

Calcul efficace du SCR modèle interne par krigeage stochastique

Nexialog Consulting, Paris, France

30 avril 2021



Résumé

La directive Solvabilité II impose aux sociétés d'assurance un montant minimum réglementaire de fonds propres intitulé *Solvency Capital Requirement* (SCR) qui est calculé par le biais d'une formule standard ou en utilisant un modèle interne intrinsèque à l'entreprise. Dans ce dernier cas, l'évaluation du SCR repose sur des approches simulatoires complexes à mettre en oeuvre. Cependant des techniques alternatives ont été développées pour réduire la complexité des calculs. Parmi ces différentes techniques, on peut citer par exemple, l'accélérateur de SDS¹, l'approche *Curve Fitting* et l'approche *Least Square Monte Carlo* (LSMC). Nous nous plaçons dans le cadre de cette dernière approche. Nous proposons dans un premier temps d'améliorer l'étape regression du LSMC en construisant un *metamodel* du *Best Estimate* à horizon un an à l'aide de la technique de krigeage stochastique. Le krigeage stochastique est une régression par processus gaussien qui est particulièrement intéressante lorsque les variables à expliquer sont données sous forme de simulations de Monte Carlo. Dans un second temps, nous proposons une démarche itérative permettant d'affiner le *metamodel* à une date ultérieure, en exploitant la connaissance acquise sur la fonction *Best Estimate* aux dates d'évaluation précédentes.

1. SDS : Simulations Dans les Simulations.

Table des matières

Introduction	5
1 Metamodelling du BE pour le calcul du SCR	6
1.1 Metamodel du BE avec krigeage stochastique	8
1.1.1 Contexte	8
1.1.2 Illustration du krigeage stochastique	8
1.1.3 Le modèle du Krigeage stochastique	11
1.1.4 Estimation des hyperparamètres du modèle	13
1.2 Metamodel du BE dans l'approche LSMC classique	13
2 Application à la modélisation ALM	14
2.1 Mise en oeuvre de l'approche LSMC	14
2.1.1 Identification des facteurs de risque	14
2.1.2 Choix de la base de régression	15
2.1.3 Base de test	16
2.2 Résultat metamodel BE avec krigeage stochastique	16
2.2.1 Effet des paramètres de la covariance sur la covariance <i>gauss</i>	17
2.2.2 Régression GP avec covariance <i>gauss</i>	18
2.2.3 Régression GP avec covariance <i>exponential</i>	19
2.2.4 Régression GP avec covariance <i>Matern 5/2</i>	20
2.2.5 Régression GP avec covariance <i>Matern 3/2</i>	21
2.2.6 Régression GP avec covariance <i>power exp</i>	22
2.3 Résultat du metamodel du BE avec l'approche LSMC classique	23
2.4 Analyse du meilleur modèle retenu	23
2.4.1 Sélection de modèle	24
2.4.2 Interprétation du meilleur modèle	24
2.4.3 Test sur les sensibilités estimées	25
3 Metamodel du pour le calcul du SCR à une date ultérieure	27
3.1 Étapes intermédiaires à la réalisation du <i>metamodel</i>	28
3.2 Construction optimale de la base d'apprentissage \mathcal{B}_{app} en t'	28
3.3 Illustration du <i>Metamodel</i> du <i>Best Estimate</i> à une date ultérieure	29
4 Metamodel du BE à une date ultérieure t''	32
4.1 Mise en oeuvre de l'approche pour la date t''	32
5 Conclusion	34
6 Annexes	37
6.1 Résultat du metamodel du BE avec l'approche LSMC classique	37
6.1.1 Régression polynomiale	37
6.1.2 Régression polynomiale <i>backward</i>	38
6.1.3 Régression polynomiale Ridge	39
6.1.4 Régression Lasso	41

6.1.5	Validation régression polynomiale	43
	Contacts	44

Abbreviations

ALM Asset Liability Management

BE Best Estimate

GP Gaussian Process

IFRS 17 International Financial Reporting Standard

LSMC Least Square Monte Carlo

NAV Net Asset Value

ORSA Own Risk and Solvency Assessment

SCR Solvency Capital Requirement

Introduction

L'une des modifications majeures apportées par Solvabilité II est que les passifs et les actifs des entreprises d'assurance doivent être évalués en valeur économique selon l'article 75 de la norme. En effet la valeur économique est le montant pour lequel un actif ou un passif pourrait être échangé entre parties bien informées, consentantes et agissant dans des conditions de concurrence normale. Elle introduit donc non seulement une vision économique pour le bilan des sociétés d'assurance mais aussi une vision harmonisée du capital économique qui se définit comme le niveau minimal des fonds propres de base de l'entreprise d'assurance ou de réassurance nécessaire pour faire face à une situation de ruine économique avec un niveau de confiance de 99.5% à horizon un an. Ce capital peut être déterminé selon une approche modulaire (formule standard) ou selon une approche modèle interne (partiel ou non). C'est cette dernière approche qui va retenir notre attention dans cet article.

En effet l'approche "modèle interne" implique une connaissance plus fine des risques de la compagnie d'assurance et nécessite de définir la distribution de la consommation des fonds propres sur un horizon d'un an. La valorisation prospective, c'est à dire la détermination de la distribution des fonds propres à horizon un an, est beaucoup plus délicate et pose de réels défis dont :

- l'absence de formule fermée pour la valorisation des options et garanties présentes dans les passifs d'assurance,
- la valeur économique du bilan est conditionnée à l'information disponible en date d'évaluation, donc elle est aléatoire.

L'objet de cet article est de présenter la mise en oeuvre opérationnelle de la modélisation prospective du bilan économique en utilisant l'approche dite Least Square Monte Carlo (LSMC) qui permet d'en estimer la valeur du capital réglementaire. Cependant nous allons améliorer l'étape de régression de cette approche par des techniques de machine learning telles que la régression par processus gaussien, plus particulièrement le **krigeage stochastique** introduit par [Ankenman et al. \(2010\)](#).

Le krigeage ou regression GP (Gaussian Process) est une technique de machine learning qui se révèle être performante pour traiter certains problèmes d'estimation de risques en finance et en actuariat, en particulier, lorsque l'information sur les risques est incomplète ou coûteuse à obtenir. Parmi les travaux qui se sont intéressés aux applications de la regression GP à l'évaluation des risques financiers ou d'assurance, nous pouvons citer par exemple [Liu and Staum \(2010\)](#), [Croix \(2015\)](#), [Cousin et al. \(2016\)](#), [Ludkovski \(2018\)](#), [Ludkovski et al. \(2018\)](#), [Risk and Ludkovski \(2018\)](#), [Crépey and Dixon \(2020\)](#).

Le krigeage stochastique est particulièrement adapté dans le cas où les sorties de la fonction à estimer sont connues sous forme de simulation de Monte Carlo. Cette approche permet, dans notre contexte d'application, de prendre en compte :

- l'incertitude liée à la régression encore appelée erreur de prédiction,
- l'incertitude liée aux simulations de Monte Carlo des flux futurs actualisés à un an,
- l'estimation des sensibilités afin de distinguer les facteurs de risque qui impactent le plus les flux futurs actualisés à un an.

Les différents objectifs qui motivent la rédaction de cet article sont :

- d'une part, de proposer une méthode de calcul du Solvency Capital Requirement (SCR) prenant en compte l'erreur intrinsèque liée aux simulations de Monte Carlo en utilisant donc le krigeage stochastique ;
- d'autre part, de proposer une estimation du capital économique à des dates ultérieures tout en utilisant la connaissance acquise au cours des évaluations antérieures.

Ce deuxième objectif permet de satisfaire aux exigences réglementaires qui imposent des exercices réguliers tels que le calcul du SCR modèle interne ou formule standard, les exercices Own Risk and Solvency Assessment (ORSA), le calcul des états financiers sous la nouvelle norme International Financial Reporting Standard (IFRS 17). L'idée est d'enrichir, au fil des exercices, une base d'apprentissage, à partir de laquelle, il est possible d'affiner le meta-modèle de krigeage pour

réaliser les exercices suivants avec davantage de précision.

Afin d'illustrer les résultats obtenus au cours de la mise en oeuvre de notre approche, nous avons utilisé un modèle Asset Liability Management (ALM) simplifié mis en place dans le cadre de la valorisation de contrats d'épargne en euros.

1 Metamodelling du BE pour le calcul du SCR

Le but de cette section est d'utiliser l'approche LSMC² avec différentes méthodes d'inférence à savoir des régressions polynomiales et la régression par processus gaussien (Krigage stochastique) afin de déterminer un *proxy* du Best Estimate à horizon un an en fonction des facteurs de risques calibrés et diffusés en monde réel.

Rappelons que nous disposons d'un modèle ALM, auquel nous faisons appel tout au long de cette étude.

Introduction

L'algorithme LSMC (Least Square Monte Carlo) est une méthode de *Monte Carlo* visant à estimer des espérances conditionnelles via un ensemble de fonctions de base (polynômes de Laguerre, polynômes d'Hermite, fonctions trigonométriques).

Cette méthode a été introduite par Longstaff et Schwartz en 2001 afin d'évaluer les espérances conditionnelles relatives aux temps d'arrêt des options bermudiennes dans leur article « *Valuing American Options by Simulation : A Simple Least Square Approach* ». Les assureurs se sont inspirés de cette approche pour adapter la méthode LSMC au calcul du SCR modèle interne.

Afin de faciliter la mise en oeuvre de cette approche, une meilleure connaissance de la cartographie des risques auxquels est soumise la compagnie d'assurance est primordiale pour anticiper l'évolution du bilan économique en fonction des différents risques identifiés. En effet, une composante Y du passif comme les fonds propres de la compagnie ou le *Best Estimate* a une valeur, qui à une date d'évaluation t , peut être vue comme une espérance conditionnelle :

$$\begin{aligned} Y(F_t) &= \mathbb{E}_{\mathbb{Q}} \left[\sum_{u=t+1}^T (CF_u^{out} - CF_u^{in}) \times \delta_{t,u} | F_t \right] \\ &= \mathbb{E}_{\mathbb{Q}} \left[\sum_{u=t+1}^T f_u(Y) \times \delta_{t,u} | F_t \right] \end{aligned} \tag{1}$$

où :

- F_t représente le vecteur des facteurs de risque à l'instant t ,
- T est l'horizon de couverture,
- CF_u^{out} : flux sortants au cours d'année u ,
- CF_u^{in} : flux entrants au cours d'année u ,
- $\delta_{t,u}$: déflateur stochastique entre les dates t et u ,
- $f_u(Y)$: la valeur des flux nets de l'instant u liés à la composante Y
- \mathbb{Q} représente la mesure risque neutre.

En pratique, il n'existe pas de formule analytique pour cette espérance conditionnelle du fait de l'existence des options aux passifs liées aux caractéristiques des contrats (TMG³, participation aux bénéfices, rachats dynamiques)

2. LSMC : *Least Square Monte Carlo*

3. TMG : Taux Minimum Garanti

La composante $Y(F_t)$ peut être estimée par *Monte Carlo* en utilisant l'approximation suivante :

$$Y(F_t) \approx Y_N(F_t) = \frac{1}{N} \sum_{n=1}^N \left[\sum_{u=t+1}^T f_u^n(Y) \times \delta_{t,u}^n | F_t \right] \quad (2)$$

Cette approche est souvent utilisée pour une valorisation du bilan en $t=0$ car :

- la valeur économique des actifs est observée sur le marché.
- le passif est valorisé par *Monte Carlo* en utilisant la formule explicitée ci dessus.

Cependant, lorsque $t \geq 1$, F_t et $Y(F_t)$ deviennent aléatoires. Donc pour déterminer une distribution de $Y(F_t)$, on utilise l'approche SDS dite «Simulation dans les simulations» représentée par le schéma suivant :

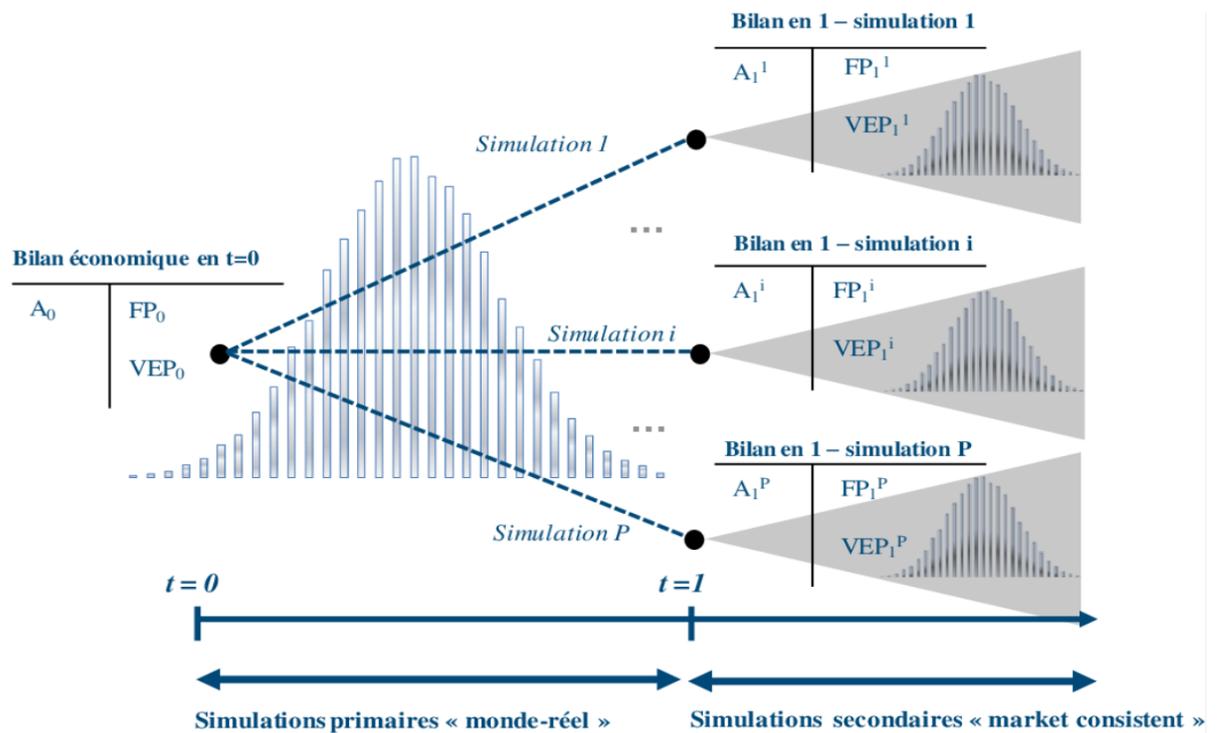


Figure 1 Méthode SDS source : Decupère (2011)

Cette approche présente les avantages suivants :

- elle fournit une première estimation empirique du bilan économique.
- elle donne des résultats précis pour les scénarios monde réel analysés.

Cependant, elle a des inconvénients tels que :

- coût opérationnel important.
- pas de robustesse des queues de distribution.
- pas d'information sur les points situés entre les simulations primaires.

Plusieurs alternatives ont été proposées pour palier à ses difficultés de l'approche SDS. Parmi ces alternatives, l'approche LSMC s'impose progressivement comme la technique de référence pour le modèle interne en Assurance Vie pour plusieurs raisons :

- elle permet d'estimer le SCR ;
- elle peut être utilisée pour les exercices de l'ORSA : couverture du risque, calcul des indicateurs de l'appétence au risque ;
- elle peut être utilisée en ALM pour l'allocation du portefeuille d'actifs.

1.1 Metamodel du BE avec krigeage stochastique

1.1.1 Contexte

Soit $\mathcal{F}_t = \sigma(F_1, \dots, F_t)$, la filtration naturelle des facteurs de risque $(F_t)_{t \geq 0}$. Dans un cadre simplifié, le SCR en t peut s'écrire en fonction de la Net Asset Value (NAV) :

$$SCR_t = NAV_t - VaR_{0.5\%}^{\mathbb{P}}(\delta_{t,t+1} * NAV_{t+1} | \mathcal{F}_t)$$

où

$$\begin{aligned} NAV_{t+1} &= Actif_{t+1} - BE_{t+1} \\ &= Actif_{t+1} - \mathbb{E}_{\mathbb{Q}}(\text{flux futurs}[t+1, \infty] | \mathcal{F}_{t+1}) \\ &= Actif_{t+1} - \mathbb{E}_{\mathbb{Q}}(\text{flux futurs}[t+1, \infty] | F_{t+1}) \quad \text{dans un cadre markovien} \\ &= Actif_{t+1} - Y(t+1, F_{t+1}) \end{aligned}$$

$$SCR_t = NAV_t - VaR_{0.5\%}^{\mathbb{P}}(\delta_{t,t+1} * (Actif_{t+1} - Y(t+1, F_{t+1})) | \mathcal{F}_t)$$

avec \mathbb{P} et \mathbb{Q} respectivement les mesures historique et risque neutre.

- L'objectif est donc de proposer une méthodologie d'estimation de la fonction :

$$Y : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}$$

$$x := (t, F_t) \rightarrow \mathbb{E}_{\mathbb{Q}}(\text{flux futurs actualisés}[t, \infty] | F_t)$$

- Pour parvenir à notre but, nous avons utilisé l'approche LSMC.

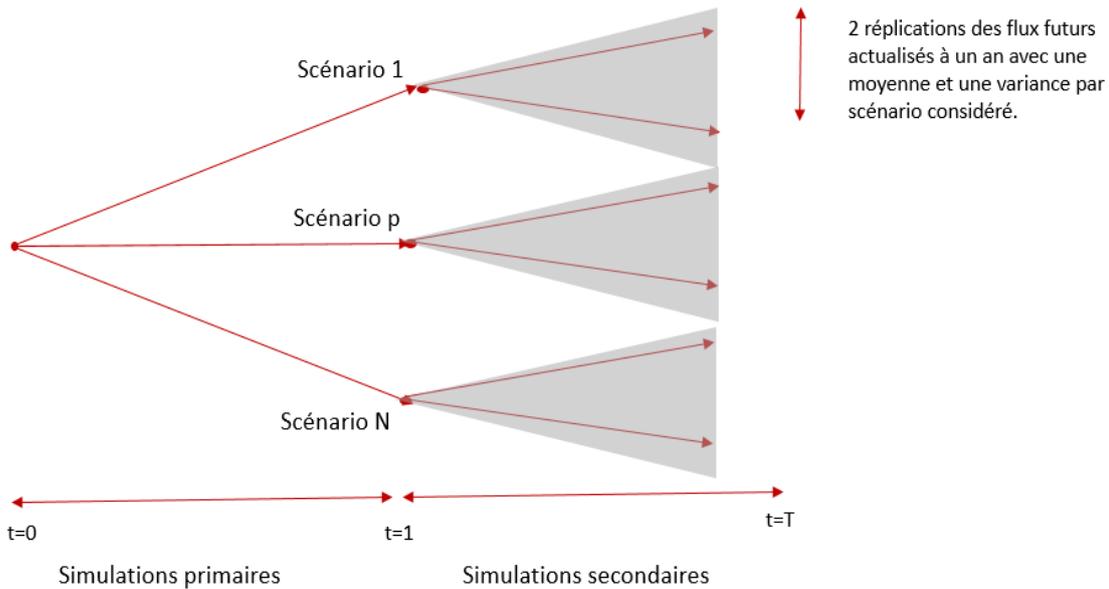


Figure 2 Alternative à la méthode SDS utilisée

- La variable à expliquer (les flux futurs actualisés à un an) est observée sous forme de sorties de simulation Monte Carlo. Nous avons considéré le **krigeage stochastique** afin de prendre en compte l'incertitude associée au caractère aléatoire des flux futurs actualisés.

1.1.2 Illustration du krigeage stochastique

La régression par processus gaussien trouve de plus en plus d'application dans les domaines de l'assurance et de la finance (voir, par exemple, les références citées en introduction). Le krigeage stochastique constitue une sous-famille d'approches de regression GP adaptée à des observations

issues de simulation de Monte Carlo.

Dans notre contexte d'application, nous cherchons à estimer la fonction

$$Y : \begin{aligned} & \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R} \\ & x := (t, F_t) \rightarrow \mathbb{E}_{\mathbb{Q}}(\text{flux futurs actualisés}[t, \infty] | F_t) \end{aligned}$$

sachant que nous disposons d'une information sur la fonction Y donnée sous la forme suivante.

- Pour M valeurs distinctes $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_M)^\top$ de (t, F_t) ;
- Et, pour chaque \mathbf{x}_i , $i = 1, \dots, M$ on observe n_0 réalisations indépendantes des flux futurs actualisés :

$$Y(\mathbf{x}_i) = (Y_1(\mathbf{x}_i), \dots, Y_{n_0}(\mathbf{x}_i)), \quad i = 1, \dots, M$$

Dans notre cas, les points $\mathbf{x}_i = (t_i, F_i)$, $i = 1, \dots, M$, représentent des réalisations distinctes de la variable (t, F_t) et $Y(\mathbf{x}_i)$ représente un vecteur de n_0 réalisations i.i.d. des flux futurs actualisés à partir de la date d'évaluation $t = t_i$, sachant qu'en t , le niveau des facteurs de risque F_t est égal à F_i .

Pour une date t donnée, estimer la fonction Y définie comme une espérance conditionnelle revient à réaliser une régression de la variable "flux futurs actualisés" sur la variable F_t , donc une régression de $Y(\mathbf{x}_i)$ sur $\mathbf{x}_i = (t_i, F_i)$, $i = 1, \dots, M$. Dans cette étude, nous allons comparer différentes approches de régression et montrer qu'il est avantageux d'incorporer la variable t dans la base d'apprentissage.

Afin de bien comprendre l'intérêt du krigeage stochastique pour estimer une fonction dont les sorties sont données sous forme de simulation de Monte Carlo, nous avons comparé plusieurs approches de krigeage, en les appliquant sur un exemple très simple. Nous avons considéré :

- l'estimation d'une fonction Y de \mathbb{R} dans \mathbb{R} ,
- dont les sorties sont observées pour $M = 5$ valeurs distinctes \mathbf{x}_i , $i = 1, \dots, 5$ de la variable d'entrée x
- pour chaque point d'entrée \mathbf{x}_i , $i = 1, \dots, 5$, nous observons $n_0 = 2$ réalisations de la sortie, soit, $Y_1(\mathbf{x}_i)$ et $Y_2(\mathbf{x}_i)$.
- Nous notons $(X, Y_k(X))$ la base d'apprentissage associée à la réalisation k , pour $k = 1, 2$.
- Nous notons $(X, \bar{Y}(X))$ la base d'apprentissage correspondant à la moyenne des réalisations, i.e., $\bar{Y}(X) = \frac{1}{2}(Y_1(X) + Y_2(X))$.

Krigeage classique sur les bases d'apprentissage $(X, Y_1(X))$ et $(X, Y_2(X))$

On réalise ici l'estimation de la fonction Y par krigeage classique, en considérant, d'une part la base d'apprentissage issue de la première réalisation $(X, Y_1(X))$ et, d'autre part, la base d'apprentissage issue de la seconde réalisation $(X, Y_2(X))$. Les résultats sont illustrés ci-dessous. Les points noirs correspondent, pour chaque point d'observation, aux 2 réalisations connues de la fonction. Les courbes en pontillées correspondent à l'estimation fournie par le krigeage classique (moyenne de krigeage) sur chacune de ces 2 bases d'apprentissage. Les zones grise entourant ces estimations représentent les intervalles de confiance ponctuels à 95%.

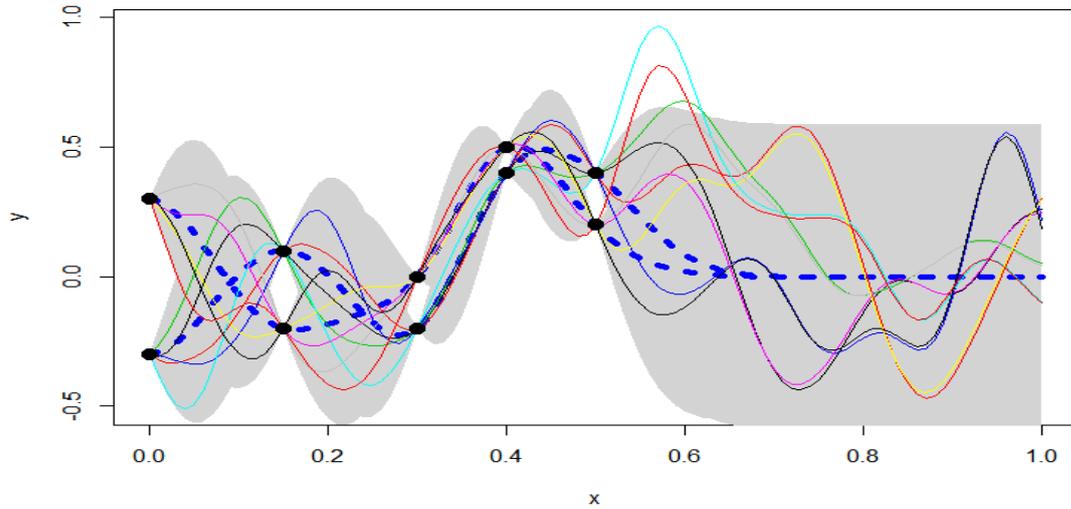


Figure 3 Krigage classique sur chacune des 2 réplifications

On remarque que les fonctions estimées sur chaque base d'apprentissage (chaque réplification) sont très différentes l'une de l'autre. Ce qui est logique du fait de la différence entre les 2 réalisations de Y aux points d'observation.

Krigage classique sur la moyenne des réplifications

Nous réalisons maintenant un exercice de krigage classique sur la base formée par la moyenne des 2 réalisations à chaque point d'observation x_i , $i = 1, \dots, 5$, i.e., sur la base d'apprentissage $(X, \bar{Y}(X))$. Les résultats sont illustrés ci-dessous. La courbe en pontillée correspond à l'estimation fournie par le krigage classique (moyenne de krigage). La zone grise entourant ces estimations représentent les intervalles de confiance ponctuels à 95%. Les points en rouge correspondent aux 2 réalisations de la fonction aux points d'observation.

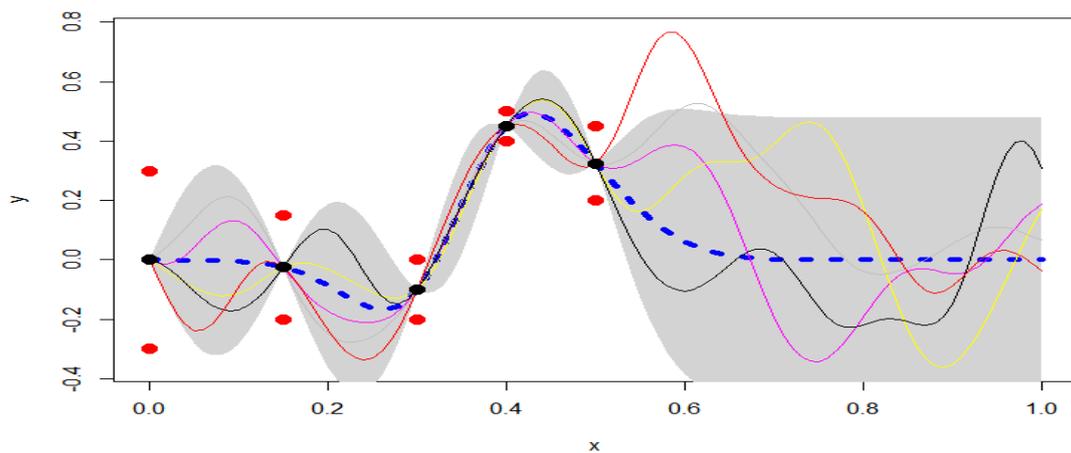


Figure 4 Krigage classique sur la moyenne des réplifications

Dans notre situation, la moyenne, à chaque point d'observation, correspond à l'estimateur de Monte Carlo d'une espérance. Elle est donc entachée d'une incertitude, qui décroît avec le nombre de réalisations retenues. Comme nous pouvons le voir, cette incertitude (qu'il est possible de quantifier dans le cadre des simulations de Monte Carlo) n'est pas prise en compte ici. En effet,

l'estimateur du krigeage interpole bien les valeurs moyennes à chaque point d'observation mais la zone d'incertitude en gris n'est pas du tout réaliste : elle se ressert fortement autour de la moyenne aux points d'observation, alors que nous avons seulement considéré 2 réalisations de Monte Carlo, donc une forte incertitude pour l'estimateur de la moyenne empirique.

Krigeage stochastique sur la moyenne

Nous réalisons maintenant un exercice de krigeage stochastique sur l'ensemble des données. Comme nous allons le voir plus bas, cela revient à réaliser un krigeage classique sur la base d'apprentissage formée par la moyenne $(X, \bar{Y}(X))$ en considérant que la moyenne est une observation bruitée de Y , et que la variance du bruit est fixée a priori à la variance empirique de l'estimateur de Monte Carlo. La courbe en pontillée correspond à l'estimation fournie par le krigeage stochastique. La zone grise entourant ces estimations représentent les intervalles de confiance ponctuels à 95%. Les points en rouge correspondent aux 2 réalisations de la fonction aux points d'observation.

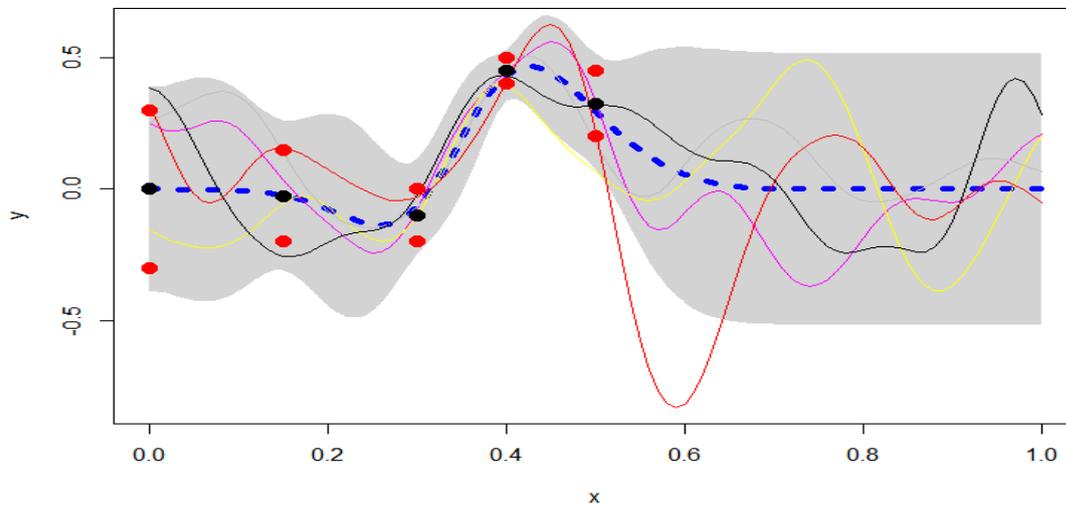


Figure 5 Krigeage stochastique sur la moyenne des réplifications avec bruit hétérogène

Comme nous pouvons le voir, le krigeage stochastique permet de remédier au problème de quantification d'incertitude rencontré par le krigeage classique sur la moyenne. Par ailleurs, la variance du bruit peut varier selon les points d'observation, ce qui n'est pas le cas du krigeage classique qui considère un bruit homoscédastique (donc de variance homogène). A noter qu'il est important, dans ce cadre d'étude, de considérer une approche capable de quantifier l'incertitude de manière réaliste afin de mettre en place une stratégie efficace d'apprentissage, par exemple, en complétant la base d'apprentissage par des points du support où l'incertitude est la plus forte ou bien en simulant davantage de réalisations de Monte Carlo à certains points clés.

1.1.3 Le modèle du Krigeage stochastique

Nous présentons maintenant le modèle de krigeage stochastique. Les développements ci-dessous sont inspirés des travaux de [Liu and Staum \(2010\)](#). Dans le cadre du krigeage stochastique, les hypothèses concernant le modèle Gaussian Process (GP) sont les suivantes :

- La fonction Y à estimer est modélisé par un processus Gaussien $f \sim \mathcal{GP}(m(\cdot), k(\cdot, \cdot))$ où
 - $m(\cdot)$ est la fonction de moyenne
 - $k(\cdot, \cdot)$ est la fonction de covariance : $k(\mathbf{x}, \mathbf{x}')$ mesure la covariance entre $Y(\mathbf{x})$ et $Y(\mathbf{x}')$ en fonction de la distance entre \mathbf{x} et \mathbf{x}' .

- Pour tout point \mathbf{x}_i du champ d'expérience, les réalisations de $Y(\mathbf{x}_i)$ sont des observations bruitées de f :

$$Y_j(\mathbf{x}_i) = f(\mathbf{x}_i) + \varepsilon_j(\mathbf{x}_i), \quad j = 1, \dots, n_0$$

où $\varepsilon_j(\mathbf{x}_i)$, $j = 1, \dots, n_0$ sont i.i.d. de loi $\mathcal{N}(0, \sigma_i^2)$, c'est l'incertitude intrinsèque aux simulations de Monte Carlo.

Le krigeage stochastique consiste à :

- Déterminer la loi a posteriori de f sachant

$$\frac{1}{n_0} \sum_{j=1}^{n_0} Y_j(\mathbf{x}_i) = \frac{1}{n_0} \sum_{j=1}^{n_0} (f(\mathbf{x}_i) + \varepsilon_j(\mathbf{x}_i)), \quad i = 1, \dots, M$$

$$\iff \bar{Y}_i = f(\mathbf{x}_i) + \frac{1}{n_0} \sum_{j=1}^{n_0} \varepsilon_j(\mathbf{x}_i), \quad i = 1, \dots, M$$

$$\iff \bar{Y}_i = f(\mathbf{x}_i) + \bar{\varepsilon}_i, \quad i = 1, \dots, M$$

avec $\bar{\varepsilon}_i = \frac{1}{n_0} \sum_{j=1}^{n_0} \varepsilon_j(\mathbf{x}_i)$

- Déterminer les paramètres de la loi de $\bar{\varepsilon}_i$, $i = 1, \dots, M$.

La variable aléatoire $\bar{\varepsilon}_i$ peut s'écrire sous la forme matricielle $\bar{\varepsilon}_i = A\varepsilon_i$ avec $A = (\frac{1}{n_0}, \dots, \frac{1}{n_0})$ et $\varepsilon_i = (\varepsilon_1(\mathbf{x}_i), \dots, \varepsilon_{n_0}(\mathbf{x}_i))$, ε_i étant un vecteur gaussien car ces composantes sont des lois normales i.i.d., on en déduit facilement que $\bar{\varepsilon}_i$ est aussi un vecteur gaussien car s'écrivant comme combinaison linéaire de vecteur gaussien. On peut donc calculer son espérance et sa variance. soit $\mathbb{E}(\bar{\varepsilon}_i)$ et $VAR(\bar{\varepsilon}_i)$ respectivement l'espérance et la variance de $\bar{\varepsilon}_i$. Alors

$$\begin{aligned} \mathbb{E}(\bar{\varepsilon}_i) &= \mathbb{E}\left(\frac{1}{n_0} \sum_{j=1}^{n_0} \varepsilon_j(\mathbf{x}_i)\right) \\ &= \frac{1}{n_0} \sum_{j=1}^{n_0} \mathbb{E}(\varepsilon_j(\mathbf{x}_i)) \quad \text{linéarité de l'espérance} \\ &= 0 \quad \text{car} \quad \varepsilon_j(\mathbf{x}_i) \sim \mathcal{N}(0, \sigma_i^2) \end{aligned} \tag{3}$$

$$\begin{aligned} VAR(\bar{\varepsilon}_i) &= A * COV(\varepsilon_i) * A^\top \\ &= \left(\frac{1}{n_0}, \dots, \frac{1}{n_0}\right) * \text{diag}(\sigma_i^2, \dots, \sigma_i^2) * \left(\frac{1}{n_0}, \dots, \frac{1}{n_0}\right)^\top \\ &= \left(\frac{\sigma_i^2}{n_0}, \dots, \frac{\sigma_i^2}{n_0}\right) * \left(\frac{1}{n_0}, \dots, \frac{1}{n_0}\right)^\top \\ &= \frac{\sigma_i^2}{n_0} \end{aligned} \tag{4}$$

Donc $\bar{\varepsilon}_i \sim \mathcal{N}(0, \frac{\sigma_i^2}{n_0})$

- Estimation de la variance σ_i^2 avec les répliquions par point d'expérience \mathbf{x}_i par

$$\hat{\sigma}_i^2 = \frac{1}{n_0 - 1} \sum_{j=1}^{n_0} (Y_j(\mathbf{x}_i) - \bar{Y}_i)^2, \quad i = 1, \dots, M$$

Après l'entraînement du modèle sur la base \mathbf{X} et $\bar{\mathbf{y}} = (\bar{Y}_i, i \in [1, M])$, on obtient la formule de la loi a posteriori de f qui s'énonce de la manière suivante :

$$p(f|\bar{\mathbf{y}}, \mathbf{X}) = \frac{p(\bar{\mathbf{y}}|f, \mathbf{X})p(f)}{p(\bar{\mathbf{y}}|\mathbf{X})}$$

avec :

- Loi a priori : $p(f) \sim \mathcal{GP}(m(\cdot), k(\cdot, \cdot))$
- La vraisemblance du modèle : $p(\bar{\mathbf{y}}|f, \mathbf{X}) \sim \mathcal{N}(f(\mathbf{X}), V)$, $V = \text{diag}(\frac{\sigma_1^2}{n_0}, \dots, \frac{\sigma_M^2}{n_0})$
- La vraisemblance marginale : $p(\bar{\mathbf{y}}|\mathbf{X}) = \int_f p(\bar{\mathbf{y}}|f, \mathbf{X})p(f|\mathbf{X})df$

1.1.4 Estimation des hyperparamètres du modèle

- **Etape 1** : on estime V par $\text{diag}(\frac{\hat{\sigma}_1^2}{n_0}, \dots, \frac{\hat{\sigma}_M^2}{n_0})$
- **Etape 2** : on estime les autres hyperparamètres θ en maximisant la vraisemblance marginale :

$$p(\bar{\mathbf{y}}|\mathbf{X}, \theta) = \int_{\theta} p(\bar{\mathbf{y}}|f, \mathbf{X})p(f|\mathbf{X}, \theta)d\theta$$

Ainsi on cherche les hyperparamètres θ du GP tels que :

$$\theta^* \in \arg \max_{\theta} \log p(\bar{\mathbf{y}}|\mathbf{X}, \theta)$$

où

$$\log p(\bar{\mathbf{y}}|\mathbf{X}, \theta) = -\frac{1}{2}(\bar{\mathbf{y}} - \mathbf{m}_{\theta}(\mathbf{X}))^{\top}(\mathbf{K}_{\theta} + V)^{-1}(\bar{\mathbf{y}} - \mathbf{m}_{\theta}(\mathbf{X})) - \frac{1}{2} \log |\mathbf{K}_{\theta} + V| + \text{const}$$

avec :

- \mathbf{m}_{θ} la fonctionne de moyenne de la loi à priori conditionnée aux paramètres θ
- \mathbf{K}_{θ} : la fonctionne de covariance de la loi à priori conditionnée aux paramètres θ

1.2 Metamodel du BE dans l'approche LSMC classique

Cette technique consiste à déterminer un proxy de la fonction Best Estimate (BE), vue comme une espérance conditionnelle, par une combinaison linéaire de fonctions des facteurs de risques. Rappelons que la fonction Y à estimer peut se mettre sous la forme :

$$Y(X) = \mathbb{E}_{\mathbb{Q}} \left[\sum_{u=t+1}^T f_u(Y) \times \delta_{(t,u)} | X \right]$$

Notons que cette espérance conditionnelle peut s'interpréter comme la projection orthogonale de la variable aléatoire $Z = \sum_{u=t+1}^T f_u(Y) \times \delta_{(t,u)}$ sur l'espace des variables aléatoires $\sigma(X)$ -mesurables et de carré intégrable. La regression LSMC consiste à :

- Simuler par Monte Carlo, des réalisations du couple (X, Z) ;
- Considérer une approximation de la fonction Y comme une combinaison linéaire (de taille finie) de fonctions de base sur X :

$$Y^N(X) = \sum_{i=1}^N \alpha_i \times L_i(X) \tag{5}$$

où N est le nombre de fonctions de base considérés, $L_i, i = 1, \dots, N$, la famille de fonctions de base retenue pour la régression, $\alpha_i, i = 1, \dots, N$, des coefficients réels ;

- Déterminer les $\alpha_i, i = 1, \dots, N$ qui minimisent l'écart quadratique entre les réalisations de Z et de $Y^N(X)$.

Ainsi, en considérant les notations suivantes :

- $N \in \mathbb{N}^*$: est le nombre de régresseurs ;
- $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_M)^{\top} \in \mathbb{R}^M \times \mathbb{R}^K$: la matrice constituée des réalisations de la variable aléatoire X (composée de K facteurs de risque), avec $\mathbf{x}_i = (x_i^1, \dots, x_i^K)$, les réalisations des différents facteurs de risque pour l'observation $i, i = 1, \dots, M$;
- $Y(\mathbf{x}_i), i = 1, \dots, M$: la réalisation de Z pour $X = \mathbf{x}_i$;
- $L(X) = (L_1(X), L_2(X), \dots, L_N(X))$: une série de fonctions appelée base de régression ;
- $\alpha = (\alpha_1, \dots, \alpha_N)$: le vecteurs des coefficients de régression donnant la sensibilité de Y aux fonctions de base $L_i(X)$;

les coefficients α sont estimés par une régression linéaire des $Y_N(\mathbf{x}_i)$ sur les $Y(\mathbf{x}_i)$, soit :

$$\hat{\alpha} := \arg \operatorname{Min}_{\alpha} \sum_{i=1}^M \left(Y(\mathbf{x}_i) - L(\mathbf{x}_i)^{\top} \cdot \alpha \right)^2 \quad (6)$$

Comme nous avons plusieurs réplifications de Z pour chaque valeur d'entrée \mathbf{x}_i , nous avons choisi dans l'approche LSMC classique, de considérer la moyenne de ces réplifications comme observations de la valeur de sortie, soit :

$$\hat{\alpha} := \arg \operatorname{Min}_{\alpha} \sum_{i=1}^M \left(\bar{Y}_i - L(\mathbf{x}_i)^{\top} \cdot \alpha \right)^2 \quad (7)$$

avec $\bar{Y} = (\bar{Y}_1, \dots, \bar{Y}_M)$ la vecteur moyenne des n_0 réplifications par point d'expérience \mathbf{x}_i , $i = 1, \dots, M$.

2 Application à la modélisation ALM

Dans cette section, nous allons mettre en oeuvre la méthode LSMC en considérant 2 approches de régression pour la construction du proxy :

- le krigeage stochastique
- la régression polynomiale

Concernant cette dernière approche, nous considérons dans la suite :

- $L(\mathbf{x}_i) = (L_1(\mathbf{x}_i), L_2(\mathbf{x}_i), \dots, L_N(\mathbf{x}_i))$ base de régression de l'approche LSMC classique avec régression polynomiale
- $L_j(\mathbf{x}_i)$ pour $1 \leq j \leq N$ une fonction polynomiale de degré inférieure ou égale à 3 du vecteur \mathbf{x}_i

2.1 Mise en oeuvre de l'approche LSMC

La mise en oeuvre de l'approche LSMC nécessite plusieurs étapes intermédiaires dont :

- identification et calibration des facteurs de risque en monde réel et en risque neutre.
- simulation des scénarios des facteurs de risque en monde réel et en risque neutre.
- lancer le modèle de projection ALM pour obtenir la réalisation des flux futurs actualisés.
- choisir la base de régression pour les flux futurs actualisés.
- effectuer l'apprentissage statistique des flux futurs actualisés sur cette base de régression.
- valider l'apprentissage sur une base de validation.
- utiliser le proxy pour calculer la distribution du BE_1 et par conséquent le SCR.

2.1.1 Identification des facteurs de risque

Les facteurs de risque pris en compte au cours de la mise en place du modèle ALM dans le contexte des contrats d'Épargne Euro sont :

- **Facteurs de risques action** : les log rendements du CAC 40 calibrés au niveau du GSE monde réel action. En effet, nous avons simulé des réalisations de l'évolution du cours du CAC à horizon un an, ainsi nous avons pu déduire les log rendements associés à chaque simulation considérée.
- **Facteurs de risques Taux** : les trois composantes calibrées avec la méthode de Nelson Siegel au cours de la calibration monde réel. En effet, on distingue :
 1. Le facteur de niveau (pc1)
 2. Le facteur de pentification (pc2)
 3. Le facteur de convexité (pc3)

- **Facteurs de risque mortalité** : Pour calibrer le facteur de risque mortalité, nous nous sommes appuyés sur l’approche développée dans [Nexialog Consulting \(2021\)](#). Les différentes étapes sont décrites ci dessous :
 - Projection des taux de mortalité pour tous les âges à horizon un an grâce à GP⁴.
 - Calcul de la variation des taux de mortalité entre la table centrale (réglementaire) et les taux projetés sur l’horizon considéré.
 - Calibration d’une loi sur cette variation afin de choquer tout le portefeuille du même facteur.
- **Facteurs de risque rachat** : On a retenu un cas très simple où on a simulé une loi uniforme entre 0 et 50% qui représente les différentes évolutions possibles de notre loi de rachat à horizon un an . Ainsi, ces rachats sont appliqués sur notre portefeuille.

On dispose donc de $K=6$ facteurs de risque constituant la matrice des observations X et qui seront utiles pour les étapes de calibration du *proxy* avec les flux futurs actualisés.

2.1.2 Choix de la base de régression

Nous avons fait le choix de deux bases de régression totalement différentes pour la régression polynomiale et pour le Krigeage stochastique. En effet le krigeage stochastique étant une régression non linéaire, elle pourrait capter l’effet non linéaire des interactions actif et passif sans modification des variables, cependant une transformation des variables explicatives est nécessaire dans le cas de la régression polynomiale.

Soit $x_i = (Action_i, PC1_i, PC2_i, PC3_i, Mortalite_i, Rachat_i)$ une observation de notre base de données.

Rappelons que ces observations sont des projections à un an issues de la calibration de l’évolution historique des différents facteurs de risque. En effet pour :

- pour les actions, nous avons réalisé nos travaux de calibration sur les données mensuelles cac 40 collectées sur le site <https://fr.investing.com/indices/france-40-historical-data> allant de la période de Août 1987 à Avril 2020.
- pour les facteurs de Nelson Siegel sur les taux, les données zéro coupon utilisées pour la calibration de l’évolution historique des taux ont été extraites sur le site <http://www.cnofrance.org/fr/courbes-des-taux-zéro-coupon.cfm>. Ce sont des données mensuelles allant de la période de Février 2012 à Avril 2020. Ces données ont été retraitées pour les maturités non liquides car la méthode d’extrapolation des zéros coupons pratiquée sur le site n’est la même que celle préconisée par l’EIOPA.
- pour la mortalité, nous avons utilisé les données empiriques sur la mortalité de la France sur le site HMD allant de la période 1995 à 2017 et couvrant les ages de 0 à 110.
- le risque de rachat structurel, nous avons fait le choix d’une loi uniforme sur $[0,0.5]$ sur l’horizon d’un an

Au niveau des différentes régressions réalisées, les bases choisies sont les suivantes :

- pour la régression linéaire on a choisit comme base orthogonale de projection des flux futurs actualisés, la fonction :

$$L_j : \begin{matrix} \mathbb{R}^6 \rightarrow \mathbb{R} \\ x_i \rightarrow L_j(x_i) \end{matrix} \quad j \in [1, N], \quad i \in [1, M]$$

avec L_j comme une combinaison des éléments de x_i de degré inférieur ou égal à 3. En procédant de cette manière, $N=120$ variables explicatives composées des combinaisons des éléments de x_i .

- pour le krigeage stochastique, on a choisi de conserver $x_i, \quad i \in 1, \dots, M$ comme étant la base de régression. Ainsi on a donc six (6) variables explicatives dans ce cas.

4. GP : *Gaussian process*

2.1.3 Base de test

Nous disposons d'une base de test \mathcal{B}_{test} qui nous permet d'apprécier la performance des différents *proxy* calibrés. Elle a été construite de la manière suivante :

- simulation de M=100 trajectoires des facteurs de risque monde réel à horizon un an différentes de celles de la base d'apprentissage.
- utilisation du modèle ALM afin de déterminer les flux futurs actualisés à un an associés à ces différentes trajectoires.

2.2 Résultat metamodel BE avec krigeage stochastique

Dans le cadre de l'estimation de la fonction $Y(t, F_t)$, nous avons testé plusieurs modèles de régression GP avec différentes fonctions de covariance afin de déduire le modèle qui maximise la vraisemblance sur nos données.

La base de régression du GP est x_i , $i \in 1, \dots, M$ avec,

$$x_i = (Action_i, PC1_i, PC2_i, PC3_i, Mortalité_i, Rachat_i)$$

En effet nous avons utilisé une seule fonction de moyenne couplée avec plusieurs fonctions de covariance.

La fonction de moyenne est exprimée de la manière suivante :

$$\begin{aligned} m(x) &= \beta_0 + x^\top * \beta \\ &= \beta_0 + \beta_1 * Action + \beta_2 * PC1 + \beta_3 * PC2 + \beta_4 * PC3 + \beta_5 * Mortalite + \beta_6 * Rachat \end{aligned}$$

Afin de bien présenter les fonctions de covariance, nous allons d'abord introduire la fonction distance d telle que :

$$\begin{aligned} &\mathbb{R}^6 * \mathbb{R}^6 \rightarrow \mathbb{R} \\ d : (x, x') &\rightarrow \sum_{k=1}^6 \left(\frac{x_k - x'_k}{\ell_k} \right)^2 \end{aligned}$$

où ℓ_k , $1 \leq k \leq 6$ sont des paramètres de corrélation positifs.

Les différentes fonctions de covariance testées sont :

- Fonction de covariance *gauss* :

$$\begin{aligned} &\mathbb{R}^6 * \mathbb{R}^6 \rightarrow \mathbb{R} \\ k_{Gauss} : (x, x') &\rightarrow \sigma_f^2 \exp \left(-\frac{1}{2} d(x, x') \right) \end{aligned}$$

- Fonction de covariance *exponential*

$$\begin{aligned} &\mathbb{R}^6 * \mathbb{R}^6 \rightarrow \mathbb{R} \\ k_{exp} : (x, x') &\rightarrow \sigma_f^2 \exp \left(-d(x, x')^{\frac{1}{2}} \right) \end{aligned}$$

- Fonction de covariance *matern 3/2*

$$\begin{aligned} &\mathbb{R}^6 * \mathbb{R}^6 \rightarrow \mathbb{R} \\ k_{Mat,3/2} : (x, x') &\rightarrow \sigma_f^2 \left(1 + \sqrt{3} * d(x, x')^{\frac{1}{2}} \right) \exp \left(-\sqrt{3} * d(x, x')^{\frac{1}{2}} \right) \end{aligned}$$

- Fonction de covariance *matern 5/2*

$$\begin{aligned} &\mathbb{R}^6 * \mathbb{R}^6 \rightarrow \mathbb{R} \\ k_{Mat,5/2} : (x, x') &= \sigma_f^2 \left(1 + \sqrt{5} * d(x, x')^{\frac{1}{2}} + \frac{5}{3} * d(x, x') \right) \exp \left(-\sqrt{5} * d(x, x')^{\frac{1}{2}} \right) \end{aligned}$$

- Fonction de covariance *power exp*

$$k_{\text{exp_puissance}} : \mathbb{R}^6 * \mathbb{R}^6 \rightarrow \mathbb{R}$$

$$(x, x') = \sigma_f^2 \exp\left(-d(x, x')^{\frac{\ell}{2}}\right)$$

2.2.1 Effet des paramètres de la covariance sur la covariance *gauss*

On va discuter des effets des paramètres de la fonction de covariance sur la fonction f .

Paramètres de corrélation ℓ : comme nous montre la figure 6, on s'aperçoit que plus ℓ augmente plus la fonction se rapproche de 1, c'est à dire que les observations $f(x_i)$ sont très corrélés. Le paramètre d'échelle nous donne une idée sur la distance à parcourir sur un axe(variables explicatives) pour que deux points ne soient plus corrélés, ainsi si ℓ est donc grand, il faut donc parcourir plus de distance sur cet axe afin de trouver des points non corrélés.

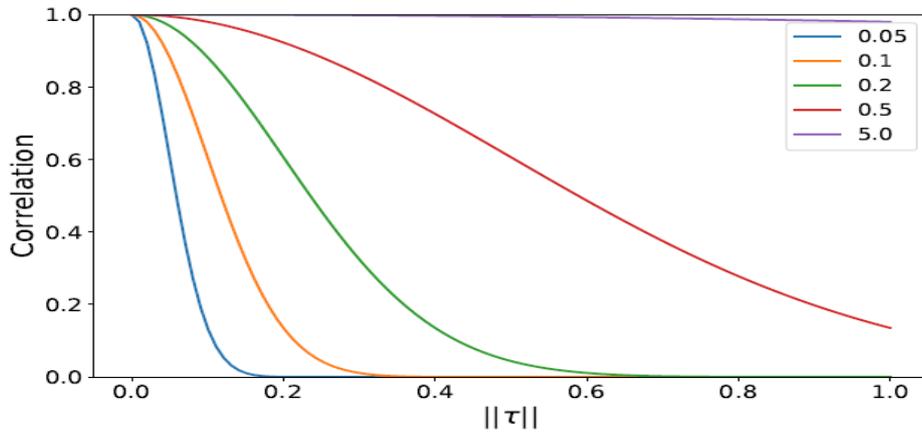
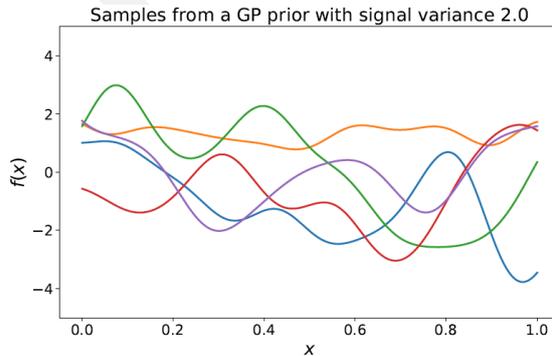


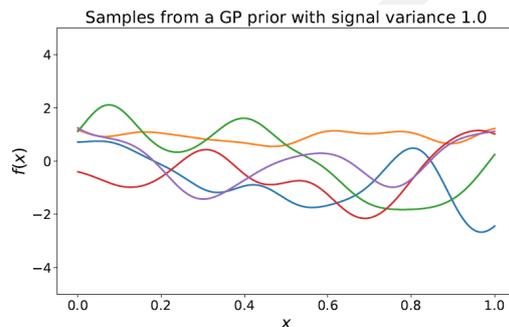
Figure 6 Effet du paramètre de corrélation

avec $\|\tau\| = \exp\left(-d(x, x')\right)$

signal variance σ_f On remarque qu'avec un σ_f petit (7b), on observe une faible volatilité de la fonction, tandis que la volatilité est plus grande avec un σ_f grand (7a). On peut donc en déduire que l'erreur de prédiction est faible pour σ_f est petit.



(a) Effet de σ_f



(b) Effet de σ_f

2.2.2 Régression GP avec covariance *gauss*

Les résultats obtenus avec le krigeage gaussien pour l'estimation de la fonction du Best Estimate $Y(t, F_t)$ sont renseignés dans le tableau suivant :

Variables	β	ℓ
Intercept	101112817.918	
Action	13668409.274	2.9299
PC1	-119444717.068	0.0001
PC2	89543544.167	0.0002
PC3	158338.250	0.0561
Mortalité	-620673.620	1.325
Rachat	-101235205.927	0.644

GP <i>gauss</i>	σ_f	Log Vraisemblance
Valeurs	0.007	-15703.16

Table 1 Hyperparamètres *gauss*

En traçant sur le même graphe, les prédictions du modèle, les observations et l'intervalle de confiance à 95% eu utilisant la matrice de covariance à posteriori sur la base de test \mathcal{B}_{test} , on obtient le schéma suivant :

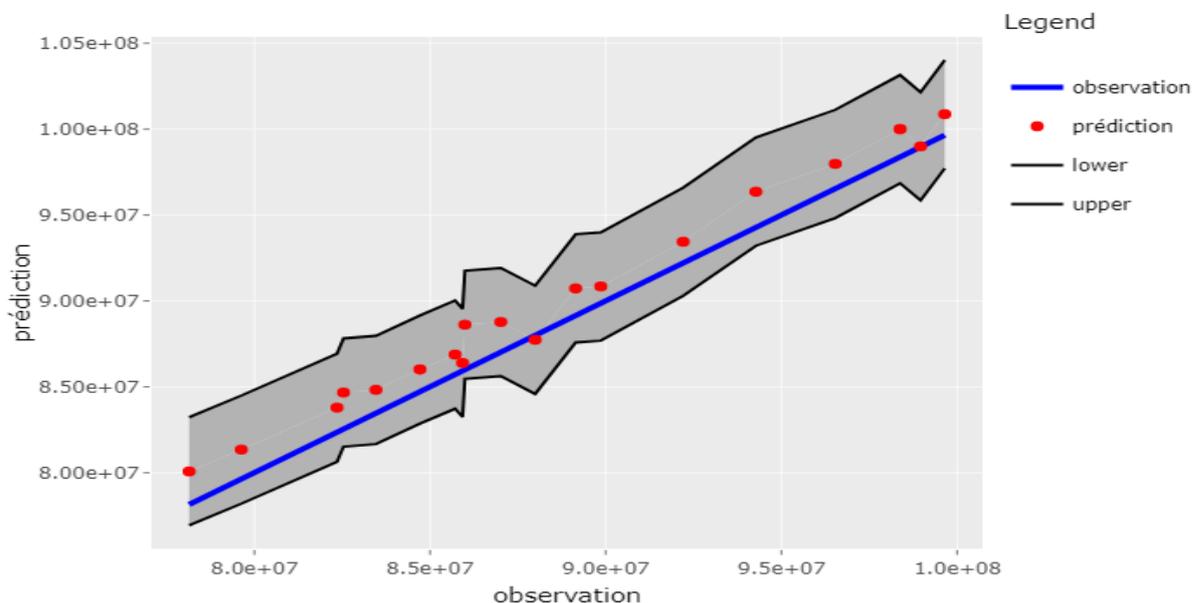


Figure 8 Ajustement et intervalle de confiance pour le Krigeage gaussien

On remarque que l'ajustement est de bonne qualité et que toutes les observations sont contenues dans l'intervalle de confiance. Rappelons que cet intervalle inclut à la fois l'erreur de prédiction et l'erreur liée aux simulations de Monte Carlo.

Pour apprécier la qualité de l'ajustement, nous avons calculé l'erreur relative moyenne sur \mathcal{B}_{test} . Ces résultats sont donnés dans le tableau suivant :

Régression GP <i>gauss</i>	valeurs
erreur relative moyenne \mathcal{B}_{test}	0.015

Table 2 Erreur \mathcal{B}_{test} Krigeage gaussien

2.2.3 Régression GP avec covariance *exponential*

Les hyperparamètres calibrés pour l'estimation de la fonction du Best Estimate $Y(t, F_t)$ sont consignés dans le tableau suivant :

Variables	β	ℓ
Intercept	101300834.761	
Action	14020052.863	0.112
pc1	-123834182.734	0.043
pc2	99624147.336	0.051
pc3	1014650.673	0.056
Mortalité	-586664.907	3.269
Rachat	-101120393.388	0.996

GP <i>exponential</i>	σ_f	Log Vraisemblance
Valeurs	0.002	-15704.16

Table 3 Hyperparamètres *exponential*

En traçant sur le même graphe, les prédictions du modèle, les observations et l'intervalle de confiance à 95% en utilisant sur la base de test, on obtient le schéma suivant

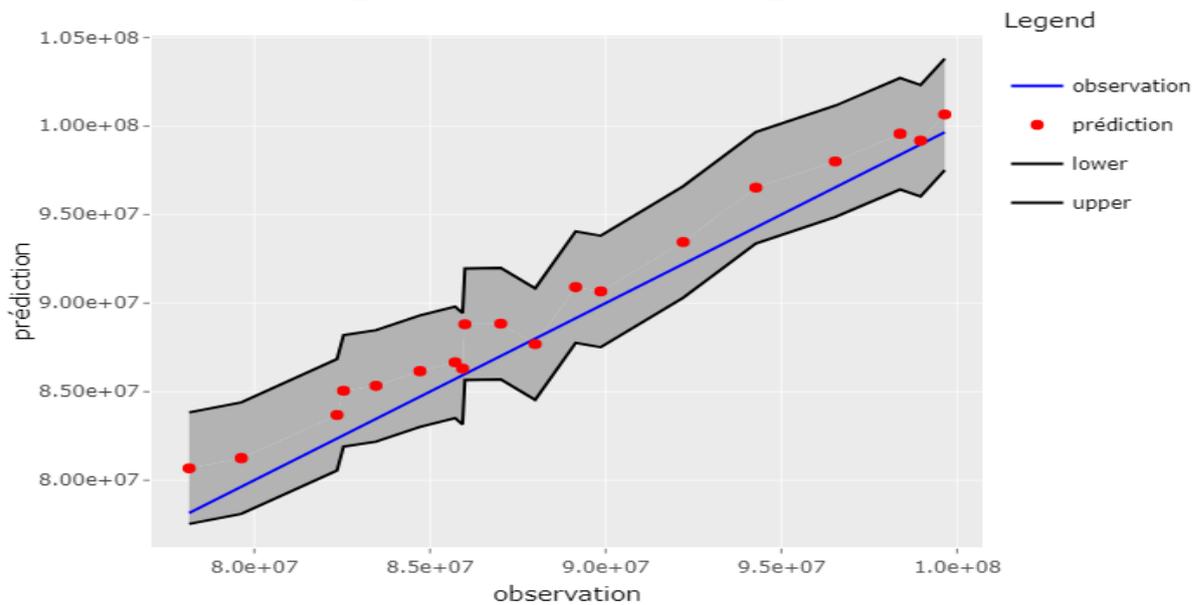


Figure 9 Ajustement et intervalle de confiance du krigeage *exponential*

Pour apprécier la qualité de l'ajustement, nous avons calculé l'erreur relative moyenne sur \mathcal{B}_{test} . Ces résultats sont donnés dans le tableau suivant :

Régression GP <i>exponentiale</i>	valeurs
erreur relative moyenne \mathcal{B}_{test}	0.017

Table 4 Erreur \mathcal{B}_{test} Krigeage *exponentiale*

2.2.4 Régression GP avec covariance *Matern 5/2*

Les hyperparamètres calibrés pour l'estimation de la fonction du Best Estimate $Y(t, F_t)$ sont consignés dans le tableau suivant :

Variabes	β	ℓ
Intercept	101097682.754	
Action	13670864.537	2.928
pc1	-118776031.293	0.0001
pc2	88673803.542	0.0003
pc3	-270293.834	0.056
Mortalité	-619670.564	1.307
Rachat	-101229769.241	0.661

GP <i>Matern 5/2</i>	σ_f	log vraisemblance
Valeurs	0.007	-15703.53

Table 5 Hyperparamètres *Matern 5/2*

En traçant sur le même graphe, les prédictions du modèle, les observations et l'intervalle de confiance à 95% eu utilisant la matrice de covariance à posteriori sur la base de test \mathcal{B}_{test} , on obtient le schéma suivant :

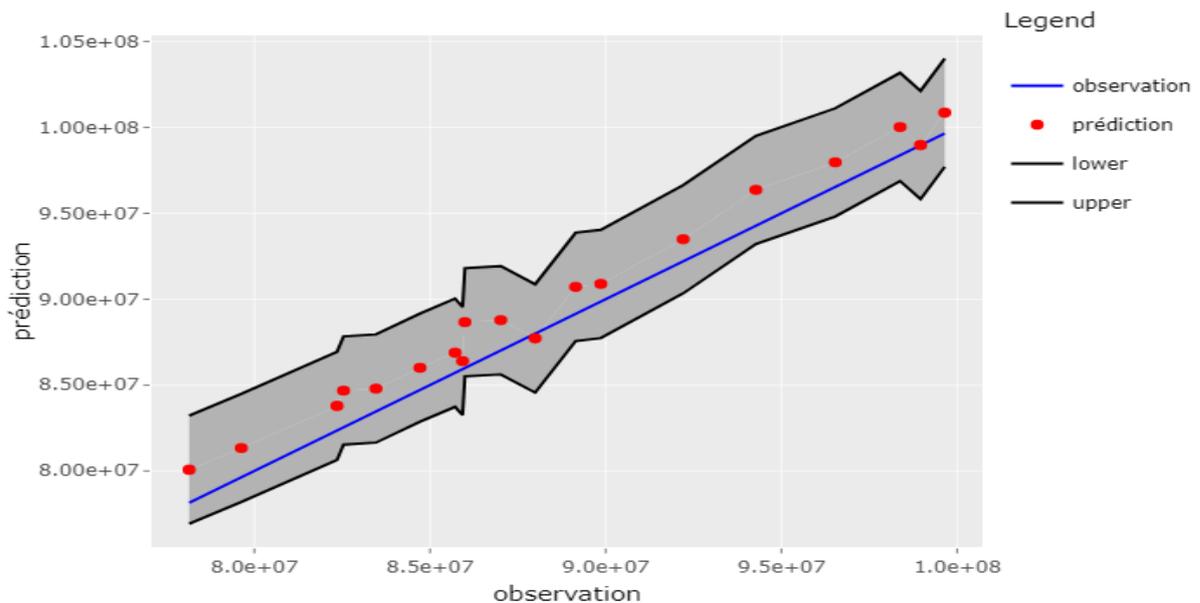


Figure 10 Ajustement et intervalle de confiance du krigeage *Matern 5/2*

Pour apprécier la qualité de l'ajustement, nous avons calculé l'erreur relative moyenne sur \mathcal{B}_{test} . Ces résultats sont donnés dans le tableau suivant :

Régression GP <i>Matern 5/2</i>	valeurs
erreur relative moyenne \mathcal{B}_{test}	0.015

Table 6 Erreur \mathcal{B}_{test} Krigeage *Matern 5/2*

2.2.5 Régression GP avec covariance *Matern 3/2*

Les hyperparamètres calibrés pour l'estimation de la fonction du Best Estimate $Y(t, F_t)$ sont consignés dans le tableau suivant :

Variables	β	ℓ
Intercept	101290599.656	
Action	14151869.253	0.148
pc1	-124100041.922	0.043
pc2	98227153.843	0.051
pc3	477997.091	0.056
Mortalité	-588585.726	3.269
Rachat	-101128536.6060	0.996

GP <i>Matern 3/2</i>	σ_f	Log vraisemblance
Valeurs	0.002	-15702.83

Table 7 Hyperparamètres *Matern 3/2*

En traçant sur le même graphe, les prédictions du modèle, les observations et l'intervalle de confiance à 95% eu utilisant la matrice de covariance à posteriori sur la base de test \mathcal{B}_{test} , on obtient le schéma suivant :

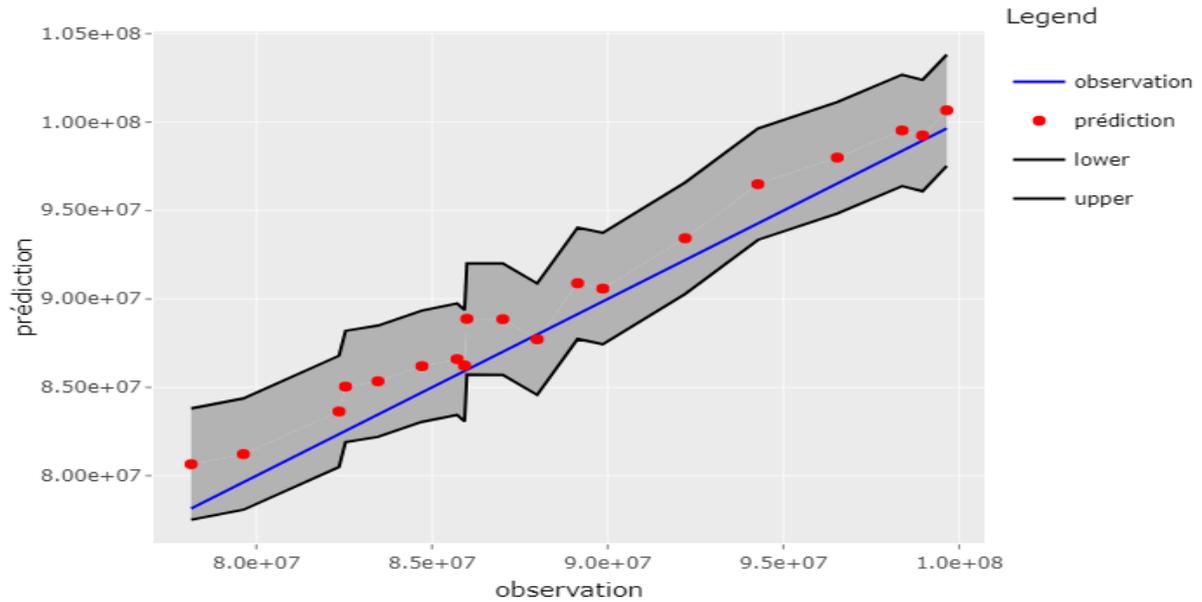


Figure 11 Ajustement et intervalle de confiance du krigeage *Matern 3/2*

Pour apprécier la qualité de l'ajustement, nous avons calculé l'erreur relative moyenne sur \mathcal{B}_{test} . Ces résultats sont donnés dans le tableau suivant :

Regression GP <i>Matern 3/2</i>	valeurs
erreur relative moyenne \mathcal{B}_{test}	0.017

Table 8 Erreur \mathcal{B}_{test} krigeage *Matern 3/2*

2.2.6 Régression GP avec covariance *power exp*

Les hyperparamètres calibrés pour l'estimation de la fonction du Best Estimate $Y(t, F_t)$ sont consignés dans le tableau suivant :

Variables	β	ℓ	p
Intercept	101301728.286		
Action	14276956.283	0.215	2.000
pc1	-126170757.905	0.0432	2.000
pc2	100920970.657	0.0515	2.000
pc3	344451.313	0.0005	0.000
Mortalité	-586237.890	1.486	0.000
Rachat	-101070079.184	0.973	0.114

GP <i>power exp</i>	σ_f	Log vraisemblance
Valeurs	0.011	-15701.610

Table 9 Hyperparamètres modèle *power exp*

En traçant sur le même graphe, les prédictions du modèle, les observations et l'intervalle de confiance à 95% eu utilisant la matrice de covariance à posteriori sur la base de test \mathcal{B}_{test} , on obtient le schéma suivant :

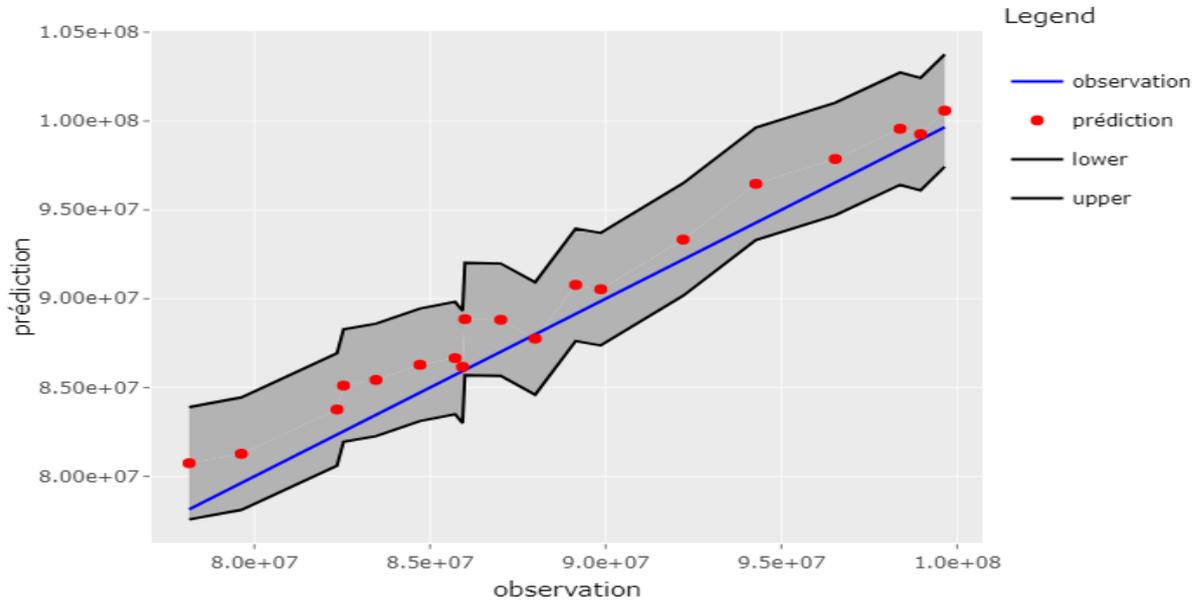


Figure 12 Ajustement et intervalle de confiance du krigeage *power exp*

Pour apprécier la qualité de l'ajustement, nous avons calculé l'erreur relative moyenne sur \mathcal{B}_{test} . Ces résultats sont donnés dans le tableau suivant :

Regression GP <i>power exp</i>	valeurs
erreur relative moyenne \mathcal{B}_{test}	0.017

Table 10 Erreur \mathcal{B}_{test} krigeage *power exp*

2.3 Résultat du metamodel du BE avec l'approche LSMC classique

Afin d'évaluer la performance du krigeage stochastique, nous avons également considéré les approches de régression classique dont :

- la régression polynomiale
- la régression *backward* avec critère sélection de variables AIC
- la régression Ridge
- la régression Lasso

Les résultats obtenus pour ces différentes régressions sont inscrites dans l'annexe du moment ce sont des méthodes connues de tout le monde.

2.4 Analyse du meilleur modèle retenu

Nous allons d'abord sélectionner le meilleur modèle par rapport à différents critères bien spécifiques, ensuite nous interpréterons les sensibilités estimées et enfin nous réaliserons une comparaison des sensibilités du proxy avec les sensibilités trouvées au niveau du modèle complet pour des chocs *stand alone*

2.4.1 Sélection de modèle

Nous allons donc essayer de retenir le modèle dont le *proxy* offre de meilleures performances sur \mathcal{B}_{test} et sur \mathcal{B}_{app}

Modèle	Erreur relative moyenne \mathcal{B}_{test}	Log Vraisemblance \mathcal{B}_{app}
Régression polynomiale	0.020	
Régression backward	0.020	
Régression Ridge	0.019	
Régression Lasso	0.020	
Régression GP <i>gauss</i>	0.015	-15703.16
Régression GP <i>exponential</i>	0.017	-15704.16
Régression GP <i>matern 5/2</i>	0.015	-15703.53
Régression GP <i>matern 3/2</i>	0.017	-15702.83
Régression GP <i>power exp</i>	0.017	-15701.61

Dans un premier temps, nous avons remarqué que tous les modèles de régression par *Gaussian Process* sont plus performants que les modèles de régression utilisée dans l'approche LSMC classique.

Cependant, la régression GP *gauss* et celle *matern 5/2* performent mieux que les autres car elles réalisent le minimum sur l'erreur relative moyenne \mathcal{B}_{test} .

Afin de sélectionner un modèle parmi les deux, nous avons dû regarder celui qui maximise la vraisemblance de la loi à priori sur \mathcal{B}_{app} . Ainsi, notre choix s'est donc logiquement porté sur la régression GP avec matrice de covariance *gauss*.

2.4.2 Interprétation du meilleur modèle

Nous allons essayer d'expliquer les sensibilités estimées par le modèle de krigeage *gauss* au cours de la détermination du proxy de la fonction $Y(t, F_t)$. Rappelons les sensibilités estimées avec le modèle :

Variables	β	ℓ
Intercept	101112817.918	
Action	13668409.274	2.9299
PC1	-119444717.068	0.0001
PC2	89543544.167	0.0002
PC3	158338.250	0.0561
Mortalité	-620673.620	1.325
Rachat	-101235205.927	0.644

GP <i>gauss</i>	σ_f	Log Vraisemblance
Valeurs	0.007	-15703.16

Table 11 Hyperparamètres *gauss*

Impact du facteur du rendement action sur le BE à un an : la sensibilité estimée pour le rendement action est positive comme on pouvait s’y attendre. En effet une hausse de 1% du rendement action augmente le BE de $1\% * \beta_1$.

Impact du facteur du facteur de niveau des taux sur le BE à un an : le niveau général des taux a un effet négatif sur le BE. En effet si les taux augmentent (variation parallèle) alors les rachats dynamiques et l’effet du facteur d’actualisation font diminuer l’engagement des assureurs. Ainsi une variation parallèle de 1% du niveau des taux diminue le BE de $1\% * \beta_2$.

Impact du facteur du facteur de pentification des taux sur le BE à un an : Rappelons que la pentification est le phénomène au cours duquel le niveau des taux de maturité élevée augmente. Elle a un impact positif sur le BE dans la mesure où la sensibilité estimée est positive. Cela se justifie par le fait que la revalorisation des contrats pour des a un impact plus grand que le taux d’actualisation sur les horizons élevés. Ainsi cela a pour effet d’augmenter l’engagement des assureurs. Ainsi une hausse de 1% de ce facteur entraîne une augmentation de $1\% * \beta_3$ du BE.

Impact du facteur du facteur de convexité des taux sur le BE à un an : La sensibilité estimée pour le facteur de convexité est positive ce qui veut dire que ce dernier évolue dans le même sens que le BE. L’explication se trouve dans le fait que le BE est plus sensible à une baisse de taux qu’à une hausse de taux. Ainsi une hausse de 1% du facteur de convexité entraîne une hausse de $1\% * \beta_4$ du BE.

Impact de la mortalité sur le BE à un an : On devait s’attendre à ce que la mortalité ait un impact négatif sur le BE cependant la sensibilité estimée est négative et négligeable comparé aux autres facteurs. Ce qui montre que la mortalité n’est pas le risque le plus important pour les contrats de type Euro. le ℓ estimé est grand ce qui montre que le BE est une fonction linéaire de la mortalité ainsi contribue moins à sa variabilité.

Impact du rachat sur le BE à un an : le rachat a logiquement un impact négatif sur le BE, en effet une hausse de 1% des rachats implique une diminution de $1\% * \beta_5$. C’est l’un des risques les plus importants en Épargne. Cela se ressent dans l’estimation de la sensibilité

2.4.3 Test sur les sensibilités estimées

Nous ferons une analyse comparative des sensibilités issues du *proxy* et celles issues du modèle complet. Le tableau ci dessous résume bien cette analyse :

Choc taux	BE ALM	BE proxy	Ecart
Baisse de 1% pc1 (niveau des taux)	103742702	102307265	1.38%
Hausse de 1% pc1 (niveau des taux)	98665434	99918371	1.26%

Table 12 Comparaison sensibilités taux

Choc Rachat	BE ALM	BE proxy	Ecart
Hausse de 25% des rachats	75804016	76092072	0.38%
Hausse de 75% des rachats	50495215	50949672	0.90%
Hausse de 90% des rachats	10001133	10046138	0.45%

Table 13 Comparaison sensibilités rachat

On a réalisé des chocs de formule standard pour les risques élémentaires de taux et de rachat. En effet on a choqué le niveau des taux à la hausse et à la baisse, et de la même manière, on a appliqué différents chocs de hausse pour le risque rachat. On remarque globalement que le proxy est en adéquation avec le modèle complet étant donné que les écarts au niveau des valeurs du BE sont négligeables.

Les figures ci-dessous représentent les différentes valeurs du proxy estimé pour différents chocs croisés réalisés sur l'évolution de certains facteurs de risque à horizon un an.

Sur la Figure 13, on a représenté le proxy estimé en fonction des chocs croisés sur le rachat et le facteur de niveau des taux.

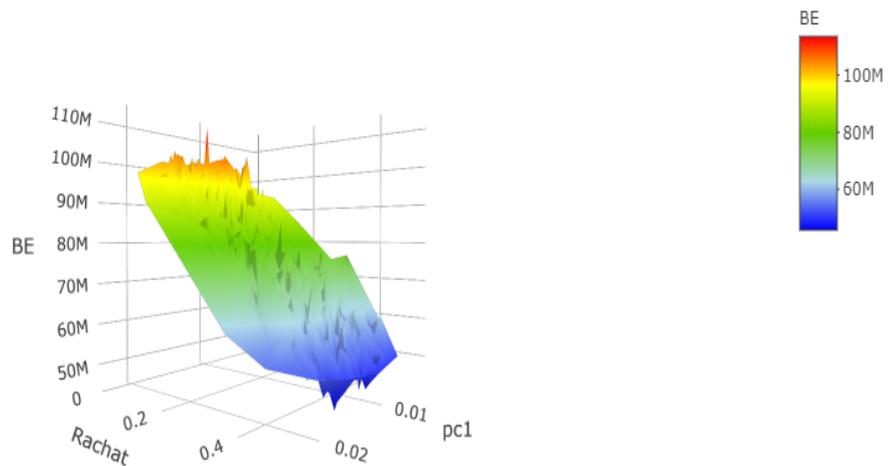


Figure 13 BE en fonction de rachat et taux

Sur la figure 14, on a représenté le proxy estimé en fonction des chocs croisés sur le rachat et l'indice action.

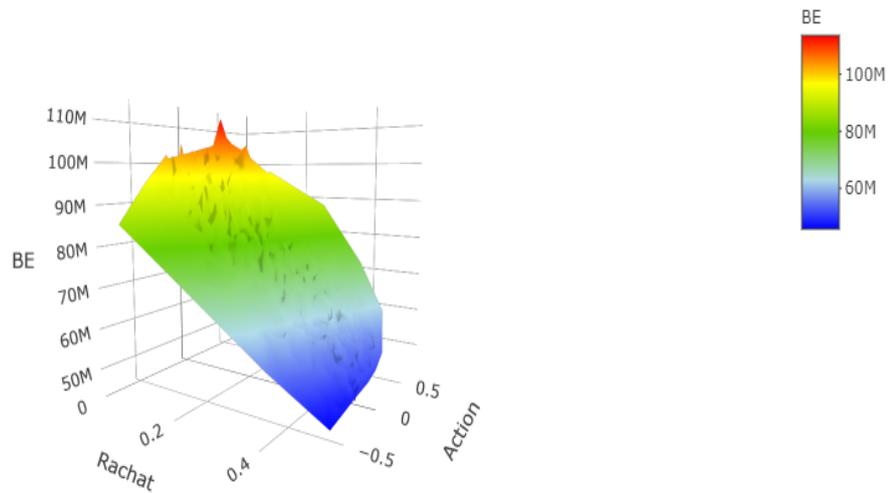


Figure 14 BE en fonction de rachats et rendement action

Sur la figure 15, on a représenté le proxy estimé en fonction des chocs croisés sur l'indice action et le facteur de niveau des taux.

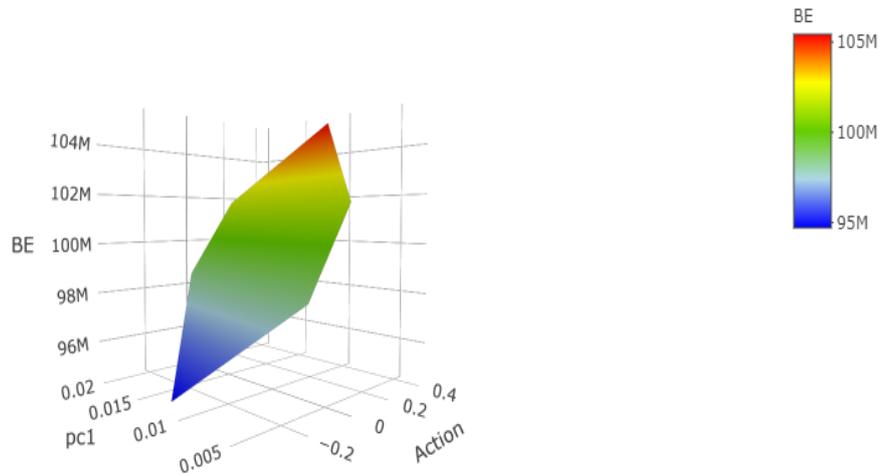


Figure 15 BE en fonction de l'indice action et le facteur de niveau des taux

3 *Metamodel* du pour le calcul du SCR à une date ultérieure

Dans cette section, nous allons essayer d'estimer la fonction $Y(t' + 1, F_{t'+1})$ dans le cadre de la valorisation du bilan économique à une date ultérieure $t' \geq t$. L'idée dans notre approche est donc d'utiliser la connaissance antérieure de la fonction à la date t à laquelle on va compléter la connaissance actuelle.

Le schéma ci dessous résume en effet notre approche :

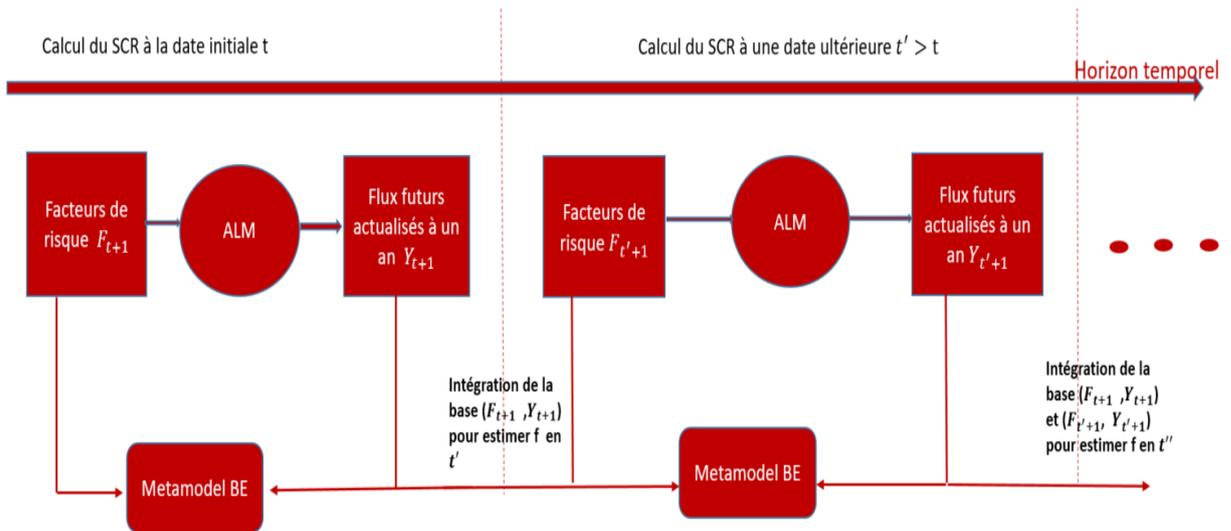


Figure 16 Etapes itératives de valorisation du bilan

L'étape de valorisation initiale a été complètement réalisée dans le chapitre précédent. En effet, pour rappel, à la première date de valorisation, nous avons régressé les flux futurs actualisés sur les facteurs de risque en t en utilisant le krigage stochastique qui convient mieux à notre variable cible. Cette méthode de régression permet non seulement de prendre en compte l'erreur de prédiction mais aussi l'erreur intrinsèque des simulations de Monte Carlo.

L'étape suivante est de se demander, comment peut-on utiliser la connaissance lors de l'exercice précédent pour estimer cette même fonction à une date ultérieure t' . C'est donc à cette question qu'on va essayer de proposer une idée de réponse dans la suite de notre développement.

3.1 Étapes intermédiaires à la réalisation du *metamodel*

Avant de procéder à la valorisation du bilan à un instant ultérieur t' , certaines étapes intermédiaires doivent être réalisées. On distingue :

- Formulation des hypothèses sur l'évolution des facteurs de risque $F_{t'+1}$
- Simulation d'un nombre N' avec ($N' \ll M$) de facteurs de risque liés aux hypothèses émises.
- Détermination des flux futurs actualisés à un an liés aux scénarios des facteurs de risque.
- **Création d'une variable temps** qui vaut :

$$\begin{cases} 0 & \text{pour les facteurs risques calibrés en } t \\ 1 & \text{pour les facteurs risques calibrés en } t' \end{cases}$$

Il faut donc construire une base d'apprentissage \mathcal{B}_{app} en t' de la manière la plus optimale possible en combinant la connaissance antérieure et celle actuelle.

3.2 Construction optimale de la base d'apprentissage \mathcal{B}_{app} en t'

Nous allons compléter la base antérieure X constituée au cours de la date de valorisation initiale t par un certain nombre d'observations bien choisies parmi les N' observations des facteurs de risque $F_{t'+1}$.

La démarche utilisée est la suivante :

- On se fixe l'objectif de compléter N ($N \leq N'$) facteurs de risque en t' à la base antérieure X
- Initialement, on utilise la fonction estimée $Y(t+1, F_{t+1})$ avec laquelle, on prédit tous les N' nouveaux facteurs de risque $F_{t'+1}$

- On regarde parmi les N' observations, celle (s) sur laquelle (lesquelles) l'erreur relative est maximale
- On intègre cette observation (\mathbf{x}_{M+1}) dans la base \mathbf{X} , tout en estimant $\hat{\sigma}_{M+1}$ avec les n_0 répliques de \mathbf{x}_{M+1} en t'
- Intégration de la variable temporelle dans \mathbf{X}
- Le modèle récalibre les hyperparamètres et on refait la procédure jusqu'à compléter la dernière observation \mathbf{x}_{M+N}

En procédant de cette manière, nous choisissons juste les observations qui nous permettent d'explorer des domaines d'expérience inconnus de l'exercice précédent.

3.3 Illustration du *Metamodel* du *Best Estimate* à une date ultérieure

Nous allons mettre en oeuvre cette approche sous différentes hypothèses formulées sur les nouveaux facteurs de risque $F_{t'+1}$. Ces hypothèses sont les suivantes :

- Action : Les nouveaux paramètres obtenus sont :

Estimateurs	Valeur
$\hat{\mu}$	0.036
$\hat{\sigma}$	0.15

Table 14 Paramètres facteur action pour le calcul ultérieur

- Taux : Les paramètres obtenus pour la modélisation de la courbe des taux sont :

paramètres N S	λ	μ	σ
β_1	0.091702405	-0.020900848	0.002553749
β_2	0.147388552	0.021555723	0.003124971
β_3	0.130651122	0.018430310	0.003648313

Table 15 Paramètres facteur taux pour le calcul ultérieur

Nous avons réalisé $N' = 50$ simulations des facteurs de risque $F_{t'+1}$. Nous disposons aussi des réalisations des flux futurs actualisés à chaque scénario en utilisant le modèle ALM. Ces observations nous permettront de sélectionner les champs d'expérience intéressants pour l'estimation de notre fonction Y en t' .

Seuil optimal pour le nombre d'observations à compléter

Après avoir appliqué notre approche pour construire la base optimale en utilisant les N' observations en t' , nous avons constaté une évolution de l'erreur qui est la suivante :

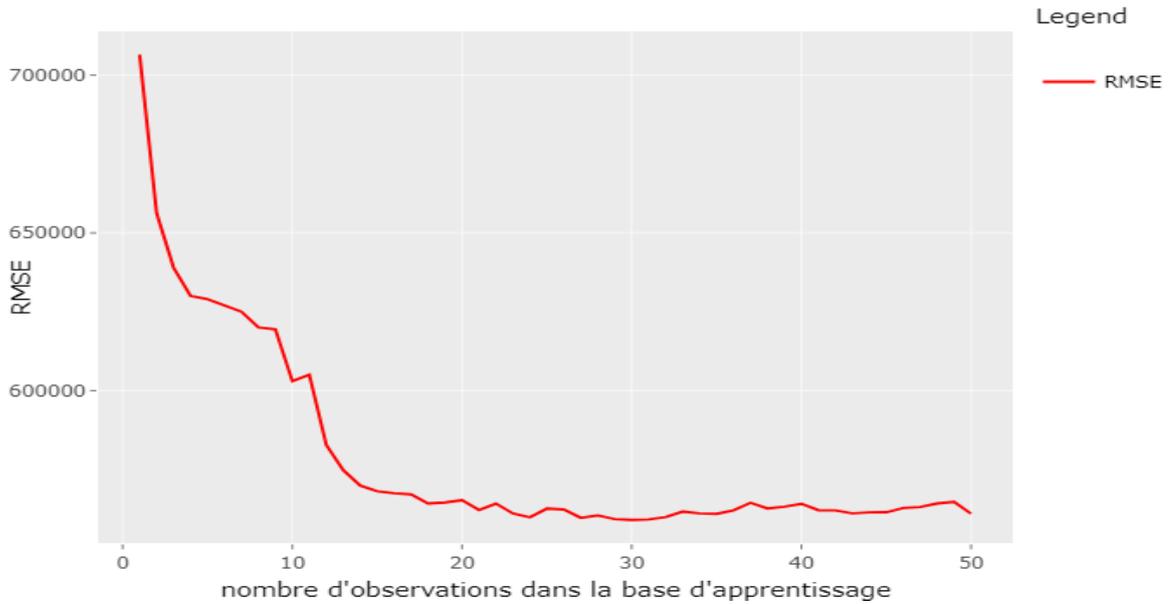


Figure 17 Évolution du RMSE

Ce schéma montre l'évolution du RMSE sur les $N' - i$ observations restantes après avoir complété i observations dans la base d'apprentissage antérieure X .

On observe qu'après avoir complété 20 observations dans l'ancienne base, les $N' - 20$ autres observations restantes n'améliorent plus le RMSE, ainsi ce sont donc des domaines déjà explorés. Dans notre cas, nous avons donc retenu $N=20$ observations de $F_{t'+1}$ parmi les N' observations simulées. Elles ont été retenues car elles nous permettent d'entraîner le modèle sur des domaines inconnus à l'étape précédente afin de bien estimer la fonction Y en t' . On dispose ainsi d'une base d'apprentissage \mathcal{B}_{app} avec $\mathcal{B}_{app} = (x_i, \dots, x_{M+N})^\top$, et

$$x_i = (Action_i, PC1_i, PC2_i, PC3_i, Mortalité_i, Rachat_i, Temps_i)$$

Rappelons que la fonction moyenne et la matrice de covariance de notre *Gaussian Process* ont changé car nous avons incorporé la variable temps.

- fonction de moyenne :

$$\begin{aligned} m(x) &= \beta_0 + x^\top * \beta \\ &= \beta_0 + \beta_1 * Action + \beta_2 * PC1 + \beta_3 * PC2 + \beta_4 * PC3 + \beta_5 * Mortalite + \beta_6 * Rachat + \\ &\quad \beta_7 * Temps \end{aligned}$$

- fonction de covariance

$$\begin{aligned} &\mathbb{R}^7 * \mathbb{R}^7 \rightarrow \mathbb{R} \\ k_{\text{Gauss}} : & (\mathbf{x}, \mathbf{x}') \rightarrow \sigma_f^2 \exp \left(-\frac{1}{2} \sum_{k=1}^{k=7} \left(\frac{\mathbf{x}_k - \mathbf{x}'_k}{\ell_k} \right)^2 \right) \end{aligned}$$

Les hyperparamètres calibrés sur \mathcal{B}_{app} sont renseignés dans le tableau suivant :

Variables	β	ℓ
Intercept	101252807.219	
Action	14198226.512	0.123
pc1	124074546.685	0.043
pc2	96437297.634	0.051
pc3	-196442.568	0.056
Mortalité	-569009.578	3.266
Rachat	-101126984.243	0.998
temps	-284322.129	1.923

GP <i>gauss</i>	σ_f	Log vraisemblance
Valeurs	0.002	-16463.27

Table 16 Hyperparamètres calibrés metamodel

La sensibilité β au facteur temps est négative car le Best Estimate tend à diminuer avec le temps (il y a une année de flux en moins) sous l'hypothèse selon laquelle le portefeuille est en *run off*. L'interprétation sur notre cas est que le Best Estimate diminue de 284322.129 entre les dates $t + 1$ à $t' + 1$.

Pour valider les performances de notre metamodel, nous avons testé son efficacité sur une base de test composée de 100 scénarios obtenus à partir de simulations des facteurs de risque $F_{t'+1}$. Nous avons représenté sur le même graphe, les prédictions du modèle, les observations de la base de test et l'intervalle de confiance prenant en compte l'erreur de prédiction et l'erreur de Monte Carlo.

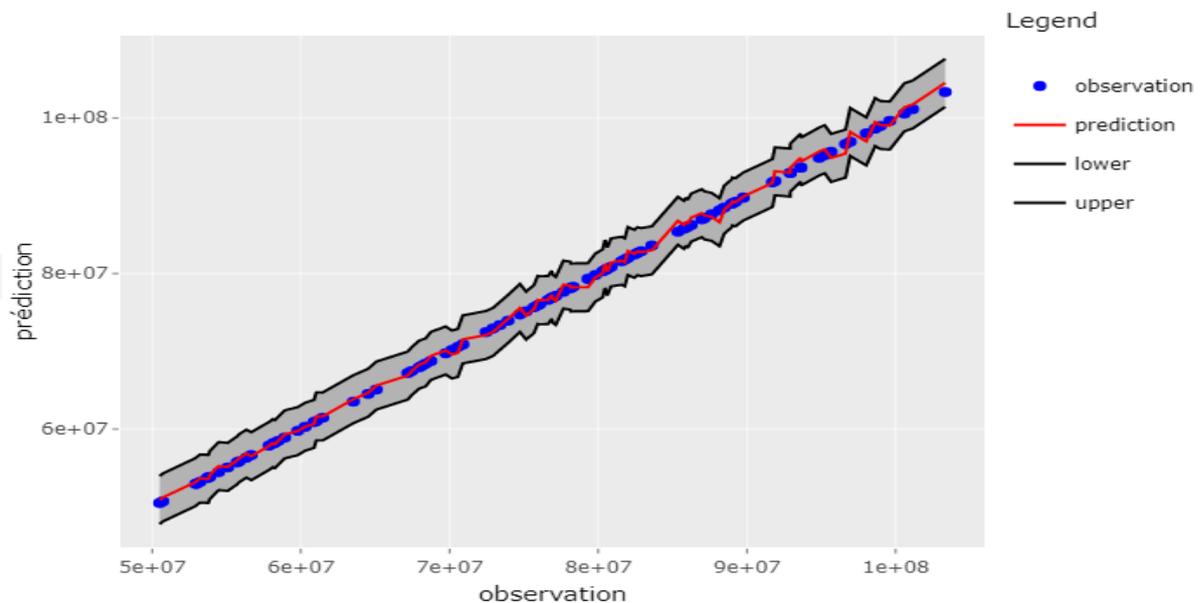


Figure 18 Prédiction vs observation

On remarque bien que le fait de combiner les informations de différentes périodes améliore la qualité de l'ajustement. En effet les observations sont pratiquement confondues avec les prédictions

et elles sont toutes dans l'intervalle de confiance.
Le tableau ci dessous regroupe les différentes erreurs réalisées sur la base de test

Regression GP <i>metamodel</i>	valeurs
erreur relative moyenne \mathcal{B}_{test}	0.0055
RMSE \mathcal{B}_{test}	561017

Table 17 Erreur moyenne relative et RMSE sur \mathcal{B}_{test}

On remarque en effet que l'on fait moins d'erreur sur la base de test en t' même si les bases de test ne sont pas les mêmes d'une année à une autre.

4 Metamodel du BE à une date ultérieure t''

Notre objectif dans cette section est de savoir comment évolue l'erreur si on réalise le metamodel sur une autre date t'' avec $t'' > t'$.

Concrètement, on va vérifier si l'erreur diminue sur la base de test lorsqu'on complète la base d'apprentissage à chaque nouvelle date.

4.1 Mise en oeuvre de l'approche pour la date t''

Les différentes étapes à réaliser sont :

- Hypothèses d'évolution des facteurs de risque sur $[t'', t'' + 1]$
- Création d'une manière optimale la base d'apprentissage en t'' en avec les facteurs de risque $F_{t+1}, F_{t'+1}, F_{t''+1}$ en procédant comme décrit ci-dessus
- Prise en compte de la variable temporelle

$$\begin{cases} 0 & \text{pour les facteurs de risque calibrés en } t \\ 1 & \text{pour les facteurs de risque calibrés en } t' \\ 2 & \text{pour les facteurs de risque calibrés en } t'' \end{cases}$$

Les hypothèses retenues sur l'évolution des facteurs de risque $F_{t''+1}$ sont :

- Action : Les nouveaux paramètres obtenus sont :

Estimateurs	Valeur
$\hat{\mu}$	0.024
$\hat{\sigma}$	0.1

Table 18 Paramètres facteur action pour le calcul ultérieur

- Taux : Les paramètres obtenus pour la modélisation de la courbe des taux sont :

paramètres N S	λ	μ	σ
β_1	0.061	-0.021	0.001
β_2	0.098	0.022	0.002
β_3	0.087	0.018	0.002

Table 19 Paramètres facteur taux pour le calcul ultérieur

Les hyperparamètres calibrés à la date t'' sont les suivants :

Variables	β	ℓ
Intercept	101213040.749	
Action	14201136.732	0.124
pc1	121234139.135	0.043
pc2	93395559.249	0.051
pc3	-796559.049	0.056
Mortalité	-572184.934	3.266
Rachat	-101064216.190	0.998
temps	-234782.620	2.460

GP <i>gauss</i>	σ_f	Log vraisemblance
Valeurs	0.002	-17223.48

Table 20 Hyperparamètres calibrés metamodel

La variable temps nous renseigne sur le fait que le BE diminue d'une année à une autre. Ainsi pour la date $t'' + 1$ le BE chute de -234782.620 multiplié par 2 par rapport à $t + 1$.

Pour valider donc les performances du modèle en t'' , nous avons testé son efficacité sur une base de test composée de 100 scénarios obtenus à partir de simulations des facteurs de risque $F_{t''}$. Nous avons donc représenté sur un même graphe, les prédictions du modèle, les observations de la base de test et l'intervalle de confiance prenant en compte l'erreur de prédiction

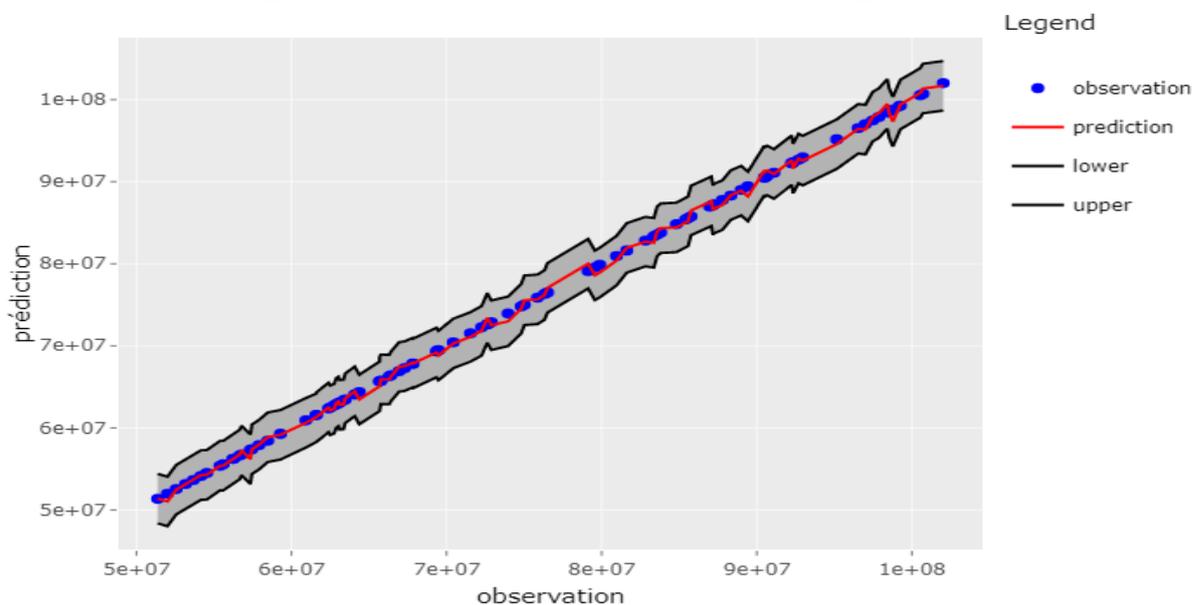


Figure 19 Prédiction vs observation

On remarque bien que le fait de combiner les informations de différentes périodes améliore la qualité de l'ajustement. En effet les observations sont pratiquement confondues avec les prédictions et elles sont toutes dans l'intervalle de confiance.

Le tableau ci dessous regroupe les différentes erreurs réalisées sur la base de test

Regression GP <i>metamodel</i>	valeurs
erreur relative moyenne \mathcal{B}_{test}	0.0053
RMSE \mathcal{B}_{test}	513787

Table 21 Erreur relative moyenne et RMSE sur \mathcal{B}_{test}

5 Conclusion

Régression GP <i>metamodel</i>	t	t'	t''
erreur relative moyenne \mathcal{B}_{test}	0.015	0.0055	0.0053

Table 22 Évolution de l'erreur relative moyenne pour différentes dates

Dans cette partie, nous avons clairement montré que le fait d'utiliser des données antérieures en complément avec les facteurs de risque du moment permet de calibrer un meilleur *proxy* d'estimation du BE. Par rapport au nombre minimal d'observations à retenir pour les facteurs du moment, cela dépend de l'objectif d'erreur qu'on se fixe.

Conclusion

Dans cet article, nous nous sommes intéressés à la mise en place de l'approche LSMC dans le cadre du modèle interne en assurance vie afin de déterminer la valeur du capital économique. L'objectif étant de proposer une alternative permettant d'améliorer l'étape de régression de l'approche LSMC en utilisant des algorithmes de machine learning adaptées à des sorties de simulation de Monte Carlo.

Pour ce fait, nous avons réalisé le metamodelling du Best Estimate à un an avec plusieurs approches de régression dont notamment la régression par processus gaussien et plus particulièrement le Krigeage stochastique. C'est une technique de régression utilisée en géostatistique qui trouve de plus en plus d'application dans le domaine de la finance et de l'assurance. Nous nous sommes donc inspirés de ces différents travaux pour calibrer le proxy du Best Estimate à un an avec le krigeage stochastique. Nos résultats démontrent que le proxy offre de meilleures performances sur note base de test contrairement aux autres approches classiques utilisées. C'est une technique qui offre beaucoup d'avantages dont :

- la quantification de l'incertitude de prédiction pour chaque observation
- la prise en compte des incertitudes liées aux simulations de Monte Carlo
- l'estimation efficace des sensibilités des facteurs de risque en sortie du modèle ALM.

Cependant cette méthode comporte une principale limite qui n'est rien d'autre que l'inversion de la matrice de covariance de la loi à priori pour la calibration des paramètres. En effet le nombre d'observations de la base d'apprentissage augmente la complexité de la calibration des hyperparamètres du modèle.

Dans un second temps, nous avons proposé une approche itérative basée sur le krigeage stochastique permettant de revaloriser le SCR à des dates ultérieures afin d'aider les compagnies d'assurance dans la production des états quantitatifs trimestriels exigés par Solvabilité II. Il s'agit d'exploiter l'information issue des calculs aux dates d'évaluation précédentes pour rendre plus efficace l'estimation du SCR aux dates d'évaluation futures en réalisant un nombre limité de simulations bien choisies.

Nous avons donc mis en place un algorithme permettant de cibler d'une manière optimale les facteurs de risque nous permettant d'acquérir de nouvelles informations à la nouvelle date de recalcul afin de ne pas complexifier la calibration des hyperparamètres. Nous avons donc logiquement remarqué que le fait de constituer une base d'apprentissage comportant les informations à différentes dates d'évaluation améliore la qualité d'ajustement du proxy.

Cependant notre approche souffre d'une certaine simplicité au niveau des facteurs de risque. En effet, nous avons considéré six facteurs de risque dans nos développements. Ce qui est très simpliste car en réalité on peut en dénoter près de 150 facteurs de risque. Ce nombre élevé de facteurs de risque augmenterait logiquement la complexité dans la calibration des hyperparamètres du krigeage stochastique. Ce qui nous pousse à réfléchir à des techniques de réduction de dimension dans d'autres études pour la prise en compte d'un grand nombre de facteurs de risque. Cela faciliterait donc la mise en place du modèle auprès des acteurs du marché de l'assurance.

Références

- Ankenman, B., B. L. Nelson, and J. Staum (2010). Stochastic kriging for simulation metamodeling. *Operations Research* 58(2), 371–382.
- Cousin, A., H. Maatouk, and D. Rullière (2016). Kriging of financial term-structures. *European Journal of Operational Research* 255(2), 631–648.
- Crépey, S. and M. F. Dixon (2020). Gaussian process regression for derivative portfolio modeling and application to credit valuation adjustment computations. *Journal of Computational Finance* 24(1).
- Croix, J.-C. (2015). Méta-modélisation par krigeage pour la mesure du risque de taux en ALM. *Mémoire d'actuariat, ISFA*.
- Decupère, S. (2011). Agrégation des risques et allocation de capital sous solvabilité II. *Mémoire d'actuariat, ENSAE*.
- Devineau, L. and S. Loisel (2009). Construction d'un algorithme d'accélération de la méthode des «simulations dans les simulations» pour le calcul du capital économique solvabilité II. *Bulletin Français d'Actuariat* 10(17), 188–221.
- Liu, M. and J. Staum (2010). Stochastic kriging for efficient nested simulation of expected shortfall. *Journal of Risk* 12(3), 3.
- Ludkovski, M. (2018). Kriging metamodels and experimental design for Bermudan option pricing. *Journal of Computational Finance* 22(1).
- Ludkovski, M., J. Risk, and H. Zail (2018). Gaussian process models for mortality rates and improvement factors. *ASTIN Bulletin : The Journal of the IAA* 48(3), 1307–1347.
- Nexialog Consulting (2021). IFRS 17 Risk Adjustment : une étude comparative sur le périmètre des contrats d'épargne en euros.
- Risk, J. and M. Ludkovski (2018). Sequential design and spatial modeling for portfolio tail risk measurement. *SIAM Journal on Financial Mathematics* 9(4), 1137–1174.

6 Annexes

6.1 Résultat du metamodel du BE avec l'approche LSMC classique

Dans cette section, nous présentons les résultats obtenus pour la calibration des différents *proxy* en utilisant l'approche LSMC classique.

6.1.1 Régression polynomiale

C'est dans le but de matérialiser la relation non linéaire existante entre le passif et l'actif que nous avons décidé de faire la régression sur les $L_i \quad i \in [1, 120]$.

La qualité de l'ajustement est représentée sur le graphique suivant :

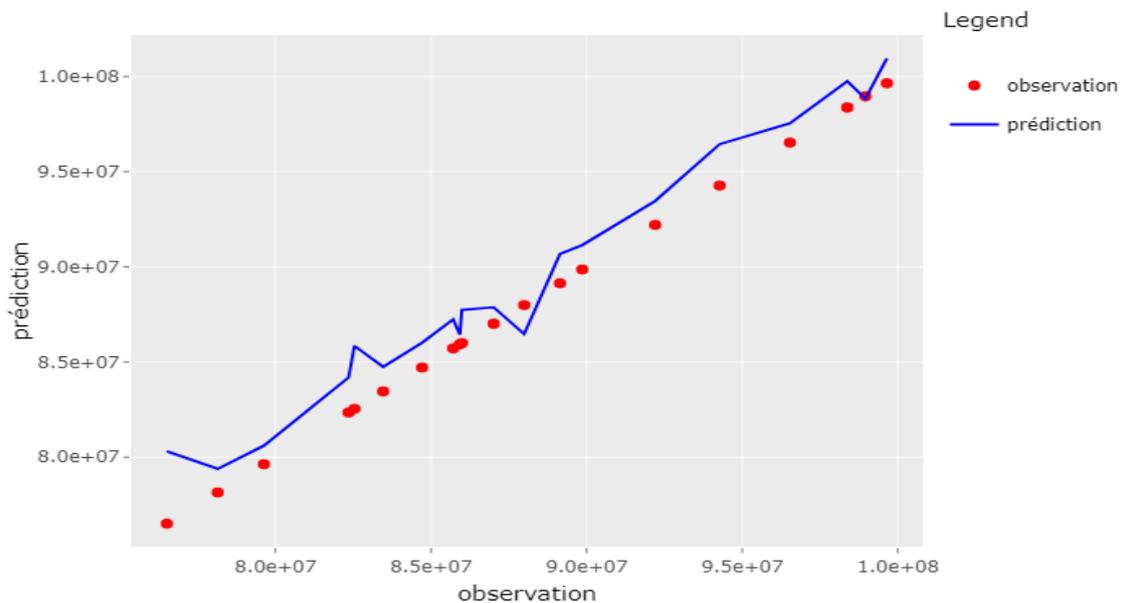


Figure 20 Ajustement régression polynomiale

En abscisse du graphe 20, nous avons les flux futurs actualisés issus du modèle ALM de la base de test \mathcal{B}_{test} tandis qu'en ordonnée, nous avons les prédictions données par la régression polynomiale. On remarque que notre ajustement est acceptable. Nous avons tenu à ressortir le R^2 ajusté qui donne une idée de la qualité de la régression et l'erreur relative moyenne réalisée sur la base d'apprentissage et sur la base de test.

Régression polynomiale	valeurs
R^2 ajusté	0.988
erreur relative moyenne \mathcal{B}_{app}	0.015
erreur relative moyenne \mathcal{B}_{test}	0.020

Table 23 Erreur \mathcal{B}_{test} et R^2 ajusté du modèle polynomial

Malgré la qualité de la régression, on s'est aperçu que certains paramètres n'ont pas convergé. Les raisons qui expliquent cette divergence peuvent être :

- les variables $L_i(x) \quad i \in [1, 120]$ sont très corrélées entre elles, ce qui est logique si on se réfère à la manière dont on les a construites.

- problème d'inversibilité de matrice, le déterminant de la matrice des variables explicatives est proche de 0 dû à la corrélation existante entre les variables, ce qui ne permet pas l'estimation robuste des paramètres.

Au regard de toutes ces raisons évoquées, nous avons jugé mieux de remplacer l'estimateur des moindres carrés par les estimateurs de *Ridge* et de *Lasso* qui permettent de palier à ce problème d'inversibilité. Avant d'utiliser les modèles de *Ridge* et *Lasso*, nous avons d'abord appliqué une méthode de régression avec sélection de variables *backward* ayant comme critère de sélection l'AIC.

6.1.2 Régression polynomiale *backward*

Les variables $L_i(x)$ $i \in [1, 120]$ retenues dans ce modèle sont celles qui expliquent le plus au sens de l'AIC les valeurs des flux futurs actualisés à un an.

En appliquant une sélection de variables *backward* avec critère de sélection AIC, on a le graphe d'ajustement suivant :

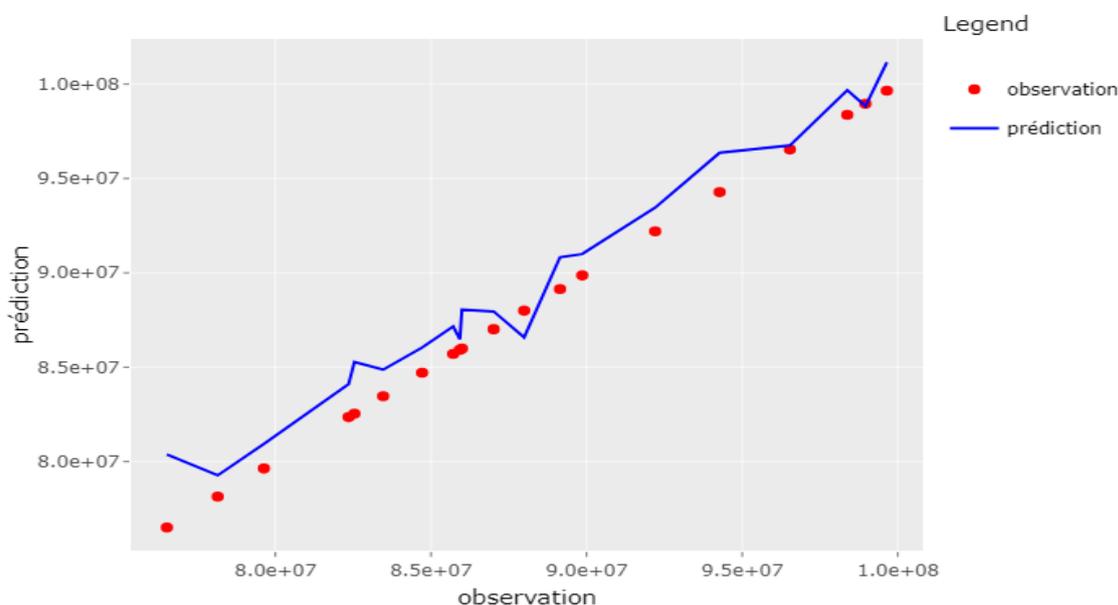


Figure 21 Ajustement regression backward

Malgré le fait que la régression backward soit un sous modèle de la régression polynomiale, l'ajustement n'a pas beaucoup changé. Ce qui est une bonne chose car on a la même qualité de régression avec moins de variables explicatives. L'erreur moyenne relative sur la base d'apprentissage et sur la base de test n'a pas non plus changé aussi.

Regression polynomiale	valeurs
R^2 ajusté	0.989
erreur relative moyenne \mathcal{B}_{app}	0.015
erreur relative moyenne \mathcal{B}_{test}	0.020

Table 24 Erreur \mathcal{B}_{test} et R^2 ajusté du modèle backward

6.1.3 Régression polynomiale Ridge

Les régressions Ridge et Lasso sont utiles lorsqu'on travaille sur des bases dont les variables sont très corrélées entre elles. En effet, on introduit une fonction de pénalité afin de contrôler la variance des estimateurs. La fonction à minimiser est de la forme suivante :

$$\min_{\beta_1, \dots, \beta_p} \sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p \beta_j^2 \quad (8)$$

avec :

- y_i : l'observation i de la variable cible y
- x_{ij} l'observation i de la variable explicative x_j
- λ est le coefficient de pénalité qui permet de contrôler l'impact de la pénalité.

L'estimateur Ridge s'écrit alors

$$\hat{\beta}_{\text{Ridge}} = (X'X + \lambda I_p)^{-1} X'y \quad (9)$$

- On peut donc avoir une estimation même si $(X'X)$ n'est pas inversible
- On voit bien que si $\lambda = 0$, alors on a l'estimateur des MCO
- si $\lambda \rightarrow \infty$ alors $\beta_j \rightarrow 0$

la méthode *Ridge* tend à attribuer aux variables fortement corrélées le même coefficient. Les coefficients estimés sont renseignés en annexe. Nous avons remarqué que tous les coefficients estimés sont cohérents par rapport aux estimations effectuées au niveau de la régression polynomiale.

On peut ainsi afficher une *Ridge path coefficients* permettant de montrer l'évolution des différents coefficients en fonction du facteur de pénalité λ . Cette évolution se présente comme suit :

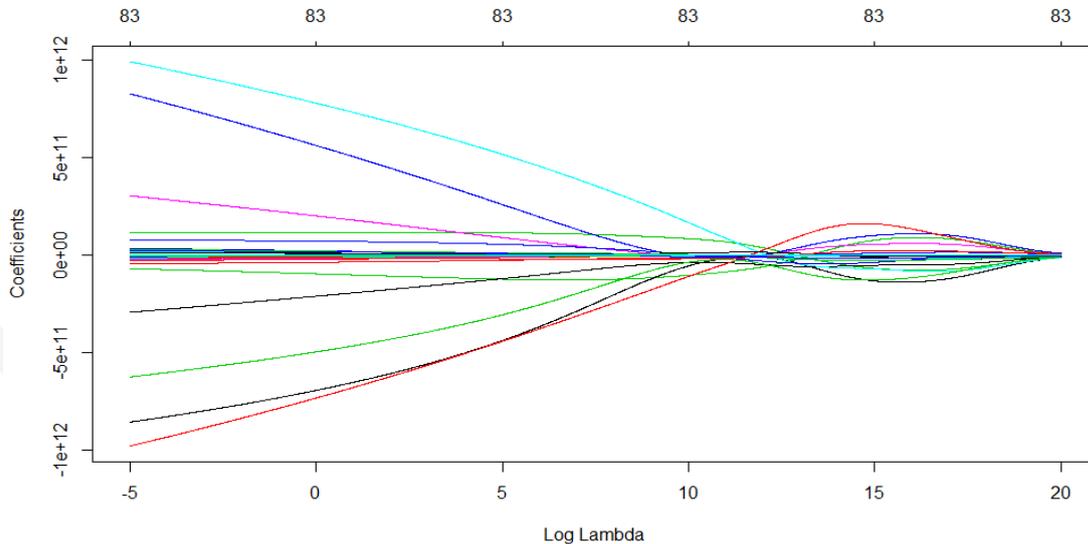


Figure 22 Évolution des paramètres calibrés en fonction de lambda

On remarque que pour des valeurs très petites de λ , on tend vers l'estimateur des moindres carrés ordinaires classiques ayant de grandes valeurs tandis que pour λ grand, on tend vers des valeurs plus petites comme on l'a montré ci dessus.

De plus, pour différentes valeurs de λ , nous pouvons utiliser différents modèles, le but maintenant est de déterminer pour quelle valeur de λ , on minimise la MSE sur la base d'apprentissage.

Nous allons donc afficher l'évolution de la MSE en fonction de λ . Pour cela nous avons fait une cross validation sur chaque modèle étudié afin de ressortir une erreur moyenne et un intervalle de confiance de l'erreur. Le schéma ci dessus résume bien les résultats obtenus.

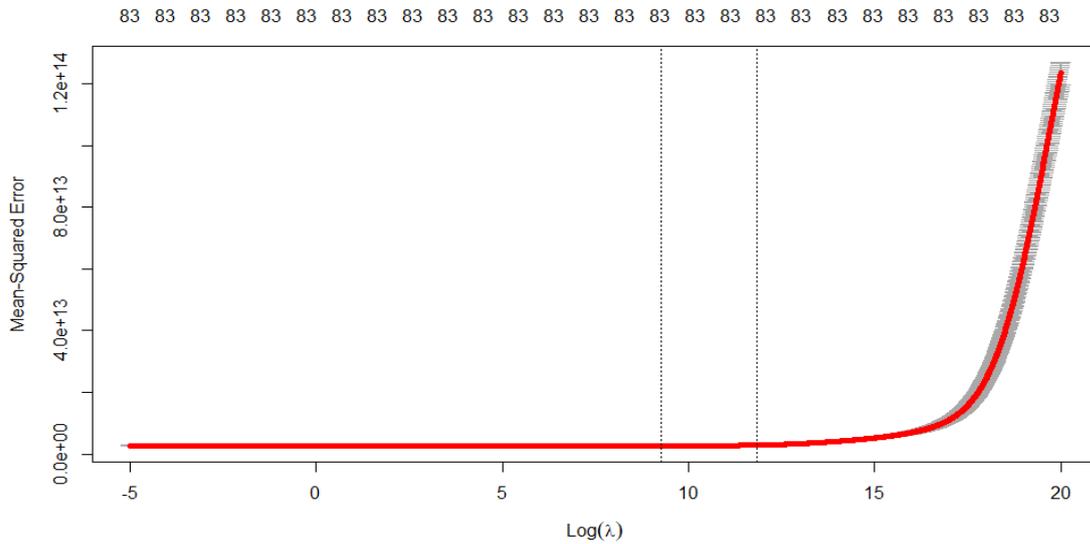


Figure 23 Évolution du mean square error en fonction de lambda

On a trouvé que $\log(\lambda) = 9.27$ permettait de minimiser la MSE⁵ sur la base d'apprentissage en utilisant la Cross validation.

C'est donc la régression obtenue avec cette valeur de λ qui nous donne les résultats suivants :

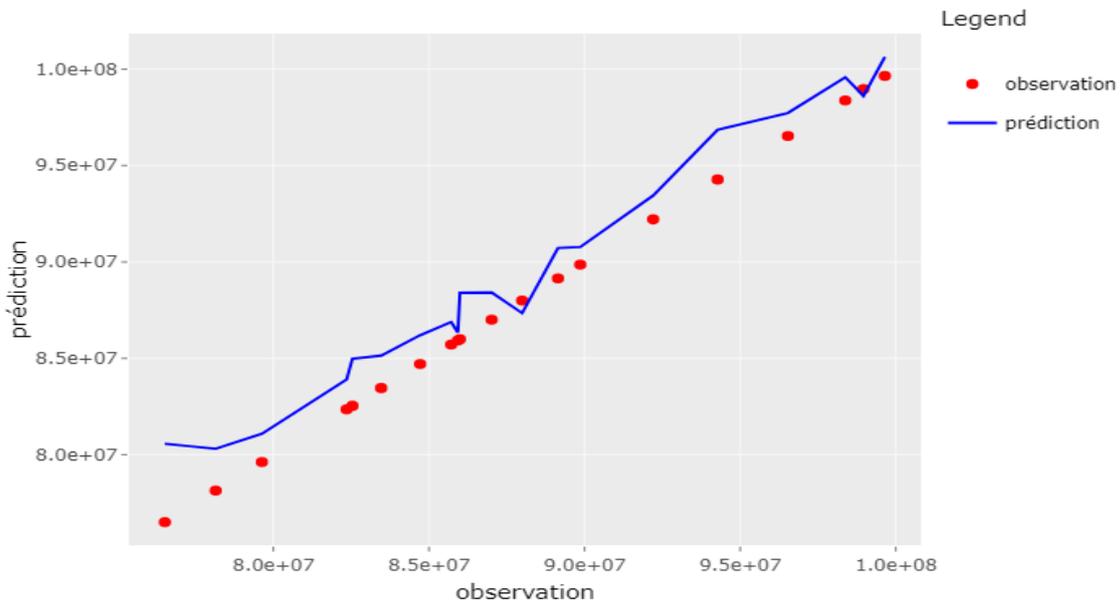


Figure 24 Ajustement prédiction et observation

5. MSE : Mean Square Error

Ridge	valeurs
R^2 ajusté	0.989
erreur relative moyenne \mathcal{B}_{app}	0.015
erreur relative moyenne \mathcal{B}_{test}	0.019

Table 25 Erreur \mathcal{B}_{test} et R^2 ajusté du modèle Ridge

La qualité de l'ajustement ne diffère pas trop des ajustements précédents, cependant les coefficients sont quand même bien estimés.

6.1.4 Régression Lasso

La fonction à minimiser dans la régression Lasso est similaire à celle de ridge. Cependant il y a une légère différence au niveau de la fonction de pénalité. En effet, la fonction à optimiser est de la forme suivante :

$$\min_{\beta_1, \dots, \beta_p} \sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (10)$$

Contrairement à Ridge, Lasso tend à mettre les coefficients de certaines variables très corrélées à 0 afin d'attribuer toute l'explication à une variable. Cette régression est d'une importance capitale en termes de sélection de variables.

Comme dans le cas de Ridge, on peut aussi afficher une *Lasso path coefficients* qui montre l'évolution des coefficients en fonction de la valeur de λ

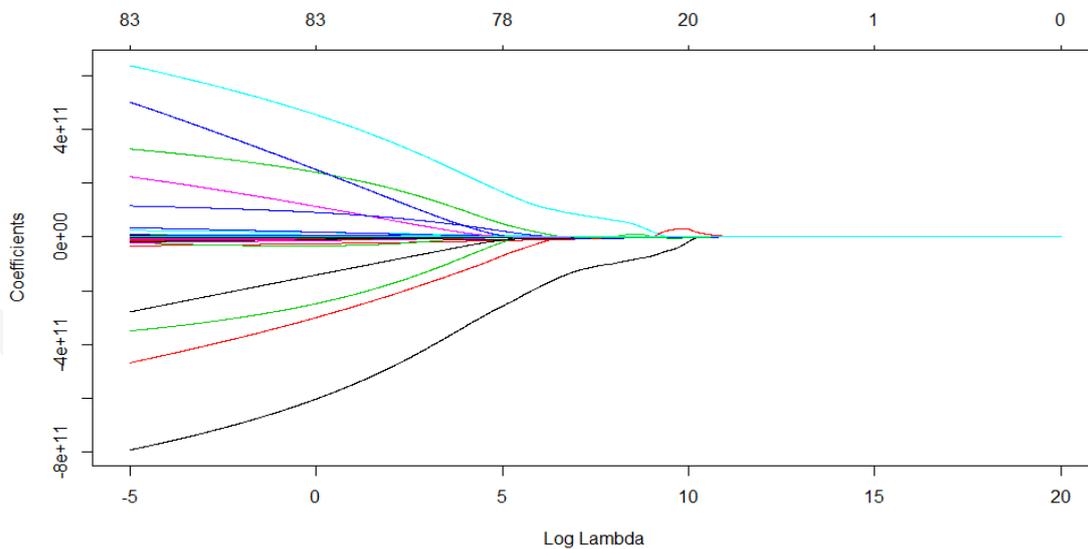


Figure 25 Évolution des paramètres calibrés en fonction de lambda

On observe que certains coefficients sont mis à 0 à partir d'une certaine valeur de lambda.

Comme dans le cas ridge, on va représenter aussi l'évolution de l'erreur sur la base d'apprentissage en fonction des valeurs de λ

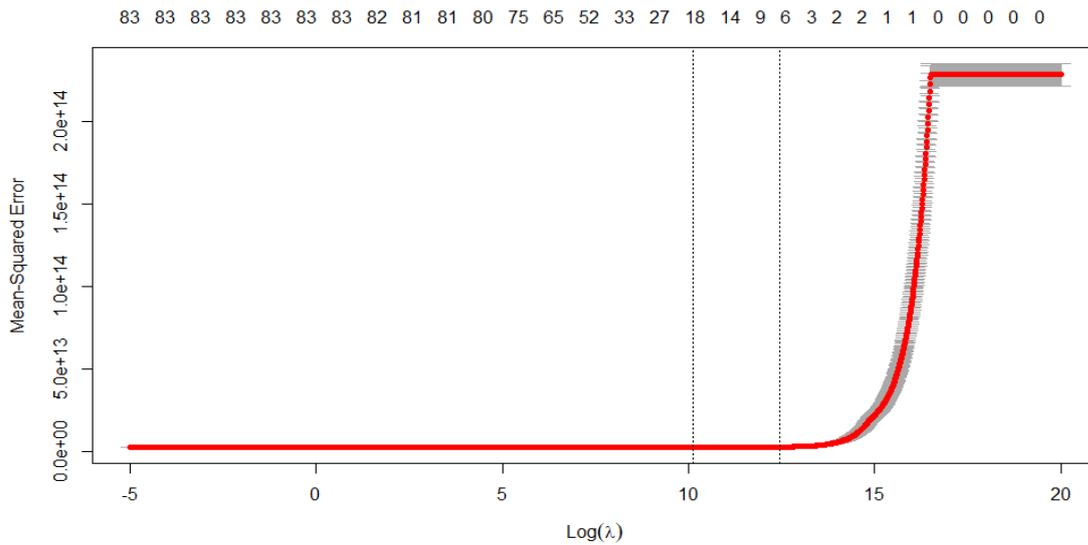


Figure 26 Évolution du mean square en fonction de lambda

On a trouvé que $\log(\lambda) = 10.14$ permettait de minimiser la MSE sur la base d'apprentissage en utilisant la Cross validation.

C'est donc la régression obtenue avec cette valeur de λ qui donne les résultats suivants :

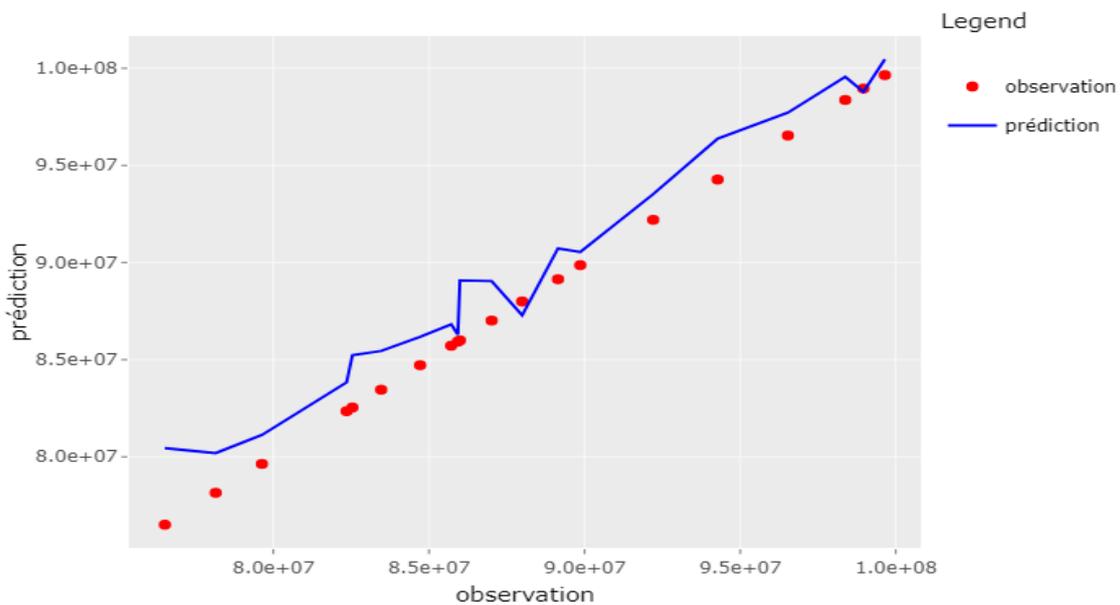


Figure 27 Évolution du mean square en fonction de lambda

Lasso	valeurs
R^2 ajusté	0.988
erreur relative moyenne \mathcal{B}_{app}	0.015
erreur relative moyenne \mathcal{B}_{test}	0.02

Table 26 Erreur \mathcal{B}_{test} et R^2 ajusté du modèle Lasso

On retrouve une qualité d'ajustement nettement proche de celle de ridge avec les mêmes ordres d'erreur sur la base d'apprentissage et sur la base de test.

6.1.5 Validation régression polynomiale

Les graphiques qui permettent de valider les hypothèses d'une régression linéaire sont :

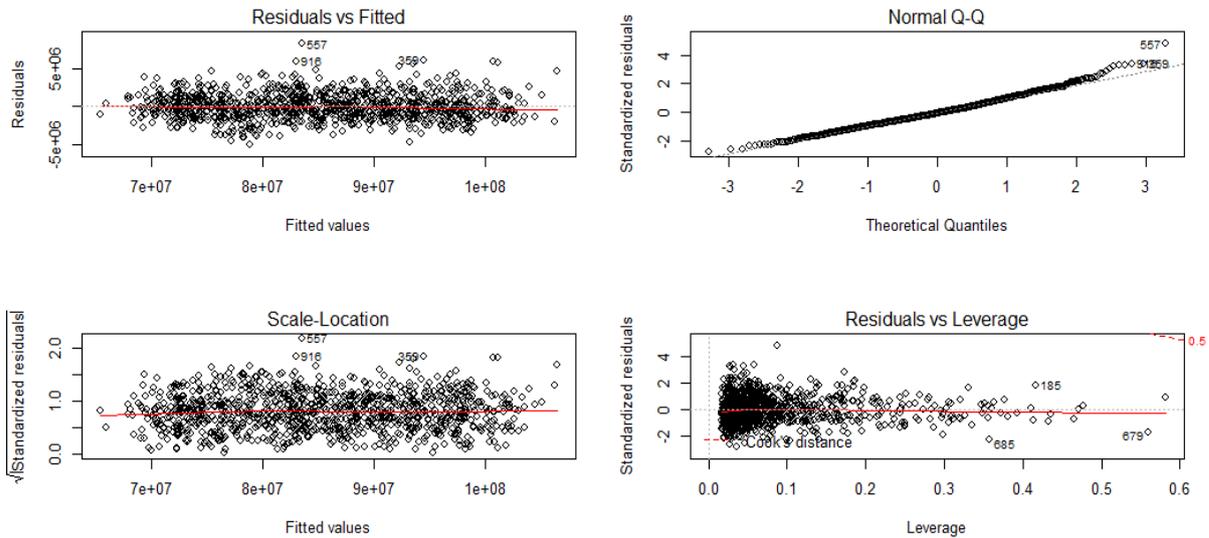


Figure 28 Validation modèle linéaire

En effet l'une des hypothèses principales qu'il faut vérifier est l'homoscédasticité des résidus. Cette hypothèse stipule que la variance des résidus ne dépend pas des valeurs prédites, on peut le voir comme une indépendance entre les résidus et les variables prédites. Ainsi, en observant le schéma 1 et le schéma 3, on peut clairement conclure que les résidus sont homoscédastiques car le nuage de points a une variance constante quelle que soit la valeur des prédictions. La seconde hypothèse importante à vérifier est la normalité des résidus du modèle, si les résidus étaient normaux, le nuage de points du QQ plot devrait se situer sur la première bissectrice. Ainsi, en observant le schéma 2 on peut affirmer sans doute que les résidus sont gaussiens. Donc nos hypothèses de régression linéaire sont validées.

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-nexialog/>

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

Habib FAYE
Manager R&D actuariat
Email : hfaye@nexialog.com

Areski COUSIN
Expert scientifique
Email : acousin-ext@nexialog.com