentons différents cas d'usages du proxy en explicitant leur implémentation, les modèles et méthodes retenus, ainsi qu'en analysant les résultats. Les cas d'usages du proxy que nous avons étudiés sont au nombre de quatre :

- Prédiction du niveau de BE (Agrégé, Euros ou UC)
- Prédiction des chocs Solvabilité II
- Prédiction des sensibilités à un arrêté en ayant déjà calculé les chocs règlementaires Solvabilité II
- Prédiction du BSCR

3.1 Prédiction du niveau de BE

La prédiction du niveau de BE est l'objectif principal de cette étude. En effet, cette évaluation est essentielle pour déterminer avec précision les provisions techniques. Elle permet aux assureurs de comprendre les engagements financiers qu'ils ont pris envers les assurés. Compte tenu de la nature du PERP et des besoins de l'assureur, la prédiction du BE a été réalisée à trois niveaux :

- BE UC
- BE Euros
- BE agrégé Euros + UC

L'objectif principal était de pouvoir prédire les BE Euros et UC. Comme mentionné précédemment, le BE UC est entièrement décrit par la PM UC. En conséquence, l'ensemble des modèles utilisés ont fourni d'excellents résultats, et ce, sans effort particulier.

En revanche, la prédiction du BE Euros s'est avérée plus difficile, nécessitant des efforts supplémentaires pour obtenir des résultats satisfaisants. En effet, parmi l'ensemble de nos modèles, seuls deux fournissent des résultats décents : *XGBoost* et *Random Forest*, le reste des modèles ne fonctionnant pas (R^2 négatif et/ou erreur très importante) ou donnant des résultats insatisfaisants. Bien que les résultats obtenus via *XGBoost* et *Random Forest* soient acceptables, ils restent perfectibles. En ce sens, nous avons tenté plusieurs approches : nouvelle sous-sélection de variables explicatives, ajout du BE Euros de l'arrêt précédent en tant que variable explicative (méthode d'ajout du BE(t-1)), prédiction du BE agrégé et du BE UC puis soustraction.

Dans cette perspective, nous avons donc consacré une partie de l'étude à la prédiction du BE agrégé (ou BE). Nous avons constaté que prédire cette quantité directement a permis d'obtenir d'excellents résultats, avec une précision et une robustesse accrues, ce qui confirme l'efficacité de notre approche méthodologique.

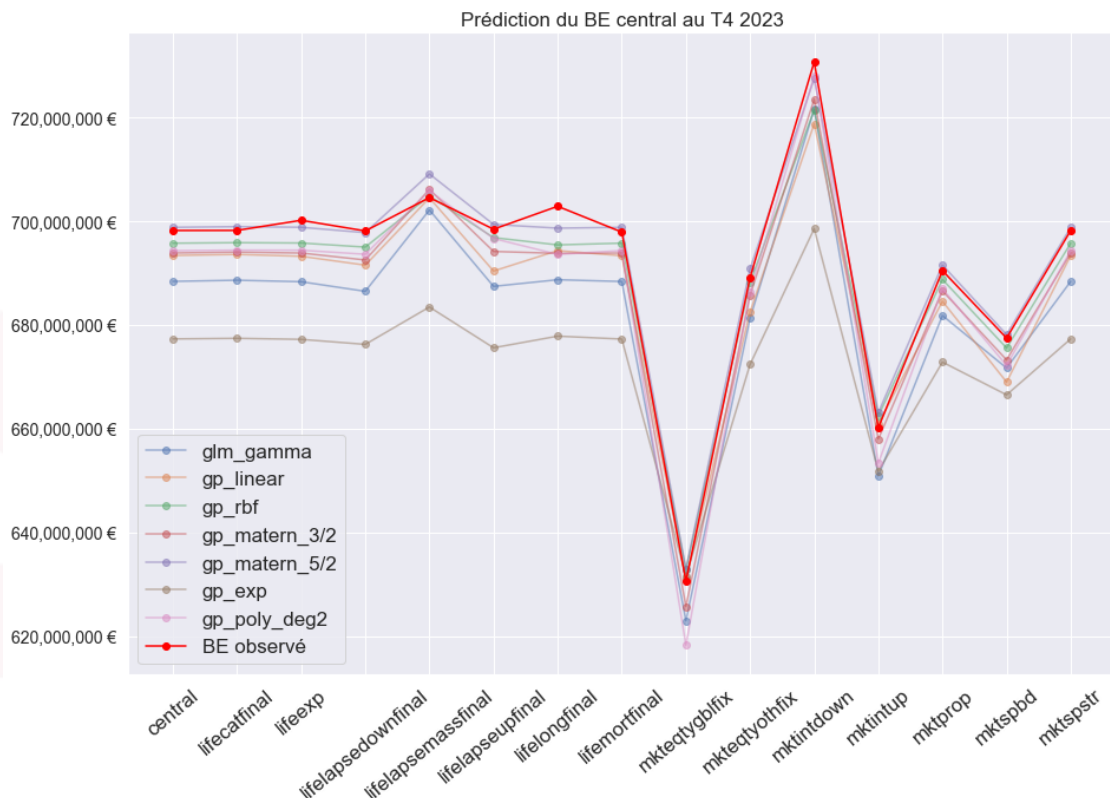


Figure 8 Prédiction du scénario central pour le BE

La Figure 8 souligne l'efficacité potentielle de notre approche pour déduire le BE Euros, ici, au T4 2023. En effet, nous retrouvons ci-dessous les prédictions du scénario central, où le GP se distingue par ses performances supérieures, surpassant les autres méthodes en termes de précision

(cf Figure 12 et 13), ce qui en fait le choix optimal pour notre étude. En utilisant ces résultats et en soustrayant les résultats de prédiction sur la partie UC, nous obtenons le BE Euros.

Nous avons aussi exploré une nouvelle sélection de variables pour le BE Euros. Pour cette sélection, nous nous sommes appuyés sur les résultats des modèles *XGBoost* et *Random Forest* issus de la première sélection (sur la base de l'expertise). En effet, ces modèles permettent de calculer la *feature importance*, ce qui quantifie la contribution de chaque variable explicative à la performance du modèle en mesurant comment la précision de la prédiction est affectée par les variations dans chaque caractéristique. Cela aide à identifier les variables les plus influentes pour la prédiction et à optimiser le modèle en conséquence. En croisant les résultats des importances des caractéristiques des deux modèles, nous avons obtenu un sous-ensemble de variables explicatives pertinent à la fois pour les modèles et d'un point de vue métier.

L'ensemble des résultats de prédiction du BE Euros sont regroupés dans la Table 4 ci-dessous.

BDD utilisée	BDD d'origine	BDD Origine : Agrégé - UC	BDD avec BE(t-1)	Nouvelle sous-sélection 3.1	Nouvelle sous-sélection avec BE(t-1)
Modèle retenu	RF	GP Matern 5/2 & GP Matern 3/2	XGBoost	GP Matern 3/2	GLM Gamma
RMSE Global (en k€)	7 900	4 300	11 000	2 700	3 200
MRE (%)	1,5	0,7	2	0,44	0,52
R ²	0,95	0,98	0,9	0,99	0,99
R ² ajd	0,94	-	0,89	0,99	0,99

Table 4 Meilleurs résultats de prédiction du BE Euro T4 2023 par approche (base) considérée

D'après le tableau ci-dessus, l'approche la plus efficace a été celle liée à une nouvelle sous-sélection de variables.

La prédiction du BE Euros est désormais faite à l'aide des variables explicatives suivantes :

- Taux 1 an
- Taux 10 ans
- Volatilité taux agrégé
- Valeur de marché
- Duration poche obligataire actif
- Duration passif
- PM Euros

La Table 5 présente les résultats finaux de nos prédictions, illustrant la performance de notre approche et mettant en évidence notre proximité avec le seuil de fuite du modèle prédéfini.

Cible	BE Euros	BE UC	BE Agrégé
BDD utilisée	BDD origine nouvelle sous-sélection	BDD d'origine	BDD d'origine
Modèle retenu	GP Matern 3/2	Matern 3/2	GP Matern 5/2
RMSE Global (en k€)	2 700	280	4 400
MRE (%)	0,44	0,25	0,59
R ²	0,99	1	0,99
R ² ajd	0,99	1	0,98
Critère fuite de modèle (Voir 1)	0,59%	0,33%	0,81%

Table 5 Résultats finaux pour le premier cas d'usage

Nous nous concentrons ensuite sur la prédiction des chocs réglementaire S2. Dans l'optique de limiter l'utilisation du modèle ALM infra-annuellement, nous avons exploré la capacité de notre

proxy à prédire uniquement les chocs S2 (voir Annexes A.1 et A.2), en ayant accumulé toute l'information disponible à la date d'utilisation. Les résultats suivants sont issus de la prédiction du BE Agrégé.

BDD utilisée	BDD d'origine	BDD avec BE(t-1)
Modèle retenu	GP Matern 5/2	GLM Gamma
RMSE Global (en k€)	1 700	3 600
MRE (%)	0,25	0,61
R ²	0,99	0,96
R ² ajusté	0,93	0,88

Table 6 Résultats globaux pour le second cas d'usage

Nous constatons que les mêmes modèles utilisés produisent les meilleurs résultats, ce qui renforce leur légitimité. Concernant les résultats, nous observons qu'ils sont satisfaisants et cohérents avec nos attentes. Lorsque nous examinons les résultats pour chaque choc, le constat reste identique, comme illustré dans le Tableau 7.

Chocs S2	MRE (%)
central	0,09
lifecatfinal	0,10
lifeexp	0,19
lifelapsedownfinal	0,05
lifelapsemassfinal	0,65
lifelapseupfinal	0,13
lifelongfinal	0,60
lifemortfinal	0,13
mkteqtygblfix	0,38
mkteqtyothfix	0,24
mktintdown	0,44
mktintup	0,44
mktprop	0,16
mktspsbd	0,10
mktspsstr	0,09

Table 7 Résultats par chocs avec le modèle GP Matern 5/2

Nous observons d'excellents résultats avec une erreur relative moyenne qui, même à son maximum, ne dépasse pas 1%. Cela montre l'adaptation du modèle aux différents scénarios issus des chocs.

Pour conclure sur ce premier cas d'usage, il est important de noter que cette étape a constitué une part substantielle de notre travail. Cependant, les résultats obtenus sont satisfaisants et presque exploitables, car nous nous rapprochons du seuil de tolérance retenu pour la fuite du modèle qui avait été défini, confirmant ainsi la pertinence de notre étude.

3.2 Prédiction des sensibilités connaissant les chocs S2

Dans le même contexte que le cas d'usage précédent, nous avons examiné la capacité du modèle proxy à prédire les sensibilités, en tenant compte des chocs S2 calculés au préalable.

Contrairement aux cas d'usage précédents, la base d'apprentissage a été modifiée. En effet, puisque nous disposons des chocs S2 calculés à l'aide du modèle ALM, nous avons pu intégrer ces informations supplémentaires dans nos modèles durant la phase d'apprentissage.

Pour ce cas d'usage, ce sont une fois de plus les mêmes familles de modèles qui donnent les meilleurs résultats, à l'exception du modèle GP, où nous avons sélectionné le noyau Matern 3/2 plutôt que le noyau Matern 5/2, la différence en termes de MRE n'étant que de 0,06%. En effet, le noyau Matern 5/2 est généralement mieux adapté aux fonctions plus lisses, tandis que le noyau Matern 3/2 est souvent plus approprié pour des fonctions présentant des irrégularités plus marquées². Étant donné la nature des sensibilités observées, ce choix ne nous a pas semblé contre-intuitif.

BDD utilisée	BDD d'origine	BDD avec BE(t-1)
Modèle retenu	GP Matern 3/2	GLM Gamma
RMSE Global (en k€)	3 300	3 900
MRE (%)	0,5	0,58
R ²	0,99	0,99
R ² ajusté	0,99	0,99
RMSE/BE Agrégé Moyen (%)	0,61%	0,72%

Table 8 Résultats globaux pour le troisième cas d'usage

Nous analysons ensuite le détail par sensibilité 9.

Sensibilité	MRE (%)
eqtydown voleqtyup	0,18
intdown	0,63
intdown JST	0,27
intdown200bps	0,17
intup	1,01
intup JST	0,91
intup200bps	0,50
spreadup	0,15
voleqtyup	0,65

Table 9 Résultats par sensibilité pour le troisième cas d'usage

Nous constatons que les sensibilités aux taux à la hausse sont parmi les moins bien prédites. Cela est la conséquence de la récente remontée des taux. En effet, les niveaux de taux les plus récents ne sont pas observés dans l'historique passé, ce qui conduit les modèles à extrapoler davantage, et peut expliquer les prévisions moins précises.

Afin de comparer l'intérêt d'ajouter des chocs S2 à la base d'entraînement, nous avons décidé de tracer les prédictions obtenues avec le GP Matern 3/2.

Note : L'axe des abscisses n'a ici aucune importance car toutes les prédictions correspondent à la même date.

Il est évident que l'ajout des informations sur les chocs réglementaires de Solvabilité II améliore significativement la prédiction des différentes sensibilités. Cela se manifeste sur la Figure 9 :

- Une distance moins importante entre les prédictions et les observations (diminution de la MRE de 0.1%), visible à l'œil nu.
- Une réduction de la taille de l'intervalle de confiance (division par 5 de la demi-largeur moyenne de l'IC), indiquant une diminution de l'incertitude associée aux prédictions et, par conséquent, une précision accrue.

2. GP Matern 5/2 de classe C^2 , GP Matern 3/2 de classe C^1

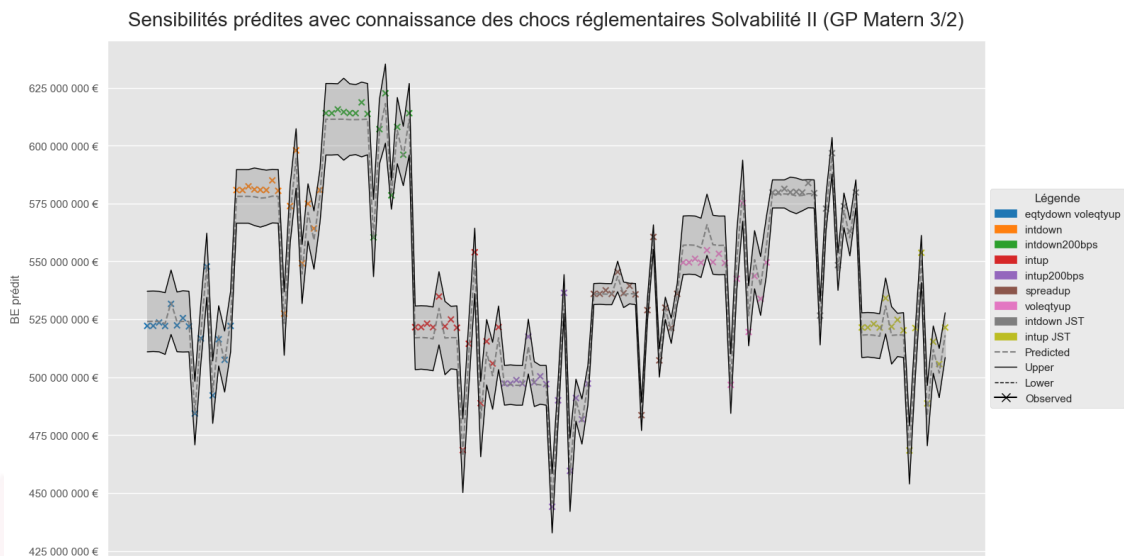
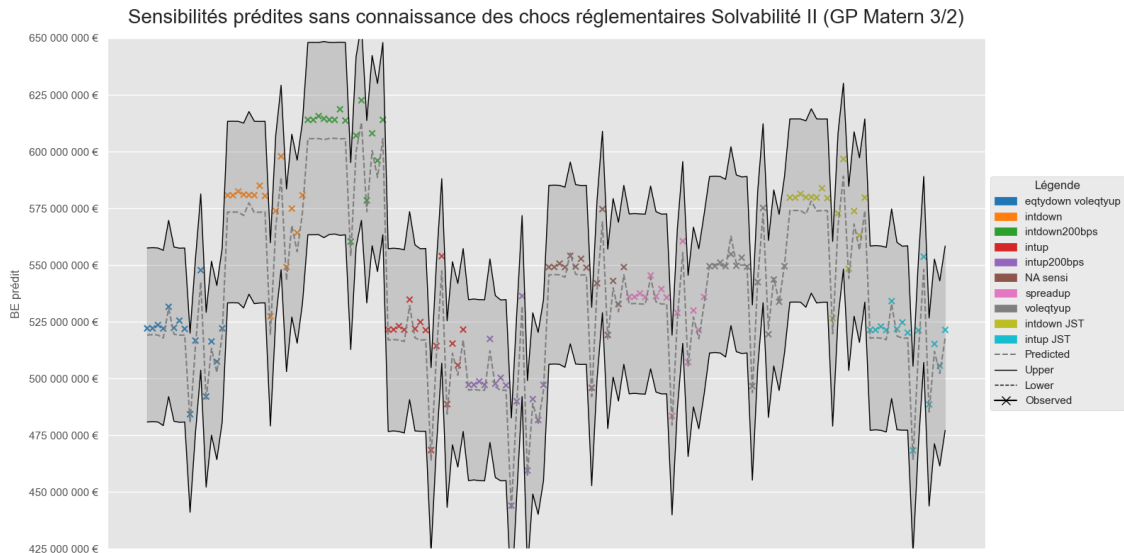


Figure 9 Comparaison de la prédiction des sensibilités au T4 2023 sans (en haut) et avec (en bas) connaissance des chocs Solvabilité II

Les analyses réalisées soulignent l'importance d'intégrer les chocs réglementaires de Solvabilité II dans la base d'entraînement pour améliorer la précision des modèles de prédiction pour le calcul des sensibilités.

En conclusion, cette intégration représente une amélioration importante dans notre approche prédictive, permettant d'obtenir des estimations plus fiables.

3.3 Prédiction du SCR

L'obtention du SCR (Solvency Capital Requirement) avec le proxy est un cas d'usage pertinent pour suivre le risque encouru au sens de Solvabilité II face aux différents chocs proposés par la directive, et plus particulièrement les SCR associés aux différentes sensibilités. C'est un indicateur clé dans le cadre du processus ORSA (Own Risk and Solvency Assessment) pour obtenir une quantification précise des risques auxquels l'assureur est exposé. Notre approche utilise la méthode standard de calcul du SCR.

Le SCR peut être déterminé à partir des Best Estimates que nous avons calculés précédemment, en utilisant les meilleurs résultats obtenus grâce au proxy basé sur les processus gaussiens. Dans un premier temps, nous nous sommes concentrés sur la variation des Best Estimates, qui se trouvent dans la partie "Liabilities" de la formule de l'actif net, et non sur la variation des actifs. En effet, la valeur des actifs n'est pas affectée par les chocs que nous appliquons ; c'est une variable explicative indépendante dans notre analyse. La formule de l'actif net est la suivante :

$$NAV = \text{Assets} - \text{Liabilities}$$

Dans ce contexte, la partie "Assets" correspond à la valeur de marché des actifs, tandis que la partie "Liabilities" correspond aux Best Estimates. Il est important de noter qu'il n'y a pas de distinction entre les fonds investis en euros et ceux investis en unités de compte (UC) dans ce calcul. Nous cherchons à prédire la variation suivante :

$$\Delta NAV_{i,j} = NAV_{\text{central},j} - NAV_{i,j}$$

où i représente les chocs au sens de Solvabilité II et j les sensibilités (voir Annexes A.1 et A.2). L'objectif est de mesurer l'impact d'une sensibilité j en comparant la NAV dans le scénario central à celle obtenue après application du choc i et de la sensibilité j .

Ensuite, nous obtenons le SCR de base, qui est calculé en agrégeant les différents modules de risque de base tels que le risque de marché, le risque de souscription, le risque de contrepartie et d'autres risques financiers, en utilisant les matrices de corrélation spécifiées par la directive Solvabilité II CEIOPS (2009) (cf Table 10 et 11). Le risque opérationnel, quant à lui, n'est pas inclus dans ce calcul de base et fait l'objet d'une évaluation séparée, conformément aux exigences spécifiques de la directive.

	Taux d'intérêt	Action	Immobilier	Spread
Taux d'intérêt	1	0	0	0
Action	0	1	0	0
Immobilier	0.5	0.75	1	0
Spread	0.25	0.25	0.25	1

Table 10 Matrice de corrélation pour le risque de marché dans QIS4.

	Taux d'intérêt	Action	Immobilier	Spread
Taux d'intérêt	1	0	0	0
Action	-0.25	1	0	0
Immobilier	0	0.25	1	0
Spread	0.25	0.25	0.5	1

Table 11 Matrice de corrélation pour le risque de vie dans QIS4.

Notre étude porte sur le calcul du SCR obtenu à partir des modules Vie et du module de marché. Ce sont les chocs que nous avons estimés à l'aide du modèle ALM classique. Enfin, en utilisant la formule standard de calcul du SCR, on obtient le SCR de base :

Pour obtenir le SCR global, il faudrait également inclure le SCR opérationnel. Néanmoins, dans le cadre de notre étude, seul le BSCR (Basic Solvency Capital Requirement) nous intéresse, car il peut être obtenu à l'aide des quantités que nous prédisons.

Comme nous l'avons vu, la quantité servant concrètement à calculer le SCR est la variation de l'actif net. Nous avons donc essayé deux approches de proxy pour le SCR : l'une en partant des prédictions du Best Estimate (BE) générées par le modèle proxy mis en place dans les parties précédentes, et l'autre par une approche directe basée sur la variation de l'actif net. Les résultats des modèles pour la prédiction des Best Estimates sont les mêmes que précédemment. Les résultats de ce cas d'usage montrent qu'un algorithme de machine learning utilisant la génération de forêts

aléatoires est mieux adapté pour prédire la variation de l'actif net, en raison des particularités de cette variable.

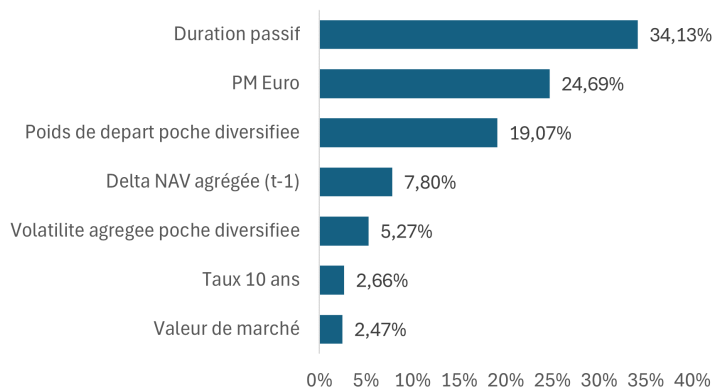


Figure 10 Importance des variables dans la prédiction de la variation de l'actif net

Pour comparer les algorithmes et évaluer la qualité de prédiction des différents modèles, nous avons utilisé des métriques classiques telles que le RMSE et le R^2 . L'approche retenue pour la prédiction de l'actif net repose sur la création d'une variable spécifique, la *Delta NAV agrégée (t-1)*, qui représente la variation de l'actif net agrégé sans aucune sensibilité ajoutée, associée à la période précédente. Cette variable a considérablement amélioré les performances du modèle.

Les résultats obtenus mettent en évidence l'importance des variables explicatives telles que la *Duration passif*, le *PM Euro*, et la *Delta NAV agrégée (t-1)*. Comme illustré dans la Figure 10, ces variables ont une influence déterminante sur la prédiction de la variation de l'actif net.

De plus, la comparaison des SCR entre les différentes approches (cf. Figure 11) montre que l'algorithme utilisant la *Delta NAV agrégée (t-1)* avec le modèle XGBoost pour prédire l'actif net se rapproche le plus du SCR réel, par rapport à l'approche utilisant les prédictions du Best Estimate avec un noyau Matern 3/2. Cela souligne l'efficacité de la méthode de prédiction de l'actif net pour en déduire le SCR. Il est également important de noter que le modèle utilisant les Best Estimate dévie beaucoup, notamment sur la hausse de la volatilité et sur la hausse des taux.

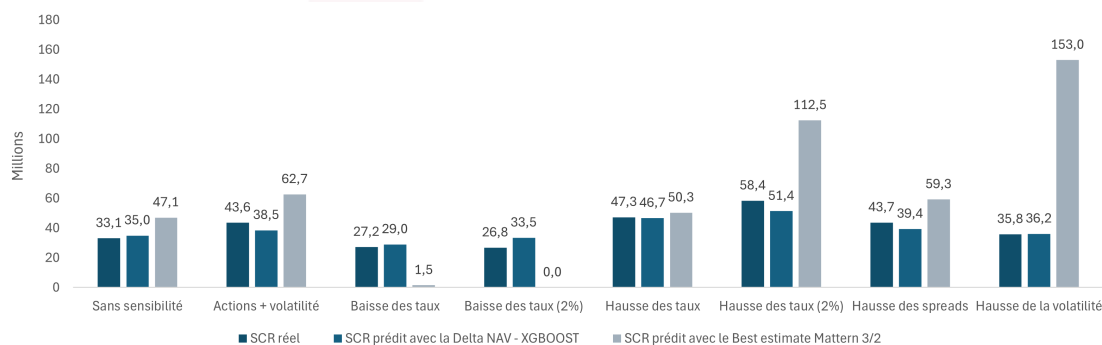


Figure 11 SCR observé et prédit au T4 2023 pour chaque sensibilité

L'analyse des résultats obtenus montre que l'algorithme de machine learning utilisant les forêts aléatoires est particulièrement bien adapté pour prédire les variations d'actifs nets. Les forêts aléatoires ont démontré une meilleure capacité à gérer la variance dans les données de variation d'actif net, fournissant des prédictions plus robustes et précises.

Conclusion

L'objectif de cette étude a été de développer, pour certain cas d'usage, un outil de prédiction du BE et du SCR comme alternative à l'exécution du modèle ALM. Les résultats de prédiction sont globalement satisfaisants, malgré la quantité limitée de données disponibles (seulement huit arrêts). Toutefois, ces conclusions doivent être considérées avec prudence, notamment si l'environnement macroéconomique était amené à évoluer, en particulier pour les taux d'intérêt. En effet, les prédictions correspondant aux sensibilités de taux à plus et moins 200 points de base étaient généralement de moindre qualité comparées à d'autres sensibilités. Il s'agit d'une des limites de notre approche : nous avons étudié une courte période pendant laquelle de nombreux changements macroéconomiques, notamment en ce qui concerne les taux, ont eu lieu. Les niveaux de taux atteints pour certains arrêts et dans certaines sensibilité n'avaient jamais été observés dans l'historique. Avec l'ajout de nouvelles données et l'élargissement de la période d'étude, il serait possible d'améliorer encore la précision des prédictions et mieux adapter le métamodèle aux variations de l'environnement économique. L'intégration de variables supplémentaires et de facteurs externes, comme des indicateurs macroéconomiques ou des données sectorielles spécifiques, pourrait également enrichir le modèle et optimiser sa performance.

D'autres limites liées à l'utilisation d'un "méta-modèle" peuvent être mentionnées. Par exemple, entre le deuxième trimestre 2021 et le quatrième trimestre 2021, nous avons constaté des changements dans les hypothèses de modélisation de la retraite, d'abord modélisée via un proxy, puis intégrée de manière plus réaliste au modèle ALM. Ce changement sur le modèle ALM nous a contraint à exclure cette période de données, réduisant ainsi notre base de données déjà limitée.

Malgré ces défis, notre étude a donné des résultats encourageants qui nous permettent d'envisager avec confiance de potentielles applications sur d'autres portefeuilles d'épargne retraite, comme le PERI.

La véritable efficacité de notre étude se confirmera à mesure que l'outil sera utilisé par les professionnels, ce qui permettra d'enrichir l'expérience de notre algorithme à travers les futurs arrêts.

Enfin, l'intégration de ce modèle dans d'autres cadres normatifs comme IFRS 17 représente une perspective intéressante, bien que cela nécessiterait des études plus approfondies pour adapter pleinement les fonctionnalités du modèle aux exigences comptables spécifiques de cette norme. La norme comptable IFRS 17 impose en effet une rigueur accrue dans la valorisation des engagements d'assurance, et la flexibilité de la méta-modélisation ALM pourrait offrir des solutions pour simuler et prédire des scénarios économiques complexes tout en réduisant le temps de calcul des provisions techniques.

Références

- Arrouy, P.-E., J. Beaudet, M. Bennouna, S. François, and A. Tonin (2023). Reducing the number of scenarios used for stochastic ALM valuation.
- Benyoussef, G. (2023). Elaboration d'un proxy machine learning pour évaluer la sensibilité aux aléas du marché en BEL et partir de sa sensibilité aux aléas du marché.
- Buzzi, A. (2017). Approximation du bilan économique sous Solvabilité II via des méthodes d'apprentissage automatique et application à l'ORSA.
- CEIOPS (2009). CEIOPS' advice for level 2 implementing measures on Solvency II : SCR standard formula.
- Cescutti, V. (2016). Estimés de solvabilité par méta-modélisation.
- Choquer, A. (2019). Méthodes de classification pour l'optimisation des calculs best estimate en assurance-vie.
- Duchesnay, , T. Löfstedt, and F. Younes (2021). *Statistics and Machine Learning in Python*. Springer.
- Ear, P. (2023). Estimation du SCR sous formule standard par des méthodes de machine learning.
- Gauville, R. (2017). Projection du ratio de solvabilité : des méthodes de machine learning pour contourner les contraintes opérationnelles de la méthode des SdS.
- Hounkonnou, P. and H. Behi (2022). Calcul dynamique du SCR en formule standard à base d'apprentissage automatique et interprétabilité des modèles utilisés.
- Kopczyk, D. (2018). Proxy modeling in life insurance companies with the use of machine learning algorithms.
- Kouadio, K. (2018). Méthodes prospectives de calcul de SCR et applications prévoyance.
- Messoussi, A. (2017). Application d'algorithmes de machine learning pour l'estimation du ratio de couverture d'un assureur-vie détenteur d'un produit épargne.
- Mouafo, M. (2020). Estimation de la valeur d'un portefeuille épargne avec une approche machine learning.
- Rasmussen, C. E. and C. K. I. Williams (2006). *Gaussian Processes for Machine Learning*. The MIT Press.

A Annexes

A.1 Définition des chocs S2 considérés dans l'étude

Cette section présente les différents chocs de Solvabilité II et leurs descriptions. Ces chocs sont utilisés pour évaluer la résilience des compagnies d'assurance face à divers scénarios de stress. Les chocs incluent des événements tels que des catastrophes naturelles, des hausses des dépenses en vie, des variations de la longévité et de la mortalité des assurés, ainsi que des chocs sur les marchés financiers.

Choc	Description du choc
Lifecatfinal	Choc catastrophe naturel en assurance vie
Lifeexp	Hausse des dépenses en vie (hausse des frais généraux et de l'inflation)
lifelapsedownfinal	Baisse des rachats au cours de la projection : les assurés restent plus longtemps en portefeuille
lifelapsemassfinal	Choc de rachat massif : rachat très important des contrats en début de projection
lifelapseupfinal	Hausse des rachats au cours de la projection : les assurés restent moins longtemps en portefeuille
Lifelongfinal	Hausse de la longévité des assurés : ils restent plus longtemps en portefeuille
Lifemortfinal	Hausse de la mortalité des assurés : ils restent moins longtemps en portefeuille
mkteqtygblfix	Choc à la baisse des actions type 1 (choc sur la valeur de marché et PM UC action de -39% environ)
mkteqtyothfix	Choc à la baisse des actions type 2 (choc sur la valeur de marché et PM UC action de -49% environ)
Mktprop	Choc à la baisse des actifs immobiliers (choc sur la valeur de marché et PM UC immo de -25%)
Mktspbd	Choc de spread sur les obligations corporate : choc de la valeur de marché selon le rating et la durée
Mktspstr	Choc de spread sur les actifs de titrisation : peu d'impact pour le PERP
Mktintdown	Choc à la baisse de la courbe des taux de départ : valorisation de la valeur de marché des obligations sur cette nouvelle courbe
Mktintup	Choc à la hausse de la courbe des taux de départ : valorisation de la valeur de marché des obligations sur cette nouvelle courbe

Table 12 Descriptif des chocs

A.2 Définition des sensibilités

Les différentes sensibilités sont décrites en Table 13. Ces sensibilités sont utilisées pour évaluer l'impact de variations spécifiques sur les résultats financiers des compagnies d'assurance. Les sensibilités incluent des variations des taux d'intérêt, des spreads, de la volatilité des actions, et des combinaisons de ces variations. A noter que, à partir du T3 2022, les sensis **Intup** et **Intdown** intègrent une hausse de la volatilité des taux de +25% en plus du choc sur la courbe des taux (ce qui les distingue des sensis JST).

Sensibilité	Description
Intdown	Baisse de la courbe des taux de 100 bps ou 50 bps (selon l'arrêté).
Intup	Hausse de la courbe des taux de 100 bps ou 50 bps (selon l'arrêté).
Spreadup	Hausse des spreads de +100 bps.
Voleqtyup	Hausse de la volatilité des actions de +25%.
eqtydown voleqtyup	Baisse de la valeur de marché des actions et hausse de la volatilité des actions de +25%.
intup200bps	Hausse de la courbe des taux de 200 bps.
intdown200bps	Baisse de la courbe des taux de 200 bps.
intup JST	Hausse de la courbe des taux de 100 bps.
intdown JST	Baisse de la courbe des taux de 100 bps.

Table 13 Description des sensibilités.

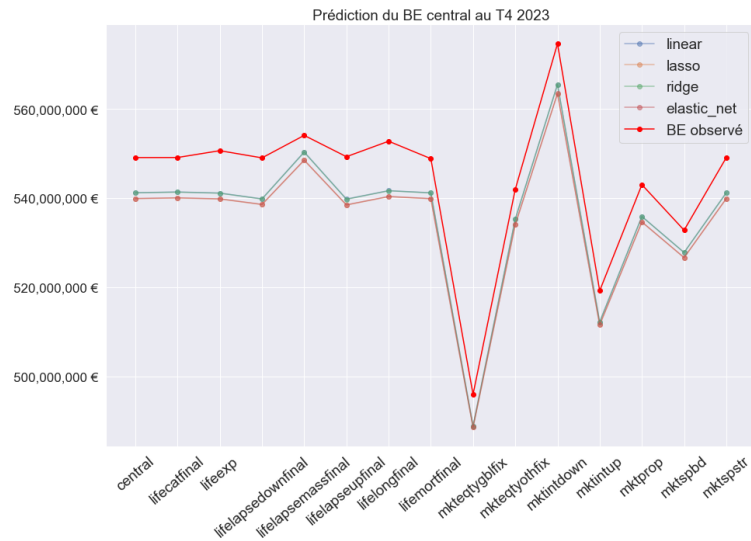


Figure 12 Prédiction du scénario central pour le BE

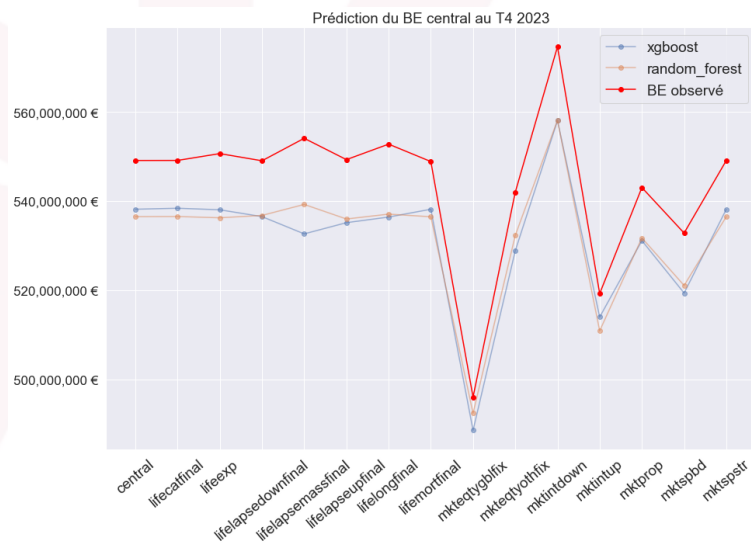


Figure 13 Prédiction du scénario central pour le BE

B Glossaire

ALM	Asset Liability Management
BE	Best Estimate
SCR	Solvency Capital Requirement
PERP	Plan d'Épargne Retraite Populaire
PER	Plan d'Épargne Retraite
NAV	Net Asset Value
S2	Solvabilité II
RMSE	Root Mean Square Error
MRE	Mean Relative Error
R²	Coefficient de détermination
PM	Provision Mathématique
PPE	Provision pour Participation aux Excédents
RC	Réserve de Capitalisation
LSMC	Least Squares Monte Carlo
GP	Gaussian Process
XGBoost	Extreme Gradient Boosting
UC	Unités de Compte
SdS	Simulation dans les simulations
QIS	Quantitative Impact Study
ORSA	Own Risk and Solvency Assessment

Nexialog Consulting est un cabinet de conseil spécialisé en Banque et en Assurance. Organisés autour de 3 domaines d'activité - Risques Bancaires, Financiers & Assurantiels - nous intervenons au sein des équipes métiers afin de les accompagner depuis le cadrage jusqu'à la mise en œuvre de leurs projets. Associant innovation et expertise, le savoir-faire de notre cabinet a permis de consolider notre positionnement sur ce segment et de bénéficier d'une croissance forte et régulière.

Les besoins de nos clients étant en constante évolution, nous nous adaptons continuellement pour proposer le meilleur accompagnement. Le département R&D de Nexialog Consulting se donne pour objectif de proposer des solutions innovantes à des problématiques métier ou d'actualité. Pour cela, nous nous appuyons sur des bibliothèques internes et sur le travail de nos consultants. Le pôle R&D Nexialog a également pour mission de former les collaborateurs sur l'évolution des techniques et la réglementation en lien avec leur activité.

Site web du cabinet : <https://www.nexialog.com>

Publications : <https://www.nexialog.com/publications/>

Contacts

Ali BEHBAHANI
Associé, Fondateur
Tél : + 33 (0) 1 44 73 86 78
Email : abehbahani@nexialog.com

Christelle BONDOUX
Associée, Directrice commerciale
Tél : + 33 (0) 1 44 73 75 67
Email : cbondoux@nexialog.com

Areski COUSIN
Directeur scientifique R&D
Email : acousin@nexialog.com