



Mémoire présenté devant le jury de l'EURIA en vue de l'obtention du  
Diplôme d'Actuaire EURIA  
et de l'admission à l'Institut des Actuares

le 6 septembre 2023

Par : Flora AUTER

Titre : Intégration du risque mortalité dans un générateur de scénarios économiques : évaluation et retraitement de l'impact de la COVID-19 sur la mortalité

Confidentialité : Non

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

**Membre présent du jury de l'Institut  
des Actuares :**

Marine HABART

Faris ROUCHATI

Signature :

**Entreprise :**

Milliman

Signature :

**Membres présents du jury de l'EURIA :** **Directeur de mémoire en entreprise :**

Franck VERMET

Eve Elisabeth TITON

Signature :

Amal ELFASSIHI

Signature :

**Invité :**

Signature :

**Autorisation de publication et de mise en ligne sur un site de diffusion  
de documents actuariels**

*(après expiration de l'éventuel délai de confidentialité)*

Signature du responsable entreprise :

Signature du candidat :



# Résumé

Un générateur de scénarios économiques (GSE) est un ensemble de modèles (taux, actions, spreads de crédit, inflation, etc.) permettant de simuler aléatoirement des scénarios d'évolution possibles des marchés financiers, qui ont une incidence sur la valeur et les performances de l'actif d'un organisme d'assurance. Néanmoins, en général les GSE n'intègrent pas le risque de mortalité, qui représente un risque majeur en assurance vie.

Dans la première partie du mémoire, le risque de mortalité est intégré à un GSE monde-réel. Pour cela, les interactions entre le risque de mortalité et les risques de marché sont étudiées. L'intégration de la mortalité permet la génération de tables de scénarios économiques simulant conjointement les différents facteurs de risque, incluant la mortalité.

La seconde partie du mémoire porte sur la modélisation de la mortalité en considérant différents facteurs de risque. La surmortalité exceptionnelle des années 2020 et 2021 due à la pandémie de la COVID-19 suscite des interrogations pour les assureurs et réassureurs, notamment dans le cadre de l'utilisation de leurs modèles internes. Ce mémoire propose des méthodes de retraitement des taux de mortalité impactés par la COVID-19. Par ailleurs, on se propose également dans cette partie d'étendre une approche de modélisation de la mortalité en considérant le facteur de risque climatique.

En dernière partie du mémoire, les tables de mortalité générées par le GSE sont testées en entrée d'un modèle de gestion actif-passif (ALM) afin de discuter l'intérêt de l'usage d'une table de mortalité stochastique par rapport à une table déterministe.

**Mots-clefs** — Mortalité, Assurance-Vie, Générateur de scénarios économiques, Monde-réel, *Asset Liability Management*, COVID-19, Pandémie, Retraitement de données

# Abstract

An economic scenario generator (ESG) is a set of models (interest rates, equities, credit spreads, inflation, etc.) used to randomly simulate possible financial market scenarios that have an impact on the value and performance of an insurance company's assets. However, ESG do not generally include mortality risk, which is a major risk in life insurance.

In the first part of the dissertation, mortality risk is integrated into a real-world ESG. To this end, the interactions between mortality risk and market risks are studied. The integration of mortality allows the generation of economic scenario tables that simulate the different risk factors jointly.

The second part of the dissertation deals with the modelling of mortality considering different risk factors. The exceptional excess mortality in 2020 and 2021 due to the COVID-19 pandemic raises questions for insurers and reinsurers, particularly when using their internal models. This dissertation proposes methods for adjusting mortality rates impacted by COVID-19. In addition, this section also proposes to extend a mortality modelling approach by considering the climatic risk factor.

In the final part of the dissertation, the mortality tables generated by the ESG are tested as inputs to an asset liability management (ALM) model in order to discuss the benefits of using a stochastic mortality table compared with a deterministic table.

**Keywords** — Mortality, Life insurance, Economic scenario generator, Real world, Asset Liability Management, COVID-19, Pandemic, Data adjustment

# Note de synthèse

## Contexte

Dans le cadre de la réglementation Solvabilité II, la gestion des risques est centrale dans le pilotage des compagnies d'assurance et de réassurance. Par ailleurs, les risques auxquels les assureurs sont soumis interagissent entre eux. Parmi ces risques, se trouve le risque de mortalité, objet de ce mémoire.

La première partie de ce mémoire porte sur l'intégration du facteur de risque de mortalité dans un GSE. La seconde partie est dédiée à la modélisation de la mortalité tenant compte de différents facteurs de risque. La surmortalité causée par la pandémie de la COVID-19 soulève des problématiques pour les assureurs qui motivent le retraitement de la mortalité impactée par la COVID-19.

## Intégration du facteur de risque de mortalité dans un GSE

Les GSE permettent la modélisation conjointe des facteurs de risque en fonction de corrélations afin de projeter des variables financières dans le temps. Nous intégrons le facteur de risque de mortalité dans un GSE.

L'article de Hanna et Devolder (2023) montre l'intérêt de considérer la corrélation entre le risque de mortalité et les risques financiers modélisés de manière stochastique dans la tarification d'un produit d'assurance vie. Or, la simulation conjointe de facteurs de risque est permise par les GSE. C'est pourquoi les corrélations entre le facteur de risque de mortalité et les facteurs de risque financiers modélisés usuellement pour les besoins des assureurs sont étudiées afin d'être prises en compte dans la modélisation de la mortalité par un GSE.

La mortalité est modélisée par le modèle de Lee-Carter qui est implémenté dans un GSE. Le modèle est calibré sur les données historiques de mortalité afin de projeter la mortalité et obtenir des tables de scénarios de mortalité. L'enjeu de l'intégration du facteur de risque de mortalité dans un GSE est l'étude des interactions entre le facteur de risque de mortalité et les facteurs de risque financiers.

Les corrélations entre les aléas des modèles stochastiques associés aux facteurs de risque rendent compte des interactions entre ces facteurs. Pour déduire ces corrélations, la struc-

ture des modèles stochastiques permet de déterminer la relation liant :

- la corrélation entre le facteur de risque de mortalité et les facteurs de risque financiers retenus,
- les corrélations entre les aléas des modèles de mortalité et des modèles financiers sous-jacents.

Parmi l'ensemble des facteurs de risque modélisés dans les GSE, les facteurs de risque corrélés au risque de mortalité sont sélectionnés. Pour cela, une recherche bibliographique et une analyse des corrélations des variables financières avec les taux de mortalité annuels et mensuels sont réalisées. Les risques d'inflation et de crédit sont sélectionnés en sachant que ce mémoire impose une contrainte de parcimonie.

L'indice inflation est modélisé par un modèle de Black-Scholes. On montre que la corrélation entre les log-incréments des taux de mortalité et les log-rendements de l'indice action est égale à la corrélation entre l'aléa du modèle de Lee-Carter et l'aléa du modèle de Black-Scholes.

Les spreads de crédit sont modélisés par le modèle de PCA-Vasicek. Le résultat asymptotique de la relation dépend des paramètres des modèles et de la corrélation entre les aléas du modèle de PCA-Vasicek avec l'aléa du modèle de Lee-Carter.

On considère les variables financières associées à l'économie américaine et les taux de mortalité américains pour l'application numérique de cette méthodologie. A partir de la corrélation entre les données historiques (tableau 1), les corrélations entre les aléas des modèles stochastiques (tableau 2) sont déterminées. Une optimisation numérique est nécessaire pour le risque de crédit.

<b>Facteur de risque observable</b>	<b>Corrélation empirique</b>
Indice inflation	0,35
Spread de crédit AA maturité 1 an	0,80
Spread de crédit AA maturité 5 an	0,64
Spread de crédit AA maturité 10 an	0,52

TABLE 1 – Corrélations empiriques des facteurs de risque observables avec le risque de mortalité

<b>Aléa des modèles</b>	<b>Corrélation entre aléas</b>
Black-Scholes - indice inflation	0,35
PCA-Vasicek Facteur 1 - Spread de crédit	-0,38
PCA-Vasicek Facteur 2 - Spread de crédit	0,38
PCA-Vasicek Facteur 3 - Spread de crédit	-0,45

TABLE 2 – Corrélations des aléas des modèles avec l'aléa du modèle de Lee-Carter

Le facteur de risque de mortalité a été intégré à un GSE en considérant les corrélations du facteur de risque de mortalité avec les facteurs de risque de l'inflation et du crédit. Pour cela, la formulation théorique permettant d'obtenir ces corrélations a été réalisée. Les tables de scénarios de mortalité produites par le GSE pourraient être utilisées et testées dans des modèles ALM réalistes.

## Modélisation de la mortalité tenant compte de l'expérience COVID-19

La COVID-19 a causé plus de 7 millions de morts dans le monde engendrant une surmortalité exceptionnelle en 2020 et 2021. L'introduction des données de 2020 dans le calibrage des modèles de mortalité stochastiques classiques tels que le modèle de Lee-Carter a plusieurs conséquences. En effet, la volatilité et la tendance du modèle augmentent avec l'introduction des données de 2020 dans le calibrage comme l'illustrent les graphiques en figure 1. En conséquence, l'espérance de vie projetée, présentée en table 3, est diminuée de plus de 2 ans pour les personnes de 47 ans aux États-Unis par exemple. Les incidences varient en fonction de l'ampleur de la pandémie dans les différents pays.

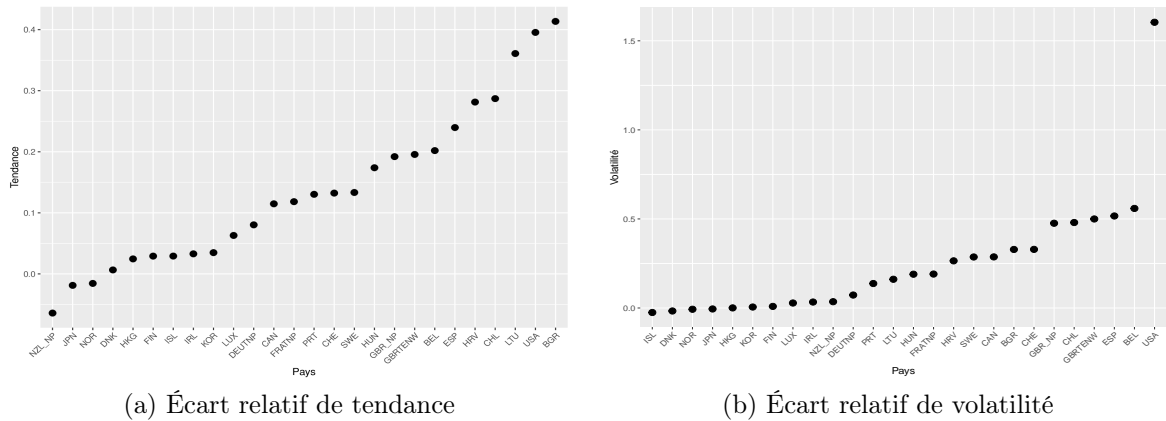


FIGURE 1 – Écart relatif des tendance (a) et volatilité (b) de  $\kappa_t^{(1)}$  des modèles de Lee-Carter dont les données de calibrage s'arrêtent en 2019 ou 2020

	USA		FRANCE		JAPON	
	2019	2020	2019	2020	2019	2020
<b>47 ans</b>	35,8	33,5	39,4	38,5	40,7	40,8
<b>67 ans</b>	17,7	16,3	20,2	19,6	20,8	20,9
<b>87 ans</b>	4,1	3,9	4,6	4,4	4,8	4,9

TABLE 3 – Espérances de vie (en années) projetées en 2021 avec le modèle de Lee-Carter

Dans le cadre de l'utilisation de modèles internes, le calibrage des modèles stochastiques de mortalité est mis à jour afin de calculer des chocs de mortalité. La correction des données

de mortalité utilisées dans le calibrage de modèles de mortalité est une option envisageable notamment dans le cadre d'un modèle interne, afin d'éviter tout double comptage entre le risque de mortalité/longévité et le risque de pandémie. Les effets immédiats de la pandémie sur la mortalité sont traités tandis que les effets à long terme de la pandémie COVID-19 sur la mortalité non pandémique ne sont pas pris en compte dans notre approche.

Deux méthodes de retraitement sont appliquées aux taux de mortalité annuels :

- Réajustement des données en excluant les décès liés à la COVID-19 : les décès attribués à la COVID-19 sont soustraits du décompte global des décès survenus au cours de l'année 2020.
- Remplacement des taux de mortalité de l'année 2020 par la moyenne des taux de mortalité annuels des années précédentes.

Trois méthodes de retraitement sont appliquées aux taux de mortalité hebdomadaires :

- *Week-specific averages* : le taux de référence hebdomadaire pour la semaine  $w$  est la moyenne des taux de mortalité pour les  $N$  années précédentes sur la semaine  $w$ .
- *Week-specific trends* : le taux de référence hebdomadaire pour la semaine  $w$  est déterminé en effectuant une régression linéaire des taux de mortalité des  $N$  années précédentes sur la semaine  $w$ .
- *Week-specific lower quartiles* : le taux de référence hebdomadaire pour la semaine  $w$  est déterminé en faisant la moyenne des quartiles des taux de mortalité des  $N$  années précédentes sur la semaine  $w$ .

Des méthodes alternatives ont été développées par Shkolnikov et al. (2022) pour répondre à la nécessité de déterminer la surmortalité causée par la COVID-19. Parmi ces méthodes alternatives, la méthode *alternative within-year* est basée sur la moyenne des 13 taux de mortalité hebdomadaires les plus bas au cours de l'année. Les taux de surmortalité calculés à partir de ces niveaux de référence ne sont pas faussés par les pics de mortalité passés et ne considèrent pas les excès de mortalité hivernale non pandémiques comme inévitables.

Les méthodes précitées sont appliquées afin de déterminer la surmortalité attribuable à la COVID-19, présentée dans le tableau 4. L'année 2020 se distingue par une surmortalité exceptionnelle due à la pandémie de COVID-19. De plus, les résultats de ces méthodes ne sont pas homogènes. En particulier, la méthode *week-specific lower quartiles* et la méthode *alternative within-year* conduisent à des résultats de surmortalité plus élevés qu'avec les autres méthodes.



	USA	France
<b>Décès COVID-19 recensés</b>	107	96
<b>Moyenne annuelle</b>	131	132
<b>Week-specific average</b>	161	90
<b>Week-specific lower quartiles</b>	178	124
<b>Week-specific trends</b>	106	130
<b>Alternative within-year</b>	134	193
<b>Alternative retrospective</b>	205	125

TABLE 4 – Surmortalité (pour 100 000 habitants) de l’année 2020 en fonction des méthodes de retraitement

Les chocs de mortalité sont calculés sur la base des projections des taux de mortalité. La prise en compte des données de mortalité retraitées conduit à des chocs de mortalité plus faibles, plus conformes à ceux obtenus en excluant l’année 2020 de la période de calibrage. Les résultats varient en fonction de la méthode d’ajustement. Les chocs de mortalité obtenus pour les États-Unis sont présentés en figure 2.

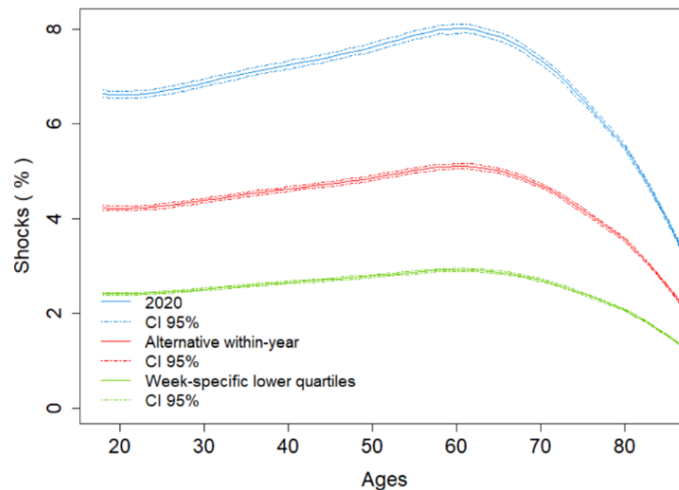


FIGURE 2 – Chocs de mortalité aux États-Unis avec ou sans méthode de retraitement

Le choc de mortalité moyen (en %) obtenu en retraitant les données de mortalité via la méthode *week-specific lower quartiles* aux données des États-Unis est  $2,46\% \pm 0,12$  tandis que le choc moyen avec le modèle sans retraitement est  $5,90\% \pm 0,30$ .

Les données de mortalité de l’année 2021 nécessitent également un retraitement. Les méthodes développées sur l’année 2020 sont appliquées à l’année 2021 afin de déterminer les chocs de mortalité à partir des données corrigées. Pour les États-Unis, le choc moyen de mortalité est  $6,96\% \pm 0,32$  lorsque les données des années 2020 et 2021 ne sont pas corrigées alors qu’en appliquant la méthode *week-specific lower quartiles* aux données de 2020 et 2021, le choc moyen de mortalité est  $2,32\% \pm 0,11$ .

## Utilisation de la mortalité stochastique dans un modèle ALM

Les modèles de gestion actif-passif (ALM) sont employés pour modéliser les interactions entre les actifs et les passifs du bilan d'une compagnie d'assurance ou de réassurance. Ces modèles prennent en entrée une loi de mortalité. Usuellement, une table de mortalité déterministe est employée. L'intérêt est de discuter de l'utilisation d'une table de mortalité stochastique par rapport à une table de mortalité déterministe dans un modèle ALM. De plus, les tables de mortalité stochastiques utilisées dans un modèle ALM peuvent être construites en prenant en compte les méthodes de retraitement des données impactées par la COVID-19. Il ressort que l'introduction d'une table de mortalité stochastique dans le modèle ALM considéré implique la diminution du *Best Estimate* (BE) bien que l'impact soit faible. Il est cependant envisageable que la mortalité stochastique ait une influence plus élevée dans le cadre de modèles ALM plus réalistes.

### Conclusion

Le facteur de risque de mortalité a été intégré dans un GSE. Le GSE a permis de prendre en compte ces interactions dans la projection de la mortalité en corrélant les aléas des modèles stochastiques associés aux facteurs de risque.

Dans la continuité de l'approche mise en œuvre dans ce mémoire, des travaux futurs sont envisageables pour améliorer l'intégration du facteur de risque de mortalité dans un GSE. Il est par exemple possible d'intégrer d'autres modèles de mortalité stochastiques ou de corrélérer le facteur de risque de mortalité à d'autres facteurs de risque financiers en suivant la même méthodologie.

Des méthodes de retraitement de la mortalité COVID-19 ont été implémentées. Il ressort que les approches les plus robustes et les plus fiables pour quantifier la surmortalité sont basées sur l'estimation de la surmortalité hebdomadaire. Ces méthodes de retraitement permettent d'obtenir des chocs de mortalité en prenant les années 2020 et 2021 dans le calibrage des modèles. Les chocs obtenus sont du même ordre de grandeur que les chocs calculés en arrêtant la modélisation de la mortalité en 2019.

# Executive summary

## Context

In the context of Solvency II regulations, risk management is central to the management of insurance and reinsurance companies. Furthermore, the risks to which insurers are subject interact with each other. One of these risks is mortality risk, which is the subject of this dissertation.

The first part of this dissertation deals with the integration of the mortality risk factor into an ESG. The second part is dedicated to mortality modelling taking into account different risk factors. The excess mortality caused by the COVID-19 pandemic raises issues for insurers that motivate the reprocessing of mortality impacted by COVID-19.

## Integrating the mortality risk factor into an ESG

ESG allows the combined modelling of risk factors based on correlations to project financial variables over time. We include the mortality risk factor in an ESG.

The article by Hanna and Devolder (2023) shows the value of considering the correlation between mortality risk and stochastically modelled financial risks when pricing a life insurance product. However, the joint simulation of risk factors is made possible by ESG. This is why the correlations between the mortality risk factor and the financial risk factors usually modelled for the needs of insurers are studied so that they can be taken into account when modelling mortality using an ESG.

Mortality is modelled by the Lee-Carter model which is implemented in an ESG. The model is calibrated on historical mortality data to project mortality and obtain stochastic mortality scenario tables. The challenge of integrating the mortality risk factor into an ESG is to study the interactions between the mortality risk factor and the financial risk factors.

The correlations between the random numbers driving the dynamic of the stochastic models associated with the risk factors reflect the interactions between these factors. To deduce these correlations, the structure of the stochastic models is used to determine the relationship between :

- the correlation between the mortality risk factor and the financial risk factors selected,

- the correlations between the random numbers driving the mortality models and the underlying financial models.

Among the risk factors modelled in an ESG, those correlated with mortality risk are selected. To do this, a literature search and an analysis of the correlations between financial variables and annual/monthly mortality rates are carried out. Inflation and credit risks are selected, bearing in mind that this dissertation imposes a parsimony constraint.

The inflation index is modelled by a Black-Scholes model. We show that the correlation between the log-increments of mortality rates and the log-returns of the equity index is equal to the correlation between the random numbers driving the dynamic of the Lee-Carter model and the random numbers driving the dynamic of the Black-Scholes model. Credit spreads are modelled by the PCA-Vasicek model. The asymptotic result of the relationship depends on the parameters of the models and the correlation between the randomness of the PCA-Vasicek model and the randomness of the Lee-Carter model.

Financial variables associated with the US economy and US mortality rates are considered for the numerical application of this methodology. From the correlation between the historical data (table 5), the correlations between the random numbers driving the dynamic of stochastic models (table 6) are determined. Numerical optimization is required for credit risk.

<b>Observable risk factor</b>	<b>Empirical correlation</b>
Inflation index	0.35
Credit spread AA maturity 1 year	0.80
Credit spread AA maturity 5 years	0.64
Credit spread AA maturity 10 years	0.52

TABLE 5 – Empirical correlations of observable risk factors with mortality risk

<b>Model random numbers</b>	<b>Correlation between random numbers</b>
Black-Scholes - Inflation index	0.35
PCA-Vasicek Factor 1 - Credit spread	-0.38
PCA-Vasicek Factor 2 - Credit spread	0.38
PCA-Vasicek Factor 3 - Credit spread	-0.45

TABLE 6 – Correlation between the random numbers driving the Lee-Carter model and the random numbers driving Black-Scholes and PCA-Vasicek models

The mortality risk factor was integrated into an ESG by considering the correlations of the mortality risk factor with the inflation and credit risk factors. To do this, the theoretical formulation allowing these correlations to be obtained was carried out. The

mortality scenario tables produced by the ESG could be used and tested in realistic ALM models.

## Mortality modelling based on the COVID-19 pandemic

COVID-19 has caused more than 7 million deaths worldwide, leading to exceptional excess mortality in 2020 and 2021. The introduction of 2020 data into the calibration of classic stochastic mortality models such as the Lee-Carter model has several consequences. The volatility and trend of the model increase with the introduction of 2020 data into the calibration, as illustrated by the graphs in figure 3. As a result, the projected life expectancy, presented in table 7, is reduced by more than 2 years for people aged 47 in the United States, for example. Incidence varies according to the scale of the pandemic in different countries.

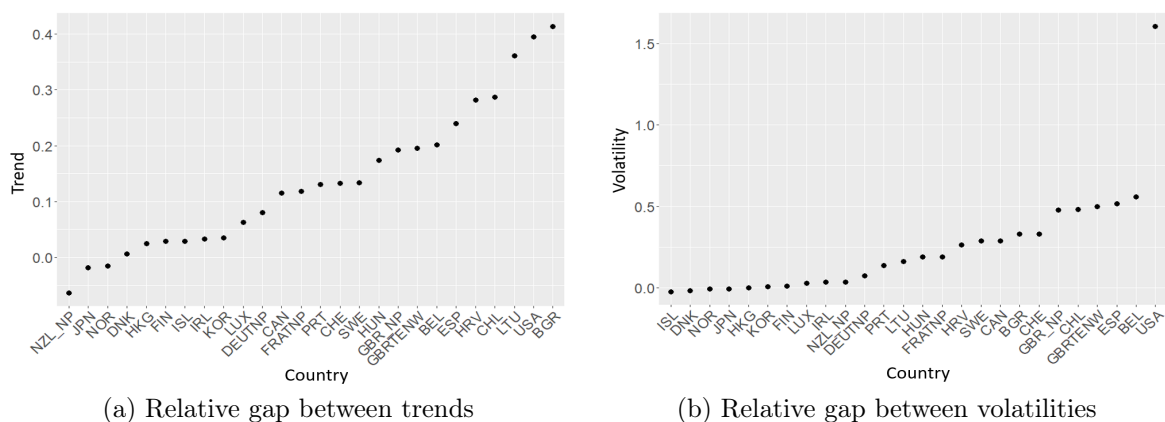


FIGURE 3 – Relative gap of trend (a) and volatility (b) of  $\kappa_t^{(1)}$  for Lee-Carter models with calibration data ending in 2019 or 2020

	USA		FRANCE		JAPON	
	2019	2020	2019	2020	2019	2020
<b>47 years old</b>	35,8	33,5	39,4	38,5	40,7	40,8
<b>67 years old</b>	17,7	16,3	20,2	19,6	20,8	20,9
<b>87 years old</b>	4,1	3,9	4,6	4,4	4,8	4,9

TABLE 7 – Life expectancy (in years) projected to 2021 using the Lee-Carter model

When using internal models, the calibration of stochastic mortality models is updated to calculate mortality shocks. Adjusting the mortality data used in the calibration of mortality models is a possible option, particularly in the context of an internal model, to avoid any double counting between mortality/longevity risk and pandemic risk. The immediate effects of the pandemic on mortality are addressed, while the long-term effects of the COVID-19 pandemic on non-pandemic mortality are not taken into account in our

approach.

Two adjustment methods are applied to annual mortality rates :

- Adjusting the data by removing the COVID-19 deaths : Deaths attributed to COVID-19 are subtracted from the overall death count for 2020.
- Replacing the mortality rates of the index year 2020 by the average of the annual mortality rates over the last previous  $N$  years.

Three adjustment methods are applied to weekly mortality rates :

- Week-specific averages : The weekly reference rate for week  $w$  is the average of the mortality rates for the  $N$  years before the index year  $y$  in week  $w$ .
- Week-specific trends : The weekly reference rate for week  $w$  is determined by performing a linear regression of the mortality rates of the  $N$  previous years on week  $w$ .
- Week-specific lower quartiles : The weekly reference rate for week  $w$  is determined by averaging the quartile mortality rates of the  $N$  previous years over the week  $w$ .

Alternative adjustment methods have been developed by Shkolnikov and others (2022) to address the need to determine the excess mortality caused by COVID-19. Among these alternative methods, the alternative within-year method is based on the average of the 13 lowest weekly mortality rates during the year. The excess mortality rates calculated from these reference levels are not distorted by past mortality peaks and do not consider non-pandemic excess winter mortality as inevitable.

All the above methods are applied to determine the excess mortality attributable to COVID-19, presented in table 8. 2020 stands out for its exceptional excess mortality due to the COVID-19 pandemic. Moreover, the results of these methods are not homogeneous. In particular, the week-specific lower quartiles method and the alternative methods lead to higher excess mortality results than the other methods.

	USA	France
<b>Official recorded COVID-19 deaths</b>	107	96
<b>Average historical annual mortality</b>	131	132
<b>Week-specific average</b>	161	90
<b>Week-specific lower quartiles</b>	178	124
<b>Week-specific trends</b>	106	130
<b>Alternative within-year</b>	134	193
<b>Alternative retrospective</b>	205	125

TABLE 8 – Excess mortality (per 100,000 inhabitants) in 2020 according to adjusting methods

Mortality shock values are calculated based on projected mortality rates. Taking into account adjusted mortality data leads to lower mortality shocks, more aligned with those obtained by excluding the year 2020 from the calibration period. Results vary according

to the adjustment method. The mortality shocks obtained for the United States are shown in figure 4.

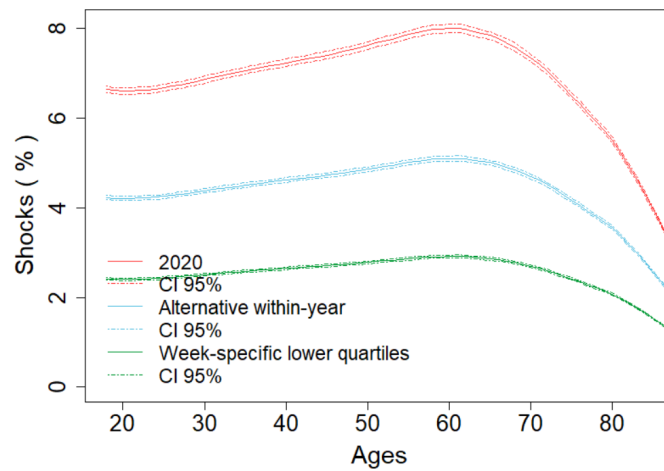


FIGURE 4 – Mortality shocks in the United States with and without adjustment method

The average mortality shock (in %) obtained by reprocessing mortality data using the week-specific lower quartiles method to U.S. data is  $2.46\% \pm 0.12$ , while the average shock with the model without reprocessing is  $5.90\% \pm 0.30$ .

Mortality data for 2021 also require adjustment. The methods developed for 2020 are applied to 2021 to determine mortality shocks based on adjusted data. For the United States, the average mortality shock is  $6.96\% \pm 0.32$  when the data are not adjusted, whereas by applying the week-specific lower quartiles method to the 2020 and 2021 data, the average mortality shock is  $2.32\% \pm 0.11$ .

## Using stochastic mortality in an ALM model

Asset and liability management (ALM) models are used to model the interactions between assets and liabilities on the balance sheet of an insurance or reinsurance company. These models take into account a mortality law. Usually, a deterministic mortality table is used. The interest is to discuss using a stochastic versus a deterministic mortality table in an ALM model. Furthermore, stochastic mortality tables used in an ALM model can be constructed by considering methods applied to adjust data impacted by COVID-19. We notice that introducing a stochastic mortality table in the ALM model used implies a reduction in the Best Estimate (BE), although the impact is small. However, it is conceivable that stochastic mortality could have a greater influence on more realistic ALM models.

## Conclusion

The mortality risk factor was integrated into an ESG. The ESG made it possible to take these interactions into account in the mortality projection by correlating the random

numbers driving the stochastic models associated with the risk factors. Following the approach implemented in this adjustment, future studies are feasible to improve the integration of the mortality risk factor in an ESG. For example, other stochastic mortality models could be integrated, or the mortality risk factor could be correlated with other financial risk factors using the same methodology.

Methods for adjusting COVID-19 mortality have been implemented. They show that the most robust and reliable approaches to quantifying excess mortality are based on estimates of weekly excess mortality. These adjusting methods make it possible to obtain mortality shocks by taking the years 2020 and 2021 into account in model calibration. The shocks obtained are of the same order of magnitude as those calculated by stopping mortality modelling in 2019.



# Remerciements

Je tiens à remercier particulièrement mes tuteurs en entreprise, Eve TITON, senior consultante chez Milliman, et Amal ELFASSIHI, consultant chez Milliman, pour m'avoir encadrée tout au long de la réalisation de ce mémoire. Je les remercie tant pour leurs précieux conseils sur le plan technique et leur expertise que pour leur disponibilité, ainsi que leur relecture.

Je remercie grandement le cabinet de conseil en actuariat Milliman pour m'avoir donné l'opportunité de réaliser ce mémoire dans un cadre enrichissant. Je suis reconnaissante envers Alexandre BOUMEZOUED, directeur du département Recherche & Développement, pour ses conseils avisés et son suivi au cours du mémoire.

Mes remerciements vont également à l'équipe Recherche & Développement de Milliman pour leur accueil, leurs conseils et les idées tout au long de la réalisation de ce mémoire.

Je remercie Franck VERMET, directeur de l'EURIA et tuteur académique pour son suivi et son investissement dans mes travaux ainsi que pour sa relecture.

Je remercie plus généralement les équipes pédagogiques de l'EURIA et de l'INSA de Rennes pour la qualité de la formation et pour m'avoir fourni les connaissances scientifiques nécessaires à la réalisation de ce mémoire.

Merci à Tinhinane et Céline pour m'avoir supportée pendant cette année et à Paul pour son aide précieuse.

Enfin, je remercie ma famille pour son soutien et son intérêt dans ce projet.

# Table des matières

Résumé	i
Abstract	ii
Note de synthèse	iii
Executive summary	ix
Remerciements	xv
Introduction	2
<b>I Intégration du facteur de risque de mortalité dans un générateur de scénarios économiques</b>	<b>3</b>
<b>1 Contexte et définition du facteur de risque mortalité</b>	<b>4</b>
1.1 Définition et usage des générateurs de scénarios économiques . . . . .	4
1.2 Quelques généralités sur les GSE . . . . .	6
1.3 Présentation des modèles de mortalité . . . . .	8
1.4 Définition du facteur de risque de mortalité dans un GSE . . . . .	13
<b>2 Corrélation du facteur de risque de mortalité avec les autres facteurs de risque</b>	<b>16</b>
2.1 Motivations . . . . .	16
2.2 Étude des corrélations entre les facteurs de risque . . . . .	17
2.3 Calculs théoriques des corrélations . . . . .	23
2.4 Calcul des corrélations entre mouvements browniens . . . . .	27
<b>3 Rapport automatique de calibrage</b>	<b>32</b>
3.1 Calibrage du modèle de mortalité . . . . .	32
3.2 Simulation de la mortalité . . . . .	35

<b>II</b>	<b>Modélisation de la mortalité en considérant différents facteurs de risque</b>	<b>37</b>
<b>4</b>	<b>Modélisation de la mortalité en tenant compte de l'expérience COVID-19</b>	<b>38</b>
4.1	Motivations . . . . .	38
4.2	Les données de mortalité . . . . .	39
4.3	Impact de la pandémie de COVID-19 sur les taux de mortalité projetés par les modèles . . . . .	42
4.4	Retraitement des points de données COVID-19 pour l'année 2020 . . . . .	50
4.5	Applications des méthodes à l'année 2021 . . . . .	67
4.6	Modèle CMI . . . . .	75
4.7	Autres impacts de la COVID-19 sur la mortalité . . . . .	76
<b>5</b>	<b>Autres facteurs de risque impactant la mortalité</b>	<b>79</b>
5.1	Modélisation de la mortalité en tenant compte du risque climatique . . . . .	79
5.2	Le risque des technologies portables . . . . .	86
<b>III</b>	<b>Utilisation de la mortalité stochastique dans un modèle ALM</b>	<b>89</b>
<b>6</b>	<b>Application dans un modèle ALM</b>	<b>90</b>
6.1	Présentation du modèle ALM . . . . .	90
6.2	Résultats du modèle ALM . . . . .	93
	<b>Conclusion</b>	<b>98</b>
	<b>Annexes</b>	<b>100</b>
<b>A</b>	<b>Outils mathématiques</b>	<b>100</b>
<b>B</b>	<b>Corrélation entre le facteur de risque de mortalité et d'autres facteurs de risque financiers</b>	<b>102</b>
<b>C</b>	<b>Impact de la pandémie de COVID-19 sur la modélisation avec les modèles APC, CBD et M7</b>	<b>109</b>
<b>D</b>	<b>Application des méthodes de retraitement</b>	<b>112</b>
<b>E</b>	<b>Application du modèle de Lee-Carter tenant compte du risque climatique</b>	<b>113</b>
<b>F</b>	<b>Résultats ALM</b>	<b>115</b>
	<b>Bibliographie</b>	<b>121</b>
	<b>Table des figures</b>	<b>124</b>



# Introduction

La réforme Solvabilité II impose aux compagnies d'assurance européennes une réglementation afin de garantir leur solvabilité. Parmi les exigences composant les trois piliers de Solvabilité II, le capital de solvabilité requis (SCR) permet de s'assurer de la bonne santé économique de la compagnie. Le calcul du SCR est fait au regard du profil de risque de l'assureur et correspond au capital nécessaire pour faire face à une ruine économique à horizon d'un an avec un niveau de confiance 99,5% impliquant la réalisation de projections. La réglementation suggère une formule standard mais les assureurs peuvent aussi développer leur modèle interne.

En plus des exigences en capital, les assureurs apportent une attention particulière à la gestion d'actifs afin de piloter la compagnie. En effet, le principe de la gestion actif-passif est de modéliser les interactions entre les actifs et les passifs du bilan d'une compagnie d'assurance ou de réassurance. C'est dans ce contexte que les générateurs de scénarios économiques (GSE) sont devenus des outils incontournables des compagnies d'assurance.

Les GSE en monde réel permettent la projection stochastique de variables financières à un horizon donné représentant les facteurs de risque auxquels l'assureur est confronté. De manière générale, les risques de taux, de crédit, action, inflation, immobilier sont simulés. En revanche, ces générateurs n'incluent pas le risque de mortalité dans la construction de ces scénarios.

La mortalité est pourtant un enjeu majeur pour les assureurs. En effet, en assurance vie particulièrement, les assureurs doivent quantifier les risques que leurs portefeuilles subissent. Par exemple, la mortalité est un risque important pour des portefeuilles de rente. C'est pourquoi des scénarios de mortalité, déterministes ou stochastiques, sont construits afin de modéliser la mortalité future. Ces scénarios de mortalité sont utilisés comme entrée des modèles de gestion actif-passif et sont également pris en compte pour mesurer la solvabilité des assureurs dans le cadre du calcul du SCR.

Puisque la mortalité est une problématique des assureurs, il est nécessaire de pouvoir l'analyser et la modéliser. La mortalité présente une structure particulière selon l'âge de la population considérée. De plus, les événements tels que les guerres, les pandémies, les phénomènes climatiques influent également sur la mortalité. Le dernier en date est la pandémie COVID-19 touchant un grand nombre de pays à travers le monde, affectant inégalement les populations d'un point de vue géographique, social et selon leur âge. La pandémie a eu un

impact sur la mortalité que les assureurs cherchent à mesurer. En effet, l'augmentation de la mortalité a des conséquences directes sur les portefeuilles des assureurs mais également sur leurs calculs actuariels impliquant des projections de la mortalité dans le temps tel que l'augmentation des chocs de mortalité et donc l'augmentation du SCR. A ce jour, la plupart des acteurs sur le marché arrêtent leur historique de mortalité en 2019 car la prise en compte des années impactées par la pandémie est en cours d'investigation dans différentes études actuarielles. C'est dans ce contexte que des méthodes de retraitement des données de mortalité des années COVID-19 représentent un enjeu pour les assureurs afin d'obtenir des projections qui ne soient pas impactées par la pandémie.

Ce mémoire se décompose en trois parties. La première partie porte sur l'intégration du risque de mortalité dans un GSE. Après un rappel des principes de construction d'un GSE et des risques qui le composent, une recherche bibliographique et une étude quantitative des corrélations entre le risque de mortalité et les autres risques sont menées. Ainsi, le GSE fournit des scénarios de projection de la mortalité cohérents avec les conditions de marché. Les sorties du GSE sont présentées et interprétées dans un rapport automatique.

La seconde partie est consacrée à la modélisation de la mortalité prenant en compte l'impact de la COVID-19. L'impact de la pandémie sur la mortalité de 2020 est analysé quantitativement afin d'envisager des méthodes de retraitement des données pandémiques. Une recherche bibliographique a mis en évidence les méthodes de retraitement usuelles ou alternatives. L'objectif de l'implémentation de ces méthodes est de pouvoir mesurer l'impact global de la COVID-19 au delà du nombre de décès recensés par les organismes de santé gouvernementaux et d'obtenir des projections de la mortalité cohérentes avec les données historiques sans l'impact d'une donnée aberrante. Ces méthodes sont implémentées sur différents pays et les résultats sont comparés.

Enfin, les scénarios stochastiques de mortalité sont utilisés en entrée d'un modèle de gestion d'actif-passif (ALM). Les modèles ALM sont habituellement utilisés avec une table de mortalité déterministe. L'intérêt est de discuter de l'utilisation d'une table de mortalité stochastique par rapport à une table de mortalité déterministe dans un modèle ALM.

Première partie

Intégration du facteur de risque de  
mortalité dans un générateur de  
scénarios économiques

# Chapitre 1

## Contexte et définition du facteur de risque mortalité

Ce chapitre fixe le cadre d'étude de l'intégration de la mortalité dans un GSE. D'une part, les GSE sont définis ainsi que leurs usages. D'autre part, les modèles de mortalité stochastiques usuels utilisés dans ce mémoire sont rappelés. Ce chapitre précise la structure d'un GSE et comment le facteur de risque de mortalité est intégré dans cet outil.

### 1.1 Définition et usage des générateurs de scénarios économiques

L'assurance est une activité de service dans laquelle l'assureur s'engage à verser des prestations lors de la survenance d'un risque aux particuliers ou entités économiques contre le paiement de primes. L'objectif principal de l'assurance est de couvrir les risques auxquels les individus, biens ou patrimoines économiques sont soumis. Parmi les risques que les assureurs cherchent à analyser et maîtriser, se trouvent, par exemple, les risques biométriques, climatiques et les risques financiers. Dans ce contexte, la gestion des compagnies d'assurance ne peut échapper au pilotage des risques financiers.

L'assurance vie est particulièrement concernée par le risque financier dans la mesure où un contrat d'assurance vie engage l'assuré au versement de primes contre des rémunérations garanties par l'assureur sur une période allant du moyen au long terme. Ces contrats incluent des garanties et options financières qui peuvent être activées sur la durée de vie du contrat. La valorisation des contrats nécessite des simulations permettant de prendre en compte les différents comportements des assurés en fonction des conditions économiques.

Les GSE sont des outils numériques indispensables aux compagnies d'assurance et aux différents acteurs du marché. Le principe général est la projection à un horizon donné de variables économiques et financières à l'aide de modèles stochastiques sous-jacents se rattachant aux risques impactant le bilan des compagnies. Un GSE fournit des simulations aléatoires des risques économiques qui reflètent les conditions des marchés. On retrouve l'utilisation de simulations des GSE dans diverses applications. On en détaille quelques exemples ci-dessous.



Tout d'abord, les simulations sont utilisées pour le calcul des provisions techniques. Ces provisions correspondent à la meilleure estimation (*best estimate*) des engagements de l'assureur à laquelle une marge de prudence est ajoutée, appelée la marge pour risque (*risk margin*). Le *best estimate* correspond à "la moyenne pondérée par leur probabilité des flux des trésoreries futurs, compte tenu de la valeur temporelle de l'argent (valeur actuelle attendue des flux futurs), estimée sur la base de la courbe des taux sans risque pertinents" selon le Code des assurances, Article R351-2 (2015). La marge pour risque est "calculée de manière à garantir que la valeur des provisions techniques prudentielles est équivalente au montant qu'une entreprise agréée pour pratiquer les opérations d'assurance ou de réassurance demanderait pour reprendre et honorer les engagements d'assurance et de réassurance" [Code des assurances, Article R351-2 (2015)]. Ces provisions sont des éléments importants du bilan de solvabilité de la compagnie d'assurance ou de réassurance et permet d'en déduire la prime d'assurance exigée par l'assureur au souscripteur. Ce calcul peut s'avérer complexe dans un contexte où les contrats d'assurance sont de plus en plus sophistiqués. Les GSE interviennent dans le processus de projection de ces contrats jusqu'à leur échéance en fonction des conditions économiques des marchés.

Un autre usage des GSE est leur utilisation dans le calcul du capital de solvabilité requis (SCR). Les compagnies d'assurance ou de réassurance doivent disposer d'un montant de fonds propres suffisant pour respecter les exigences énoncées dans la réglementation Solvabilité II. Parmi ces exigences, on trouve le SCR qui correspond au montant de capital qui doit être immobilisé par la compagnie afin de pouvoir faire face à des situations de crises économiques extrêmes. Par définition, le SCR est le montant de capital dont doit disposer une compagnie pour faire face à une ruine économique au niveau de confiance 99.5% à horizon un an. Le SCR correspond au quantile de la distribution des pertes à horizon un an, à la manière de la *Value at Risk*. Deux méthodes sont proposées aux assureurs : l'utilisation de la Formule Standard développée par l'EIOPA ou la construction d'un modèle interne permettant d'évaluer la distribution des pertes. Il est aussi possible d'utiliser une méthode hybride, un modèle interne partiel.

Le calcul du SCR via la Formule Standard est une approche modulaire. Cette démarche permet le calcul d'un SCR par facteur de risque. Le SCR est obtenu par l'agrégation des SCR par facteur de risque à l'échelle intra-modulaire puis inter-modulaire.

L'utilisation d'un modèle interne requiert le développement de celui-ci, son maintien dans le temps et sa justification auprès des autorités de contrôle. Son utilisation est plus complexe mais offre une manière de calculer le SCR plus cohérente avec le profil de risque de la compagnie d'assurance ou de réassurance. Une approche connue mais très coûteuse est la méthode des simulations dans les simulations, qui consiste à projeter le passif depuis la date actuelle de la valorisation sous les conditions de marchés simulées par le GSE.

Enfin, citons également l'utilisation des GSE en gestion des risques. La gestion des risques est une des quatre fonctions clés composant le système de gouvernance établi dans le pilier 2 de Solvabilité II. Les compagnies cherchent à évaluer leur profil de risque et les possibles dérives de celui-ci afin de pouvoir respecter de manière permanente les exigences de solvabilité. Les simulations établies par les GSE sont utilisées pour évaluer le profil de

risque de la compagnie de manière prospective. En effet, des anticipations sur les variables financières telles que les variations de l'inflation ou des taux d'intérêt peuvent être intégrées dans un GSE fournissant des simulations cohérentes avec les hypothèses fixées. Les GSE sont également utilisés pour la gestion de la compagnie d'assurance à plus court terme pour garantir l'équilibre des portefeuilles d'actifs et des stratégies de couverture. Les GSE sont donc des outils numériques utilisés pour la gestion des risques. Ils prennent en compte le contexte économique actuel et à venir en simulant les variables financières de manière cohérente avec les interactions entre les différents facteurs de risque.

## 1.2 Quelques généralités sur les GSE

L'objectif d'un GSE est la simulation de variables financières se rattachant à des sous-modules de risques financiers modélisés dans l'outil faisant partie du risque de marché. Les facteurs de risque intégrés dans un GSE dépendent de la manière dont l'assureur veut projeter son passif. Nous nous intéressons à ces facteurs de risque :

- Le **taux nominal** est le taux définissant les intérêts à payer lors de la souscription d'un emprunt.
- Le **taux réel** est le taux nominal corrigé de l'inflation et de la prime de risque.
- Le **taux de change** est le cours d'une devise par rapport à une autre, pour les GSE modélisant plusieurs économies.
- Le **spread de crédit** est la différence entre taux risqué et taux sans risque, et représente la probabilité qu'une compagnie fasse défaut en fonction de sa notation.
- L'**indice boursier** est l'indice calculé à partir de la valeur d'un ensemble d'actions permettant d'évaluer la croissance économique sur un secteur.
- L'**indice immobilier** est l'indice permettant de rendre compte de l'évolution du cours de l'immobilier sur un secteur économique.
- L'**indice inflation** est l'indice permettant de rendre compte de l'évolution de l'inflation sur un secteur économique.

Pour obtenir les tables de scénarios économiques, plusieurs étapes sont indispensables et certaines relèvent des choix de l'utilisateur en fonction du cadre d'application des tables de simulations :

- La première étape est le choix de l'univers de probabilité. Deux types de GSE sont à distinguer : les GSE sous la probabilité risque-neutre et les GSE sous la probabilité monde-réel. Le choix de la probabilité est en relation directe avec les objectifs finaux de l'emploi des simulations.
- La seconde étape est la sélection des facteurs de risque modélisés conjointement. Certains facteurs de risque peuvent être écartés selon le profil de risque de la compagnie ou en fonction des analyses envisagées. A chaque facteur de risque, des données et des modèles sont associés.

- Le calibrage des modèles est la troisième étape. Les modèles sont ajustés aux données pour en déterminer les paramètres optimaux. Cette étape inclut un travail sur la théorie des modèles et l'application numérique de ces modèles.
- La dernière étape est la simulation. Les trajectoires sont générées à partir des modèles construits à l'étape précédente tout en prenant en compte les corrélations entre facteurs de risque qui sont déterminées en amont de l'utilisation du GSE.
- Au cours des étapes de calibrage et de simulation, des processus de validation sont mis en place incluant une validation quantitative avec des tests statistiques et qualitative telle qu'une validation graphique laissée à la main de l'utilisateur. Ces tests permettent de s'assurer de la cohérence des trajectoires générées. Dans l'article de Hibbert et al. (2001), les critères qualitatifs que doit satisfaire un GSE sont énoncés. Ils permettent de donner un aperçu des contraintes à respecter :
  - **La représentativité et l'interprétation économique** : les modèles du GSE doivent reproduire correctement les comportements des actifs financiers en capturant leurs caractéristiques et comportements tout en respectant les principes économiques en vigueur.
  - **La parcimonie** : la complexité d'un GSE doit être mesurée, l'objectif étant de trouver un compromis entre simplicité et précision.
  - **La transparence** : les modèles doivent être explicables et une bonne présentation des résultats est requise pour assurer la qualité de communication quant à l'utilisation du GSE et de ses résultats.

Ainsi, les apports de ce mémoire à un GSE sont mis en oeuvre en respectant ces bonnes pratiques.

Il existe deux familles de GSE, les GSE en univers risque-neutre et les GSE en univers monde-réel.

Les générateurs de scénarios économiques en univers risque-neutre (*Risk-neutral economic scenarios generator*, GSE RN) sont employés pour l'évaluation de produits financiers, *pricing*, selon les probabilités extraites des marchés financiers. Dans le cadre de l'univers risque-neutre, le calibrage du GSE se fait sur des données de marché afin de pouvoir répliquer le prix des produits dérivés cotés sur les marchés quand cela est possible. On dit alors que le GSE est *market consistent*, la valorisation fournie par le GSE est cohérente avec les marchés financiers. Ce type de GSE est basé sur les attentes pures du marché. L'hypothèse essentielle de l'univers risque-neutre est l'Absence d'Opportunité d'Arbitrage (AOA), c'est à dire que sur le marché, avec un investissement de départ nul, il est impossible de réaliser un gain sans risque, le marché est supposé parfait [Buckdhan (2023)]. La complétude des marchés est aussi supposée, c'est à dire que tous les flux sont répliquables par un portefeuille d'actifs sans risque et d'actifs risqués. La réalisation de ces hypothèses implique l'existence de la mesure de probabilité risque-neutre qui est l'unique mesure de probabilité sous laquelle la valeur des actifs actualisée est une martingale.

La gestion du risque est le cadre général de l'utilisation des GSE en univers monde-réel et est aussi une différence majeure avec les GSE sous la probabilité risque-neutre.

De manière intuitive, les GSE risque-neutre se veulent neutres face au risque alors que le principe de l'approche monde-réel est d'intégrer le risque pour pouvoir optimiser la stratégie d'investissement, piloter les risques ou faire des prévisions financières. De ce fait, les simulations en monde-réel sont réalisées à partir des données historiques prenant en compte les rendements historiques du portefeuille et permettant de capter les attentes du marché à propos des fluctuations futures. Les modèles utilisés dans un GSE monde-réel sont différents des modèles employés en risque-neutre puisque ces modèles se réfèrent à des distributions de prix dans lesquelles les préférences de risque du marché sont intégrées. C'est une approche réaliste de l'évolution des marchés financiers. Contrairement aux GSE risque-neutre, il n'y a pas d'exigence équivalente à la martingalité ou à la *market consistency* pour les GSE monde-réel.

Les générateurs de scénarios économiques monde-réel (*Real-world economic scenarios generator*, GSE RW) sont utilisés pour divers objectifs :

- Étude de la gestion actif-passif (ALM) pour trouver par exemple une allocation optimale du portefeuille de la compagnie.
- Etape de calcul du SCR dans le cadre d'un modèle interne, les scénarios monde-réel sont réalisés à horizon un an.
- Application des processus ORSA (*Own Risk and Solvency Assessment*) avec des scénarios en monde-réel sur l'horizon d'un plan d'affaires.
- Valorisation des compagnies d'assurances dans le cadre d'opérations de fusion et acquisition. Pour la vente ou l'achat, les scénarios monde-réel permettent l'évaluation des actifs ou passifs incluant une prime de risque.

Les GSE fournissent des simulations des risques économiques qui reflètent les conditions des marchés. Les GSE monde-réel sont des outils indispensables pour la gestion des risques des compagnies d'assurance ou de réassurance. C'est dans ce contexte que la mortalité stochastique est intégrée à un GSE monde-réel.

### 1.3 Présentation des modèles de mortalité

Les notations liées aux tables de mortalité et aux grandeurs associées sont définies dans cette partie [Vermet (2023)]. Les données de mortalité sont des données annuelles âge par âge :

- l'année civile notée  $t$  s'étend du temps exact  $t$  au temps exact  $t + 1$
- $x$  désigne l'âge de l'individu

$d(x, t)$  est défini comme le nombre de décès enregistrés d'individus d'âge exact  $x$  au dernier anniversaire au cours de l'année civile  $t$ . L'exposition centrale au risque  $E(x, t)$  est définie comme la somme des durées de vie de chaque individu d'âge exact  $x$  au dernier anniversaire au cours de l'année civile  $t$ .

La mortalité peut être caractérisée par un modèle à deux états : l'état "vivant" et l'état "décédé", notés respectivement  $V$  et  $D$ . On parle ainsi du processus de vie et de

mort tel que  $Z = (Z_u)_{u \geq 0}$  représente l'état de l'individu à l'instant  $u$ ,  $Z_u \in \{V, D\}$ . Soit  $T$  une variable aléatoire positive caractérisant la durée de séjour de l'individu dans l'état vivant : c'est la durée de vie de l'individu. La force de mortalité d'un individu d'âge  $x$ , qui s'interprète comme la probabilité de décès par unité de temps au voisinage de  $x$ , est définie telle que :

$$\mu_x = \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{P}(Z_{x+h} = D | Z_x = V) = \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{P}(x \leq T \leq x+h | T > x) \quad (1.1)$$

Puisque la force de mortalité dépend du temps, une composante temporelle est ajoutée à l'équation (1.1). Soit  $\mu(x, t)$  la force de mortalité d'un individu d'âge  $x$  au temps  $t$  qui peut se décomposer comme  $t = x + d$  avec  $d$  l'année de naissance de l'individu :

$$\mu_{x,t} = \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{P}(x \leq T \leq x+h | T > x, d)$$

Les taux de mortalité  $q_{x,t}$  sont définis et correspondent à la probabilité qu'un individu d'âge  $x$  décède au cours de l'année  $t$ , c'est à dire entre le temps  $t$  et  $t+1$  :

$$q_{x,t} = \mathbb{P}(T < x+1 | T \geq x, d)$$

La fonction de survie à  $u$  années correspondant à la probabilité qu'un individu de date de naissance  $d$  survive  $u$  années s'exprime en fonction de la force de mortalité :

$$S(u) = \exp\left(-\int_0^u \mu(x, x+d) dx\right)$$

Dans un objectif de simplification, l'hypothèse d'intensité de décès fractionnaire constante est émise. C'est à dire que le risque de mortalité est supposé constant sur des intervalles de la forme  $[x, x+1[ \times [t, t+1[$ , ainsi  $\forall x, \forall t, \forall dx \in [0, 1[, \forall dt \in [0, 1[, \mu_{x+dx, t+dt} = \mu_{x,t}$ . Sous cette hypothèse, il existe une relation explicite entre l'estimateur du maximum de vraisemblance du taux de mortalité  $\widehat{q}_{x,t}$  et celui de la force de mortalité  $\widehat{\mu}_{x,t}$  :

$$\widehat{q}_{x,t} = 1 - \exp(-\widehat{\mu}_{x,t}) \text{ et } \widehat{\mu}_{x,t} = \frac{d(x, t)}{E(x, t)}$$

Sous cette hypothèse, le nombre de décès observés  $d(x, t)$  suit une loi de Poisson :  $D(x, t) \sim \mathcal{P}(\mu_{x,t} E(x, t))$  en supposant l'exposition centrale  $E_{x,t}$  fixée. Les grandeurs considérées dans les tables de mortalité et utilisées dans la modélisation de la mortalité sont les taux de mortalité  $q_{x,t}$ , c'est pourquoi l'exposition initiale au risque est définie :

$$E^0(x, t) = E(x, t) + \frac{1}{2}d(x, t)$$

De ce fait,  $D(x, t)$  suit une loi Binomiale :  $D(x, t) \sim \mathcal{B}(E^0(x, t), q_{x,t})$ , et l'estimateur du maximum de vraisemblance de  $q_{x,t}$  devient :

$$\widehat{q}_{x,t} = \frac{d(x, t)}{E^0(x, t)}$$

L'espérance de vie est une grandeur largement utilisée car elle donne une mesure concrète en unité de temps contrairement aux taux de mortalité qui peuvent paraître plus

abstraits. L'espérance de vie cohorte prend en compte de futures améliorations possibles pour chaque âge en calculant l'espérance de vie pour une cohorte spécifique :

$$e_x(t) = \sum_{u=1}^{110-x} {}_u p_x(t) \quad \text{avec} \quad {}_u p_x(t) = \prod_{k=0}^{u-1} [1 - m(x+k, t+k)] \quad (1.2)$$

Les modèles de mortalité considérés dans ce mémoire sont les modèles GAPC, les modèles généralisés de mortalité âge-période-cohorte (*Generalized Age-Period-Cohort*) [Villegas et al. (2015)]. Quatre éléments caractérisent cette famille :

1. **La composante aléatoire** : la variable aléatoire  $D_{xt}$  qui correspond au nombre de décès suit une distribution de Poisson ou une distribution Binomiale, ainsi :

$$D_{xt} \sim \mathcal{P}(E_{xt}\mu_{xt}) \quad \text{ou} \quad D_{xt} \sim \mathcal{B}(E_{xt}^0, q_{xt})$$

avec respectivement  $\mathbb{E}(D_{xt}/E_{xt}) = \mu_{xt}$  et  $\mathbb{E}(D_{xt}/E_{xt}^0) = q_{xt}$ .

2. **La composante systématique** : les effets à l'âge  $x$ , l'année  $t$  et l'année de naissance, correspondant à la cohorte  $c = t - x$  sont captés par le prédicteur  $\eta_{xt}$  défini par :

$$\eta_{x,t} = \alpha_x + \sum_{i=0}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x}$$

avec :

- $\alpha_x$  : composante en âge.
- $N \geq 0$  : un entier correspondant au nombre de termes âge-période décrivant les tendances de mortalité. Chaque  $\kappa_t^{(i)}$ ,  $i = 1, \dots, N$  précise la tendance de la mortalité et  $\beta_x^{(i)}$  module l'effet selon l'âge.
- $\beta_x^{(0)} \gamma_{t-x}$  : effet de cohorte avec le paramètre de cohorte  $\gamma_{t-x}$  et  $\beta_x^{(0)}$  modulant l'effet selon l'âge.

$\kappa_t^{(i)}$ ,  $i = 1, \dots, N$ , et  $\gamma_{t-x}$  sont des processus stochastiques. Ils permettent la projection stochastique de la mortalité.

3. **La fonction de lien** : notée  $g$ , elle fait le lien entre la composante aléatoire et la composante systématique telle que :

$$g\left(\mathbb{E}\left(\frac{D_{xt}}{E_{xt}}\right)\right) = \eta_{xt}$$

Plusieurs fonctions de lien sont possibles. Généralement, la fonction de lien logarithmique est couplée à la distribution de Poisson et la fonction de lien logit est couplée à la distribution binomiale.

4. **L'ensemble de contraintes sur les paramètres** : il est nécessaire d'ajouter des contraintes sur les paramètres des modèles stochastiques de mortalité pour les rendre identifiables. Une fonction de contrainte fait correspondre un vecteur arbitraire de paramètres

$$\theta = (\alpha_x, \beta_x^{(1)}, \dots, \beta_x^{(N)}, \kappa_t^{(N)}, \beta_x^{(0)}, \gamma_{t-x})$$

à un vecteur de paramètres transformés

$$v(\theta) = \tilde{\theta} = (\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, \dots, \tilde{\beta}_x^{(N)}, \tilde{\kappa}_t^{(N)}, \tilde{\beta}_x^{(0)}, \tilde{\gamma}_{t-x}).$$

Plusieurs modèles particuliers de cette famille sont décrits dans les paragraphes suivants. La nomenclature des modèles en M1 à M8 est introduite par Cairns et al. (2009).

### 1.3.1 Modèle de Lee-Carter

Lee et Carter (1992) ont développé un modèle sur des données américaines, il s'agit d'un modèle de prévision de la mortalité. Ce modèle, appelé modèle M1, est une approche à la fois robuste et simple de la projection de la mortalité partant du principe que le nombre de décès suit une distribution de Poisson de fonction de lien logarithmique appliquée à la force de mortalité. Le modèle est construit de cette manière :

$$\eta_{x,t} = \log(\mu_{x,t}) = \alpha_x + \beta_x \kappa_t \quad (1.3)$$

avec :

- $\mu_{x,t}$  : force de mortalité à l'âge  $x$ , l'année  $t$ .
- $\alpha_x$  : structure statique en âge et s'interprète comme la valeur moyenne des  $\log(\mu_{x,t})$  au cours du temps.
- $\beta_x$  : paramètre de sensibilité à l'âge  $x$  par rapport à l'évolution dans le temps des  $\kappa_t$ .
- $\kappa_t$  : évolution temporelle de la mortalité.  $\kappa_t$  est modélisé et projeté à l'aide d'un processus ARIMA, *Auto-Regressive Integrated Moving Average*. Une marche aléatoire avec dérive est souvent utilisée pour des résultats satisfaisants. La série temporelle est construite de cette manière :

$$\kappa_t = \kappa_{t-1} + \delta + \epsilon_t \quad \text{avec} \quad \epsilon_t \sim \mathcal{N}(0, \sigma^2) \text{ iid.} \quad (1.4)$$

$\delta$  est le paramètre de dérive et  $\epsilon_t$  est un bruit blanc gaussien de variance  $\sigma^2$ .

Le modèle n'est identifiable qu'en fixant deux contraintes sur les paramètres :

$$\begin{aligned} \sum_x \beta_x &= 1 \\ \sum_t \kappa_t &= 0 \end{aligned}$$

Les estimateurs des paramètres du modèle  $(\widehat{\alpha}_x, \widehat{\beta}_x, \widehat{\kappa}_t)$  sont déterminés à l'aide du critère des moindres carrés :

$$(\widehat{\alpha}_x, \widehat{\beta}_x, \widehat{\kappa}_t) = \operatorname{argmin} \sum_{x,t} (\log(\widehat{\mu}_{x,t}) - \alpha_x - \beta_x \kappa_t)^2$$

où  $\widehat{\mu}_{x,t}$  sont les taux de mortalité historiques.

### 1.3.2 Modèle Age-Period-Cohort

Le modèle Age-Period-Cohort, communément appelé modèle APC ou modèle M3, est un modèle dont la structure est commune à celle du modèle de Renshaw-Haberman introduisant un paramètre de cohorte. La différence entre le modèle APC et le modèle de

Renshaw-Haberman est la pondération de l'effet cohorte par un coefficient dépendant de l'âge. Le modèle est défini ainsi :

$$\log(\mu_{x,t}) = \alpha_x + \kappa_t + \gamma_{t-x} \quad (1.5)$$

avec :

- $\mu_{x,t}$  : force de mortalité à l'âge  $x$ , l'année  $t$ .
- $\alpha_x$  : structure statique en âge.
- $\kappa_t$  : série temporelle construite de manière identique à la série temporelle du modèle de Lee-Carter.
- $\gamma_{t-x}$  : effet cohorte.

Pour rendre le modèle identifiable, trois contraintes sont fixées :

$$\begin{aligned} \sum_t \kappa_t &= 0 \\ \sum_c \gamma_c &= 0 \\ \sum_c c\gamma_c &= 0 \end{aligned}$$

Sous ces contraintes, l'effet cohorte varie autour de 0 sans tendance linéaire. De la même manière que pour le modèle de Lee-Carter, les paramètres sont estimés par un algorithme d'optimisation des moindres carrés.

### 1.3.3 Modèle de Cairns-Blake-Dowd

Le modèle de Cairns-Blake-Dowd, aussi appelé modèle CBD ou modèle M5, se différencie dans sa structure qui présente deux termes âge-période sans paramètre spécifique à l'âge. Le modèle est défini tel que :

$$\log(\mu_{x,t}) = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} \quad (1.6)$$

où  $\bar{x}$  est la moyenne d'âge des données observées. La prévision des taux de mortalité est obtenue en projetant les deux séries temporelles  $\kappa_t^{(1)}$  et  $\kappa_t^{(2)}$  qui sont des marches aléatoires avec dérive. Il n'est pas nécessaire de fixer de contraintes sur les paramètres, le modèle est identifiable. Haberman et Renshaw supposent que les décès suivent une loi binomiale de fonction de lien logit sur les taux de mortalité pour estimer les paramètres du modèle.

### 1.3.4 Modèle M7

Le modèle M7 est une extension du modèle CBD auquel un effet cohorte et un effet quadratique selon l'âge ont été ajoutés, la formulation du modèle devient :

$$\eta_{x,t} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + ((x - \bar{x})^2 - \widehat{\sigma_x^2})\kappa_t^{(3)} + \gamma_{t-x} \quad (1.7)$$



où  $\widehat{\sigma}_x^2$  est la valeur moyenne de  $(x - \bar{x})^2$ . Ce modèle n'est pas identifiable, c'est pourquoi des contraintes sur les paramètres sont ajoutées. Elles assurent que l'effet cohorte fluctue autour de 0 et n'a pas de tendance linéaire :

$$\begin{aligned}\sum_c \gamma_c &= 0 \\ \sum_c c\gamma_c &= 0 \\ \sum_c c^2\gamma_c &= 0\end{aligned}$$

Dans la suite de ce chapitre, les modèles présentés ci-dessus sont implémentés pour modéliser la mortalité. Le calibrage des modèles stochastiques de mortalité est réalisé en maximisant la log-vraisemblance qui dépend des hypothèses faites sur la distribution du nombre de décès. Des algorithmes d'optimisation tels que Newton-Raphson sont employés mais aussi des modèles linéaires ou non linéaires généralisés dont les modèles de mortalité stochastiques sont des exemples. Pour le calibrage des modèles, la librairie *R* utilisée est **StMoMo**, développée par Villegas et al. (2015), dans laquelle tous les modèles de mortalité standards sont implémentés.

## 1.4 Définition du facteur de risque de mortalité dans un GSE

Cette section porte sur les principes de l'intégration du facteur de risque de mortalité, de ses caractéristiques et de la manière dont il interagit avec les autres facteurs de risque existants.

### 1.4.1 Données d'entrée du GSE

La mortalité est intégrée dans un GSE monde-réel qui nécessite un historique de données en entrée. Pour le facteur de risque de mortalité, les tables de mortalité historiques constituent les données d'entrée : le nombre de décès et l'exposition. Il existe plusieurs possibilités quant aux choix des tables, ces choix impactent la modélisation :

- **La zone géographique** : les GSE modélisent une ou plusieurs économies. Trois économies sont considérées : l'euro (EUR), le dollar américain (USD) et la livre sterling (GBP). Les données de mortalité à charger doivent être cohérentes avec le choix de l'économie, mais plusieurs options sont envisageables. Pour l'euro ou la livre sterling, les tables de l'ensemble de la zone géographique, table européenne et table du Royaume-Uni, peuvent être considérées. Il est aussi possible de sélectionner un seul pays selon un critère fixé ou d'agréger plusieurs tables.
- **Les âges** : les données de mortalité peuvent être présentées âge par âge ou par classe d'âge avec une maille de taille variable.
- **La période** : la profondeur de l'historique des données est un choix dans la construction des tables de mortalité. Il est nécessaire d'avoir un historique suffisamment profond pour pouvoir calibrer les modèles.

### 1.4.2 Données de corrélation

L'intérêt d'un GSE est de pouvoir modéliser conjointement les différents facteurs de risque. Le lien entre les facteurs de risque est exprimé à l'aide de la corrélation entre ces facteurs. Les variables considérées sont les taux réels et les taux nominaux, les indices actions, les spreads de crédit, l'indice des prix de l'immobilier et l'indice d'inflation.

Dans le cadre de la modélisation et de la construction des scénarios économiques, ce sont les mouvements browniens des modèles manifestant le caractère aléatoire des scénarios, définis en annexe, qui doivent être corrélés. À partir de la matrice de corrélation empirique entre les variables économiques, est construite une matrice de corrélations entre les aléas des modèles considérés, à noter qu'un modèle peut compter plusieurs aléas. Le passage de la matrice de corrélation entre facteurs de risque à la matrice de corrélation entre les mouvements browniens est expliqué dans le chapitre 2.

Les grandeurs considérées pour le calcul des corrélations empiriques ne sont pas directement l'historique des variables financières. En effet, les calculs théoriques montrent l'utilité de considérer les incréments ou le logarithme des incréments pour le calcul des corrélations. Les grandeurs considérées sont présentées dans le tableau 1.1. En ce qui concerne la mortalité, le logarithme des taux de mortalité est modélisé.

### 1.4.3 Modélisation et rapport d'analyse

Chaque facteur de risque est modélisé dans le GSE par un modèle sélectionné. Pour le risque de mortalité, le modèle de Lee-Carter est implémenté, mais d'autres modèles comme CBD, APC ou M7 peuvent aussi l'être. Le choix du modèle dépend des caractéristiques de celui-ci. Les étapes de calibrage et de simulation sont les parties d'exécution du GSE selon les paramètres d'entrée.

Un ensemble de tests de validation est réalisé pour s'assurer de la cohérence des scénarios avec la réalité du marché économique. Ces tests de validation reposent sur des critères quantitatifs ou sur la visualisation de résultats graphiques. C'est l'objet du rapport d'analyse qui est en sortie du GSE. Les graphiques de résultats du calibrage et des simulations sont présentés. Ce rapport a plusieurs visées :

- **La transmission de l'information.** Il permet à l'utilisateur d'avoir une trace de ce qui a été implémenté dans le GSE, et de pouvoir en visualiser des résultats.
- **L'analyse des résultats.** Les résultats présentés sont analysés et interprétés. Pour le modèle de mortalité, il n'existe pas de test permettant d'affirmer si le modèle s'ajuste correctement aux données. Le calcul du score du  $R^2$  est un indicateur par exemple, d'où l'intérêt de le présenter dans ce rapport.

Le rapport automatique portant sur la modélisation de la mortalité dans le GSE est présenté dans le chapitre 3.

Classe d'actifs	Variable simulée	Variable observable	Facteur de risque observable
Taux nominaux	Taux nominal forward $F_j(t)$	Taux nominal continûment composé $(R^n(t, m))$ $= -\frac{1}{m} \ln(P^n(t, m))_{t, m}$	Variation des taux nominaux continûment composés $(R^n(t, m))$ $-R^n(t - \Delta, m)_{t, m}$
Taux réels	Taux court réel $r_t$	Taux réel continûment composé $(R^r(t, m))$ $= -\frac{1}{m} \ln(P^r(t, m))_{t, m}$	Variation des taux réels continûment composés $(R^r(t, m))$ $-R^r(t - \Delta, m)_{t, m}$
Inflation	Indice inflation $I(t)$	CPI $I(t)$	Log-rendement déflaté de l'indice $\left( \ln \left( \frac{I(t)P^n(t-\Delta, \Delta)}{I(t-\Delta)P^r(t-\Delta, \Delta)} \right) \right)_t$
Action	Indice action $S(t)$	Indice action $(S(t))_t$	Log-rendement déflaté de l'indice $\left( \ln \left( \frac{S(t)P^n(t-\Delta, \Delta)}{S(t-\Delta)} \right) \right)_t$
Immobilier	Indice immobilier $RE(t)$	Indice immobilier $(RE(t))_t$	Log-rendement déflaté de l'indice $\left( \ln \left( \frac{RE(t)P^n(t-\Delta, \Delta)}{RE(t-\Delta)} \right) \right)_t$
Crédit	Intensité de défaut $\lambda_\sigma(t)$	Spreads logarithmiques $(Spd(t, m))_{t, m=5}$ $= -\frac{1}{m} \ln \left( \frac{P^c(t, m)}{p^n(t, m)} \right)$	Variation des spreads logarithmiques $(Spd(t, m))$ $-Spd(t - \Delta, m)_{t, m=5}$
Mortalité	Logarithme des taux de mortalité $\eta(x, t)$	Logarithme des taux de mortalité $(\eta(x, t))_{x, t}$	Log-incrément des taux de mortalité $(\eta(x, t) - \eta(x, t - \Delta))_{x, t}$

TABLE 1.1 – Description des facteurs de risque

### Conclusion

L'objectif de cette partie est d'intégrer la mortalité à un GSE en univers monde-réel à la manière des modèles financiers. La mortalité est calibrée et simulée dans le GSE en prenant en compte les corrélations entre les facteurs de risque. L'objectif est d'obtenir des tables de simulations de la mortalité et de présenter les paramètres et les résultats dans le rapport d'analyse.

## Chapitre 2

# Corrélation du facteur de risque de mortalité avec les autres facteurs de risque

Dans ce chapitre, la corrélation du risque de mortalité avec les autres facteurs de risque est étudiée. L'objectif est de pouvoir simuler les scénarios de mortalité conjointement avec les facteurs de risque auxquels la mortalité est corrélée.

### 2.1 Motivations

L'étude des corrélations entre la mortalité et les risques financiers est motivée par plusieurs études qui montrent l'intérêt d'abandonner l'hypothèse d'indépendance entre les risques financiers et les risques actuariels. Tout d'abord, Hanna et Devolder (2023) intègrent la dépendance entre risques financiers et risque de mortalité dans la tarification d'un produit d'assurance-vie mixte réparti entre un fond de participation et un fond en unité de compte. La corrélation influe sur la tarification du produit montrant également la pertinence de cette corrélation. Ainsi, l'étude de la corrélation entre la mortalité et les risques financiers est un vrai sujet actuariel. L'article suppose l'existence de cette corrélation, comme nous le faisons, sur la base de plusieurs analyses.

Parmi les études appuyant l'hypothèse de l'existence de la corrélation entre risque de mortalité et risques financiers, on trouve d'abord l'article de Dacorogna et Cadena (2015). La dépendance est étudiée dans plusieurs pays et sous différents cadres. Lorsque les 10 années les plus néfastes à l'évolution de la mortalité sont sélectionnées, la corrélation entre la mortalité et les marchés financiers est importante alors que le faible effet de la mortalité sur les marchés financiers est observé lorsque l'ensemble de l'historique est considéré. Ensuite, l'étude de Maurer (2011) met en évidence que des changements démographiques peuvent engendrer des variations à long terme des taux d'intérêt et des rendements boursiers.

Cette recherche bibliographique soutient l'existence d'une corrélation entre le risque de mortalité et les risques financiers. Elle motive l'étude des corrélations réalisées dans ce mémoire. Pour cela, les formulations théoriques permettent de faire le lien entre les corrélations entre le facteur de risque de mortalité et les facteurs de risque financiers retenus et les corrélations entre les aléas des modèles de mortalité et ceux des modèles financiers. Les variables avec lesquelles la mortalité est corrélée sont mises en évidence, tous les facteurs

de risque ne sont pas corrélés entre eux.

## 2.2 Étude des corrélations entre les facteurs de risque

### 2.2.1 Analyse des corrélations empiriques françaises

Dans un premier temps, nous calculons les corrélations sur les données françaises. Les corrélations empiriques sont calculées entre chaque couple de facteurs de risque observables présentés dans le chapitre 1. Pour cela, les données historiques sont utilisées. Les différentes classes de facteurs de risque regroupent différentes variables financières pour chaque économie. La corrélation entre les taux de mortalité français, américains avec ces variables financières est déterminée. Les variables financières sont pré-sélectionnées dans un souci de visibilité des graphiques. Les maturités 1 an et 10 ans sont privilégiées donnant des informations sur le taux court et le taux long.

Les variables financières sont dépendantes du temps tandis que les taux de mortalité dépendent de l'âge et du temps. Par conséquent, pour déterminer ces corrélations, il est nécessaire d'avoir un historique des taux de mortalité indépendant de l'âge. Pour cela plusieurs méthodes d'agrégation des taux de mortalité peuvent être considérées :

- **Agrégation totale des taux de mortalité** : pour une année donnée, le taux de mortalité correspond au nombre total de décès sur l'exposition totale de la population sur cette année.
- **Agrégation avec la pyramide des âges** : les taux de mortalité par âge sont agrégés en pondérant par la proportion de personnes de cet âge dans la population totale.

Un intervalle de confiance est calculé autour des corrélations de Pearson  $\rho$ . Le calcul de l'intervalle de confiance est basé sur la z-transformation de Fisher. Pour chaque paire  $(i, j)$  où  $i$  et  $j$  sont les réalisations de deux risques distincts,  $\rho_{i,j}$  est la corrélation empirique entre les deux facteurs de risque. On définit la variable  $Z_{i,j}$  telle que

$$Z_{i,j} = \frac{1}{2} \ln \left( \frac{1 + \widehat{\rho}_{i,j}}{1 - \widehat{\rho}_{i,j}} \right) \sim \mathcal{N} \left( \frac{1}{2} \ln \left( \frac{1 + \rho_{i,j}}{1 - \rho_{i,j}} \right), sd = \frac{1}{\sqrt{n-3}} \right) \quad (2.1)$$

La z-transformation est appliquée aux simulations  $Z_{i,j}$  telle que  $\widehat{\rho}_{i,j} = \frac{e^{2Z_{i,j}} - 1}{e^{2Z_{i,j}} + 1}$  pour chaque corrélation de Pearson.

Les résultats de ces corrélations sont présentés figure 2.1. Les corrélations sont calculées sur un historique d'environ 20 ans selon les variables. La largeur de l'intervalle de Fisher dépend du nombre de points. Avec 20 points, les intervalles de confiance sont larges et comprennent 0.

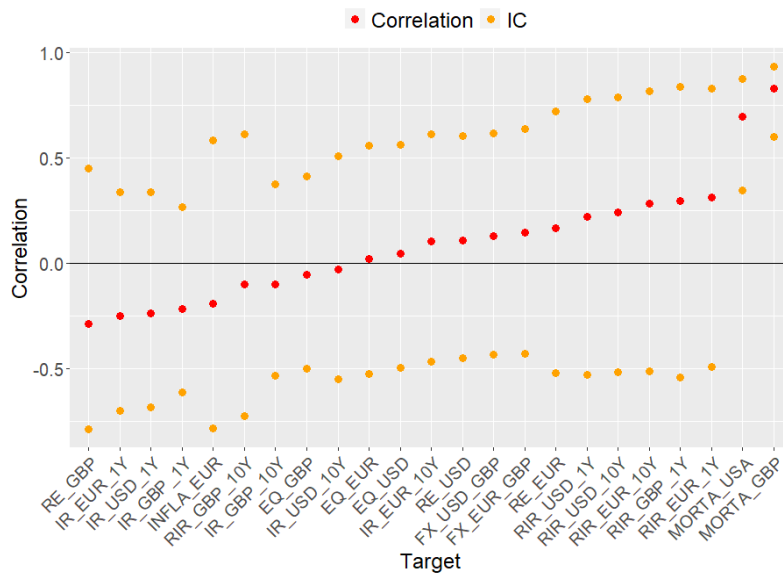


FIGURE 2.1 – Corrélation des taux de mortalité français avec les variables financières

En revanche, les taux de mortalité des différents pays sont corrélés entre eux. La corrélation entre les taux de mortalité français et américains est 0,69 et l'intervalle de confiance est  $[0,45; 0,84]$ .

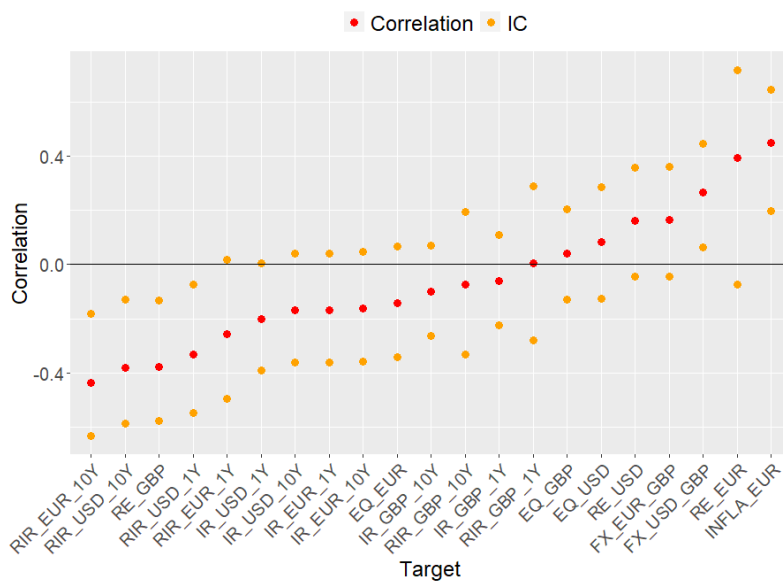


FIGURE 2.2 – Corrélation des taux de mortalité mensuels français avec les variables financières

L'INSEE (2023) fournit les taux de mortalité mensuels français. Les corrélations sur ces taux de mortalité sont calculées sur près de 200 points. La largeur des intervalles de

confiance est réduite. Calculer les corrélations à un pas de temps mensuel permet de mieux capter les variations conjointes si elles existent. Les corrélations des taux de mortalité mensuels français avec les variables financières sont présentées en figure 2.2. La corrélation avec l'inflation est de 0,45 avec un intervalle de confiance de largeur 0,44 alors qu'en considérant les taux de mortalité annuels la largeur de l'intervalle de confiance est 1,36.

Les corrélations des taux de mortalité mensuels avec les spreads de crédit sont visualisées en figure 2.3 en distinguant les notations (*ratings*). Les corrélations significatives sont des spreads de maturité 10 ans pour les notations A.

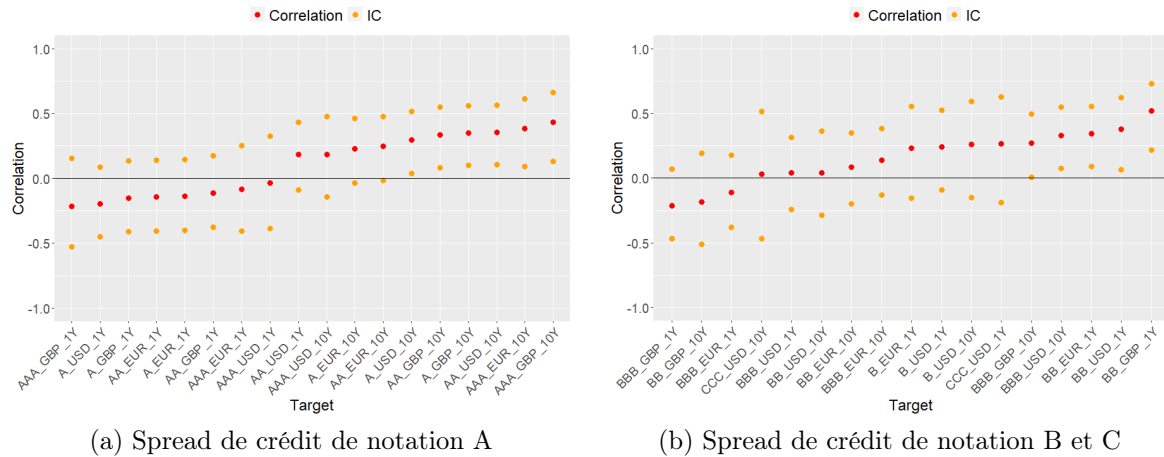


FIGURE 2.3 – Corrélation des taux de mortalité mensuels français avec les spreads de crédit

La longueur de l'historique considéré dans le calcul des corrélations a aussi un impact. Avec la pandémie, 2020 est un point atypique dans l'historique des taux de mortalité (chapitre 4). Les corrélations incluant ou pas l'année 2020 sont calculées et présentées en figure 2.4. La corrélation avec l'inflation augmente de 0,12 en incluant l'année 2020.

La corrélation entre les taux de mortalité et les variables financières n'est pas intuitive. Les graphiques montrent qu'il existe des corrélations non nulles mais l'interprétation de ces corrélations est plus complexe. Il faut distinguer l'existence d'une corrélation entre deux variables et la présence d'un lien de causalité entre celles-ci. Dans le cadre de ce mémoire, nous nous concentrons sur deux types de variables dans un souci de parcimonie : l'inflation et les spreads de crédit. Leurs corrélations avec les taux de mortalité mensuels sont significatives et l'avis d'expert appuie ce choix. Ainsi, nous supposons ici que les corrélations avec les autres variables financières sont nulles.

La sélection des facteurs de risque est basée sur une recherche bibliographique. D'une part, différentes études s'intéressent à la corrélation entre le risque de mortalité et le risque de crédit. Une étude de Buccioli et Zarri (2015) montre la corrélation entre l'aversion au risque et les événements négatifs de la vie. Les événements négatifs du vécu des individus ont une influence sur la prise de risque financier. En plus des événements collectivement vécus comme les crises financières ou les guerres, les événements spécifiques à un individu

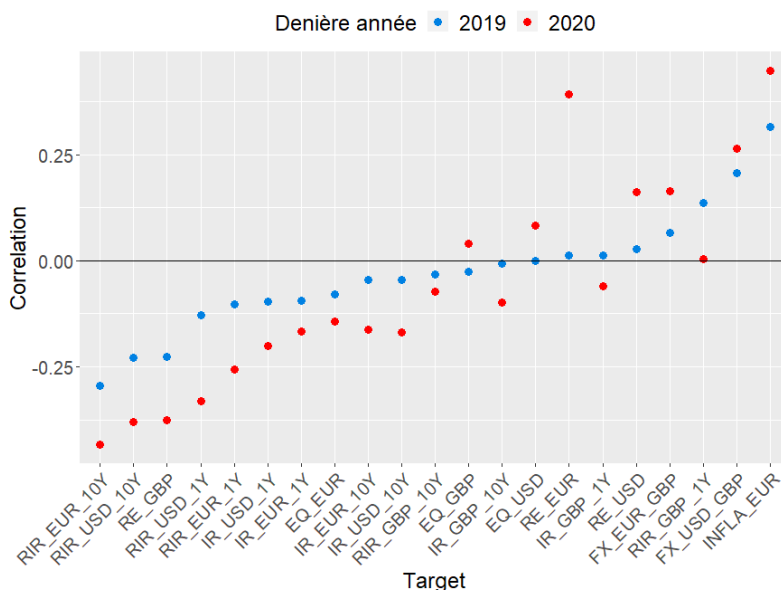


FIGURE 2.4 – Comparaison des corrélations des taux de mortalité mensuels français avec les variables financières dont les données s’arrêtent en 2019 ou 2020

ont des conséquences sur la volonté des individus à investir. Il ressort que cette volonté diminue significativement après avoir traversé des expériences négatives. Deux événements ressortent : les victimes d’une agression physique et la perte d’un enfant. De plus, la corrélation est durable dans le temps dans le cas d’un décès traumatisant. Il a été montré que l’effet de la perte d’un enfant sur la probabilité de détenir des actions est similaire à l’effet de la crise de 2008. Cette étude appuie l’existence d’une corrélation entre la mortalité et les variables financières relatives à la confiance accordée aux marchés.

Les spreads de crédit représentent la possibilité qu’une compagnie puisse faire défaut, il s’agit d’une comparaison entre le rendement d’une obligation d’une maturité fixée émise par une compagnie d’une notation donnée et une obligation équivalente sans risque. Ces obligations ne sont pas destinées aux particuliers. En revanche, les spreads de crédit représentent la confiance accordée aux compagnies qui émettent ces produits financiers.

D’autre part, l’inflation est la seconde variable d’intérêt. Un article publié dans la revue *Harvard Business Review*, Duffy (2022), montre que l’augmentation de l’inflation aux Etats-Unis a causé une dégradation de la santé. En effet, dans un contexte de crise économique, le coût des nécessités quotidiennes augmente, ne laissant plus autant d’argent à allouer aux frais de santé. On peut émettre l’hypothèse que l’inflation et la mortalité sont corrélées.

De plus, des études montrent le lien entre mortalité pandémique et variables financières qui appuient la sélection des facteurs de risque financiers pour l’étude des corrélations. Premièrement, l’article de Byström (2021) montre la relation entre le risque de crédit et la pandémie. En effet, le risque de crédit augmente avec la propagation de la COVID-19. En-



suite, l'endettement des entreprises a été largement accentué par la pandémie de COVID-19 et l'article de Chițu et al. (2022) montre l'impact de la pandémie sur les côtes de crédit. Enfin, l'article de Bonam et Smădu (2021) traite des liens entre pandémie et inflation en montrant les effets à long terme des pandémies sur l'inflation en Europe. Selon l'étude, les pandémies entraînent une baisse significative de l'inflation tendancielle qui tombe en dessous de son niveau initial pendant plus de dix ans. Ainsi, ces études montrent la relation entre risque de mortalité pandémique et les risques de crédit et d'inflation. Dans la suite du mémoire, ces deux facteurs de risque sont retenus pour être corrélés à la mortalité.

Les corrélations entre la mortalité et l'inflation d'une part et entre la mortalité et les spreads de crédit d'autre part sont sélectionnées sur la base de différentes recherches bibliographiques et quantitatives. Les autres corrélations sont supposées nulles dans un souci de simplification. Ce mémoire propose une méthodologie pour intégrer les corrélations précitées dans un GSE. Il est possible d'utiliser cette méthode afin de corréler la mortalité avec d'autres facteurs de risque.

### 2.2.2 Analyse des corrélations empiriques américaines

Les corrélations empiriques entre les variables financières et la mortalité américaine sont également calculées. Les paramètres calibrés des modèles financiers ne sont disponibles que pour les États-Unis. C'est pourquoi les prochaines étapes de la méthode seront donc appliquées aux États-Unis.

Les corrélations empiriques sont calculées pour les taux de mortalité annuels américains (figure 2.5). Les mêmes problématiques relevées pour la France se posent pour les États-Unis. L'historique est peu profond. Les intervalles de confiance ne permettent donc pas de vérifier la significativité des corrélations.

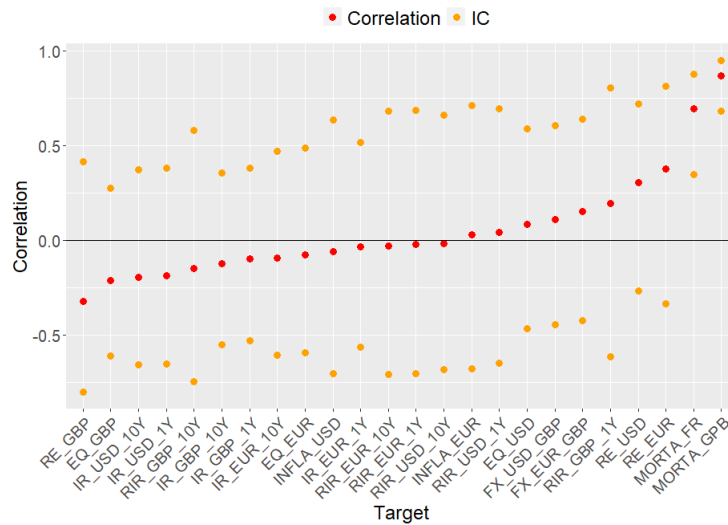


FIGURE 2.5 – Corrélacion des taux de mortalité américains avec les variables financières

L’inflation et le crédit sont sélectionnés afin de pouvoir corréliser ces facteurs de risque à la mortalité. Il serait intéressant d’approfondir les études sur la corrélation de ces facteurs de risque car les données de mortalité annuelles ne sont pas adaptées à une étude de corrélation. De plus, plusieurs pistes pourraient être explorées :

- L’inflation et la mortalité pourraient être corrélées avec un décalage de temps. Par exemple, la mortalité a fortement augmenté en 2020 alors que la hausse de l’inflation est importante à partir de 2021.
- La mortalité pour des classes d’âges spécifiques.

L’étude des corrélations présente des limites. La méthodologie d’implémentation de la mortalité dans un GSE est générique. Ainsi, elle peut être appliquée en intégrant de nouvelles études sur les corrélations pour plus de cohérence dans les simulations.

L’inflation et le crédit sont les facteurs de risque sélectionnés pour être corrélés à la mortalité américaine. L’indice inflation est modélisé pour le modèle de Black-Scholes et les spreads de crédit sont modélisés par le modèle PCA-Vasicek. Ce sont des modèles simples, répandus et facilement interprétables. De plus, ils sont suffisamment tractables pour obtenir des formules fermées pour les corrélations qui sont présentées dans les sections suivantes. Ce mémoire n’étant pas consacré à la modélisation de l’inflation et du crédit en monde-réel, il est naturel de considérer ces modèles.

## 2.3 Calculs théoriques des corrélations

L'objectif est de déterminer la relation liant :

- la corrélation entre le **facteur de risque de mortalité** et les **facteurs de risque financiers** retenus
- les corrélations entre les **aléas** des modèles de mortalité et des modèles financiers sous-jacents.

Par souci de simplification, nous parlerons de corrélation entre deux modèles. Il s'agit d'un abus de langage.

Les cinq étapes de la méthodologie sont :

1. Le calcul des corrélations empiriques : les corrélations entre les facteurs de risque observables sont calculées avec la corrélation de Pearson.
2. La définition du cadre de modélisation : les quantités modélisées sont définies ainsi que les modèles appliqués.
3. La relation entre les quantités observables et les quantités modélisées : cette étape n'est pas nécessaire lorsque les quantités modélisées sont observables.
4. La détermination des quantités homogènes à des incréments de mouvements Browniens : en utilisant la définition des modèles, l'objectif est d'obtenir une expression des facteurs de risque modélisés homogènes à des Gaussiennes.
5. Le calcul des corrélations théoriques : les corrélations empiriques entre deux facteurs de risque déterminées à l'étape 1 sont exprimées en fonction des corrélations entre les aléas des modèles sous-jacents à l'aide des expressions déterminées à l'étape 4.

Les corrélations théoriques sont calculées pour différents modèles financiers.

### 2.3.1 Corrélation entre le modèle de Lee-Carter et le modèle de Black-Scholes

La corrélation théorique entre le modèle de Lee-Carter et le modèle de Black-Scholes est déterminée dans ce paragraphe. Le modèle de Black-Scholes à volatilité constante est largement utilisé pour modéliser l'indice action. Il est également adapté à la modélisation de l'inflation. Il existe plusieurs modèles financiers dérivés du modèle de Black-Scholes.

On considère le modèle de Black-Scholes avec prime de risque en univers monde-réel :

$$dS_t = S_t(\mu dt + \theta dB_t) \quad (2.2)$$

avec :

- $S_t$  : indice action modélisé
- $\mu \in \mathbb{R}$  : prime de risque des actions par rapport au taux sans risque
- $\theta \in \mathbb{R}^+$  : volatilité constante de l'indice
- $B_t$  : mouvement Brownien, définition en annexe A

Les notations du modèle de Lee-Carter défini en section 1.3.1 sont conservées. On considère un pas de temps  $\Delta$ .

La corrélation historique entre le facteur de risque de mortalité modélisé avec le modèle de Lee-Carter et le facteur de risque action modélisé par Black-Scholes est égale à la corrélation entre les aléas des deux modèles :

$$\text{Corr}\left(\eta_{x,t} - \eta_{x,t-\Delta}, \log\left(\frac{S_t}{S_{t-\Delta}}\right)\right) = \text{corr}(\varepsilon_t, \varepsilon_t^1) \quad (2.3)$$

avec  $\varepsilon_t$  l'aléa du modèle de Lee-Carter et  $\varepsilon_t^1$  l'aléa du modèle de Black-Scholes, des variables Gaussiennes centrées réduites.

**Preuve**

Appliquer le lemme d'Itô, défini en annexe A, à l'équation (2.2) permet d'obtenir que :

$$d \log(S_t) = \left(\mu - \frac{1}{2}\theta^2\right)dt + \theta dB_t \quad (2.4)$$

On retrouve l'expression de l'incrément de l'indice action entre deux dates séparées du temps  $\Delta$  à partir de l'équation (2.4) :

$$\begin{aligned} \log\left(\frac{S_t}{S_{t-\Delta}}\right) &= \left(\mu - \frac{1}{2}\theta^2\right)\Delta + \theta(B_t - B_{t-\Delta}) \\ &= \left(\mu - \frac{1}{2}\theta^2\right)\Delta + \theta\sqrt{\Delta}\varepsilon_t^1 \end{aligned} \quad (2.5)$$

avec  $\varepsilon_t^1$  une variable aléatoire Gaussienne centrée réduite associée au mouvement Brownien B.

On réécrit le modèle de Lee-Carter :

$$\eta_{x,t} = \alpha_x + \beta_x \kappa_t \quad (2.6)$$

avec  $\forall t, \kappa_{t+1} = \kappa_t + \delta + \xi_t, \xi_t \sim \mathcal{N}(0, \sigma^2)$  indépendantes.

En développant l'expression du modèle (2.6), le logarithme des taux de mortalité s'exprime tel que :

$$\eta_{x,t} = \alpha_x + \beta_x \left( \kappa_0 + t\delta + \sum_{j=0}^{t-1} \xi_{t-j} \right) \quad (2.7)$$

On considère  $\Delta$  un intervalle de temps entre deux dates, fixé à un an. Le log-incrément des taux de mortalité sur  $\Delta$  est donc déduit de l'équation (2.7) :

$$\eta_{x,t} - \eta_{x,t-\Delta} = \beta_x \left( \Delta\delta + \sqrt{\Delta}\sigma\varepsilon_t \right) \text{ avec } \varepsilon_t \sim \mathcal{N}(0, 1) \text{ iid} \quad (2.8)$$

La corrélation entre le modèle de Lee-Carter et le modèle de Black-Scholes s'exprime telle que :

$$\begin{aligned} \text{Corr}\left(\eta_{x,t} - \eta_{x,t-\Delta}, \log\left(\frac{S_t}{S_{t-\Delta}}\right)\right) &= \frac{\text{Cov}\left(\eta_{x,t} - \eta_{x,t-\Delta}, \log\left(\frac{S_t}{S_{t-\Delta}}\right)\right)}{\sqrt{\text{Var}\left(\eta_{x,t} - \eta_{x,t-\Delta}\right)}\sqrt{\text{Var}\left(\log\left(\frac{S_t}{S_{t-\Delta}}\right)\right)}} \\ &= \frac{\Delta\beta_x\sigma\theta\text{Corr}(\varepsilon_t, \varepsilon_t^1)}{\sqrt{\left(\sqrt{\Delta}\beta_x\sigma\right)^2}\sqrt{\left(\sqrt{\Delta}\theta\right)^2}} \\ &= \text{corr}(\varepsilon_t, \varepsilon_t^1) \end{aligned}$$

### 2.3.2 Corrélation entre le modèle de Lee-Carter et le modèle de PCA-Vasicek

Le modèle de Vasicek avec analyse en composantes principales (PCA-Vasicek) est un modèle de structure par terme utilisé pour modéliser les taux d'intérêt nominaux, les taux d'intérêt réels et les spreads de crédit. Ce modèle permet de simuler la structure par terme avec une extraction des premières composantes principales et en les projetant grâce à des processus AR(1) homogènes à la dynamique de Vasicek.

On modélise  $Y(t+\Delta, t+m)$  correspondant soit à la différence des taux entre les instants  $t$  et  $t + \Delta$ , soit au log-ratio shifté des taux aux instants  $t$  et  $t + \Delta$  avec  $\Delta$  le pas de temps et  $m$  la maturité,  $M$  la maturité maximale. Pour tout  $m \in \llbracket 1, M \rrbracket$ , les estimateurs de la moyenne et de l'écart-type sont :

$$\mu(m) = \frac{1}{T} \sum_{k=1}^T Y(k, k+m) \quad \text{et} \quad \sigma(m) = \sqrt{\frac{1}{T-1} \sum_{k=1}^T (Y(k, k+m) - \mu(m))^2}$$

On définit la standardisation  $X(k, k+m) := \frac{Y(k, k+m) - \mu(m)}{\sigma(m)}$  et  $X$  est la matrice associée de dimensions  $T \times M$  tel que  $X_{k,m} = X(k, k+m)$ . Pour tout  $k \in \llbracket 1, 2, 3 \rrbracket$ . La matrice de covariance empirique est définie par :

$$\mathbf{S}_{m,n} := \frac{1}{T} \sum_{k=1}^T X_{k,m} X_{k,n}, \quad (m, n) \in \{1, \dots, M\}^2$$

$\mathbf{S}$  est semi-définie positive.  $P$  une matrice orthogonale et  $D$  une matrice diagonale telles que  $P'\mathbf{S}P = D$ . Pour tout  $m \in \llbracket 1, M \rrbracket$ ,  $P_{(n)}$  la  $n^{\text{ième}}$  colonne de la matrice  $P$ . Ainsi, pour tout  $k \in \{1, \dots, T\}$  :

$$X_k = \sum_{n=1}^M u(k, n) P_{(n)}$$

Pour  $n \in \llbracket 1, 2, 3 \rrbracket$ , correspondant aux trois premières composantes principales, la série temporelle  $u(k, n) \in \mathbb{R}^M$  est modélisée avec un modèle AR(1). On suppose que pour tout

$n \in \{1, 2, 3\}$ , il existe  $(l_n, \alpha_n, \nu_n) \in \mathbb{R}^{+3}$ , pour tout  $k \in \llbracket 2, T \rrbracket$ ,  $(\varepsilon_{k,n})_{(k,n) \in \llbracket 1, T \rrbracket \times \{1, 2, 3\}}$  :

$$u(k, n) - u(k-1, n) = -\alpha_n \Delta(u(k-1, n) - l_n) + \nu_n \sqrt{\Delta} \varepsilon_{k,n} \quad (2.9)$$

Pour  $m \in \{1, \dots, M\}$ , la variance de  $(Y_{k,k+m})_{k \in \mathbb{N}^*}$  est donné pour tout  $k \in \mathbb{N}^*$  :

$$V_k^m = \text{Var}[Y_{k,k+m}] = \sigma_m^2 \sum_{n=1}^3 P_n(m)^2 \frac{v_n^2 \Delta (1 - (1 - \alpha_n \Delta)^{2k})}{1 - (1 - \alpha_n \Delta)^2} \quad (2.10)$$

En reprenant les notations associées au modèle de Lee-Carter, et en fixant les notations pour le modèle de PCA-Vasicek :

- Lee-Carter :  $Y_{x,k}^1 = Y^1(x, k) = \eta_{x,k\Delta} - \eta_{x,(k-1)\Delta} = \beta_x(\Delta\xi + \sqrt{\Delta}\sigma\varepsilon_1(t))$
- PCA-Vasicek :  $Y_{k,m} = Y(k\Delta, k\Delta + m)$

$$\begin{aligned} \text{Cov}\left(Y_{x,k}^1, Y_{k,m}\right) &= \text{Cov}\left(\beta_x(\Delta\xi + \sqrt{\Delta}\sigma\varepsilon_1(k)), \sigma(m) \sum_{n=1}^3 u(k\Delta, n) P_n(m) + \mu(m)\right) \\ &= \text{Cov}\left(\beta_x \sqrt{\Delta} \sigma \varepsilon_1(k), \sigma(m) \sum_{n=1}^3 u(k\Delta, n) P_n(m)\right) \\ &= \beta_x \sigma \sqrt{\Delta} \sigma(m) \sum_{n=1}^3 P_n(m) \text{Cov}(\varepsilon_1(k), u(k\Delta, n)) \end{aligned} \quad (2.11)$$

De plus, en utilisant l'équation (2.9) pour tout  $n = 1, 2, 3$  :

$$\begin{aligned} \text{Cov}(u(k\Delta, n), \varepsilon_1(k)) &= v_n \sqrt{\Delta} \sum_{l=0}^{k-1} (1 - \alpha_n \Delta)^l \text{Cov}(\varepsilon_1(k), \varepsilon_2(k-l, n)) \\ &= v_n \sqrt{\Delta} \text{Cov}(\varepsilon_1(k), \varepsilon_2(k, n)) \\ &= v_n \sqrt{\Delta} \rho_n \text{ avec } \rho_n = \text{Cov}(\varepsilon_1(k), \varepsilon_2(k, n)) \end{aligned}$$

Donc,

$$\text{Cov}\left(Y_{x,k}^1, Y_{k,m}\right) = \beta_x \sigma \sigma(m) \sqrt{\Delta} \sqrt{\Delta} \sum_{n=1}^3 v_n P_n(m) \rho_n \quad (2.12)$$

De l'expression de la covariance (2.12) et de la variance (2.10), on en déduit la corrélation :

$$\text{Corr}(Y_{x,k}^1, Y_{k,m}) = \frac{\sum_{n=1}^3 P_n(m) v_n \rho_n}{\sqrt{\sum_{n=1}^3 \frac{P_n(m)^2 v_n^2 (1 - (1 - \alpha_n \Delta)^{2k})}{1 - (1 - \alpha_n \Delta)^2}}} \quad (2.13)$$

Lors du calibrage du modèle PCA-Vasicek, la contrainte  $\alpha_n \Delta < 1$  est imposée permettant d'obtenir le résultat asymptotique de la relation :

$$\text{Corr}(Y^1(x, t), Y(t, t + m)) \xrightarrow{t \rightarrow +\infty} \frac{\sum_{n=1}^3 P_n(m) v_n \rho_n}{\sqrt{\sum_{n=1}^3 \frac{P_n(m)^2 v_n^2}{1 - (1 - \alpha_n \Delta)^2}}} \quad (2.14)$$

$$\rho_n = \text{Cov}(\varepsilon_1(k), \varepsilon_2(k, n))$$

## 2.4 Calcul des corrélations entre mouvements browniens

Nous nous intéressons au calcul des corrélations entre aléas à partir des formules théoriques définies dans la section précédente. Les facteurs de risque étudiés sont l'inflation et le risque de crédit. L'objectif est de pouvoir déterminer ces corrélations, d'évaluer leur pertinence et leur sensibilité à différentes variables. Les calculs des corrélations sont réalisés sur les données des Etats-Unis.

### 2.4.1 Modélisation de l'inflation

L'indice inflation est modélisé par un modèle de Black-Scholes à volatilité constante. Les paramètres du modèle sont calibrés sur l'historique de l'indice d'inflation totale  $I_t^{CPI}$  dont le pas de temps est mensuel :

$$\widehat{\mu}^{CPI} = \frac{1}{n} \sum_{i \in [1, n]} \log \left( \frac{I_{t_i}^{CPI}}{I_{t_{i-1}}^{CPI}} \right) 12 + \frac{\widehat{\sigma}^{CPI}^2}{2}$$

$$\widehat{\sigma}^{CPI}^2 = 12 \text{Var} \left( \left\{ \log \left( \frac{I_{t_i}^{CPI}}{I_{t_{i-1}}^{CPI}} \right) \right\}_{i \in [1, n]} \right)$$

La corrélation entre l'aléa de ce modèle et l'aléa du modèle de Lee-Carter, équation (2.3), est égale à la corrélation historique entre les deux facteurs de risque observés.

### 2.4.2 Modélisation du risque de crédit

Les spreads de crédit sont modélisés avec le modèle PCA-Vasicek dont la corrélation avec le modèle de Lee-Carter est définie dans la section 2.3.2. Le modèle PCA-Vasicek est calibré sur les spreads de crédit américains par notation. De plus, la corrélation entre le modèle de Lee-Carter et le modèle PCA-Vasicek (équation (2.14)) dépend de la maturité de l'obligation associée. Or, nous cherchons à déterminer la corrélation entre les trois aléas

du modèle PCA-Vasicek et l'aléa du modèle de Lee-Carter, trois corrélations sont donc à calculer.

**Choix des maturités** : la combinaison des trois corrélations entre aléas devrait vérifier (2.14) pour chaque maturité. Trois maturités sont sélectionnées pour résoudre un système de 3 équations à 3 inconnues. Les maturités sélectionnées sont 1 an, 5 ans et 10 ans.

Les taux de mortalité sont annuels tandis que les spreads de crédit sont mensuels. Les composantes principales du modèle PCA-Vasicek étant modélisées par des AR(1), les paramètres calibrés dépendent de la fréquence d'observation des données utilisées, ce qui n'est pas le cas du modèle de Black-Scholes. Le modèle PCA-Vasicek est calibré sur des données mensuelles dont les paramètres de calibrage interviennent dans le calcul de la corrélation. On cherche à exprimer la corrélation théorique entre les variations annuelles de l'ACP (analyse en composantes principales) et les variations annuelles des taux de mortalité en fonction des corrélations entre les variations mensuelles de l'ACP et les variations annuelles des taux de mortalité. Le modèle PCA-Vasicek modélise les log-ratios shiftés, le *shift* est fixé. En pratique, le pas de temps des taux de mortalité est annuel et celui des spreads de crédit est mensuel, les valeurs sont directement introduites dans le calcul. Ainsi, la corrélation est définie avec  $S_{t,m}$  le spread de crédit pour la notation considérée de maturité  $m$  à l'instant  $t$  :

$$\text{Corr}\left(\eta_{x,t+1} - \eta_{x,t}, \log\left(\frac{S_{t+1,m} + \text{shift}}{S_{t,m} + \text{shift}}\right)\right) = \frac{\sum_{i=1}^{12} \beta_x \sigma \sigma(m) \frac{1}{12} \sum_{n=1}^3 \nu_n P_n(m) \rho_n}{\sqrt{\beta_x \sigma} \sqrt{V_{PCA}}} \quad (2.15)$$

$$V_{PCA} = \sum_{i=0}^{11} \sigma(m)^2 \sum_{n=1}^3 P_n(m)^2 \frac{\nu_n^2 \frac{1}{12} \left(1 - \left(1 - \alpha_n \frac{1}{12}\right)^{2i}\right)}{1 - \left(1 - \alpha_n \frac{1}{12}\right)^2} + \sum_{\substack{i,j=0 \\ i \neq j}}^{11} \sigma(m)^2 \sum_{n=1}^3 P_n(m)^2 \left( \left( \sum_{l_1=0}^{j-1} \left(1 - \alpha_n \frac{1}{12}\right)^{i-j+2l_1} \right) \mathbb{1}_{i>j} + \left( \sum_{l_2=0}^{i-1} \left(1 - \alpha_n \frac{1}{12}\right)^{j-i+2l_2} \right) \mathbb{1}_{j>i} \right) \quad (2.16)$$

### Preuve

Les log-ratios annuels sont exprimés en fonction des log-ratios mensuels dans l'équation (2.17) :

$$Y(t+1, t+m) = \log\left(\frac{S_{t+1,m} + \text{shift}}{S_{t,m} + \text{shift}}\right) = \sum_{i=1}^{12} \log\left(\frac{S_{t+\frac{i}{12},m} + \text{shift}}{S_{t+\frac{i-1}{12},m} + \text{shift}}\right) \quad (2.17)$$

Avec la formule de covariance (2.12) et l'expression précédente (2.17), la covariance est reformulée :

$$\text{Cov}\left(\eta_{x,t+1} - \eta_{x,t}, \log\left(\frac{S_{t+1,m} + \text{shift}}{S_{t,m} + \text{shift}}\right)\right) = \sum_{i=1}^{12} \text{Cov}\left(\eta_{x,t+1} - \eta_{x,t}, \log\left(\frac{S_{t+\frac{i}{12},m} + \text{shift}}{S_{t+\frac{i-1}{12},m} + \text{shift}}\right)\right)$$



$$= \sum_{i=1}^{12} \beta_x \sigma \sigma(m) \frac{1}{12} \sum_{n=1}^3 \nu_n P_n(m) \rho_n \quad (2.18)$$

La variance est déterminée en utilisant la même approche :

$$\begin{aligned} & \text{Var} \left( \log \left( \frac{S_{t+1,m} + \text{shift}}{S_{t,m} + \text{shift}} \right) \right) \\ &= \sum_{i=1}^{12} \sum_{j=1}^{12} \text{Cov} \left( \log \left( \frac{S_{t+\frac{i}{12},m} + \text{shift}}{S_{t+\frac{i-1}{12},m} + \text{shift}} \right), \log \left( \frac{S_{t+\frac{j}{12},m} + \text{shift}}{S_{t+\frac{j-1}{12},m} + \text{shift}} \right) \right) \\ &= \sum_{i=1}^{12} \text{Var} \left( \log \left( \frac{S_{t+\frac{i}{12},m} + \text{shift}}{S_{t+\frac{i-1}{12},m} + \text{shift}} \right) \right) + \\ & \quad \sum_{\substack{i,j=0 \\ i \neq j}}^{11} \text{Cov} \left( \log \left( \frac{S_{t+\frac{i+1}{12},m} + \text{shift}}{S_{t+\frac{i}{12},m} + \text{shift}} \right), \log \left( \frac{S_{t+\frac{j+1}{12},m} + \text{shift}}{S_{t+\frac{j}{12},m} + \text{shift}} \right) \right) \\ &= \sum_{i=0}^{11} \sigma(m)^2 \sum_{n=1}^3 P_n(m)^2 \frac{v_n^2 \frac{1}{12} \left( 1 - \left( 1 - \alpha_n \frac{1}{12} \right)^{2i} \right)}{1 - \left( 1 - \alpha_n \frac{1}{12} \right)^2} + \\ & \quad \sum_{\substack{i,j=0 \\ i \neq j}}^{11} \text{Cov} \left( \sigma(m) \sum_{n=1}^3 u \left( \frac{j}{12}, n \right) P_n(m) + \mu(m), \sigma(m) \sum_{p=1}^3 u \left( \frac{i}{12}, n \right) P_n(m) + \mu(m) \right) \end{aligned}$$

On suppose  $t = 0$  car le modèle est stationnaire.

$$\text{Cov} \left( u \left( \frac{j}{12}, n \right), u \left( \frac{i}{12}, n \right) \right) = 0 \quad \text{si } n \neq p$$

$$\begin{aligned} &= \sum_{i=0}^{11} \sigma(m)^2 \sum_{n=1}^3 P_n(m)^2 \frac{v_n^2 \frac{1}{12} \left( 1 - \left( 1 - \alpha_n \frac{1}{12} \right)^{2i} \right)}{1 - \left( 1 - \alpha_n \frac{1}{12} \right)^2} + \\ & \quad \sum_{\substack{i,j=0 \\ i \neq j}}^{11} \sigma(m)^2 \sum_{n=1}^3 P_n(m)^2 \text{Cov} \left( u \left( \frac{j}{12}, n \right), u \left( \frac{i}{12}, n \right) \right) \\ &= \sum_{i=0}^{11} \sigma(m)^2 \sum_{n=1}^3 P_n(m)^2 \frac{v_n^2 \frac{1}{12} \left( 1 - \left( 1 - \alpha_n \frac{1}{12} \right)^{2i} \right)}{1 - \left( 1 - \alpha_n \frac{1}{12} \right)^2} + \\ & \quad \sum_{\substack{i,j=0 \\ i \neq j}}^{11} \sigma(m)^2 \sum_{n=1}^3 P_n(m)^2 \sum_{l_1=0}^{j-1} \sum_{l_2=0}^{i-1} v_n^2 \left( 1 - \alpha_n \frac{1}{12} \right)^{l_1} \left( 1 - \alpha_n \frac{1}{12} \right)^{l_2} \text{Cov}(\varepsilon(j - l_1, n), \varepsilon(i - l_2, n)) \\ &= \sum_{i=0}^{11} \sigma(m)^2 \sum_{n=1}^3 P_n(m)^2 \frac{v_n^2 \frac{1}{12} \left( 1 - \left( 1 - \alpha_n \frac{1}{12} \right)^{2i} \right)}{1 - \left( 1 - \alpha_n \frac{1}{12} \right)^2} + \\ & \quad \sum_{\substack{i,j=0 \\ i \neq j}}^{11} \sigma(m)^2 \sum_{n=1}^3 P_n(m)^2 \left( \left( \sum_{l_1=0}^{j-1} \left( 1 - \alpha_n \frac{1}{12} \right)^{i-j+2l_1} \right) \mathbb{1}_{i>j} + \left( \sum_{l_2=0}^{i-1} \left( 1 - \alpha_n \frac{1}{12} \right)^{j-i+2l_2} \right) \mathbb{1}_{j>i} \right) \\ &= V_{PCA} \quad (2.19) \end{aligned}$$

Les expressions (2.18) et (2.19) donnent la formulation de la corrélation (2.15).

### 2.4.3 Résultats des corrélations

La méthode pour déterminer les corrélations entre les aléas des modèles est présentée dans les sections précédentes et est appliquée sur les données des Etats-Unis. Les corrélations empiriques entre les facteurs de risque observables sont déterminées. Elles permettent d'obtenir les corrélations entre les mouvements Browniens des modèles à l'aide des formules théoriques. La matrice de corrélation obtenue par ce processus est rendue semi-définie positive (PSD) par l'algorithme d'Higham [Higham (2002)]. Cet algorithme a été développé pour déterminer la matrice de corrélation semi-définie positive la plus proche d'une matrice de corrélation donnée. Le principe repose sur l'algorithme de Newton en minimisant la distance entre la matrice de corrélation initiale et la matrice semi-définie positive sous une norme spécifique.

Les résultats des corrélations entre les aléas du modèle de Lee-Carter et les aléas des modèles pour l'indice inflation et les spreads de crédit sont présentés dans le tableau 2.2 et sont calculés à partir des corrélations empiriques dans le tableau 2.1.

Facteur de risque observable	Corrélation empirique
Indice inflation	0,35
Spread de crédit AA maturité 1 an	0,80
Spread de crédit AA maturité 5 an	0,64
Spread de crédit AA maturité 10 an	0,52

TABLE 2.1 – Corrélations empiriques des facteurs de risque observables avec le risque de mortalité

Aléa des modèles	Corrélation entre aléas
Black-Scholes - indice inflation	0,35
PCA-Vasicek Facteur 1 - Spread de crédit	-0,38
PCA-Vasicek Facteur 2 - Spread de crédit	0,38
PCA-Vasicek Facteur 3 - Spread de crédit	-0,45

TABLE 2.2 – Corrélations des aléas des modèles avec l'aléa du modèle de Lee-Carter

L'application du retraitement en matrice semi-définie positive peut générer des variations par rapport à la matrice de corrélation initiale. Ainsi, les corrélations entre les aléas utilisées pour la simulation des modèles peuvent être impactées par ce processus. En effet, comme présenté dans la section 1.4.1, les données d'entrée du GSE peuvent varier selon les usages. Les taux de mortalité en entrée sont ceux de la population totale des Etats-Unis. L'application du retraitement aux résultats du tableau 2.2 n'entraîne pas de variation des corrélations. En revanche, si les taux de mortalité des hommes sont considérés, le traitement PSD entraîne des changements dans les corrélations entre aléas (tableau 2.3).

Aléas des modèles	Avant retraitement	Après retraitement
	PSD	PSD
Browniens de BS - indice inflation	0,40	0,31
PCA-Vasicek Facteur 1 - Spread de crédit	-0,52	-0,46
PCA-Vasicek Facteur 2 - Spread de crédit	0,81	0,65
PCA-Vasicek Facteur 3 - Spread de crédit	-0,78	-0,63

TABLE 2.3 – Corrélation des aléas des modèles avec l'aléa du modèle de Lee-Carter avant et après retraitement en matrice semi-définie positive

Cet exemple montre que le retraitement en matrice PSD peut entraîner des variations dans les corrélations entre aléas mises en entrée du GSE. Il est possible de fixer des hypothèses afin que les corrélations avant et après retraitement soient les plus proches possibles. Par exemple, dans le cadre du modèle de PCA-Vasicek, la corrélation avec les trois facteurs de risque peut être supposée égale. Il s'agit d'une piste d'étude pour améliorer la cohérence du processus de simulation.

#### Conclusion

Ce chapitre présente une méthode pour déterminer la corrélation entre le facteur de risque de mortalité et les facteurs de risque financiers pour les simuler conjointement dans un GSE.

Tout d'abord, les recherches bibliographiques et l'analyse quantitative des corrélations empiriques motivent la prise en compte des corrélations entre le facteur de risque de mortalité et certains facteurs de risque financiers dans le GSE. Dans ce mémoire, deux facteurs de risque sont corrélés à la mortalité : le risque de crédit et le risque d'inflation.

Ensuite, la relation liant la corrélation empirique entre deux facteurs de risque et la corrélation entre les aléas de leurs modèles sous-jacents respectifs est déterminée. Ces formulations théoriques permettent de déterminer les corrélations entre les aléas des différents modèles. L'optimisation numérique est utilisée pour calculer ces corrélations.

Enfin, les corrélations entre les aléas sont mises en entrée du GSE afin de pouvoir simuler conjointement les facteurs de risque.

Cette étude peut être étendue pour corréler la mortalité à d'autres facteurs de risque financiers. En effet, le processus a été réalisé pour deux risques sélectionnés mais des études de ces risques peuvent appuyer l'introduction de nouvelles corrélations. De plus, un travail sur le retraitement en matrice semi-définie positive de la matrice de corrélation entre aléas permettrait de conserver des corrélations plus justes pour les simulations.

## Chapitre 3

# Rapport automatique de calibrage

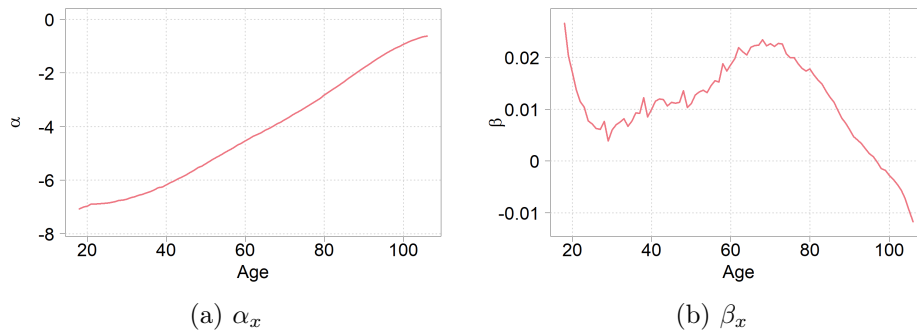
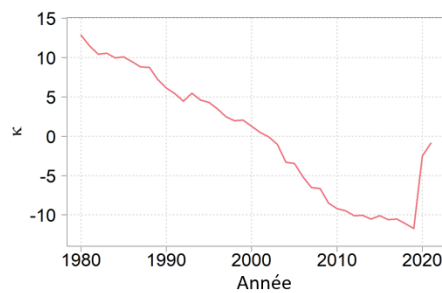
Les résultats de la modélisation de la mortalité sont présentés dans un rapport automatique généré en sortie du GSE. Chaque facteur de risque est présenté dans ce rapport automatique, ainsi l'ajout d'une partie dédiée au facteur de risque de mortalité est essentiel à son intégration dans le GSE. Les intérêts de ce rapport sont d'analyser les résultats du calibrage et de la projection de la mortalité. Il est possible de vérifier que le modèle est correctement calibré et de visualiser les résultats. Le rapport permet de comprendre ce qui a été modélisé et donne une vision de la performance du modèle. La présentation des sorties du GSE permet éventuellement de comparer les résultats à ceux d'autres modèles de mortalité. Le calibrage du modèle de mortalité est d'abord analysé puis les résultats de la simulation de la mortalité sont présentés.

### 3.1 Calibrage du modèle de mortalité

La première étape de la modélisation d'un facteur de risque dans un GSE en univers monde-réel est le calibrage sur les données historiques. Les données de mortalité sont l'exposition et le nombre de décès dépendant de l'âge et du temps.

L'économie sélectionnée pour présenter les résultats de ce chapitre est le dollar américain. C'est pourquoi le modèle de Lee-Carter est calibré sur les données de la population totale des États-Unis sur la période de 1980 à 2020.

Le modèle est calibré âge par âge à un pas de temps annuel. A cette étape, les paramètres sont déterminés en fonction de la structure du modèle défini dans la section 1.3.1, ils sont présentés dans le rapport. Avec le modèle de Lee-Carter, les paramètres  $\alpha_x$ ,  $\beta_x$  et  $\kappa_t$  permettent d'analyser la structure en âge de la mortalité, la sensibilité à l'âge par rapport à l'évolution dans le temps et l'évolution temporelle de la mortalité.

FIGURE 3.1 – Paramètres dépendants de l'âge  $\alpha_x$  et  $\beta_x$  du modèle de Lee-CarterFIGURE 3.2 – Paramètre  $\kappa_t$  du modèle de Lee-Carter

Les paramètres du modèle sont également présentés dans le rapport. Les variations de ces paramètres en fonction de changement dans les données sont des outils d'analyse de la modélisation de la mortalité.

Tendance	Volatilité
-0,34	1,68

TABLE 3.1 – Paramètres de la série temporelle  $\kappa_t$ 

L'intérêt est également de vérifier que le modèle intégré dans le GSE s'adapte correctement aux données et fournit des résultats satisfaisants. Pour cela, plusieurs outils sont utilisés :

- Les résidus : la carte de chaleur des résidus permet d'identifier les données pour lesquelles le modèle n'est pas adapté. Par exemple, sur la carte de chaleur 3.3, les résidus des années 2020 et 2021 sont supérieurs aux résidus des années antérieures.

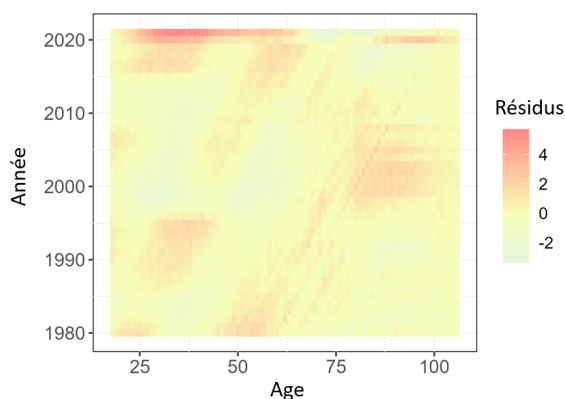


FIGURE 3.3 – Carte de chaleur des résidus du modèle de Lee-Carter sur les États-Unis

- Le coefficient  $R^2$  : ce coefficient, défini en annexe A, permet d’apprécier l’ajustement du modèle aux données. Il est calculé sur l’ensemble des âges et âge par âge. Pour les États-Unis, ce coefficient est de 99,5%. Ce résultat est intéressant lorsqu’il est comparé au  $R^2$  obtenu avec un autre modèle. Le graphique 3.4 permet de déduire les âges pour lesquels le modèle n’est pas satisfaisant et est cohérent avec les valeurs des résidus.



FIGURE 3.4 – Coefficient  $R^2$  du modèle de Lee-Carter en fonction de l’âge

- L’AIC et le BIC : ces deux indicateurs permettent de juger la performance du modèle, définis en annexe A. Ces critères peuvent être utilisés pour comparer ou sélectionner les modèles.

AIC	BIC
218391	219749

TABLE 3.2 – AIC et BIC du modèle

## 3.2 Simulation de la mortalité

Les scénarios de mortalité sont le résultat de la modélisation de la mortalité dans le GSE. La mortalité est projetée sur un horizon de 50 ans via 1000 simulations. Ces tables de simulations sont analysées afin de caractériser ces scénarios. La projection de la mortalité est visualisée à l'aide des percentiles. Cette visualisation donne une idée de l'allure des trajectoires. Les résultats pour l'âge 65 ans sont présentés en figure 3.5.

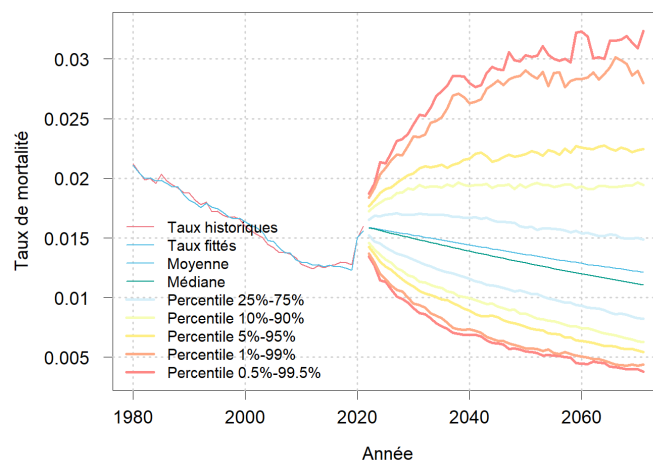


FIGURE 3.5 – Simulation des taux de mortalité à l'âge de 65 ans

Il est également intéressant de s'appuyer sur des indicateurs tels que l'amélioration de la mortalité dans le temps pour analyser le facteur de risque. L'amélioration de la mortalité  $i_{x,t}$  est définie telle que

$$i_{x,t} = \frac{\mu_{x,t-1} - \mu_{x,t}}{\mu_{x,t-1}}$$

L'amélioration de la mortalité est calculée sur les taux historiques et sur les taux fittés. Les figures 3.6 et 3.7 mettent en évidence les années au cours desquelles la mortalité a été impactée à la hausse ou à la baisse et la manière dont cette évolution est captée par le modèle. Les deux cartes de chaleur montrent que l'année 2020 présente des taux d'amélioration de la mortalité négatifs et inférieurs à  $-20\%$  pour tous les âges sur les données historiques (figure 3.6). Cette variation dans l'amélioration de la mortalité en 2020 est captée par le modèle de Lee-Carter pour les âges inférieurs à 80 ans. De plus, les cohortes sont observables sur la carte de chaleur de l'amélioration des taux de mortalité historiques sous la forme de diagonales. Ces cohortes ne sont pas prises en compte dans le modèle de Lee-Carter du fait de sa structure. Ces graphiques présentés dans le rapport automatique apportent des précisions à l'utilisateur sur les résultats fournis par le modèle calibré.

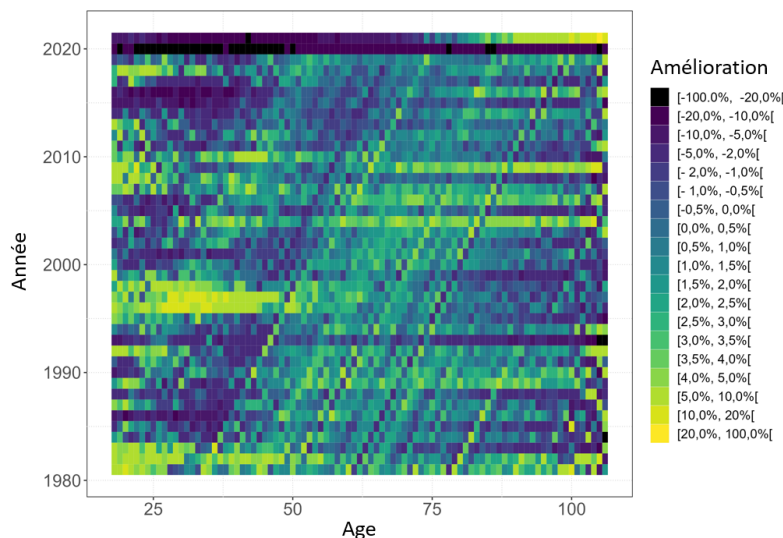


FIGURE 3.6 – Carte de chaleur de l'amélioration de la mortalité sur les taux historiques des États-Unis

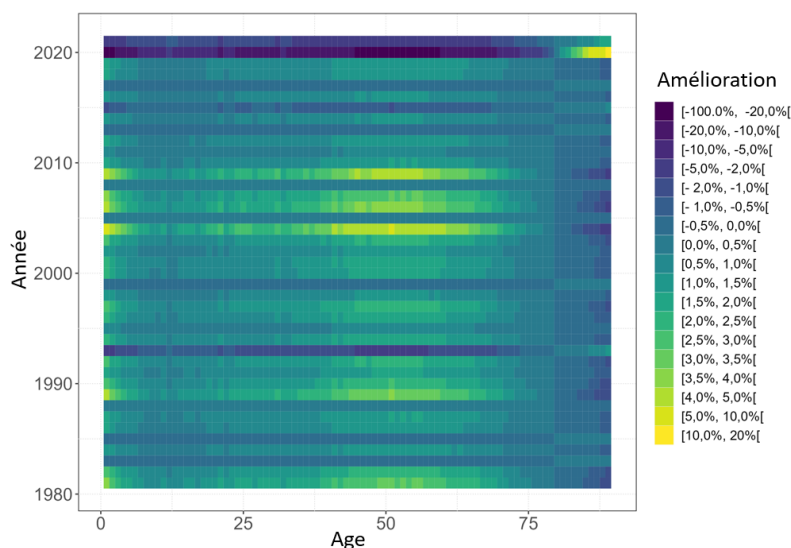


FIGURE 3.7 – Carte de chaleur de l'amélioration de la mortalité sur les taux modélisés avec le modèle de Lee-Carter calibré sur des données des États-Unis

### Conclusion

L'intégration de la mortalité dans un GSE permet d'obtenir des scénarios de mortalité stochastiques. Le rapport automatique est un outil de présentation des résultats afin de faciliter leur analyse et peut également servir à la comparaison de modèles de mortalité. Le rapport automatique de calibrage est nécessaire à la compréhension des résultats du GSE.



## Deuxième partie

# Modélisation de la mortalité en considérant différents facteurs de risque

## Chapitre 4

# Modélisation de la mortalité en tenant compte de l'expérience COVID-19

Ce chapitre traite de la modélisation de la mortalité en tenant compte de la pandémie de la COVID-19. L'objectif de cette étude est de mesurer l'impact de la pandémie sur la modélisation de la mortalité. De plus, différentes méthodes de retraitement des données de 2020 seront envisagées afin de comparer la surmortalité suggérée par ces méthodes et les données de décès dus à la COVID-19 communiquées par les organisations gouvernementales.

### 4.1 Motivations

La pandémie mondiale de la COVID-19 a débuté courant de l'année 2020 et a causé plus de 7 millions de morts dans le monde. Cette augmentation brutale de décès a des conséquences sur les taux de mortalité de l'année 2020. Les assureurs et réassureurs souhaitent prendre en compte l'expérience COVID-19 dans la modélisation de la mortalité pour plusieurs raisons.

Pour les assureurs et réassureurs utilisant un modèle interne, la mise à jour des modèles de mortalité et de longévité est un sujet d'actualité. L'objectif est d'intégrer dans les données de calibrage les années les plus récentes qui sont affectées par la COVID-19, ce qui impacte plusieurs aspects des modèles internes.

Dans le cadre des contrats d'assurance vie, les assureurs déterminent les taux de décès sur les tables d'expérience construites à partir de leur portefeuille. Ces tables permettent le calcul des chocs de mortalité utilisé pour le calcul du SCR. Certains assureurs emploient les tables de mortalité nationales afin de déterminer ces taux de mortalité qu'ils ajustent avec des facteurs d'abattement pour se ramener à leur portefeuille. Ainsi, lorsque les taux de mortalité nationaux sont impactés par un événement tel que la pandémie, les lois de *Best Estimate* et les chocs de mortalité sont impactés ainsi que le SCR de la compagnie. Une variation importante du SCR due à ce type d'événement a des conséquences sur le pilotage de la compagnie.

Les modèles internes prennent en compte le risque pandémique dans un module de

risque spécifique. Les chocs de mortalité calculés à partir des modèles calibrés sur les données affectées par la COVID-19 prendraient aussi en compte ce risque pandémique. Pour éviter un double comptage du risque pandémique dans les modèles internes, la correction des données impactées par la COVID-19 est nécessaire.

Dans ce chapitre, les conséquences de la pandémie sur les taux de mortalité et leur modélisation sont étudiées. La pandémie est un événement exceptionnel causant des effets ponctuels sur la mortalité. Les modèles de mortalité ne sont pas construits pour capturer ces irrégularités. C'est pourquoi le retraitement des données impactées par la COVID-19 a pour objectif d'éviter des distorsions dans la projection de la mortalité des portefeuilles assurés.

Les effets immédiats de la pandémie sur la mortalité sont traités dans ce chapitre. En revanche, il existe des effets longs de la COVID-19 sur la mortalité et la santé. Ces effets ne sont pas intégrés dans les modélisations et sont présentés dans la section 4.7.

## 4.2 Les données de mortalité

La pandémie s'est propagée de manière inégale dans le monde. C'est pourquoi la situation sanitaire de plusieurs pays est comparée en considérant le nombre de décès attribués à la COVID-19 en 2020. Pour ce faire, la *Human Mortality Database* est utilisée [HMD (2018)]. Il s'agit de la base de référence des estimations de la mortalité. Les données de mortalité sont disponibles pour une quarantaine de pays ou régions dans le monde. La période de recensement des données varie d'un pays à l'autre.

Les données de l'année 2020 sont au centre de cette étude afin de pouvoir étudier les impacts de la pandémie, c'est pourquoi la liste des pays étudiés est restreinte aux 27 pays pour lesquels les données sont disponibles sur le portail HMD. Ce sont majoritairement des pays d'Europe, auxquels s'ajoutent les Etats-Unis, le Canada, l'Australie, le Japon, Hong Kong, la Corée du Sud, la Nouvelle-Zélande et le Chili. Il s'agit de pays développés. Pour 13 pays, les données de l'année 2021 sont disponibles. Le tableau ci-dessous (figure 4.1) recense les pays étudiés et l'historique des données qui sera exploité dans la suite. La base HMD fournit les données pour la population des hommes, des femmes et la population totale de chaque pays.

Pays	Code HMD du pays	$t_{\min}$	$t_{\max}$
Australie	AUS	1921	2020
Belgique	BEL	1841	2021
Bulgarie	BGR	1947	2021
Canada	CAN	1921	2020
Chili	CHL	1992	2020
Croatie	HRV	2001	2020
Denemark	DNK	1835	2022
Angleterre & Pays de Galles	GBRTENW	1841	2020
Finlande	FIN	1878	2021
France	FRATNP	1878	2020
Allemagne	DEUTNP	1990	2020
Hong Kong	HKG	1936	2020
Hongrie	HUN	1950	2020
Islande	ISL	1838	2020
Irlande	IRL	1950	2020
Japon	JPN	1947	2021
Lituanie	LTU	1959	2020
Luxembourg	LUX	1960	2021
New-Zeland	NZL_NP	1948	2021
Norvège	NOR	1846	2022
Portugal	PRT	1940	2021
République de Corée	KOR	2003	2020
Espagne	ESP	1908	2020
Suède	SWE	1751	2022
Suisse	CHE	1876	2021
Royaume-Uni	GBR_NP	1922	2020
États-Unis	USA	1933	2021

TABLE 4.1 – Pays étudiés dans la base HMD au 03/03/2023

Le nombre de décès COVID-19 recensés pour chaque pays et répertoriés par Mathieu et al. (2021) est présenté dans le tableau 4.2 et sert de base de comparaison dans la suite du mémoire. A noter que ces chiffres correspondent au nombre de décès recensés et pas nécessairement au nombre de décès total dû à la COVID-19. En se restreignant aux données de l'année 2020, ces chiffres ne rendent pas nécessairement compte de la situation sanitaire actuelle en fonction de l'ampleur des vagues de pandémie survenues après 2020.

Pays	Nombre de décès COVID-19 en 2020	Nombre de décès COVID-19 en 2020 par million d'habitants
Nouvelle-Zélande	25	5
République de Corée	917	18
Hong Kong	148	20
Japon	3414	27
Australie	909	35
Norvège	433	80
Finlande	592	107
Islande	50	136
Denemark	1286	221
Canada	16685	439
Irlande	2237	448
Allemagne	50309	604
Lituanie	1800	644
Portugal	6840	663
Luxembourg	489	773
Chili	16499	863
Suisse	7523	871
Suède	9649	932
Croatie	3860	954
France	64644	957
Hongrie	9537	978
États-Unis	352055	1068
Bulgarie	7515	1085
Royaume-Uni	75198	1119
Espagne	54459	1150
Belgique	19757	1709

TABLE 4.2 – Nombre de décès COVID-19 par pays

Les données du nombre de décès par million d'habitants permettent de scinder cet échantillon de pays en trois groupes :

- "**peu ou pas touché**" : les pays comptant moins de 500 décès par million d'habitants, avec l'Australie, le Canada, le Danemark, la Finlande, Hong-Kong, l'Islande, l'Irlande, le Japon, la Nouvelle-Zélande, la Norvège et la Corée du Sud.
- "**modérément touché**" : les pays comptant entre 500 et 1000 décès par million d'habitants, avec l'Allemagne, la Lituanie, le Portugal, le Luxembourg, le Chili, la Suisse, la Suède, la Croatie, la Hongrie et la France.
- "**fortement touché**" : les pays comptant plus de 1000 décès par million d'habitants avec les USA, la Bulgarie, le Royaume-Uni, l'Espagne et la Belgique.

### 4.3 Impact de la pandémie de COVID-19 sur les taux de mortalité projetés par les modèles

Dans cette partie, l'impact de la pandémie sur la mortalité est étudié. L'objectif est d'évaluer dans quelle mesure la COVID-19 influe sur les tables de mortalité et d'analyser la manière dont les différents modèles de mortalité captent cette variation.

#### 4.3.1 Visualisation des données

Les taux de mortalité et leur logarithme ont une structure particulière en âge et en temps. Leur représentation apporte des informations sur la mortalité. Par exemple, les taux de mortalité des États-Unis en 2019 sont présentés en figure 4.1. On observe que les taux de mortalité pour les enfants de moins d'un an sont particulièrement élevés, il s'agit de la mortalité infantile. La mortalité est la plus faible autour de l'âge de 10 ans. On retrouve aussi la "bosse des accidents" qui correspond aux accidents et suicides pour les personnes entre 15 et 25 ans. Enfin, le logarithme de taux de mortalité est une fonction linéaire de l'âge pour les âges supérieurs à 40 ans.

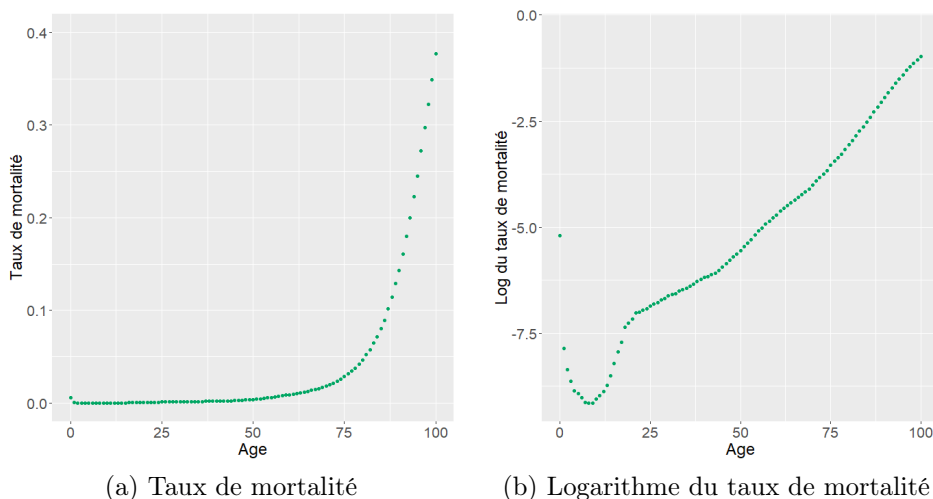


FIGURE 4.1 – Taux de mortalité (a) et leur logarithme (b) de la population totale des États-Unis en 2019 en fonction de l'âge

Il y a un écart entre les taux de mortalité des hommes et des femmes comme le montre la figure 4.2. Les femmes présentent des taux de mortalité inférieurs à ceux des hommes, cet écart est particulièrement visible à partir de l'âge de 10 ans. La "bosse des accidents" apparaît beaucoup moins prononcée sur les taux de mortalité féminins.

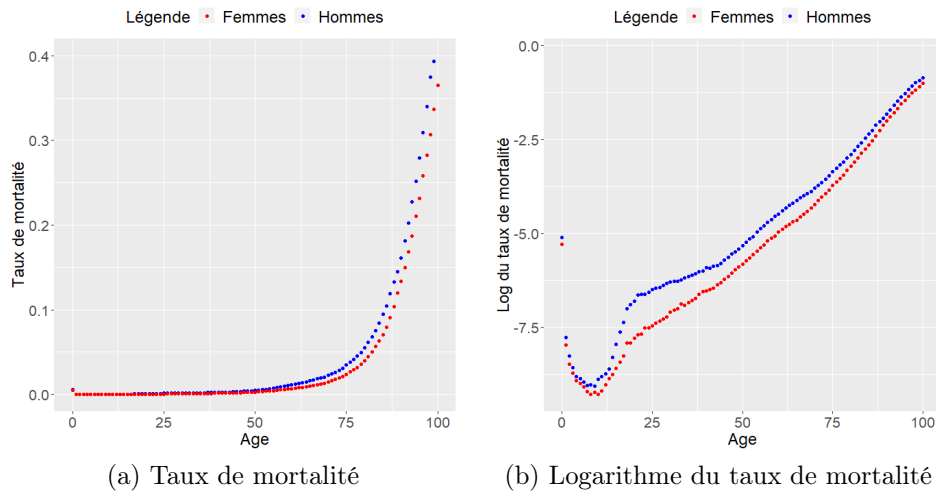


FIGURE 4.2 – Taux de mortalité (a) et leur logarithme (b) en 2019 aux États-Unis en fonction de l'âge et du genre

La mortalité varie selon les pays. Les taux de mortalité des États-Unis et du Japon sont comparés sur la figure 4.3. Le niveau de développement et l'allure à laquelle les pays se sont développés ont une influence sur les taux de mortalité actuels.

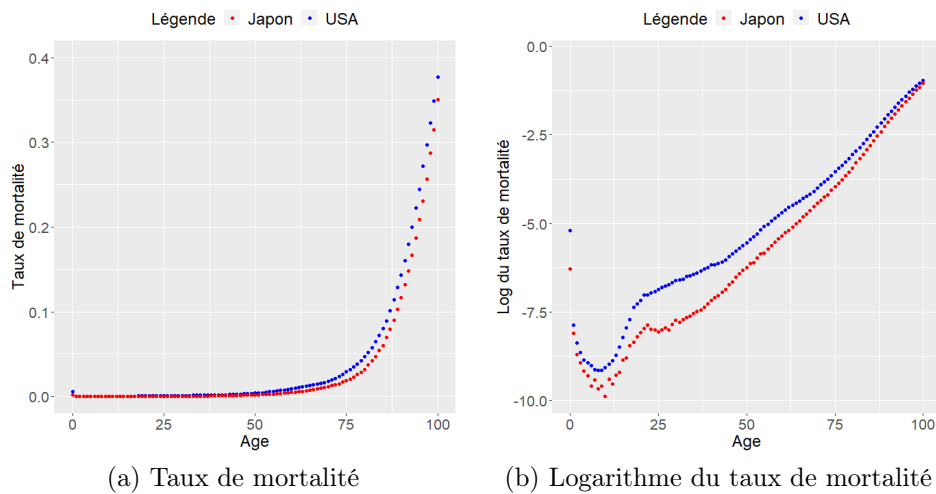


FIGURE 4.3 – Taux de mortalité (a) et leur logarithme (b) de la population totale en 2019 en fonction de l'âge et du pays (États-Unis ou Japon)

L'évolution des taux de mortalité entre les années 2019 et 2020 est visualisée. En plus des écarts observés entre les pays sur les années antérieures à la pandémie, les situations sanitaires varient d'un pays à l'autre en fonction de la propagation du virus et de la politique sanitaire menée dans chaque pays. Par exemple, aux États-Unis (figure 4.4), les taux de mortalité en 2020 sont supérieurs aux taux de mortalité en 2019 en conséquence de la pandémie. La population américaine a davantage été victime de la crise sanitaire

alors qu'au Japon (figure 4.5), une faible variation est observée.

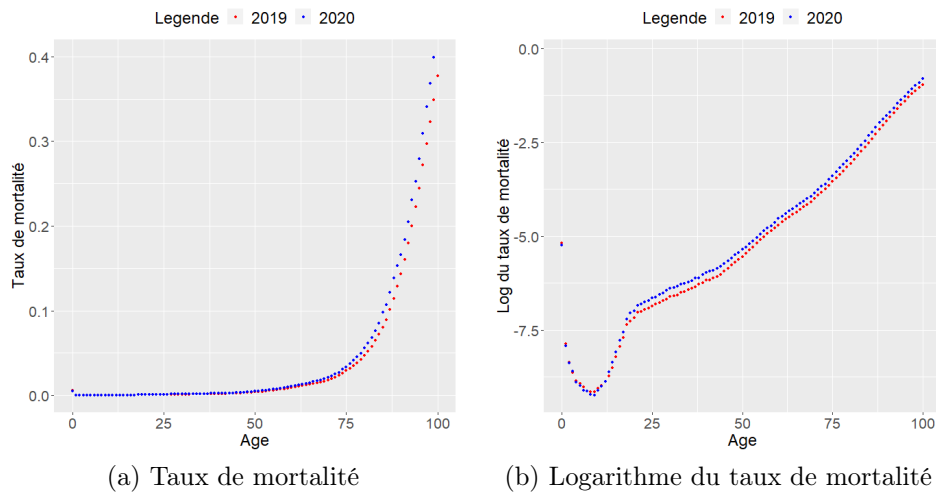


FIGURE 4.4 – Taux de mortalité (a) et leur logarithme (b) de la population totale aux États-Unis en 2019 et 2020 en fonction de l'âge

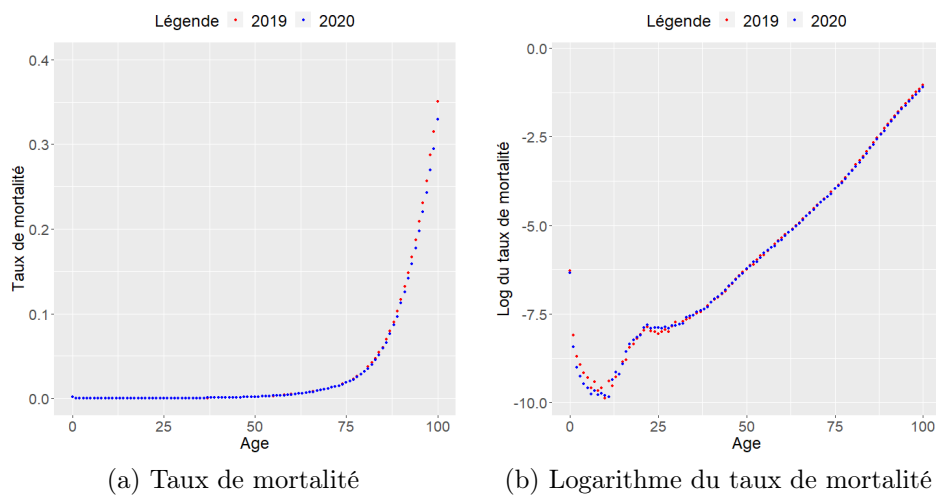


FIGURE 4.5 – Taux de mortalité (a) et leur logarithme (b) de la population totale au Japon en 2019 et 2020 en fonction de l'âge

La carte de chaleur des taux de mortalité en fonction de l'âge et du temps donne une vision plus générale de leur évolution. Pour les États-Unis (figure 4.6), l'augmentation sévère des taux de mortalité due à la COVID-19 des personnes de plus de 70 ans est visible alors que les taux de mortalité ont une tendance décroissante depuis les années 1980. De plus, les taux de mortalité des plus de 25 ans ont aussi augmenté aux États-Unis en 2020 en sachant que la tendance est à la hausse depuis 2010. Les cohortes sont identifiables sous forme de diagonales.



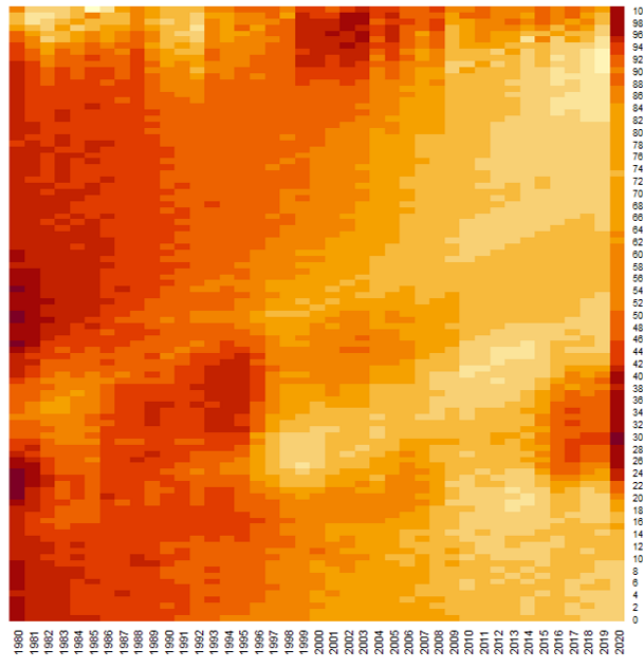


FIGURE 4.6 – Carte de chaleur du logarithme des taux de mortalité aux Etats-Unis

La visualisation des taux de mortalité montre que la pandémie a un impact différent selon les pays et selon les âges. Les variations observées sur la mortalité de l'année 2020 sont en accord avec le nombre de décès COVID-19 recensés. Sur la base du nombre de décès, les pays sont classés en trois groupes en fonction de l'impact de la COVID-19 sur la mortalité. Les Etats-Unis, fortement touchés, la France, modérément touchée, et le Japon, peu touché, sont sélectionnés pour être les représentants de leur groupe.

### 4.3.2 Application des modèles de mortalité

L'objectif de cette section est de déterminer l'impact de la pandémie sur la modélisation de la mortalité. Les modèles de mortalité sont calibrés sur les données dont le dernier point est l'année 2019 ou l'année 2020. Cela nous permet de déduire les effets de l'inclusion de l'année 2020 dans les données de calibrage.

Plusieurs indicateurs permettent de quantifier l'impact de la pandémie sur la modélisation :

- Les **paramètres de calibrage**, la tendance et la volatilité des séries temporelles  $\kappa_t^{(i)}$  définies dans l'équation 1.4.
- Les **projections des taux de mortalité** à partir des modèles ajustés.
- L'**espérance de vie projetée** calculée à partir des projections.
- L'**intervalle de confiance** autour des projections des taux de mortalité.

Les modèles sont construits sur les données de mortalité âge par âge et sont ajustés sur une période de quarante ans. Les modèles sont construits sur la population totale et en distinguant la population des femmes et des hommes pour chaque pays. Pour chaque modèle, l'ajustement prenant en compte l'année 2020 est comparé à l'ajustement s'arrêtant en 2019 afin de capter la manière dont le modèle intègre l'année 2020.

Les variations des indicateurs présentés entre le modèle dont les données de calibrage s'arrêtent en 2019 et celui dont les données s'arrêtent en 2020 sont analysées. De plus trois éléments sont comparés :

- La mortalité réelle survenue en 2020.
- La mortalité projetée en 2020 avec un modèle de mortalité dont les données de calibrage s'arrêtent en 2019.
- La mortalité projetée à un horizon de 50 ans avec les deux modèles.

Ces indicateurs, ainsi que les résultats de modélisation et projection des modèles sont mis en relation avec les connaissances à propos du nombre de décès recensés afin d'évaluer la cohérence de résultats.

Pour étudier ces impacts, plusieurs modèles de mortalité sont considérés, les modèles de Lee-Carter, APC, CBD et M7 définis dans la section 1.3. Ceci, afin d'identifier les différences sur la manière dont les modèles de mortalité s'ajustent à l'année 2020. Les résultats sont présentés pour le modèle de Lee-Carter, l'analyse des autres modèles est également réalisée.

Les résultats du calibrage et de la projection des taux de mortalité avec le modèle sont représentés pour les Etats-Unis (figure 4.7), la France (figure 4.8) et le Japon (figure 4.9).

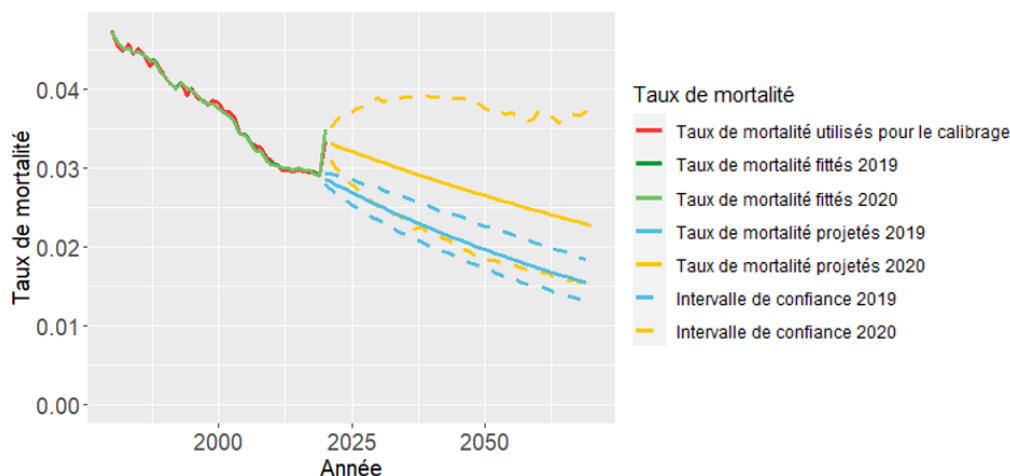


FIGURE 4.7 – Modélisation et projection des taux de mortalité à 75 ans des États-Unis avec le modèle de Lee-Carter

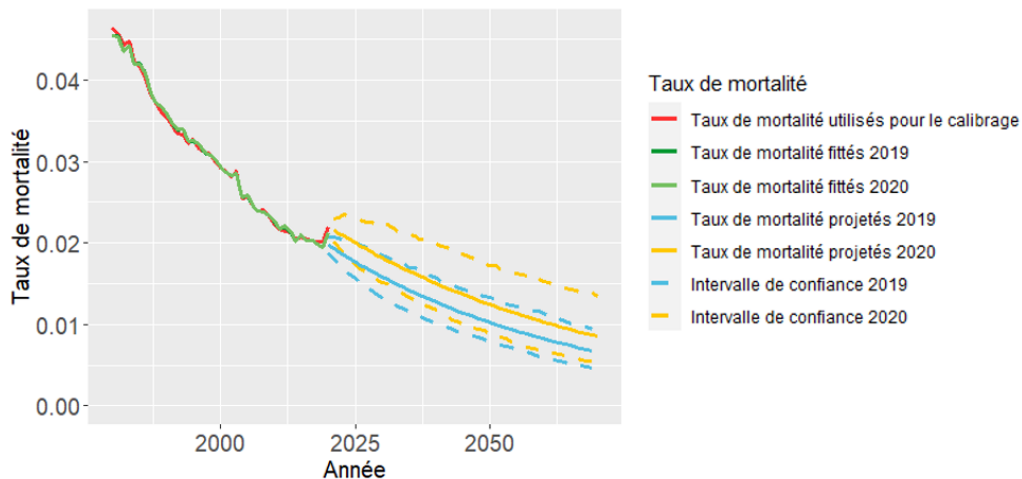


FIGURE 4.8 – Modélisation et projection des taux de mortalité à 75 ans en France avec le modèle de Lee-Carter

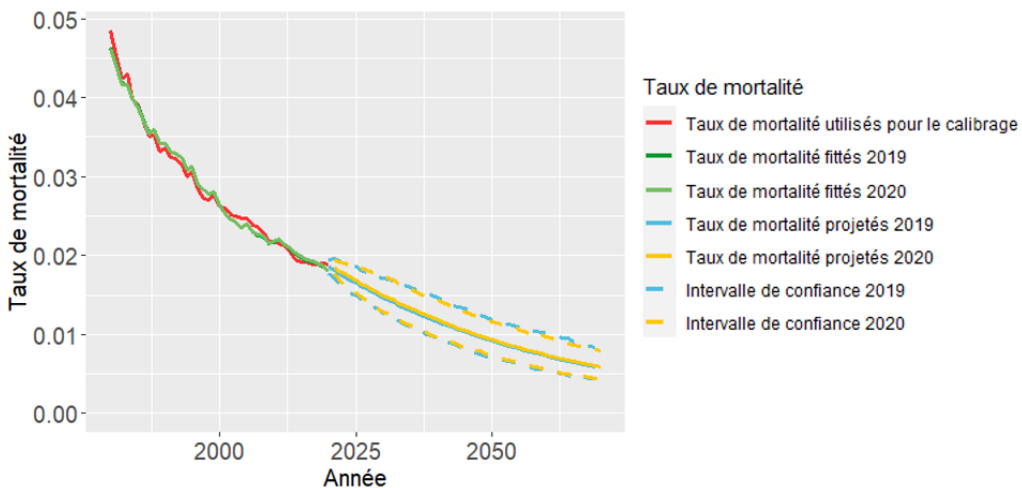


FIGURE 4.9 – Modélisation et projection des taux de mortalité à 75 ans du Japon avec le modèle de Lee-Carter

Les trois pays présentent des trajectoires de mortalité différentes, d'où la nécessité de poursuivre l'étude de l'impact de la pandémie sur les modèles. L'augmentation des taux de mortalité en 2020, visible sur ces graphiques, est cohérente avec les données de décès de la COVID-19. Ces pays sont bien représentatifs de leur groupe.

Le modèle de Lee-Carter s'ajuste aux données de 2020. **Intégrer l'année 2020 dans la modélisation implique une hausse des taux projetés et une augmentation de la largeur de l'intervalle de confiance autour des projections.** Ces conséquences sont particulièrement visibles pour les États-Unis. A l'inverse, pour le Japon, intégrer l'année 2020 à la modélisation a très peu d'impact sur les projections.

La tendance et la volatilité des séries temporelles  $\kappa_t^{(i)}$  sont extraites. Les modèles de Lee-Carter et APC sont constitués d'une unique série temporelle tandis qu'il y en a deux dans le modèle CBD et trois dans le modèle M7. Pour le modèle de Lee-Carter, l'écart relatif entre la tendance du modèle dont l'ajustement s'arrête en 2019 et celle du modèle incluant 2020 est représenté de la même manière que pour la volatilité (figure 4.10) afin de pouvoir capter les variations induites par l'introduction des données de 2020. Les valeurs pour les États-Unis, la France et le Japon sont répertoriées dans les tableaux 4.3 et 4.4.

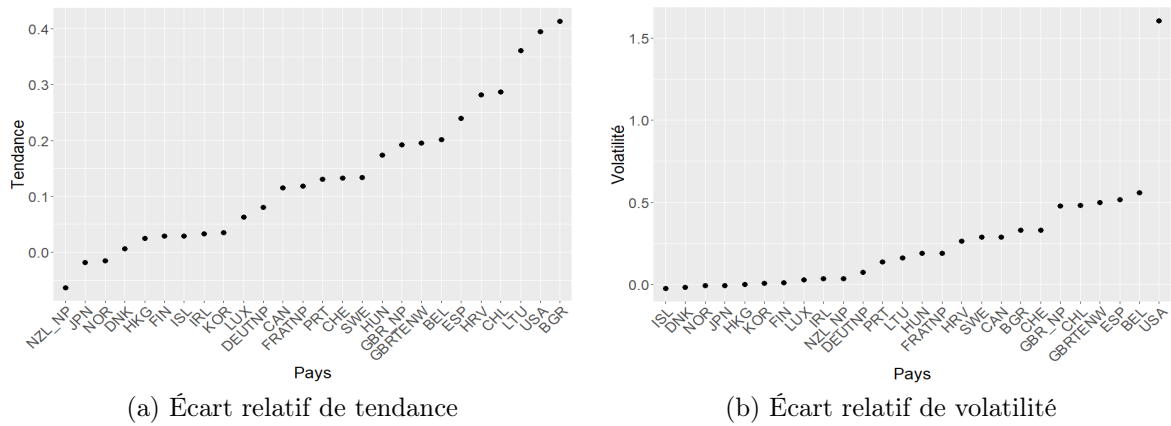


FIGURE 4.10 – Écart relatif (en %) de tendance (a) et volatilité (b) de  $\kappa_t^{(1)}$  entre les modèles de Lee-Carter avec et sans l'année 2020 en fonction du pays

	Tendance 2019	Tendance 2020	Écart relatif
<b>Japon</b>	-1,35	-1,38	-2%
<b>France</b>	-1,52	-1,34	12%
<b>USA</b>	-0,67	-0,40	40%

TABLE 4.3 – Tendances des modèles 2019 et 2020 de Lee-Carter

	Volatilité 2019	Volatilité 2020	Écart relatif
<b>Japon</b>	1,37	1,37	-0,5%
<b>France</b>	1,83	2,18	19%
<b>USA</b>	0,68	1,79	160%

TABLE 4.4 – Volatilités des modèles 2019 et 2020 de Lee-Carter

Ainsi, l'inclusion de l'année 2020 dans la modélisation entraîne l'augmentation relative de la tendance et de la volatilité. Les pays sont ordonnés en fonction de l'écart relatif observé. L'écart de la tendance varie entre  $-5\%$  et  $40\%$  selon les pays. Trois groupes de pays se distinguent sur ces graphiques et sont en cohérence avec les trois groupes constitués à partir des données de décès COVID-19 recensés. En effet, le Japon, faisant partie du

groupe 1, ne présente pas de variation dans la tendance et la volatilité contrairement à la France. Les États-Unis, membre du troisième groupe, présentent des résultats d'autant plus marqués avec une variation de la volatilité de 160%.

Puisque les groupes définis à partir des données de décès COVID-19 recensés sont retrouvés, cela suggère la cohérence des données de mortalité avec les données portant sur le nombre de décès COVID-19. **Plus un pays est impacté par la crise sanitaire, plus l'écart entre les tendances et volatilités de 2019 et 2020 est important.**

L'espérance de vie est aussi impactée par l'introduction des données 2020. Les espérances de vie projetées en 2023 se basant sur les projections du modèle de Lee-Carter sont répertoriées dans le tableau 4.5.

	USA		FRANCE		JAPON	
	2019	2020	2019	2020	2019	2020
<b>47 ans</b>	35,78	33,53	39,43	38,49	40,69	40,80
<b>67 ans</b>	17,71	16,34	20,23	19,56	20,85	20,92
<b>87 ans</b>	4,14	3,97	4,58	4,43	4,83	4,85

TABLE 4.5 – Espérances de vie projetées en 2023 avec le modèle de Lee-Carter

L'impact varie en fonction des pays. Pour les pays de groupe 1 comme le Japon, l'espérance de vie n'est pas impactée par la pandémie alors que l'espérance de vie en France et aux États-Unis est impactée à la baisse. L'impact est différent selon les âges. L'espérance de vie est diminuée de 0,15 ans pour les personnes de 87 ans en France alors qu'elle baisse de 0,67 ans pour les personnes de 67 ans. **L'intégration de l'année 2020 dans la modélisation entraîne une diminution de l'espérance de vie projetée.**

Les variations de tendance, de volatilité et d'espérance de vie sont également observées avec les modèles APC et CBD et sont cohérents avec le modèle de Lee-Carter. Les résultats sont en annexe C.

Le modèle CBD est construit à partir de deux séries temporelles. On remarque que la tendance et la volatilité de la série temporelle  $\kappa_t^{(1)}$  montrent des variations cohérentes avec les données de décès COVID-19 tandis que les variations de la tendance et de la volatilité de  $\kappa_t^{(2)}$  sont plus faibles.

L'espérance de vie et le score  $R^2$  des modèles sont comparés en annexe C. L'espérance de vie projetée est impactée de la même manière, quel que soit le modèle utilisé, alors que les scores  $R^2$  n'évoluent pas à l'intégration de l'année 2020 dans le modèle. Les scores varient aussi selon les pays. La mortalité de certains pays est plus difficile à ajuster, cela est dû aux fluctuations importantes de la mortalité au cours du temps. Il est à noter que les projections avec le modèle M7 ne sont pas cohérentes pour tous les cas d'application.

Les effets de l'introduction des données de l'année 2020 dans la modélisation de la mortalité ont été identifiés :

- L'augmentation de la tendance et de la volatilité des séries temporelles des modèles.
- La diminution de l'espérance de vie projetée.

L'ampleur des variations observées sur ces indicateurs est cohérente avec le nombre de décès recensés par pays. En comparant le calibrage des modèles de mortalité classiques, il apparaît que le modèle de Lee-Carter présente le meilleur compromis entre robustesse et précision.

## 4.4 Retraitement des points de données COVID-19 pour l'année 2020

### 4.4.1 Motivations

L'augmentation des taux de mortalité à cause de la pandémie a des conséquences sur la modélisation de la mortalité et notamment sur la projection de celle-ci. Introduire ces données dans le modèle engendre une hausse des taux de mortalité projetés. Or, pour éviter un double comptage du risque de pandémie dans le cas du modèle interne des assureurs, on cherche à retraiter la mortalité impactée par la COVID-19. Les effets à long terme de la pandémie COVID-19 sur la mortalité non pandémique ne sont pas pris en compte dans notre approche.

C'est pourquoi des méthodes de retraitement des données de l'année 2020 sont testées afin de corriger les données et obtenir des projections de taux de mortalité cohérentes. Corriger ces données signifie déterminer un taux de mortalité de référence qui aurait pu se produire en 2020 si la pandémie n'était pas survenue. La surmortalité est la différence entre la mortalité survenue et la mortalité de référence. Elle peut aussi être exprimée sous la forme d'un delta de taux de mortalité ou en nombre de décès. La méthode de calcul de la mortalité attendue est arbitraire, il en existe des standards souvent utilisées. Ces méthodes ont été challengées par de nouvelles alternatives. En effet, les méthodes usuelles, souvent très simples, sont limitées dans leurs performances.

L'objectif est de déterminer la surmortalité engendrée par la COVID-19 et de la comparer au nombre de décès de la COVID-19 recensés. Le recensement des décès COVID-19 n'est pas de qualité égale suivant les pays en fonction des politiques de dépistage et de déclaration de décès. Le calcul de la surmortalité permet d'obtenir des informations sur ce recensement qui correspond aux effets directs de la COVID-19 et permet aussi de mesurer les éventuels effets indirects de la COVID-19 qui ne sont pas mesurés dans les bases de données gouvernementales.

#### 4.4.2 Méthodes usuelles sur les taux annuels

Les méthodes pour retraiter les données de mortalité peuvent être standards ou alternatives, utilisées historiquement ou développées pour la période pandémique, elles peuvent faire appel à des données supplémentaires ou basées uniquement sur les données HMD. Ces méthodes sont définies et appliquées dans cette section.

##### 4.4.2.1 Soustraction des décès COVID-19

Retraiter les décès consiste à soustraire les décès dus au COVID-19 au nombre de décès total de l'année 2020. Il s'agit de la méthode la plus intuitive pour retrouver un niveau de référence sans COVID-19. Cette méthode requiert des données de décès COVID-19 recensés à une maille suffisamment fine pour ne pas perdre d'information. La base de données de Caporali et al. (2022) publiée par l'INED (l'Institut national d'études démographiques) répertorie le nombre de décès cumulé par semaine à partir de janvier 2020 par tranche d'âge de cinq ou dix ans, tranches d'âge variant selon les pays. Ces données sont disponibles pour dix pays : le Danemark, l'Angleterre et le Pays de Galles, l'Allemagne, le Japon, la Norvège, le Portugal, l'Espagne, la Suède, la Suisse, les USA.

La première limite de cette méthode est l'accès aux données, l'étude est restreinte aux pays dont les données sont fournies. Les résultats de cette méthode dépendent de la qualité des données fournies.

Puisque les modèles de mortalité sont appliqués âge par âge, les décès recensés par classe d'âge sont répartis entre les différents âges de la classe de deux manières possibles :

- De façon équitable entre tous les âges de la classe, la limite de cette méthode se trouve surtout au niveau des individus âgés de plus de 60 ans pour lesquels l'hypothèse selon laquelle le nombre de décès par âge est homogène au sein de la classe n'est pas vérifiée.
- De façon proportionnelle à l'exposition des individus dans la classe.

Notons par ailleurs que la pandémie a des effets directs qui sont les décès COVID-19, ainsi que des effets indirects comme la baisse du nombre d'accidents de voiture ou le report d'actes médicaux. Avec cette méthode de retraitement, les décès directs de la COVID-19 sont soustraits alors que la correction ne prend en compte que les effets directs.

##### 4.4.2.2 Moyenne de la mortalité historique

Les taux de mortalité de 2020 sont corrigés par la moyenne des taux de mortalité sur les  $N$  dernières années. Le nombre d'années d'historique pris en compte est fixé à cinq ans. Dans la littérature, cette période de cinq ans est souvent conseillée pour que les résultats ne soient pas trop dépendants d'anomalies passées tout en restant cohérents avec la structure de la mortalité actuelle.

Les taux de mortalité ont une tendance décroissante dans le temps qui n'est pas prise en compte dans cette méthode. En réalisant la moyenne des taux de mortalité historiques, la méthode reste dépendante des fluctuations passées.

#### 4.4.3 Méthodes usuelles sur les taux hebdomadaires

Pour répondre aux besoins nés pendant la crise sanitaire, le portail HMD a mis à disposition une nouvelle ressource : Short-Term Mortality Fluctuations (STMF) [Jdanov et al. (2021)]. Il s'agit du nombre de décès hebdomadaires par groupe d'âge agrégé et par genre. Ces tables sont disponibles pour une liste de pays limitée et sont intéressantes seulement si l'année 2020 est présente dans la base ainsi que dans la base HMD. Les méthodes présentées par la suite utilisent les taux de mortalité hebdomadaires afin de déterminer la mortalité de référence pour l'année 2020.

Country	First Week	Last Week
Australie	01.2015	04.2023
Belgique	01.2000	15.2023
Bulgarie	01.2010	13.2023
Canada	01.2010	52.2022
Chili	01.2016	19.2023
Croatie	01.2001	12.2023
Denemark	01.2007	18.2023
Finlande	01.1990	14.2023
France	01.2000	15.2023
Allemagne	01.2000	18.2023
Hongrie	01.2000	14.2023
Islande	01.2000	13.2023
Lituanie	01.2000	13.2023
Luxembourg	01.2000	13.2023
Nouvelle-Zélande	52.2010	18.2023
Norvège	01.2000	18.2023
Portugal	01.2000	17.2023
République de Corée	01.2010	30.2022
Espagne	01.2000	35.2022
Suède	01.2000	16.2023
Suisse	01.2000	18.2023
États-Unis	02.2015	14.2023

TABLE 4.6 – Pays étudiés dans la base STMF

Cette base de données a été créée pour fournir des données permettant une analyse des fluctuations de la mortalité toutes causes confondues par semaine au cours de chaque année civile. La décision d'ajouter cette nouvelle ressource au portail HMD a été prise à la suite de la pandémie de COVID-19. L'importance croissante des fluctuations de mortalité à court terme ou saisonnières, dues à des risques temporaires tels que les épidémies de grippe, les températures extrêmes et les catastrophes naturelles ou causées par l'homme, a également motivé l'ajout de cette nouvelle ressource. L'objectif de ces méthodes est



d'obtenir un niveau de référence de la mortalité, appelé *standard death rates* (SDR) pour une année d'indice donnée  $y$ , avec  $y = 2020$  pour le retraitement de l'année 2020.

#### 4.4.3.1 Méthode Week-specific Average

Le taux de référence hebdomadaire pour la semaine  $w$  est la moyenne des taux de mortalité des  $N$  années précédant l'année d'indice  $y$  de la semaine  $w$ , c'est à dire :

$$\overline{SDR}(T, w) = \frac{1}{N} \cdot \sum_{y \in Y} SDR(y, w) \quad (4.1)$$

avec :

- $T$  : Année cible
- $Y$  : Ensemble des années historiques considérées
- $N$  : Nombre d'années constituant  $Y$

Avec cette méthode, un taux de mortalité de référence est déterminé par semaine.

#### 4.4.3.2 Méthode Week-specific Trends

Le taux de référence hebdomadaire pour la semaine  $w$  est déterminé en effectuant une régression linéaire des taux de mortalité des  $N$  années précédentes sur la semaine  $w$ , c'est à dire :

$$\widehat{SDR}(T, w) = \hat{\alpha}_w + \hat{\beta}_w \cdot T \quad (4.2)$$

avec :

- $T$  : Année cible
- $\hat{\alpha}_w$  et  $\hat{\beta}_w$  : Coefficients de régression déterminés sur  $Y$ , ensemble des  $N$  années antérieures

#### 4.4.3.3 Méthode Week-specific Lower Quartiles

Le taux de référence hebdomadaire de la semaine  $w$  est déterminé en réalisant la moyenne des taux de mortalité constituant le quartile des taux de mortalité des  $N$  années antérieures sur la semaine  $w$ .

$$SDR^{Q_1}(T, w) = \frac{1}{|L_w|} \cdot \sum_{y \in L_w} SDR(y, w) \quad (4.3)$$

avec :

- $T$  : Année cible
- $Y$  : Ensemble des années historiques considérées
- $N$  : Nombre d'années historiques
- $L_w = \{y \in N : SDR(y, w) < \text{quantile}(\{SDR(y, w) : y \in Y\}, 0, 25)\}$

#### 4.4.3.4 Méthode Yearly Average-week

Le taux de référence est annuel et correspond à la moyenne des taux hebdomadaires sur les  $N$  années antérieures. Contrairement aux méthodes précédentes, le taux de mortalité de référence est déterminé pour l'ensemble de l'année cible indépendamment des semaines :

$$\overline{SDR}(T) = \frac{1}{S} \cdot \sum_{w,y \in W,Y} SDR(y, w) \quad (4.4)$$

avec :

- $T$  : Année cible
- $Y$  : Ensemble des années historiques considérées
- $W$  : Ensemble des semaines
- $S$  : Nombre de semaines sur les  $Y$  années
- $N$  : Nombre d'années historiques

#### 4.4.3.5 Méthode Summer Average-week

Le taux de référence est annuel et correspond à la moyenne des taux hebdomadaires en période estivale. La période estivale de l'hémisphère nord s'étend de la semaine 13 à la semaine 48, de avril à novembre, celle de l'hémisphère sud de la semaine 1 à 21 et 39 à 52 ou 53 selon les années, c'est à dire de janvier à mai et de octobre à décembre.

$$\overline{SDR}^S(T) = \frac{1}{S} \cdot \sum_{w,y \in W,Y} SDR(y, w) \quad (4.5)$$

avec :

- $T$  : Année cible
- $Y$  : Ensemble des années historiques considérées
- $W$  : Ensemble des semaines de la période estivale
- $S$  : Nombre de semaines dans la période estivale sur les  $Y$  années
- $N$  : Nombre d'années historiques

#### 4.4.4 Méthodes alternatives

Ces méthodes alternatives sont décrites dans l'article de Shkolnikov et al. (2022) dont le propos est de définir deux méthodes alternatives pour déterminer un niveau de mortalité de référence pour l'année pandémique mais aussi pour les années antérieures.

#### 4.4.4.1 Méthode alternative rétrospective

La méthode rétrospective proposée dans l'article est dérivée des taux de mortalité hebdomadaires les plus bas atteints les années précédentes :

$$\text{SDR}_{\text{areto}}^B(y, w) = \hat{\alpha}_w + \hat{\beta}y \quad (4.6)$$

Où :

- $w$  le numéro de la semaine
- $\hat{\beta}$  est une estimation de la pente de la régression linéaire de SDR annuel en fonction de l'année :  $t : \text{SDR}(t) = \beta \cdot t + \varepsilon_t, y - N \leq t \leq y - 1$
- Les effets hebdomadaires  $\hat{\alpha}_w$  sont définis comme suit  $\hat{\alpha}_w = \min_t 2[\text{SDR}(t, w) - \hat{\beta}t]$  avec  $y - N \leq t \leq y - 1$  ( $\min_t 2$  la deuxième plus petite valeur)

Pour chaque semaine  $w$  de l'année d'indice  $y$ , le taux de mortalité hebdomadaire de référence  $\text{SDR}_{\text{areto}}^B(y, w)$  est égal à la deuxième valeur la plus basse des taux de mortalité hebdomadaires pour la même semaine au cours des  $N$  années antérieures avec un décalage supplémentaire pour tenir compte de la tendance sous-jacente de la mortalité.

Le choix du nombre d'années historiques est important puisque la présence d'anomalies en nombre trop important dans l'historique a une influence sur les résultats de la méthode. Les anomalies considérées sont des pics de mortalité hivernale qui peuvent être dus à des épidémies de grippe saisonnières. Les canicules sont aussi des anomalies, dont la durée ne dépasse généralement pas un mois. L'article montre qu'une période rétrospective de sept ans est optimale, ce qui est restreignant vis à vis de la disponibilité des données. En appliquant cette méthode sur les années antérieures, la surmortalité engendrée par une épidémie de grippe saisonnière intense peut être déterminée. De ce fait, il est conseillé dans l'article d'utiliser une période supérieure ou égale à cinq ans pour obtenir des résultats cohérents.

La tendance des taux de mortalité est décroissante dans le temps, l'année 2020 déroge à cette règle. Plusieurs méthodes usuelles se basent sur la moyenne des taux de mortalité sur les années antérieures et ne prennent pas en compte la tendance des taux de mortalité, ce qui explique en partie le fait qu'elles sous-estiment la surmortalité. Cette méthode de référence est basée sur les valeurs minimales des années antérieures et n'est donc pas sensible aux pics de mortalité passés, ce qui est une différence majeure avec les méthodes standards.

#### 4.4.4.2 Méthode alternative intra-annuelle

La méthode intra-annuelle, *alternative within-year*, se base sur les taux de mortalité hebdomadaires de l'année cible. Cette méthode est indépendante des pics des taux de mortalité historiques. Pour calculer la mortalité de référence, la méthode se base sur les taux de mortalité hebdomadaires les plus bas au cours de l'année cible  $y$  :

$$\text{SDR}_{wy}^B(y) = (1/13) \cdot \sum_{w \in Q_1} \text{SDR}(y, w) \quad (4.7)$$

Où  $Q_1$  inclut les 13 semaines (un quart d'année) constituant le quartile inférieur des taux de mortalité de l'année cible.

Généralement, les taux de mortalité constituant le quartile inférieur surviennent hors de la saison hivernale. Cependant, l'année 2020 présente des pics de mortalité en dehors de la saison hivernale en lien avec les différentes vagues de COVID-19 survenues au cours de l'année. Les taux de mortalité sélectionnés pour le calcul de la valeur de référence ne sont pas tous sur les mois d'été. Contrairement à la méthode alternative rétrospective, cette approche permet de calculer un niveau de référence unique pour toute l'année cible.

Seuls les taux de l'année cible sont utilisés, donc même si la tendance n'est pas mesurée elle peut être capturée à condition que tous les taux de mortalité de l'année cible ne soient pas des anomalies. C'est le cas pour l'année 2020, puisque malgré les pics importants de mortalité, les taux reviennent à un niveau normal entre deux vagues consécutives. En considérant ce niveau de référence inférieur, les pics de surmortalité hivernale ne sont pas traités comme inévitables. C'est un des changements majeurs par rapport aux méthodes usuelles.

Les méthodes alternatives ne sous-estiment pas la surmortalité contrairement aux méthodes basées sur les moyennes historiques. De plus, à cause de la définition de ces deux méthodes alternatives, la méthode intra-annuelle présente des résultats supérieurs à la méthode rétrospective.

#### 4.4.5 Comparaison des méthodes

##### 4.4.5.1 Calcul de la surmortalité

Les méthodes de retraitement sont appliquées sur la période de 2005 à 2020 lorsque les données le permettent. L'objectif de cette application est d'approcher la surmortalité de l'année 2020 afin d'en déduire la surmortalité attribuable à la pandémie.

Analyser la surmortalité de 2020 passe également par la détermination de la surmortalité des années antérieures. Il faut pouvoir confronter les résultats de l'année 2020 à ceux des années précédentes puisque l'écart entre les surmortalités passées est le signe d'excès de mortalité.

Les données de mortalité hebdomadaires permettent de visualiser les variations périodiques de la mortalité au cours de l'année. La saison hivernale est marquée par une augmentation de la mortalité contrastant avec le creux de la période estivale. Cette augmentation entre décembre et mars est liée à la circulation de virus favorisée par les températures basses. Des événements particuliers sont à l'origine de fluctuations de la mortalité telles que les canicules ou épidémies.

Les méthodes appliquées déterminent un niveau de mortalité de référence hebdomadaire ou annuel pour finalement déduire la surmortalité annuelle correspondant à la

différence entre la surmortalité survenue et le niveau de référence. Les résultats de surmortalité sont répertoriés dans des cartes de chaleur. La surmortalité pour 100 000 habitants est considérée et une carte de chaleur est construite par méthode.

Les méthodes considérées présentent des résultats différents quant à la surmortalité annuelle. Les résultats des méthodes *week-specific lower quartiles* (figure 4.11), *week-specific trends* (figure 4.12) et *alternative within-year* (figure 4.13) sont présentées dans les cartes de chaleur.

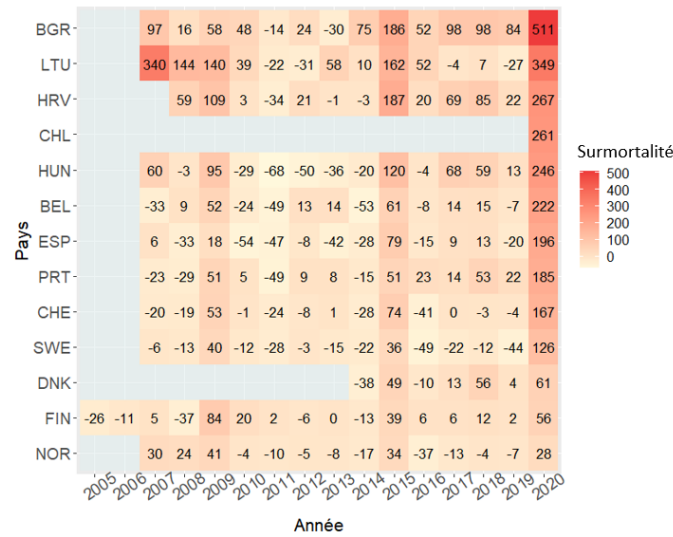


FIGURE 4.11 – Carte de chaleur de la surmortalité (pour 100 000 habitants) déterminée avec la méthode *week-specific lower quartiles*

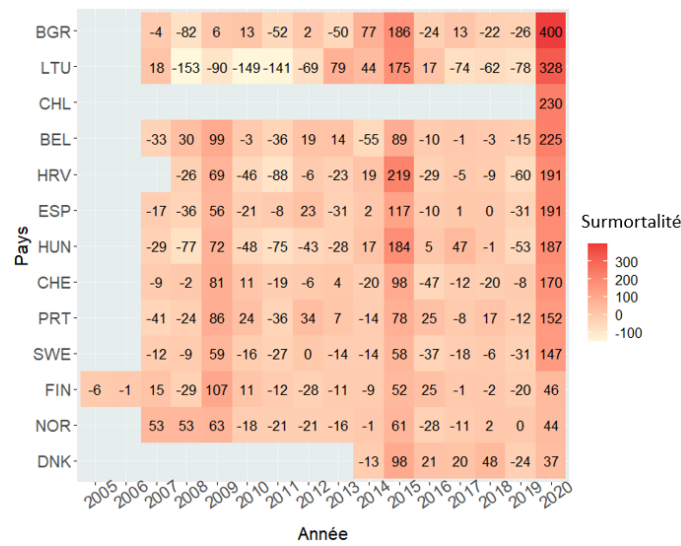


FIGURE 4.12 – Carte de chaleur de la surmortalité (pour 100 000 habitants) déterminée avec la méthode *week-specific trends*

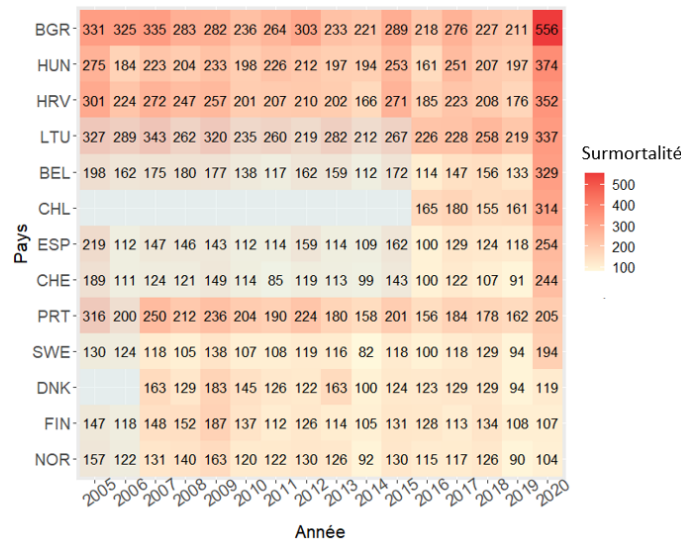


FIGURE 4.13 – Carte de chaleur de la surmortalité (pour 100 000 habitants) déterminée avec la méthode *alternative within-year*

Ces méthodes sont cohérentes entre elles puisqu'elles convergent vers une surmortalité de l'année 2020 similaire. Il est à noter que les résultats de la méthode *alternative within-year* sont supérieurs aux résultats des méthodes rétrospectives dans la mesure où les fluctuations passées n'interfèrent pas et que le seuil de référence de la méthode *alternative within-year* est déterminé à l'aide de quartiles permettant d'atteindre des seuils suffisamment bas sur l'année 2020.

La surmortalité déterminée pour chaque pays sur la période de 2005 à 2019 est différente selon les méthodes. Les méthodes *week-specific lower quartiles* et *week-specific trends* présentent des résultats négatifs signifiant que moins de décès seraient survenus par rapport au nombre de décès attendu. Il s'agit d'une des limites de ces méthodes qui considèrent des hausses de la mortalité, et notamment celle de la mortalité hivernale, comme inévitables. Cependant, il est aussi possible de considérer les épidémies de grippe hivernales comme des phénomènes exceptionnels qu'il est possible d'éviter par des campagnes de prévention et de vaccination par exemple. De ce point de vue, la surmortalité annuelle devrait refléter l'augmentation de la mortalité en hiver liée à ces épidémies comme dans les méthodes alternatives qui présentent des résultats de surmortalité nettement supérieurs sur les données historiques.

L'excès de mortalité de l'année 2020 calculé à partir des différentes méthodes est présenté dans le tableau 4.7.

	USA	France
<b>Décès COVID-19 recensés</b>	107	96
<b>Moyenne annuelle</b>	131	132
<b>Week-specific average</b>	161	90
<b>Week-specific lower quartiles</b>	178	124
<b>Week-specific trends</b>	106	130
<b>Alternative within-year</b>	134	193
<b>Alternative retrospective</b>	205	125

TABLE 4.7 – Surmortalité (pour 100 000 habitants) de l'année 2020 en fonction des méthodes

Pour les États-Unis et la France, la plupart des méthodes d'ajustement aboutissent à une surmortalité en 2020 plus élevée que les décès officiels enregistrés dans le cadre de la COVID-19. Ce résultat était attendu car ces méthodes produisent une surmortalité totale qui inclut également les effets indirects de COVID-19. Il convient de noter que les résultats de ces méthodes ne sont pas homogènes. En particulier, les méthodes *week-specific lower quartiles*, *alternative within-year* et *alternative retrospective* conduisent à des résultats de surmortalité plus élevés que les méthodes de moyenne annuelle, *week-specific average* et *week-specific trends*.

La pertinence des méthodes de retraitement est évaluée selon trois critères :

- La méthode prend-elle en compte la décroissance des taux de mortalité ?
- La méthode est-elle biaisée ?
- Quelle est la précision de la méthode ?

Le tableau 4.8 répertorie les caractéristiques des méthodes de retraitement selon ces trois critères.

Méthode	Critère 1 : Décroissance	Critère 2 : Biais	Critère 3 : Précision
Soustraction des décès COVID-19	NON	Ne prend pas en compte les effets indirects de la COVID-19	Correction cohérente pour les pays à faible intensité de COVID-19
Moyenne de la mortalité historique	NON	Biaisée par les fluctuations passées	Sous-estime la surmortalité
Week-specific Averages	NON	Basée sur les taux hebdomadaires passés contenant des fluctuations (grippe, canicule)	Sous-estime la surmortalité
Week-specific Trends	OUI	Résultats de la régression influencés par les valeurs aberrantes	Sous-estime la surmortalité
Week-specific Lower Quartiles	OUI	Atteint un niveau de référence bas Biaisée par la mortalité hivernale	Sous-estime la surmortalité
Yearly Average-week	NON	Biaisée par les fluctuations passées des taux de mortalité	Sous-estime la surmortalité Résultats proches des méthodes basées sur la moyenne historique
Summer Average-week	NON	Non biaisée par la période d'hiver qui est exclue Influencé par les canicules	Sous-estime la surmortalité
Alternative within-year	NON	Non biaisée par les pics de mortalité passés Ne traite pas comme inévitable la surmortalité hivernale non pandémique	Les fluctuations de la mortalité au cours d'une même année contribuent davantage à la surmortalité que ne le suppose la méthode alternative rétrospective Ne sous-estime pas la surmortalité, résultats supérieurs aux méthodes standards
Alternative retrospective	OUI	Non biaisée par les pics de mortalité passés Ne traite pas comme inévitable la surmortalité hivernale non pandémique	La méthode <i>alternative within-year</i> montre des résultats supérieurs à la méthode rétrospective

TABLE 4.8 – Récapitulatif des méthodes de retraitement



#### 4.4.5.2 Illustration de l'application des méthodes à la mortalité passée

Les méthodes de retraitement sont appliquées à des années spécifiques afin de déterminer la surmortalité attribuable à des événements particuliers. Cela permet de confronter les résultats des méthodes aux données officielles. Nous nous intéressons à la canicule de 2003 et aux phénomènes d'épidémie de grippe hivernale.

##### La canicule de 2003

La canicule de 2003 en France est un épisode de forte chaleur survenu en août 2003 qui a causé 19490 morts selon l'étude de World Health Organization (2008). L'application des méthodes standards et alternatives sur l'année 2003 permet de déduire la surmortalité de cette année.

Retrouver la mortalité engendrée par des événements spécifiques est difficilement réalisable en considérant un niveau de référence annuel. De plus, les trois années précédant la canicule sont marquées par des épisodes de grippe importants que les méthodes de retraitement captent. C'est pourquoi ce sont les méthodes permettant de déterminer un niveau de référence hebdomadaire qui sont appliquées et la surmortalité est déterminée en se restreignant à la période de l'événement. Les résultats sont présentés dans le tableau 4.9.

Méthodes	Surmortalité (nombre de décès) Été 2003
Week-specific averages	18071
Week-specific trends	21184
Week-specific lower quartiles	18071
Alternative within-year	18880

TABLE 4.9 – Surmortalité de l'été 2003 en France en fonction des méthodes

Les résultats des méthodes sont cohérents avec le nombre de décès attendus et convergent tous vers un résultat commun. Les méthodes rétrospectives qui attribuent un niveau de référence hebdomadaire permettent de déterminer un niveau de référence adapté à la mortalité saisonnière. De plus, la méthode *alternative within-year*, qui même si elle donne un niveau de référence annuel, donne de bons résultats. Le niveau de référence est abaissé à un niveau minimum, ce qui permet à la méthode de ne pas sous-estimer la surmortalité.

##### La grippe saisonnière

La grippe saisonnière est une des raisons de la hausse de la mortalité hivernale. L'intensité de la vague de mortalité hivernale fluctue d'une année à l'autre. L'application des méthodes afin de déterminer la surmortalité attribuable à la grippe ne donne pas de résultats satisfaisants pour plusieurs raisons :

- La mortalité due à la grippe est confondue avec la vague de mortalité hivernale.
- Les méthodes rétrospectives sont influencées par les vagues de mortalité hivernales passées.

Les résultats de l'application sont en annexe D.

#### 4.4.5.3 Application des modèles de mortalité sur les données retraitées

Le modèle de Lee-Carter est calibré avec des données allant jusqu'à l'année 2020. Nous analysons ensuite l'impact de l'utilisation de données de mortalité retraitées pour l'année 2020 sur le calibrage des modèles de mortalité, en termes de projection des taux de mortalité et des espérances de vie.

La figure 4.14 montre la projection des taux de mortalité à 70 ans aux États-Unis avec et sans données de mortalité ajustées. Les taux de mortalité ajustés en 2020 sont cohérents avec les taux de mortalité historiques. En outre, la largeur des intervalles de confiance des projections de mortalité de Lee-Carter est réduite lorsque l'on utilise des données ajustées, en raison d'une plus faible volatilité  $\sigma$ .

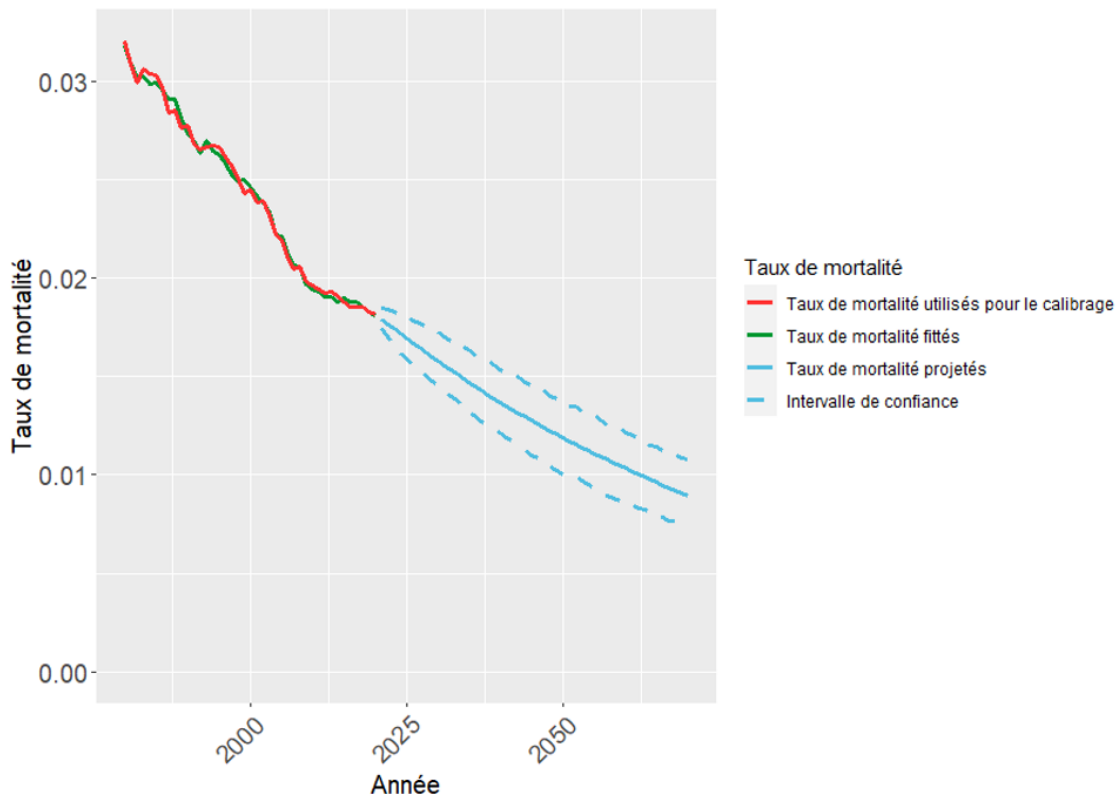


FIGURE 4.14 – Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction de l'année 2020 par *alternative within-year*

L'inclusion des données de mortalité de l'année 2020 non ajustées dans la période de calibrage du modèle de Lee-Carter entraîne une réduction significative de l'espérance de vie. Le tableau 4.10 présente les résultats obtenus en tenant compte des données de mortalité ajustées pour les États-Unis, la France pour trois âges représentant des groupes d'âge

touchés différemment par la pandémie. Les quartiles inférieurs spécifiques à la semaine et la méthode *alternative within-year* sont celles qui conduisent aux espérances de vie les plus élevées pour la France.

	USA			FRANCE		
	Week-specific lower quartiles	Alternative within-year	Historical average	Week-specific lower quartiles	Alternative within year	Historical average
<b>47 ans</b>	35,78	34,55	35,47	38,49	40,10	38,93
<b>67 ans</b>	17,72	16,96	17,52	19,56	20,71	19,88
<b>87 ans</b>	4,16	4,05	4,12	4,43	4,70	4,50

TABLE 4.10 – Espérances de vie projetées en 2023 avec le modèle de Lee-Carter avec correction

Afin de projeter la mortalité en corrigeant les impacts de la COVID-19, des méthodes de retraitement des données de mortalité sont définies et implémentées. Les méthodes indirectes basées sur les taux de mortalité hebdomadaires permettent de prendre en compte les effets indirects de la COVID-19 contrairement à la méthode directe basée sur le nombre de décès de la COVID-19.

Parmi les méthodes indirectes, les méthodes *week-specific lower quartiles*, basées sur les quartiles des taux de mortalité des années antérieures et la méthode *alternative within-year* permettent de corriger les taux de mortalité afin qu'ils soient dans la continuité de ceux de 2019.

Ces méthodes sont particulièrement intéressantes dans le cadre des modèles internes. En effet, le risque pandémique est mesuré dans un module spécifique, d'où l'intérêt de retraiter les données de mortalité pour éviter le double comptage des effets de la pandémie.

#### 4.4.6 Calcul des chocs de mortalité

Cette section a pour objectif de déterminer l'impact de la pandémie sur le calcul des chocs de mortalité et de pouvoir analyser les résultats des chocs en intégrant les corrections de l'année 2020.

La méthode de calcul des chocs de mortalité s'inspire de la méthode utilisée par l'EIOPA (2018). La méthode se base sur les projections des taux de mortalité à horizon un an obtenues avec un modèle de mortalité. Les trois étapes de la méthode de calcul des chocs sont les suivantes :

1. **Projection des tables de mortalité.** Le modèle de Lee-Carter est calibré sur les taux de mortalité historiques. Puisque l'objectif est de calculer le choc à horizon un an, la projection des taux de mortalité est réalisée à horizon un an. Au delà de la période d'un an, les composantes de la série temporelle sont déterminées de manière déterministe. De cette manière, les simulations des taux de mortalité sur 120 ans sont obtenues pour 5000 simulations.
2. **Calcul des espérances de vie résiduelles.** Pour chaque table de mortalité, l'espérance de vie résiduelle pour chaque âge  $x$  et année  $t$  est déterminée. On introduit le choc  $h$ , optimisé dans l'étape suivante.

$$e_x^h(t) = \frac{1}{2} + \sum_{k=1}^{+\infty} \prod_{s=0}^{k-1} (1 - (1+h)q_{x+s}(t+s)) \quad (4.8)$$

3. **Détermination du choc de mortalité.** Le choc de mortalité est le choc minimisant l'écart entre l'espérance de vie résiduelle choquée et l'espérance de vie résiduelle du centile 0.5% déterminé sur les 1000 simulations. Un intervalle de confiance du choc de mortalité estimé est également construit.

$$h_{\text{inf}}(x) = \operatorname{argmin}_{h \in ]-1, 1[} \left( e_x^h(t) - e_x^{0.5\%}(t) \right)^2 \quad (4.9)$$

Ainsi, les chocs sont déterminés à partir des résultats du modèle de Lee-Carter avec ou sans les méthodes de retraitement décrites. Un des intérêts est de pouvoir déduire l'influence de la pandémie sur le calcul des chocs de mortalité et les chocs attendus en employant des données retraitées. Les chocs de mortalité sont calculés âge par âge et sont agrégés par classe d'âge pour rendre compte de l'impact de la COVID-19 sur la population.

Les résultats sont présentés pour les Etats-Unis (figure 4.15), la France (figure 4.16) et le Japon (figure 4.17).

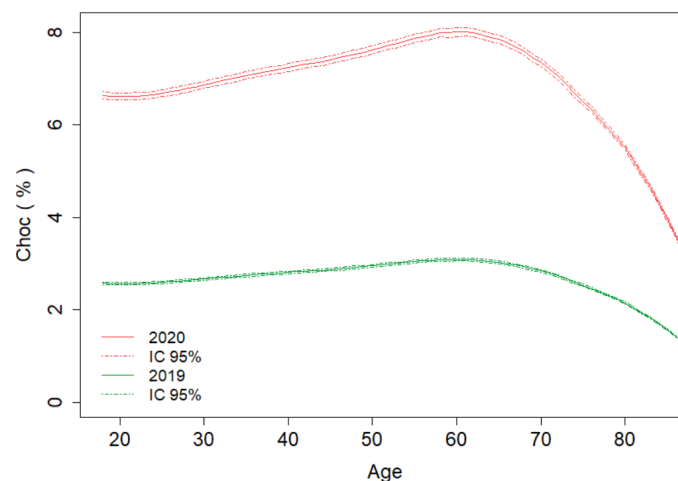


FIGURE 4.15 – Chocs de mortalité aux États-Unis avec et sans l'année 2020

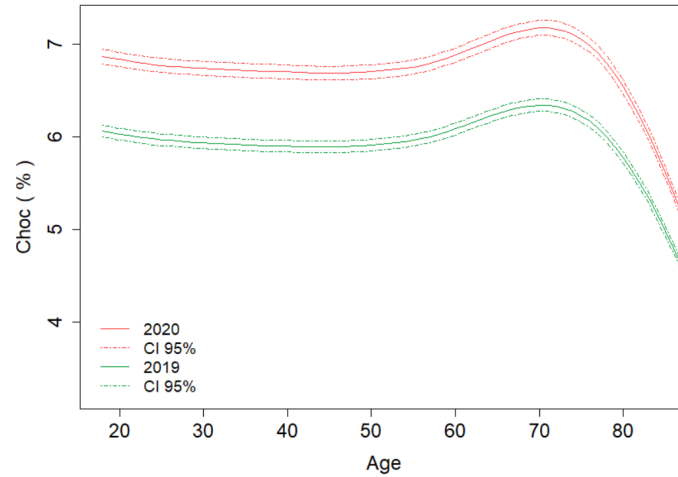


FIGURE 4.16 – Chocs de mortalité en France avec et sans l'année 2020

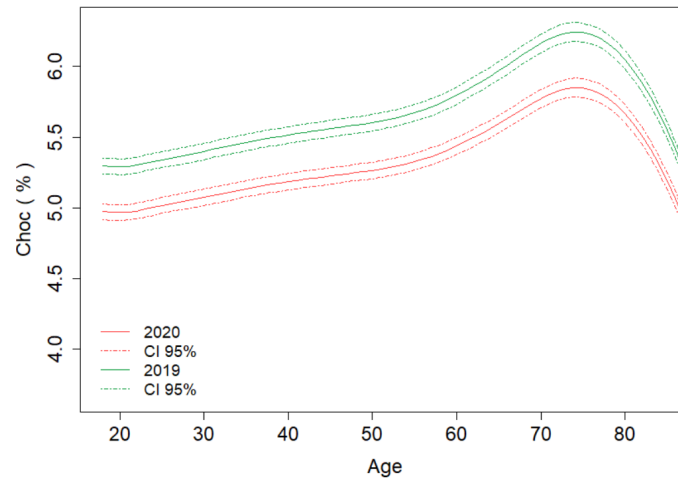


FIGURE 4.17 – Chocs de mortalité au Japon avec et sans l'année 2020

**L'inclusion de données de mortalité de 2020 entraîne une augmentation des chocs de mortalité.** En effet, les chocs de mortalité prenant en compte l'année 2020 sont supérieurs par rapport aux chocs de mortalité calculés en s'arrêtant à l'année 2019. Les chocs sont plus de deux fois supérieurs pour les Etats-Unis. En effet, ce pays a été fortement impacté par la crise et particulièrement la population de la classe d'âge 55-80 ans, d'où l'augmentation des chocs sur cette classe d'âge. L'augmentation des chocs de mortalité français est de l'ordre de 1% pour l'ensemble des âges alors que les chocs de mortalité du Japon sont très proches.

Pour le Japon, l'introduction des données de mortalité de 2020 sans retraitement dans le calibrage du modèle donne des valeurs de chocs de mortalité légèrement plus faibles qu'avec des données de calibrage s'arrêtant en 2019. Ce résultat est cohérent avec le fait que

la surmortalité due à la COVID-19 est très faible. C'est pourquoi le retraitement de l'année 2020 pour les pays dont le profil est similaire à celui du Japon ne semble pas nécessaire.

Les chocs calculés en prenant en compte les données de mortalité ajustées sont plus faibles, et plus conformes aux chocs calculés lorsque l'année 2019 est la dernière année de la période de calibrage. Ainsi, les résultats des méthodes appliquées aux États-Unis et à la France sont présentés figure 4.18.

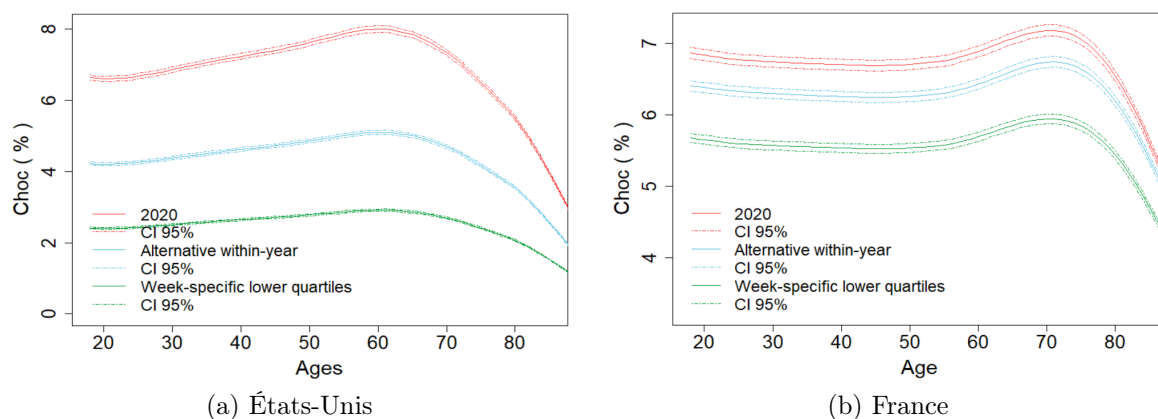


FIGURE 4.18 – Chocs de mortalité aux États-Unis (a) et en France (b) avec et sans correction de l'année 2020

Les chocs de mortalité varient en fonction des méthodes de retraitement appliquées. Par exemple aux États-Unis, la méthode *alternative within-year* fournit des chocs supérieurs d'environ 2% sur l'ensemble des classes d'âges par rapport aux chocs calculés avec les méthodes de la moyenne historique et *week-specific lower quartiles*. Les résultats de chocs sont plus proches pour la France. **Dans l'ensemble, ces méthodes d'ajustement conduisent à des chocs de mortalité plus proches de ceux obtenus en excluant l'année 2020 de la période de calibrage du modèle de Lee-Carter.** La moyenne des chocs par tranche d'âge pour les États-Unis et la France est présentée dans les tableaux 4.11 et 4.12.

Dernière année de la période de calibrage	Méthode de correction	18-35	36-55	56-80	80-94
2019	Sans correction	2,62	2,88	2,80	1,30
2020	Sans correction	6,90	7,59	7,39	3,36
2020	Moyenne historique	2,64	2,90	2,83	1,34
2020	Alternative within-year	4,31	4,75	4,63	2,13
2020	Week-specific lower quartiles	2,47	2,72	2,66	1,28

TABLE 4.11 – Moyenne des chocs de mortalité en % par tranche d'âge aux États-Unis

Dernière année de la période de calibrage	Méthode de correction	18-35	36-55	56-80	80-94
2019	Sans correction	5,97	5,90	6,16	4,57
2020	Sans correction	6,77	6,70	6,97	5,13
2020	Moyenne historique	5,94	5,88	6,13	4,55
2020	Alternative within-year	6,32	6,25	6,54	4,95
2020	Week-specific lower quartiles	5,94	5,53	5,77	4,32

TABLE 4.12 – Moyenne des chocs de mortalité en % par tranche d'âge en France

Des intervalles de confiance des chocs de mortalité ont été construits au niveau de confiance  $\alpha$  fixé à 95% sur les 1000 simulations. La construction de l'intervalle est classique, il s'agit d'un intervalle de confiance contenant la moyenne de l'échantillon au niveau de confiance 95%.

Les réalisations sont notées  $Y_n$  dont la loi est d'espérance  $m$  et d'écart-type  $\frac{\sigma}{\sqrt{n}}$ . On pose  $T_n = \frac{Y_n - m}{\frac{\sigma}{\sqrt{n}}}$  qui suit une loi normale centrée réduite. L'intervalle de confiance en est déduit avec  $u_{1-\frac{\alpha}{2}}$  le quantile de la loi normale centrée réduite de niveau  $1 - \frac{\alpha}{2}$  :

$$\left[ Y_n - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; Y_n + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] \quad (4.10)$$

La largeur moyenne de l'intervalle de confiance pour le modèle calibré à partir des données de 2019 est 0,27 pour la France et 0,11 pour les États-Unis tandis que pour le modèle incluant les données de 2020, la largeur augmente à 0,30 pour les États-Unis et 0,34 pour la France. Les données relatives à la pandémie engendrent une augmentation de la volatilité dans le modèle et donc une augmentation de la largeur des intervalles de confiance.

Les méthodes de retraitement permettent de revenir à des niveaux de chocs plus bas et de réduire la taille de l'intervalle de confiance.

- Le choc de mortalité moyen aux États-Unis avec la méthode *week-specific lower quartiles* est  $2,46 \pm 0,12$  tandis que le choc moyen avec le modèle calibré jusqu'en 2020 sans retraitement est  $5,90 \pm 0,30$  et  $2,50 \pm 0,11$  avec le modèle calibré jusqu'en 2019.
- Concernant la France, le choc de mortalité moyen avec la méthode *week-specific lower quartiles* est  $6,31 \pm 0,31$  alors que le choc est  $6,88 \pm 0,34$  avec le modèle calibré jusqu'à 2020 et  $6,33 \pm 0,27$  avec le modèle s'arrêtant en 2019.

## 4.5 Applications des méthodes à l'année 2021

L'objet de cette section est d'appliquer la stratégie d'ajustement des taux de mortalité à l'année 2021 durant laquelle la pandémie s'est poursuivie.

L'une des problématiques liées aux taux de mortalité est le décalage temporel entre la publication des données et l'année à laquelle les taux de mortalité sont enregistrés. Le portail HMD publie les données de mortalité avec une latence de deux ans environ. Ainsi, les données de 2021 sont disponibles pour un nombre limité de pays.

La pandémie de la COVID-19 a débuté en 2020 et s'est poursuivie en 2021. Plusieurs vagues sont survenues et, comme en 2020, les différents pays ont été touchés de manière inégale en 2021.

Les taux de mortalité des années 2019, 2020 et 2021 sont comparés aux États-Unis et au Japon (figure 4.19). La COVID-19 a très peu touché le Japon en 2020 et en 2021, on n'observe aucune évolution des taux de mortalité. A l'inverse, les États-Unis ont été plus durement touchés par la pandémie. Les taux de mortalité de 2021 pour la tranche d'âge 25 à 50 ans ont augmenté par rapport à 2020. De plus, à partir de 90 ans, les taux de mortalité de 2021 tendent vers ceux de 2019. Il peut s'agir d'un effet de rattrapage, les personnes fragiles de cette tranche d'âge sont déjà décédées de la COVID-19 en 2020.

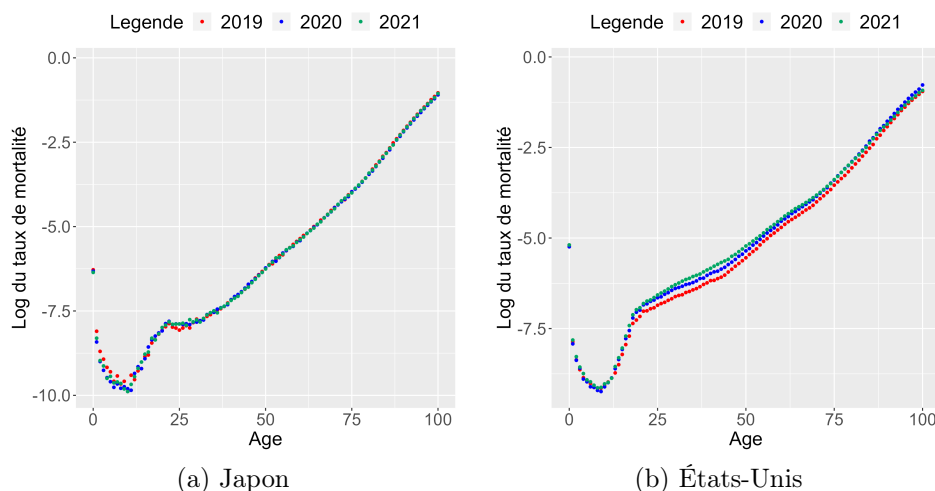


FIGURE 4.19 – Taux de mortalité de 2019 à 2021 au Japon et aux États-Unis

Les taux hebdomadaires STMF fournissent des informations plus précises sur l'évolution de la mortalité aux États-Unis et en France (figure 4.20).



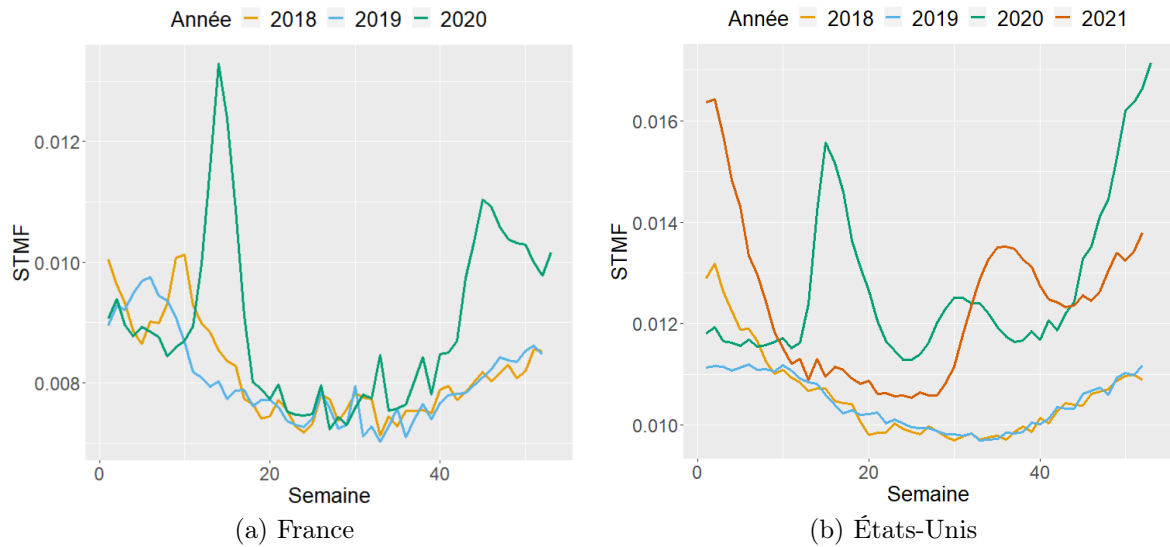


FIGURE 4.20 – Taux de mortalité hebdomadaires (STMF) de 2018 à 2021 en France et aux États-Unis

Les taux STMF des États-Unis sur les quatre dernières années montrent les vagues de COVID-19 successives. La mortalité en début d'année 2021 est importante, conséquence d'une vague ayant débuté en octobre 2020 et est marquée par une seconde vague en début d'année 2021. Entre les vagues de COVID-19, le niveau de mortalité est toujours supérieur au niveau des années précédentes.

En France, les taux de mortalité sont revenus à un niveau normal pendant une période de plus de quatre mois entre deux vagues en 2021. La capacité à revenir à un niveau de mortalité de référence a des conséquences sur l'application des méthodes de retraitement. Les résultats pour les États-Unis sont présentés dans la suite.

La pandémie a eu un impact sur les taux de mortalité en 2021, ce qui motive l'utilisation des méthodes de retraitement sur les données de 2021. En 2020, une méthode de retraitement est choisie, cette même méthode de retraitement est appliquée en 2021. Dans un souci de clarté, la même méthode est employée sur les deux années mais il est possible de combiner les méthodes entre elles dans la limite de leurs conditions d'utilisation.

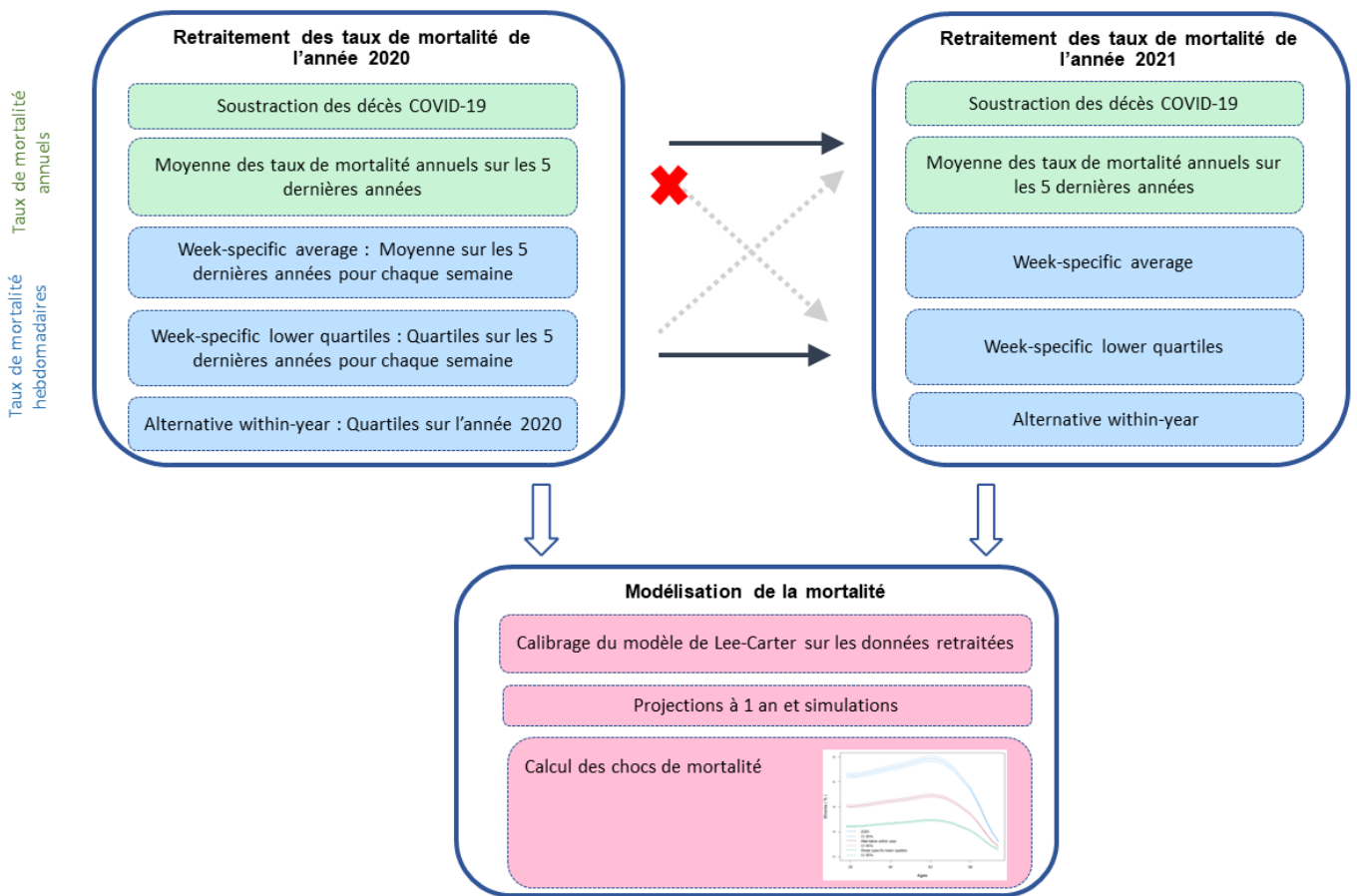


FIGURE 4.21 – Schéma de la méthodologie de retraitement et modélisation des taux de mortalité

Les méthodes de retraitement sont appliquées en 2020 et en 2021 sur les États-Unis. Les résultats pour une sélection de méthodes sont présentés. Le modèle de Lee-Carter est calibré sur la période de 1980 à 2021 sur les données préalablement retraitées.

Afin d'avoir un référentiel de comparaison, les taux de mortalité sans correction sont modélisés à l'âge de 70 ans aux États-Unis (figure 4.22). Les taux de mortalité de 2021 sont supérieurs à ceux de 2020. L'augmentation des taux génère une augmentation de la volatilité du modèle et par conséquent un intervalle de confiance à 95% large sur les projections des taux de mortalité.

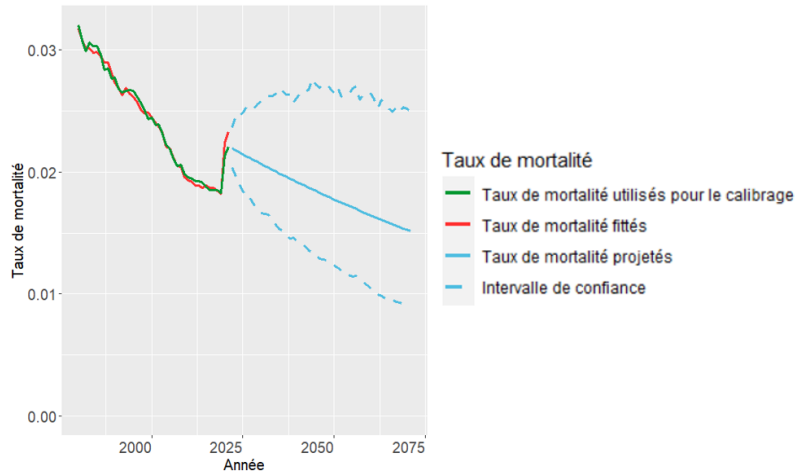


FIGURE 4.22 – Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter sans correction

Les méthodes directes qui sont le retraitement par la moyenne des taux annuels historiques et le retraitement par la soustraction des décès sont appliquées, les projections sont présentées en figure 4.23 et 4.24. Le retraitement par la moyenne permet de corriger les taux. En revanche, du fait que cette méthode ne prend pas en compte la décroissance des taux de mortalité, les taux corrigés sont supérieurs aux taux de 2019. Le calcul du taux de 2021 dépend du taux de 2020, ce qui propage l'erreur dans le temps. La soustraction des taux de mortalité ne corrige pas les effets indirects de la pandémie. C'est pourquoi les taux de 2020 et 2021 sont supérieurs aux taux de 2019. L'application de cette méthode est limitée dans les années futures.

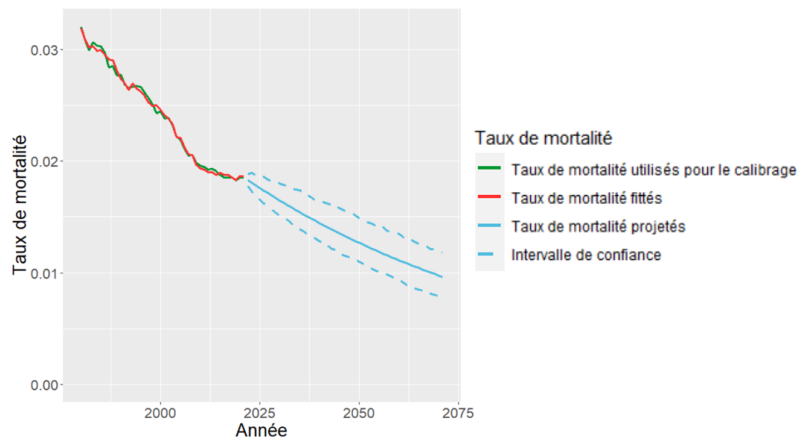


FIGURE 4.23 – Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction par la moyenne historique

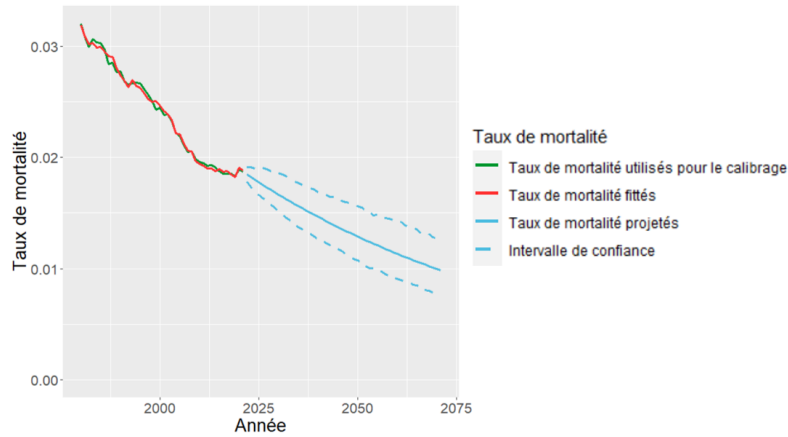


FIGURE 4.24 – Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction par soustraction des décès

Les méthodes indirectes sont également appliquées aux années 2020 et 2021. Les résultats obtenus avec les méthodes *week-specific averages* (figure 4.25), *week-specific lower quartiles* (figure 4.26) et la méthode *alternative within-year* (figure 4.27) sont présentés.

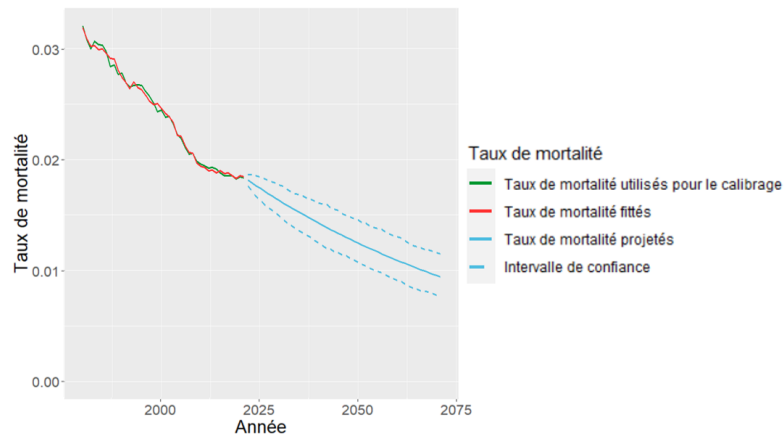


FIGURE 4.25 – Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction par la méthode *week-specific average*

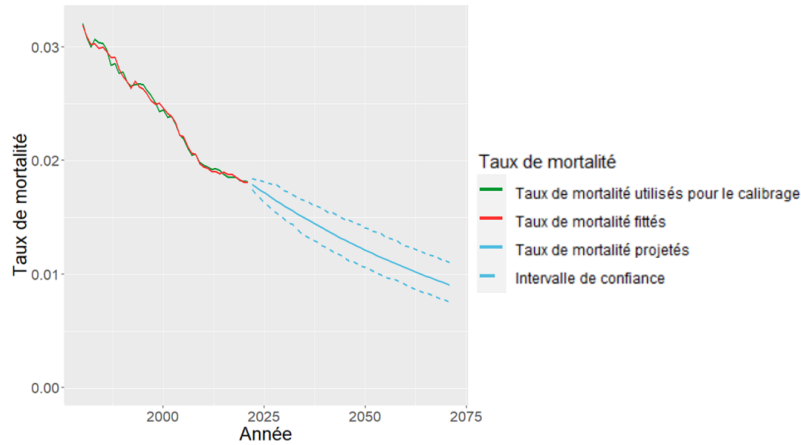


FIGURE 4.26 – Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction par la méthode *week-specific lower quartiles*

Avec les méthodes rétrospectives, le calcul du niveau de mortalité de référence en 2021 dépend des taux hebdomadaires de 2020. Elles permettent de corriger les taux de mortalité de manière à ce qu'ils paraissent dans la continuité des taux de mortalité historiques avant la pandémie. La méthode *week-specific lower quartiles* donne des taux de mortalité respectant la tendance décroissante alors qu'une erreur semble se cumuler avec la méthode *week-specific averages*. Ceci peut être dû au fait que les taux de mortalité hebdomadaires ne reviennent pas à leur niveau normal au cours de l'année.

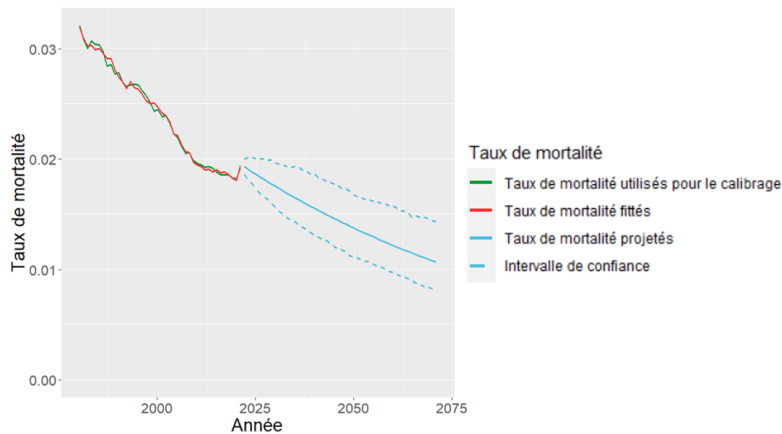


FIGURE 4.27 – Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction par la méthode *alternative within-year*

A l'inverse, le niveau de référence de la méthode *alternative within-year* est indépendant des niveaux de référence calculés les années précédentes. De cette manière, il n'y a pas de cumul d'erreur mais il y a une grande dépendance avec les taux de mortalité hebdomadaires de l'année en question.

Les espérances de vie en 2022 sont mesurées avec les différentes méthodes et sont répertoriées dans le tableau 4.13. Les méthodes de retraitement appliquées à 2020 et 2021 permettent d'obtenir des espérances de vie corrigées. En cohérence avec les graphiques de projections, la méthode *week-specific lower quartiles* est la méthode donnant des espérances de vie les plus élevées avec un écart de plus de 2,8 ans à l'âge de 47 ans par rapport à la modélisation sans retraitement.

	Sans retraite- ment	Soustraction des décès	<i>Week- specific ave- rages</i>	<i>Week-specific lower quar- tiles</i>	<i>Alternative within- year</i>
<b>47 ans</b>	33,15	33,18	35,72	35,96	34,77
<b>67 ans</b>	16,14	16,15	17,71	17,86	17,12
<b>87 ans</b>	3,94	3,95	4,12	4,14	4,05

TABLE 4.13 – Espérances de vie aux États-Unis en 2023

Les chocs de mortalité sont calculés à partir des modèles de Lee-Carter calibrés sur la période de 1980 à 2021 pour les États-Unis. Le choc moyen de mortalité est 6,98% lorsqu'il n'y a pas de retraitement (figure 4.28). Les méthodes de retraitement permettent de diminuer de 67% ces chocs de mortalité puisque les chocs avec la méthode *week-specific lower quartiles* sont de 2,32% (figure 4.29). Les chocs fournis par la méthode *alternative within-year* sont supérieurs de 2% en moyenne par rapport aux autres méthodes de retraitement sélectionnées. Ce qui témoigne des limites de la méthode sur l'année 2021. Les résultats par tranche d'âge de la méthode *week-specific lower quartiles* sont présentés dans le tableau 4.14.

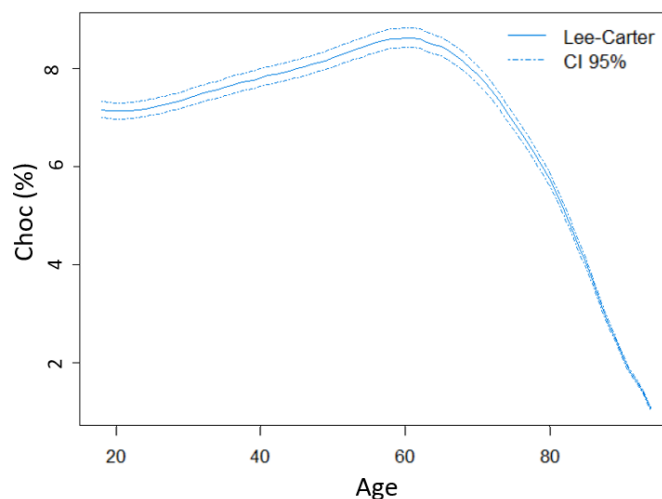


FIGURE 4.28 – Choc de mortalité aux États-Unis sans retraitement des années 2020 et 2021

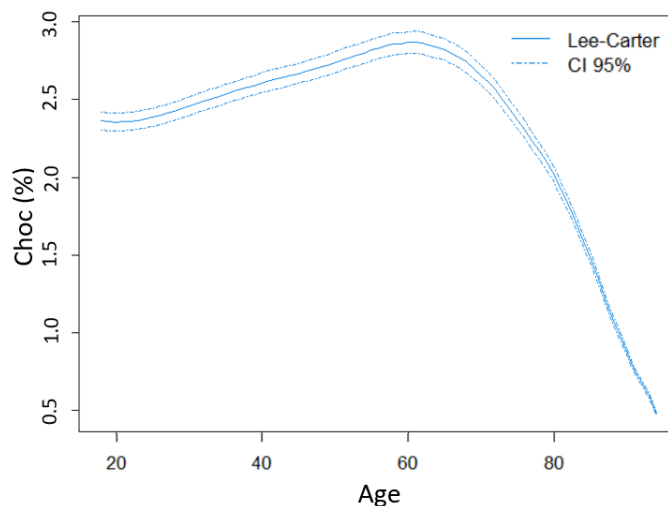


FIGURE 4.29 – Choc de mortalité aux États-Unis avec retraitement des années 2020 et 2021 par la méthode *week-specific lower quartiles*

	Sans retraitement	<i>Week-specific lower quartiles</i>
<b>18-35</b>	7,03	2,42
<b>35-55</b>	7,78	2,68
<b>55-80</b>	7,53	2,62
<b>80-110</b>	3,39	1,18

TABLE 4.14 – Moyenne des chocs de mortalité en % par tranche d'âge aux États-Unis

La largeur de l'intervalle de confiance du choc de mortalité est réduit à l'aide des méthodes de retraitement. Ainsi le choc moyen de mortalité est passé de  $6,98 \pm 0,32$  sans retraitement à  $2,32 \pm 0,11$  avec la méthode *Week-specific lower quartiles*, le graphique 4.29 représente les chocs de mortalité corrigés qui ont un niveau cohérent pour le calcul du SCR.

Il apparaît important de pouvoir retraiter des données de mortalité de l'année 2021 qui ont été largement impactées par la pandémie. La disponibilité des données permet l'application des méthodes de retraitement aux années 2020 et 2021 aux États-Unis. Les résultats varient en fonction des méthodes de retraitement mais permettent toutes d'obtenir un niveau de choc de mortalité proche de celui déterminé lorsque le calibrage des modèles s'arrête en 2019.

## 4.6 Modèle CMI

D'autres approches existent pour intégrer les données COVID-19 dans la modélisation de la mortalité. Le modèle CMI, modélisant la mortalité, a été adapté pour tenir compte de la pandémie. L'intégration des données COVID-19 repose sur une modification de la

modélisation contrairement aux méthodes appliquées précédemment, qui se basent sur le retraitement des données d'entrées. Le modèle de projection de mortalité, *Continuous Mortality Investigation*, est grandement utilisé sur le marché de l'assurance pour projeter l'amélioration de la longévité [CMI (2023)].

Le modèle CMI est d'un modèle déterministe basé sur l'hypothèse selon laquelle les taux actuels d'amélioration de la mortalité convergent vers un taux unique à un horizon long. Le modèle est mis à jour régulièrement en fonction des nouvelles données. Historiquement, pour une année donnée, le modèle est calibré jusqu'au 31 décembre de l'année en utilisant les données des 40 dernières années.

Les années 2020 et 2021 constituent une anomalie du fait de la pandémie, c'est pourquoi le modèle CMI a aussi été adapté pour prendre en compte cette hausse de la mortalité. Un nouveau paramètre a été introduit dans le modèle. L'utilisateur peut attribuer un poids aux différentes années individuellement dans le calibrage du modèle. Dans le modèle de 2022, le poids des années 2020 et 2021 est 0% tandis qu'il est de 100% sur les années antérieures de calibrage. Les effets de la COVID-19 sont de plus en plus faibles dans le temps, c'est pourquoi le CMI suggère une pondération de 25% pour les données de 2022. Dans les années à venir, le poids des données de mortalité augmentera pour atteindre 100% en 2025 selon les prévisions du CMI en fonction d'une indication plus claire de la tendance de la mortalité. Les poids sont révisés chaque année.

Puisque le modèle est mis à jour régulièrement, plusieurs stratégies de calibrage peuvent être envisagées en utilisant les versions précédentes du modèle :

- Le modèle CMI\_2019 : il projette la mortalité sous l'hypothèse que la pandémie est un événement exceptionnel et temporaire qui ne devrait pas avoir d'impact sur l'amélioration de la mortalité future.
- Le modèle CMI\_2022 avec pondération : avec cette méthode, l'impact sur l'espérance de vie est faible, les années 2020 et 2021 ne sont pas considérées comme pertinentes dans la projection de la mortalité. Selon cette méthode, les années pandémiques pourraient influencer les taux de mortalité futurs en fonction de la pondération fixée. Il est à noter que les pondérations choisies impactent grandement les projections. Cet effet est accentué lorsque plusieurs années consécutives sont retraitées, puisque qu'il y a une interaction entre les années.

## 4.7 Autres impacts de la COVID-19 sur la mortalité

Les effets de la COVID-19 à long terme sont encore méconnus. Une des conséquences de la pandémie sur la mortalité à long terme intéresse particulièrement les chercheurs, il s'agit de l'impact de la COVID-19 sur la santé mentale. On distingue deux types de santé mentale sur lesquels la pandémie a des conséquences différentes :

- La santé mentale au sens courant du terme incluant par exemple la dépression et l'anxiété.



- Les maladies neurodégénératives et particulièrement la maladie de Parkinson et la maladie d’Alzheimer.

Selon l’article de Santé publique France (2022), une dégradation de la santé mentale des français a été observée en conséquence de la pandémie. En effet, les différents confinements et la crise économique ont déclenché ou accentué des troubles psychologiques. Or, un état psychologique en dégradation peut impacter négativement la mortalité.

De plus, le rapport de Boumezoued et al. (2019) pour la SOA (*Society of Actuaries*) présente les résultats de la modélisation de la mortalité par cause de décès. Il en ressort que les projections de la mortalité due aux maladies neurodégénératives sont en hausse particulièrement chez les femmes. Cette tendance risque d’être accentuée par la pandémie. Plusieurs études telles que celle de Huang et al. (2023) s’intéressent à la relation entre la pandémie et les maladies neurodégénératives. La COVID-19 augmenterait la progression de la maladie de Parkinson et une augmentation de la mortalité due à cette maladie a été constatée depuis la pandémie. Les symptômes de la maladie d’Alzheimer seraient également accentués.

Les personnes présentant des comorbidités sont plus fragiles face au virus. Il est possible qu’un effet de rattrapage soit observé avec la diminution de la mortalité des personnes qui présentent ces comorbidités.

Ainsi, si on prend en compte ces éléments dans la projection de la mortalité, les projections ne seraient pas aussi optimistes et l’objectif ne serait pas de retraiter complètement la COVID-19 en 2020 et 2021 pour prendre en compte tous les effets sur le long-terme. Les modèles de mortalité classiques ne seraient plus adaptés. Le traitement des effets long terme de la COVID-19 évoqués dans ce paragraphe n’est pas traité dans ce mémoire.

### Conclusion

Nous avons montré que l'introduction des taux de mortalité de l'année 2020 dans les données de calibrage du modèle de mortalité a des conséquences sur la modélisation. Pour éviter un double comptage avec un risque de pandémie mesuré séparément, ainsi que pour dériver des tendances de mortalité à long terme sans instabilité à partir d'une expérience unique de pandémie, des méthodes de retraitement des données impactées par la pandémie sont nécessaires.

Des méthodes basées sur les taux de mortalité hebdomadaires permettent d'exclure les effets indirects de la COVID-19 par rapport à la méthode simple qui consiste à soustraire le nombre de décès COVID-19 enregistrés. Dans le cas de la correction des données de mortalité de 2020 pour le calibrage du modèle pour un pays où la surmortalité COVID-19 est significative, les projections de mortalité et les chocs de mortalité qui en résultent sont :

- Nettement inférieurs à ceux obtenus en calibrant le modèle sur des données de mortalité non ajustées.
- Plus conformes à ce qui aurait été obtenu en calibrant les modèles à l'aide des données de mortalité jusqu'en 2019.

La pandémie s'est poursuivie en 2021 et il est nécessaire de corriger cette année également. Les mêmes méthodes de retraitement peuvent être considérées.

L'application des méthodes de retraitement aux années 2020 et 2021 a aussi mis en évidence leurs limites. Les méthodes *alternative within-year* et *week-specific lower quartiles* sont les méthodes qui permettent d'obtenir des projections de la mortalité sans distorsion et des chocs de mortalité qui ne sont pas impactés par la mortalité pandémique. Ces méthodes peuvent être confrontées à d'autres approches. Par exemple, il est possible d'introduire des poids sur les données de calibrage de la mortalité comme c'est le cas pour la dernière version du modèle CMI.

# Chapitre 5

## Autres facteurs impactant la mortalité

La mortalité est impactée par différents facteurs de risque incluant le risque pandémique. Ce chapitre présente deux risques dont les conséquences sur la mortalité ont été mises en évidence ces dernières années : le risque climatique et les technologies portables.

Le risque climatique est beaucoup évoqué dans le contexte du changement climatique avec l'augmentation de la fréquence et de l'intensité d'événements extrêmes. Des modèles sont envisagés pour prendre en compte la composante climatique dans la modélisation de la mortalité.

A l'inverse du risque climatique, les technologies portables, qui sont des dispositifs collectant des données sur le mode de vie ou la santé, auraient potentiellement un impact à la baisse sur la mortalité. Bien qu'il soit difficile de mesurer quantitativement l'impact de ces technologies récentes sur l'espérance de vie, les acteurs du marché de l'assurance montrent un intérêt croissant pour l'exploitation de ce type de technologies.

### 5.1 Modélisation de la mortalité en tenant compte du risque climatique

#### 5.1.1 Motivations et données

Le rapport du GIEC paru en 2021 est alarmant quant au changement climatique et ses conséquences. Les assureurs sont encouragés à prendre en compte les risques liés au changement climatique en intégrant, par exemple, le risque climatique dans leur modèle de mortalité. Deux projets Milliman ont été menés en 2021 et 2022 sur l'intégration du risque de climatique dans un modèle de mortalité. L'objectif est de confronter les deux méthodes développées, de les consolider afin de développer une méthode applicable à tout risque climatique et zone géographique. Le modèle de Lee-Carter prenant en compte le risque climatique a été testé sur le cas des vagues de chaleur en France et aux États-Unis. Les résultats ont été analysés et comparés entre les deux pays.

Tout d'abord, Germain (2021) a développé dans son mémoire un modèle de Lee-Carter prenant en compte le risque climatique sur les données françaises. De plus, un groupe

d'étudiants de l'ENSAE<sup>1</sup>, sous la direction de Eve Titon, ont appliqué un modèle similaire à celui défini dans le mémoire de V. Germain sur les différents états des États-Unis.

Le risque climatique considéré ici est le risque engendré par les vagues de chaleur. V. Germain a appliqué le modèle de Lee-Carter climatique à la France à l'échelle nationale. L'objectif de cette section est de vérifier la pertinence du modèle sur les États-Unis à l'échelle de chaque état. Pour la construction d'un tel modèle, des données spécifiques sont requises :

- Les données de mortalité globale, les données HMD sont la source de référence.
- Les données de mortalité pour une cause spécifique, la base de données *Global Health Data* (GHD) est publiée par l'*Institute for Health Metrics and Evaluation* (IHME). La base de données GHD est constituée du nombre de décès par cause. La cause considérée ici est celle des vagues de chaleur.
- Les données climatiques et météorologiques, la source varie en fonction de la zone géographique étudiée. La *National Oceanic and Atmospheric Administration* (NOAA) fournit une table de variables climatiques appelée *Global Historical Climatology Network Daily database* (GHCND) aux États-Unis. La base de données contient des données climatiques journalières de plus de 10 000 stations météorologiques des États-Unis avec, par exemple, des variables portant sur la température, la pluviométrie.

### 5.1.2 Méthodologie

Les méthodologies de V. Germain et des étudiants de l'ENSAE sont comparées afin de déterminer le modèle le plus performant. Le modèle retenu permet une meilleure explication de la mortalité en fonction du risque climatique à la fois sur la France et sur les États-Unis à différentes échelles. La méthode sélectionnée et ses résultats sont présentés.

#### 5.1.2.1 Définition du modèle de Lee-Carter climatique

Le modèle développé est un modèle dérivé de Lee-Carter tenant compte du risque climatique. L'objectif du modèle est de pouvoir capter la mortalité due à une cause climatique.

$$\ln(\mu_{x,t}) = \alpha_x + \beta_x \kappa_t + \delta_x C_t \quad (5.1)$$

Le terme  $\delta_x C_t$  est ajouté à la structure du modèle de Lee-Carter classique :

- $C_t$  : indicateur climatique construit à partir de variables climatiques sélectionnées.
- $\delta_x^c$  : paramètre de calibrage qui s'apparente au paramètre  $\beta_x$  du modèle de Lee-Carter classique.

L'indicateur climatique est construit pour un risque climatique afin d'expliquer la mortalité liée à cette cause climatique. Le terme  $\beta_x \kappa_t$  explique la mortalité globale avec la cause climatique exclue.

1. Eurydice Maucuer, Noureddine Oulid Azouz, Benjamin Patra, Margaux Regnault, projet d'approfondissement

### 5.1.2.2 Construction de l'indicateur climatique

Les variables climatiques permettent de construire l'indicateur climatique  $C_t$ . Le principe est d'expliquer au mieux les taux de mortalité climatiques à l'aide des variables climatiques. La méthode utilisée pour construire l'indicateur climatique est une régression linéaire des taux de mortalité climatique en fonction des variables climatiques. La sélection des variables pertinentes est importante dans la construction de l'indicateur à l'aide du critère de l'AIC. Un nombre restreint de variables est sélectionné pour une meilleure interprétabilité du modèle.

L'indicateur climatique  $C_t$  est défini tel que :

$$C_t = a + b^T X_t \quad (5.2)$$

avec

- $a, b$  : les paramètres de la régression linéaire
- $X_t$  : la matrice des variables climatiques sélectionnées, où  $t$  désigne l'année.

Pour un risque climatique donné, les variables sélectionnées dépendent du territoire étudié.

### 5.1.2.3 Effet de rattrapage

L'effet de rattrapage se manifeste à la suite d'un événement fragilisant une population. Une surmortalité est observée dans la période de l'événement mais engendre également une baisse de la mortalité dans la période qui suit. Ce phénomène peut être observé avec un décalage dans le temps par rapport à la survenance de l'événement. Les gripes saisonnières, pics de pollution ou encore vagues de chaleur comme la canicule de 2003 en France peuvent être à l'origine d'un effet de rattrapage. Ce type d'événements peut aussi accentuer la vulnérabilité avec la propagation ou l'aggravation de maladies, il s'agit d'effets indirects.

Pour prendre en compte l'effet de rattrapage, l'indicateur climatique est corrigé de manière à ce qu'il dépende des variables climatiques de l'année en cours et de l'année précédente. Une baisse de la mortalité est observée l'année suivant une vague de chaleur mais il ne s'agit pas d'une amélioration de la mortalité. Deux paramètres sont inclus dans le modèle :

- $\alpha$  : impact du climat de l'année  $t$ .
- $\gamma$  : effet de rattrapage pour l'année  $t - 1$

$$CI_t = a + \alpha(b^T X_t) + \gamma(b^T X_{t-1}) \quad (5.3)$$

### 5.1.2.4 Calibrage du modèle

Le modèle de Lee-Carter classique est construit habituellement âge par âge. Le risque climatique considéré impacte de manière différente la population en fonction de l'âge. C'est

pourquoi les données sont regroupées en trois classes d'âge afin d'expliquer correctement l'impact de la mortalité climatique par âge tout en ayant suffisamment de décès par classe d'âge pour pouvoir appliquer cette méthodologie. On définit  $x_0 = 0$ ,  $x_1$ ,  $x_2$ , et  $x_3 = 111$ , avec les classes d'âge définies telles que  $\forall i \in [0, 2]$ ,  $c_i = [x_i, x_{i+1} - 1]$ . Le modèle est d'abord construit par classe d'âge avant de passer à l'échelle âge par âge.

Les étapes du calibrage du modèle sont décrites synthétiquement dans le tableau 5.1.

Etape du calibrage	Description de l'étape
Calibrage de $\alpha_{c_i}$	Modèle de Lee-Carter classique sur les taux de mortalité HMD par classe d'âge
Calibrage de $\delta_{c_i}$	$\operatorname{argmin}_{\delta_{c_i}} \sum_{i=0}^2 (\ln(\mu_{c_i,t}) - \alpha_{c_i} - \delta_{c_i}^c C_t)^2$ <ul style="list-style-type: none"> <li>— <math>\omega_{i,t}</math> : poids permettant de mettre en exergue les augmentations de la mortalité climatique.</li> <li>— <math>\mu_{c_i,t}</math> : taux de mortalité.</li> </ul>
Intégration de l'effet de rattrapage	$\operatorname{argmin}_{\alpha, \gamma} \sum_{i=0}^2 (\ln(\mu_{c_i,t}) - \alpha_{c_i} - \delta_{c_i}^c C I_t)^2$ <ul style="list-style-type: none"> <li>— <math>\mu_{c_i,t}</math> : taux de mortalité.</li> <li>— <math>C I_t = a + \alpha(b^T X_t) + \gamma(b^T X_{t-1})</math></li> </ul> L'algorithme de Brent est utilisé pour la minimisation.
Calibrage du modèle final âge par âge	Les paramètres construits par tranche d'âge $\delta_{c_i}$ sont convertis en paramètres âge par âge. Les coefficients $\alpha_x$ , $\beta_x^o$ et $\kappa_t^o$ sont estimés par un modèle de Lee-Carter construit sur les résidus finaux dont l'exponentielle est homogène à des taux de mortalité : $R_{x,t} = \ln(\mu_{x,t}) - \delta_x^c C I_t$

TABLE 5.1 – Étapes de calibrage du modèle de Lee-Carter climatique

### 5.1.3 Résultats sur les États-Unis

Le modèle de Lee-Carter prenant en compte le risque climatique est appliqué sur les États-Unis à l'échelle des états. Le climat des États-Unis n'est pas homogène sur tout le pays. La zone géographique qui s'étend sur plusieurs états au sud-est des États-Unis est celle sujette aux vagues de chaleur. C'est pourquoi les résultats du modèle de Lee-Carter incluant le risque de vagues de chaleur ne donne pas des résultats homogènes sur tout le pays.

L'indicateur climatique est construit pour chaque état. La carte 5.1 montre les scores

$R^2$  de la régression de l'indicateur climatique par état. Les scores dépassant les 80% sont ceux des états au sud-est entre le Texas et la Floride en remontant vers le nord. De plus, même si le score  $R^2$  paraît satisfaisant, le modèle ne fournit pas des résultats cohérents pour les états en dehors de la zone affectée par le risque climatique étudié.

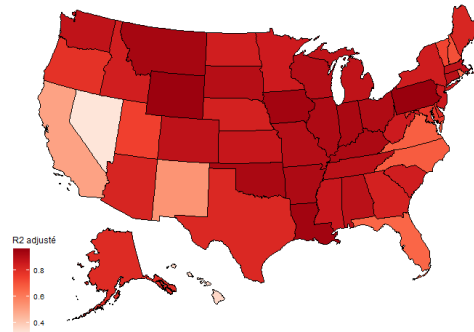


FIGURE 5.1 – Score  $R^2$  de la régression de l'indicateur climatique par état

Les variables climatiques considérées sont relatives à la température et sont listées en annexe dans le tableau E.1. Une sélection de variables est réalisée à l'échelle de chaque état pour construire l'indicateur climatique, 8 variables sont retenues. Certaines variables sont récurrentes, comme le montre le graphique en annexe E.

Afin de présenter les résultats du modèle, un état sujet aux vagues de chaleur est sélectionné, l'**Oklahoma**. La régression du taux de mortalité climatique en fonction des variables climatiques donne un  $R^2 \approx 92\%$ , l'indicateur climatique s'ajuste correctement aux taux de mortalité climatique (figure 5.2). Les variables climatiques sélectionnées sont détaillées en annexe E.

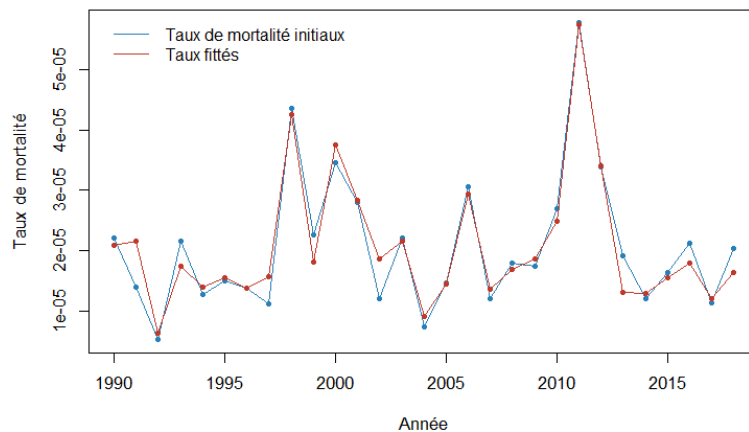


FIGURE 5.2 – Indicateur climatique et taux de mortalité climatique en Oklahoma

Les classes d'âge utilisées pour le calibrage du modèle sont  $[0, 49]$ ,  $[50, 69]$ ,  $[70, 110]$ . Les taux de mortalité climatiques par classe d'âge sont présentés figure 5.3, montrant le risque croissant avec l'âge de la mortalité climatique.

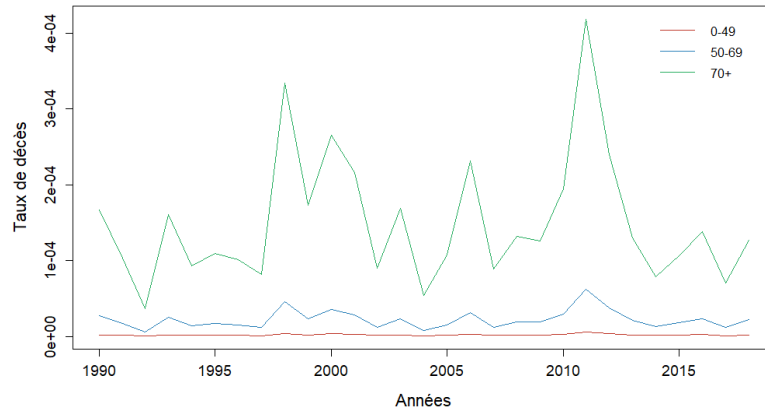


FIGURE 5.3 – Mortalité climatique par classe d'âge en Oklahoma

On compare les taux de mortalité enregistrés dans la base HMD, les taux de mortalité modélisés par un modèle de Lee-Carter classique et ceux modélisés à l'aide du modèle de Lee-Carter climatique. Par exemple, pour l'âge de 75 ans (figure 5.4), les résultats des deux modèles de Lee-Carter sont proches, le coefficient  $R^2$  est de 73,94% pour le modèle de Lee-Carter climatique contre 73,77% pour le modèle de Lee-Carter classique.

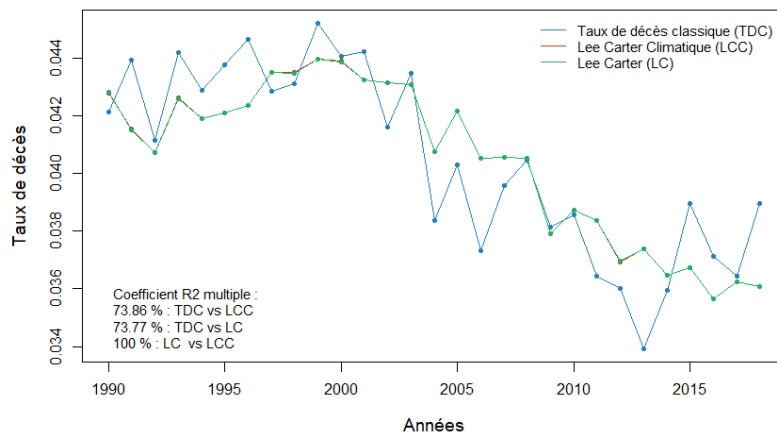


FIGURE 5.4 – Comparaison des modélisations à l'âge de 75 ans en Oklahoma

Les coefficients  $R^2$  entre les taux de mortalité originaux et les modèles Lee-Carter sont représentés figure 5.5. L'ajustement des deux modèles est similaire, mais pour les âges élevés, le modèle climatique est meilleur que le modèle original de Lee-Carter. Pour 60% des âges, le modèle climatique est meilleur que le modèle classique en termes de  $R^2$ . En considérant une marge de tolérance de 5%, le modèle climatique est plus performant



dans 75% des cas. Le tableau 5.2 met en évidence l'amélioration sur une tranche d'âge particulière, les 60-80 ans, pour le cas d'étude de l'Oklahoma.

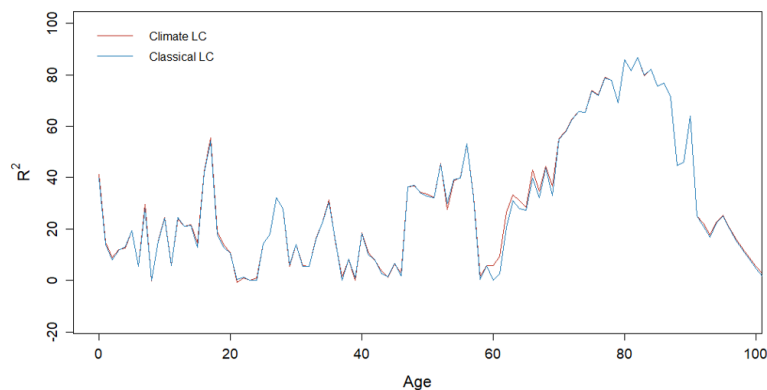


FIGURE 5.5 – Coefficient d'ajustement  $R^2$  des deux modèles sur les données d'Oklahoma

	LC classique	LC climatique
<b>0-24</b>	16,13 %	16,63 %
<b>25-59</b>	19,91 %	19,50 %
<b>60-79</b>	44,87 %	46,33 %
<b>80+</b>	30,34 %	30,90 %

TABLE 5.2 – Scores AIC et BIC des modèles de Lee-Carter classique et climatique

Pour comparer les performances des modèles, des critères d'information AIC et BIC définis en annexe A sont utilisés. Les résultats des critères AIC et BIC des modèles de Lee-Carter classique et climatique appliqués à l'Oklahoma sont présentés dans le tableau 5.3. Ainsi, l'intégration du risque climatique permet une amélioration peu significative des critères d'information.

	LC classique	LC climatique
<b>AIC</b>	34792,37	34715,49
<b>BIC</b>	36305,5	36228,62

TABLE 5.3 – Scores AIC et BIC des modèles de Lee-Carter classique et climatique

#### 5.1.4 Comparaison avec la France

Dans le mémoire de Germain (2021), le modèle a été développé sur les données françaises. Il est intéressant de comparer les résultats du modèle sur l'Oklahoma et la France. Une distinction majeure est l'échelle à laquelle le modèle a été appliqué, à l'échelle des états pour les États-Unis tandis que l'ensemble du territoire national a été considéré pour la France, ce qui soulève deux limites du modèle :

- L'hétérogénéité du climat sur le territoire : en considérant les données du territoire national français, on suggère que le climat est homogène sur tout le pays au même titre que le risque.
- La taille de la population : seuls quatre états des États-Unis ont une population supérieure à douze millions d'habitants. Les modèles de mortalité ne sont pas stables lorsqu'ils sont appliqués à des modèles de mortalité sur des populations dont la taille n'est pas suffisante.

On compare les performances des modèles entre la France et l'état d'Oklahoma. Les modèles de Lee-Carter climatiques s'ajustent beaucoup mieux au taux de mortalité français que pour l'Oklahoma. Le coefficient  $R^2$  du modèle de Lee-Carter climatique appliqué à la France est 99,75% contre 99,73% pour le modèle de Lee-Carter classique. Ces résultats sont largement supérieurs à ceux des États-Unis. Finalement, le modèle de Lee-Carter climatique améliore de manière similaire le modèle de Lee-Carter classique pour la France et l'état d'Oklahoma. Cependant, la modélisation de la mortalité sur cet état des États-Unis est moins performante que pour la France.

L'intégration du risque de vague de chaleur dans le modèle de Lee-Carter permet d'améliorer la modélisation de la mortalité avec une augmentation du coefficient  $R^2$  pour les personnes âgées. L'amélioration est relativement peu significative. Cependant, l'intégration du risque climatique dans le modèle de Lee-Carter permet de modéliser la mortalité aussi bien qu'avec un modèle de Lee-Carter classique tout en expliquant la part de mortalité attribuable au risque climatique. Cette modélisation est intéressante dans le cadre de la projection de la mortalité climatique en fonction de scénarios climatiques.

Les risques climatiques sont multiples et impactent spécifiquement des zones géographiques. L'étude de ce modèle renforce la pertinence de l'intégration du risque climatique dans les modèles de mortalité.

## 5.2 Le risque des technologies portables

Les technologies portables, appelées *wearables*, font partie intégrante de notre mode de vie. Elles font l'objet d'études pour déterminer leurs impacts sur la mortalité et l'espérance de vie. L'article de Soulas et al. (2023) présente un état de l'art des connaissances sur les *wearables*.

Une technologie portable est un « dispositif informatique doté d'une application qui collecte et traite les données ». Ces technologies sont développées au sein d'un secteur d'activité pour un besoin précis et remplissent différents rôles. Il existe quatre objectifs principaux : la surveillance et l'assistance destinées aux personnes malades, la prévention et la communication pour les personnes en bonne santé. Cette démarche s'inscrit dans la volonté de prendre conscience de son mode de vie et de l'améliorer. Ces technologies pourraient s'inscrire dans la progression de l'espérance de vie puisqu'elles sont employées dans le secteur de la santé et par le système de soins. Pour l'affirmer, il faut pouvoir

quantifier son impact sur l'amélioration de la mortalité.

Les données sont au centre de ces technologies. Elles en font la richesse et l'intérêt mais soulèvent aussi des questionnements autour de l'usage de ces technologies. Les données collectées par les *wearables* sont plus riches que les données relatives à la santé collectées dans le passé. De nouvelles variables peuvent être mesurées, et les données sont collectées plus fréquemment et en temps réel. Cependant, des problématiques se posent sur l'usage de ces données. D'une part, le cycle des données des technologies portables est coûteux. D'autre part, l'utilisation des données est soumise à la législation.

Pour mesurer l'impact sur l'espérance de vie, différentes méthodes sont envisagées. Une approche possible, définie par Boumezoued et al. (2019) est la modélisation des taux de mortalité par cause de décès. Ces taux sont projetés avec un modèle de Lee-Carter en prenant en compte les interactions entre les causes de décès. Les aléas des modèles de Lee-Carter sont corrélés entre eux. Ainsi, cette approche nécessite de données de santé spécifiques aux causes de décès considérées que les *wearables* peuvent fournir.

La réduction des risques intéresse particulièrement les assureurs. Ces données fournies par les technologies portables peuvent être intégrées dans les contrats d'assurance :

- Tarification des produits d'assurance.
- Ajustement des caractéristiques des contrats d'assurance en fonction des données récoltées tel que des programmes de bonus/malus en fonction de la réalisation d'un objectif.

L'intégration de ces données dans les contrats d'assurance permet à des clients à haut niveau de risque d'accéder à des assurances en prenant en compte des données relatives à leur santé. Cette dimension des contrats d'assurance santé et d'assurance vie est en cours de développement et vise à encourager les assurés à améliorer leur mode de vie. Par exemple, une assurance santé propose un bonus si l'activité physique de l'assuré mesurée par un tracker d'activité est suffisante.

Les *wearables* s'inscrivent dans la dynamique d'amélioration de la mortalité. Ces résultats encouragent la poursuite de recherches avec le déploiement de ces technologies et l'augmentation des sources de données exploitables.

### Conclusion

Deux facteurs aux impacts opposés sur la mortalité ont été décrits dans cette section.

- Le risque climatique est pris en compte dans la modélisation de la mortalité via un modèle de Lee-Carter. Les réflexions sur ce modèle s'inscrivent dans la recherche de modèles permettant d'inclure ce risque dans l'explication de la mortalité. Le risque climatique impactant également les risque financiers, des études conjointes sont envisageables ainsi que la prise en compte de ce risque dans les GSE.
- Les *wearables* s'inscrivent dans l'amélioration de l'espérance de vie au cours du temps. Bien que leur modélisation soit complexe, il reste intéressant de pouvoir analyser leur impact sur la mortalité.

La mortalité est soumise à divers risques qui interagissent entre eux. La modélisation de ces risques reste un sujet complexe.

Troisième partie

Utilisation de la mortalité  
stochastique dans un modèle ALM

# Chapitre 6

## Application dans un modèle ALM

Les modèles de gestion actif-passif, *Asset Liability management* (ALM), prennent en entrée une loi de mortalité. L'objectif de ce chapitre est de discuter l'intérêt de l'utilisation d'une table de mortalité stochastique par rapport à une table de mortalité déterministe dans un modèle ALM.

### 6.1 Présentation du modèle ALM

Les modèles ALM sont employés pour modéliser les interactions entre les actifs et les passifs du bilan d'une compagnie d'assurance ou de réassurance. La modélisation ALM doit se conformer à la réglementation, elle doit aussi prendre en compte les directives de gestion spécifiques de l'entreprise et les caractéristiques de son activité. Un modèle ALM peut être utilisé pour définir une stratégie d'allocation d'actifs en fonction du portefeuille de la compagnie. Il peut aussi servir pour projeter les flux futurs de la compagnie afin de déterminer la meilleure estimation du passif, *Best Estimate* (BE). Les modèles ALM sont de complexité variable en fonction des besoins.

Peu importe l'objectif, la modélisation se base sur les flux financiers de la compagnie. Les principaux flux sont des échanges avec les assurés pour le paiement des cotisations et des sinistres et avec les investisseurs auxquels l'assureur transfère des liquidités. La plupart des flux se font à partir de comptes bancaires. L'assureur réalise des transferts de liquidités entre ses comptes bancaires et ses portefeuilles d'investissement. Pour déterminer les montants que l'assureur va investir au cours d'une période définie, l'assureur doit estimer deux éléments :

- Les sorties de trésorerie qu'il va avoir sur la période à venir afin de conserver les liquidités nécessaires et ne pas vendre d'actifs à un moment qui ne serait pas opportun.
- Les ressources qu'il va recevoir au cours de cette même période telles que les cotisations, les intérêts d'obligations, les dividendes d'actions, les loyers d'immeubles.

L'organisation de ces flux est représentée dans le schéma figure 6.1. Pour ce faire, l'assureur transmet le montant à investir et les consignes d'investissements en fonction de son profil de risque. De cette manière, l'entité responsable de la gestion de l'investissement gère le portefeuille avec l'achat et la vente de titres selon les consignes transmises. Ainsi, l'un des enjeux des modèles ALM est de pouvoir estimer le plus justement ces flux afin de

déterminer une stratégie d'investissement optimale.

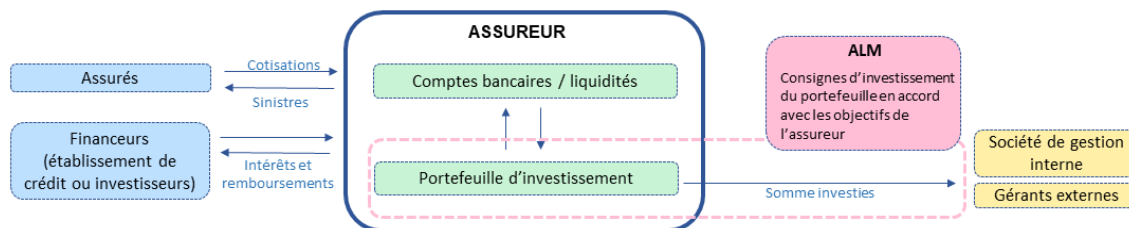


FIGURE 6.1 – Schéma de l'organisation des flux financiers au sein d'une compagnie

Les caractéristiques d'un modèle ALM sont définies en fonction de la stratégie d'allocation d'actifs, du passif de la compagnie tel que les caractéristiques du portefeuille assuré ou des produits d'assurance modélisés.

Les assureurs cherchent à optimiser leurs allocations d'actifs en fonction de leur passif. **La mortalité est un facteur de risque qui va principalement impacter le passif de l'assureur. Les modèles ALM sont généralement utilisés avec une table de mortalité déterministe. Or, la mortalité déterministe dans un modèle ALM ne rend pas compte des variations du passif dans le temps en fonction des variations de la mortalité. En corrélant la mortalité à des facteurs financiers, le lien est fait entre le passif et les conditions de marché qui influent sur l'actif de l'assureur. L'objectif est donc d'utiliser les résultats de la projection des taux de mortalité avec le GSE en entrée d'un modèle ALM.**

Le modèle ALM utilisé est un *toy model* développé en interne par Milliman et dont la construction se base sur le modèle développé par Floryszczak, Le Courtois et Majri (2016). Ce modèle n'a pas vocation à être utilisé sur le marché de l'assurance et sert d'outil d'application. L'objectif de l'utilisation de ce *toy model* est de modéliser les *cash-flow* d'assurance en prévoyant le comportement du portefeuille. Pour ce faire, les flux d'entrée et de sortie sont modélisés pour les contrats émis. Le résultat analysé à la sortie du modèle est donc le BE.

Ce modèle ALM repose sur plusieurs hypothèses simplificatrices :

- Un seul produit est modélisé, un contrat épargne.
- La stratégie d'allocation ne change pas sur la durée de la projection.
- Les assurés sont tous considérés du même âge.
- L'assureur ne vend pas de contrat au cours de la projection, il est en *run-off*.
- Il n'y a pas d'unité dans les grandeurs manipulées dans le *toy model*.

Les éléments principaux du modèle sont présentés ci-dessous.

### Caractéristiques des contrats épargne

On considère des contrats d'épargne avec option de rachat où l'assuré verse son épargne

et la compagnie garantit un taux minimum, qui donne droit à un revenu minimal annuel à l'assuré. L'épargne de l'assuré est investie sur divers supports tels que les actions et les obligations dont on évalue leur valeur de marché. Ainsi, l'assuré se voit accorder une participation aux bénéfices qui correspond à un revenu supplémentaire. La participation au bénéfice est une proportion du gain acquis par l'ensemble du portefeuille. Si l'assuré décède ou rachète son contrat, alors le contrat s'achève. Ainsi, le contrat est caractérisé par le montant initial investi, le taux minimum garanti, le taux de participation aux bénéfices et l'âge de l'assuré.

### **Modélisation des flux de sortie du contrat**

Il y a deux types de sorties :

- Le décès de l'assuré : on considère une table de mortalité donnant la probabilité de décès par âge. Cette table est déterministe ou stochastique. Une table de mortalité stochastique est composée des scénarios de mortalité générés à l'aide d'un modèle de mortalité stochastique. Il peut s'agir des tables de sortie d'un GSE. Contrairement à une table déterministe, la table stochastique dépend du temps. Une table déterministe considère la mortalité constante dans le temps.
- Le rachat du contrat : le taux de rachat sur une période fixée est la somme du taux de rachat structurel et du taux de rachat conjoncturel dépendant du taux proposé par l'assureur.

### **Flux de liquidité**

Les flux entrants de liquidité sont les coupons et les flux sortants sont les sorties de contrat. Ces flux sont payés en utilisant le compte de liquidité dont le taux d'intérêt est connu. Lorsque l'assureur n'a plus assez de liquidité, il vend ses actifs, actions ou obligations.

### **Modèle financier**

Puisque la compagnie investit sur les marchés, un modèle financier stochastique est nécessaire pour modéliser l'indice action et les taux d'intérêt. Le modèle de Black-Scholes modélise l'indice action. Un modèle dérivé du modèle de Vasicek modélise quant à lui les taux d'intérêt (modèles couramment utilisés dont la description détaillée n'est pas fournie dans ce mémoire). Le modèle de Black-Scholes est également employé pour la modélisation de l'indice immobilier.

### **Gestion d'actifs**

Puisque l'adossement de l'actif et du passif est un des objectifs du modèle ALM, la gestion d'actifs est modélisée de manière à être cohérente avec les échéances du passif. Pour cela, l'actif est composé d'actifs de différentes natures :

- des obligations qui ne sont pas exposées aux risques de marché, donc uniquement des obligations sans risque émises au pair,
- d'un indice action qui peut offrir de meilleurs rendements,
- d'un indice immobilier,
- des liquidités disponibles.



On suppose les investissements cycliques, lorsqu'une obligation arrive à maturité, les fonds sont réinvestis sur de nouvelles obligations.

### Gestion de passifs et bilan de Solvabilité II

Dans le bilan de Solvabilité II, on retrouve l'actif en valeur de marché et le passif de la compagnie. L'évaluation du passif est le sujet complexe et se décompose en deux parties :

- La meilleure estimation du passif, *Best-Estimate Liabilities* (BEL), est la valeur actualisée des flux de trésorerie futurs sortants. Il représente les engagements de l'assureur auprès des assurés.
- Les fonds propres de la compagnie, appelé *Net Asset Values* (NAV), est le montant restant à la compagnie après qu'elle aura payé ses dettes. Il correspond à la différence entre la valeur de l'actif et du BEL.

De ces étapes peut être déduit le SCR tel que le montant minimum qui doit être couvert par les fonds propres pour faire face à une ruine économique à horizon un an au niveau de confiance 95%. Les engagements de l'assureur sur ces contrats épargne sont projetés grâce à des simulations de trajectoires économiques, d'une loi de rachat et d'une table de mortalité.

Les modèles ALM sont utilisés pour optimiser les allocations d'actifs des assureurs en fonction de leur passif. La mortalité est généralement déterministe dans les modèles ALM. Cependant, avec une table de mortalité déterministe, les modèles ALM ne prennent pas en compte les variations du passif dans le temps en fonction des variations de la mortalité. C'est pourquoi l'objectif est de discuter de l'utilisation de tables de mortalité stochastiques dans le cadre d'un *toy model* simple où les flux d'un contrat épargne sont modélisés pour un groupe d'assurés d'âge homogène.

## 6.2 Résultats du modèle ALM

Le modèle ALM est utilisé avec une table de mortalité américaine déterministe The United States Social Security Administration (2023) et les scénarios de mortalité américains projetés à l'aide du GSE. Pour étudier les sorties du modèle ALM, on considère le bilan économique simplifié de l'assureur dont le passif est composé du BE et des fonds propres. Trois groupes d'assurés d'âge 30, 50 et 75 ans sont étudiés, les âges sont choisis afin de représenter des générations distinctes.

Les distributions du *Best Estimate* en fonction de l'âge avec une table de mortalité déterministe sont présentées en figure 6.2. On observe une augmentation du BE avec l'âge. Puisqu'il s'agit d'un contrat épargne, le BE est une fonction croissante de la probabilité de décès.

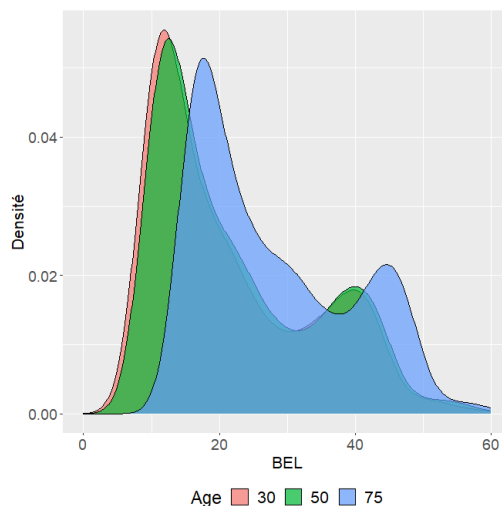


FIGURE 6.2 – Distribution de *Best Estimate* avec une table de mortalité déterministe pour différents âges

Les distributions du BE sont comparées en figure 6.3 en utilisant une table de mortalité déterministe et une table de mortalité stochastique. Les graphiques complémentaires des distributions de BE et de NAV sont en annexe F.

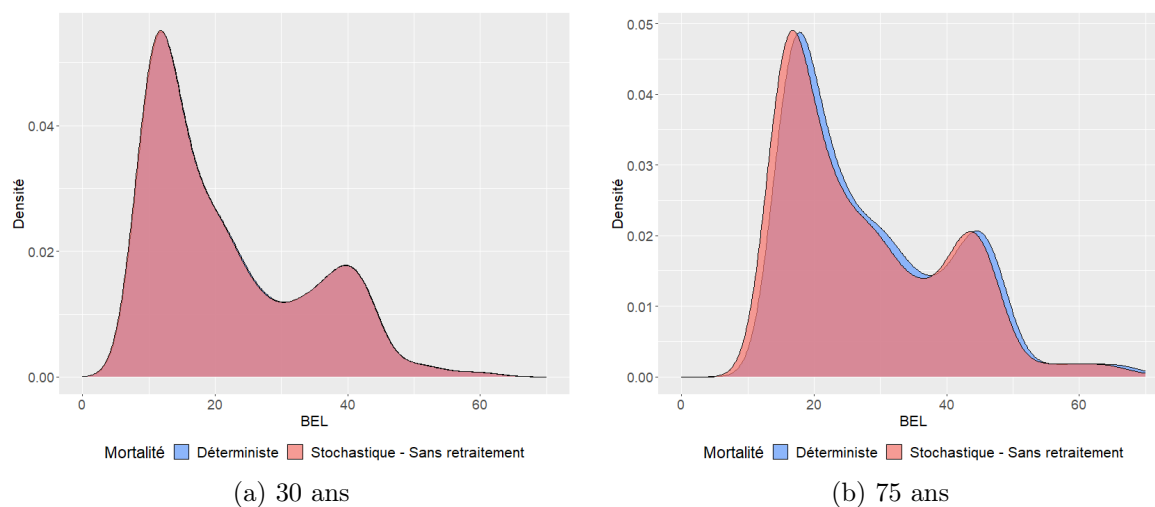


FIGURE 6.3 – Impact de la mortalité stochastique sur le *Best Estimate* pour des groupes d'assurés d'âge 30 ans et 50 ans

La table de mortalité stochastique engendre une variation dans la distribution du BE. Il est à noter que les taux de mortalité des années 2020 et 2021, années impactées par la COVID-19, sont dans les données de calibrage du modèle de Lee-Carter. La table de mortalité stochastique sans retraitement permet de prendre en compte l'amélioration de la mortalité dans le temps bien qu'il y ait l'effet COVID-19.

Il est envisageable de retraiter les données des années 2020 et 2021 avec les méthodes

présentées dans le chapitre 4 pour calibrer le modèle de Lee-Carter dans le GSE.

Les moyennes du BE et de la NAV sont présentées en figure 6.4 pour l'âge 75 ans afin de construire le passif du bilan simplifié de la compagnie d'assurance. Les graphiques des distributions de BE et de NAV avec le retraitement des années impactées par la COVID-19 par la méthode *week-specific lower quartiles* sont en annexe F.

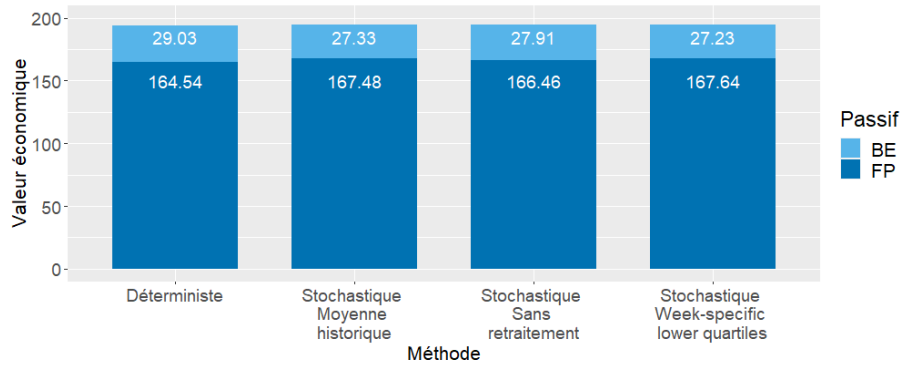


FIGURE 6.4 – Impact des méthodes de projection de la mortalité sur le passif pour un groupe d'assurés d'âge 75 ans

Avec l'utilisation de tables stochastiques, le BE diminue et d'autant plus lorsque des méthodes de retraitement sont appliquées pour corriger la mortalité COVID-19. Les variations sont cependant faibles, ce qui est dû à la construction du modèle qui est très simple. En effet, la sensibilité du modèle ALM à la mortalité est mesurée en appliquant une variation de 5% de la table de mortalité déterministe. Ce choc entraîne une augmentation de 0,3 du BE, soit une augmentation de 1%. Relativement à ce résultat de sensibilité, les variations du BE avec l'utilisation de la table stochastique sont cohérentes.

La diminution du BE avec l'introduction des tables de mortalité stochastiques permet de prendre en compte les projections de la mortalité future. L'estimation des engagements auprès des assurés diminue, tandis que les fonds propres augmentent. En revanche, ce *toy model* ALM est très simple puisqu'il y a un seul type de contrat, les assurés sont tous du même âge. De plus, la loi de rachat ne prend pas en compte le fait que l'exposition au risque de rachat augmente si la mortalité diminue.

### Conclusion

Dans le cadre du *toy model* utilisé, l'introduction de tables de mortalité stochastiques dans un modèle ALM implique la diminution de BE en considérant un contrat épargne pour des assurés de même âge. L'engagement vis-à-vis de ces contrats est sensible aux variations de la mortalité.

Cependant, nous avons montré que ce *toy model* n'est pas particulièrement sensible à l'utilisation des tables de mortalité stochastiques. L'impact sur le BE est tout de même faible et le coût d'implémentation plus élevé qu'avec une table déterministe, ce qui limite l'intérêt de l'utilisation des tables de mortalité stochastiques dans ce type de modèle. Il est cependant envisageable que la mortalité stochastique ait une influence dans le cadre de modèles ALM plus réalistes.

# Conclusion

La modélisation de la mortalité est un enjeu important pour les compagnies d'assurance vie et de réassurance. Le risque de mortalité est pris en compte dans le pilotage des risques des compagnies ainsi que dans le calcul des exigences prudentielles. Ce mémoire d'actuariat a porté sur l'intégration du facteur de risque de mortalité dans un GSE tout en évaluant l'impact et le retraitement de la COVID-19 sur la modélisation de la mortalité.

Dans une première partie, le facteur de risque de mortalité a été intégré dans un GSE. Le modèle de Lee-Carter a été implémenté dans un GSE en univers monde-réel. L'étude de Hanna et Devolder (2023) motive la prise en considération de la corrélation entre le risque de mortalité et les risques financiers. Le GSE a permis de prendre en compte ces interactions dans la projection de la mortalité en corrélant les aléas des modèles stochastiques associés aux facteurs de risque. Pour cela, la relation liant la corrélation empirique entre deux facteurs de risque et la corrélation entre les aléas de leurs modèles sous-jacents respectifs est déterminée. Par exemple, pour les États-Unis, nous avons mis en évidence la corrélation avec le risque de crédit et le risque d'inflation. La méthodologie définie dans ce mémoire nous a permis de déterminer les corrélations entre les aléas associés aux facteurs de risque sélectionnés dans le cadre de ce cas d'étude. La sensibilité des corrélations du facteur de risque de mortalité avec les autres facteurs a été étudiée en fonction des données de mortalité employées. Enfin, le rapport automatique est essentiel pour analyser et interpréter les sorties de la modélisation de la mortalité dans un GSE.

Des travaux futurs sont prévus pour améliorer l'intégration du facteur de risque de mortalité dans un GSE. Par exemple, il est envisageable d'intégrer d'autres modèles de mortalité stochastiques tels que les modèles classiques ou des modèles de mortalité plus sophistiqués. L'intégration de nouveaux modèles implique la détermination des corrélations. De plus, une contrainte de parcimonie a imposé la sélection de deux facteurs de risque à corrélérer avec la mortalité. Sur la base d'autres études, il est possible de corrélérer le facteur de risque de mortalité à d'autres facteurs de risque financiers en suivant la même méthodologie. Enfin, les facteurs de risque peuvent être modélisés par différents modèles. Par exemple, en termes d'alternatives, les modèles à changement de régime pour l'inflation et le modèle Jarrow-Lando-Turnbull pour le crédit qui sont des modèles plus sophistiqués permettent de réaliser des simulations plus réalistes de ces facteurs de risque.

La seconde partie du mémoire a porté sur la modélisation de la mortalité en considérant différents facteurs de risque. Les conséquences de l'inclusion des données de mortalité

impactées par la COVID-19 dans la modélisation ont été constatées. En effet, l'augmentation de la tendance et de la volatilité des modèles a été relevée ainsi que la diminution de l'espérance de vie. Les projections de la mortalité sont impactées par la pandémie de manière inégale selon les pays. En conséquence, les chocs de mortalité déduits de ces projections sont également impactés à la hausse.

Ainsi, des méthodes de retraitement ont été appliquées aux données impactées par la pandémie. On en déduit que les approches les plus robustes et les plus fiables pour quantifier la surmortalité due à des facteurs de risque à court terme tels que la COVID-19 sont basées sur l'estimation de la surmortalité hebdomadaire, contrairement à celles qui se fondent uniquement sur les décomptes officiels des décès dus à la COVID-19. Les chocs de mortalité calculés à partir des données retraitées sont du même ordre de grandeur que les chocs calculés en arrêtant la modélisation de la mortalité en 2019. Ainsi, la correction de ces chocs de mortalité est un enjeu important de ces méthodes de retraitement.

Il y a encore beaucoup à faire en ce qui concerne la modélisation de la mortalité. En effet, des facteurs externes ont diverses conséquences sur la mortalité et nécessitent une modélisation particulière. Il existe d'autres effets de la COVID-19 sur la pandémie qui ne sont pas pris en compte dans les projections de la mortalité présentées dans ce mémoire tels que l'évolution des maladies neurodégénératives. Ces effets peuvent être étudiés en modélisant la mortalité par cause de décès, faisant l'objet d'une autre étude.

De plus, il est aussi possible d'étudier d'autres méthodes d'ajustement des données. Il serait intéressant d'envisager des scénarios de mortalité pour les années 2020 et suivantes en supposant que la pandémie COVID-19 n'ait jamais eu lieu, en tenant compte de toutes les autres particularités de ces années affectant la mortalité, par exemple les épisodes grippeux ou les fortes chaleurs. D'autres méthodes peuvent aussi être utilisées pour prendre en compte l'expérience COVID-19 lors de le calibrage des modèles de mortalité tel que l'extension du modèle à effet de période de Lee-Carter avec des sauts transitoires.

Ainsi, la modélisation de la mortalité reste un sujet vaste et complexe qui interagit avec les autres risques assurantiels et qui est impactée par de nombreux facteurs externes.

# Annexes

# Annexe A

## Outils mathématiques

### Mouvement Brownien uni-dimensionnel

Un mouvement brownien uni-dimensionnel  $(B_t)_{t \geq 0}$  est un processus stochastique dépendant du temps  $t$  défini sur l'espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$  tel que :

- L'accroissement  $B_t - B_s$ ,  $t > s$  est indépendant de  $(B_u)_{0 \leq u \leq s}$ .
- $B_t - B_s$ , pour tous  $t$  et  $s$  tels  $t > s$  est une variable aléatoire Gaussienne centrée de variance  $t - s$ .
- $(B_t)_{t \geq 0}$  est presque sûrement continu.
- $B_0 = 0$  est supposé, il s'agit d'un mouvement Brownien standard.

### Lemme d'Itô unidimensionnel

Soit  $(B_t)_{t \geq 0}$  un mouvement brownien uni-dimensionnel défini sur l'espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ .

Le processus d'Itô uni-dimensionnel  $X_t$  est un processus stochastique vérifiant l'équation différentielle stochastique suivante :

$$\begin{cases} dX_t = \mu_t dt + \sigma_t dB_t \\ X_0 = x \end{cases}$$

Avec  $x$  un réel (condition initiale déterministe),  $\mu$  un processus adapté tel que  $\int_0^t |\mu_s| ds < \infty$  et  $\sigma$  un processus adapté cadlag tel que  $\mathbb{E} \left[ \int_0^t \sigma_s^2 ds \right] < \infty, \forall t > 0$

Soit  $f : (x, t) \rightarrow f(x, t)$  une fonction de  $\mathbb{R} \times \mathbb{R}^{+*}$  dans  $\mathbb{R}$ , de classe  $\mathcal{C}^2$  par rapport à  $t$  et  $\mathcal{C}^2$  par rapport à  $x$  :

$$df(X_t, t) = \frac{\partial f}{\partial t}(X_t, t)dt + \frac{\partial f}{\partial x}(X_t, t)dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(X_t, t)\sigma_t^2 dt \quad (\text{A.1})$$

### Coefficient $R^2$

Le coefficient  $R^2$  est défini comme suit :

$$R^2 = 1 - \frac{SSR}{SST}$$



avec SSR la somme des carrés des résidus et SST la somme des carrés des écarts à la moyenne des valeurs observées.

### AIC et BIC

Le critère d'information d'Akaike, *Akaike Information Criterion*, AIC, est un défini tel que :

$$AIC = 2k - 2\ln(L) \quad (\text{A.2})$$

Le critère d'information bayésien, *Bayesian Information Criterion*, BIC, est défini tel que :

$$BIC = 2\ln(L) - k\ln(n) \quad (\text{A.3})$$

avec :

- $k$  : le nombre de paramètres du modèle
- $L$  : le maximum de la vraisemblance du modèle
- $n$  : le nombre d'observations dans l'échantillon

## Annexe B

# Corrélation entre le facteur de risque de mortalité et d'autres facteurs de risque financiers

### Corrélation entre deux modèles de Lee-Carter

L'objectif est de déterminer la relation liant :

- la corrélation entre deux facteurs de risque de mortalité de deux pays différents
- les corrélations entre les aléas des modèles de Lee-Carter modélisant ces taux de mortalité.

Les paramètres et les taux de mortalité sont numérotés en référence aux deux pays dont on étudie la corrélation. Le cadre de modélisation est le modèle de Lee-Carter tel que :

$$\eta_{x,t}^i = \alpha_x^i + \beta_x^i \kappa_t^i \quad (\text{B.1})$$

avec  $\forall t, \kappa_{t+1}^i = \kappa_t^i + \delta^i + \xi_t, \xi_t \sim \mathcal{N}(0, (\sigma^i)^2)$

En développant l'expression du modèle (B.1), le taux de mortalité s'exprime tel que :

$$\eta_{x,t}^i = \alpha_x^i + \beta_x^i \left( \kappa_0^i + t\delta^i + \sum_{j=0}^{t-1} \xi_{t-j} \right) \quad (\text{B.2})$$

On considère  $\Delta$  un intervalle de temps entre deux dates, fixé à un an. Le log-incrément des taux de mortalité sur un intervalle de temps  $\Delta$  est donc déduit de l'équation (B.2) :

$$\eta_{x,t}^i - \eta_{x,t-\Delta}^i = \beta_x^i \left( \Delta\delta^i + \sqrt{\Delta}\sigma^i \varepsilon_t^i \right) \text{ avec } \varepsilon_t^i \sim \mathcal{N}(0, 1) \quad (\text{B.3})$$

Les log-incréments des taux de mortalité sont homogènes à des Gaussiennes dans (B.3). La corrélation théorique entre les taux de mortalité de deux pays distincts est définie telle que :

$$\text{Corr} \left( \eta_{x,t}^1 - \eta_{x,t-\Delta}^1, \eta_{x,t}^2 - \eta_{x,t-\Delta}^2 \right) = \frac{\text{Cov} \left( \eta_{x,t}^1 - \eta_{x,t-\Delta}^1, \eta_{x,t}^2 - \eta_{x,t-\Delta}^2 \right)}{\sqrt{\text{Var} \left( \eta_{x,t}^1 - \eta_{x,t-\Delta}^1 \right)} \sqrt{\text{Var} \left( \eta_{x,t}^2 - \eta_{x,t-\Delta}^2 \right)}}$$

$$\begin{aligned}
&= \frac{\Delta \beta_x^1 \beta_x^2 \sigma^1 \sigma^2 \text{Corr}(\varepsilon_t^1, \varepsilon_t^2)}{\sqrt{\left(\sqrt{\Delta} \beta_x^1 \sigma^1\right)^2} \sqrt{\left(\sqrt{\Delta} \beta_x^2 \sigma^2\right)^2}} \\
&= \text{corr}(\varepsilon_t^1, \varepsilon_t^2)
\end{aligned}$$

La corrélation historique entre les facteurs de risque est égale à la corrélation entre les aléas :

$$\text{Corr}\left(\eta_{x,t}^1 - \eta_{x,t-\Delta}^1, \eta_{x,t}^2 - \eta_{x,t-\Delta}^2\right) = \text{corr}(\varepsilon_t^1, \varepsilon_t^2) \quad (\text{B.4})$$

### Corrélation entre le modèle AR-GARCH et le modèle de Lee-Carter

Le modèle AR-GARCH, *Autoregressive with Generalized Autoregressive Conditional Heteroscedasticity*, est utilisé pour modéliser l'indice action. L'objectif est de déterminer la relation liant :

- la corrélation entre le facteur de risque de mortalité et le facteur de risque action
- les corrélations entre les aléas du modèle de Lee-Carter modélisant les taux de mortalité et du modèle AR-GARCH modélisant les rendements de l'indice action.

On note  $Y$ , le log-rendement de l'indice action :

$$Y(t + \Delta) = \ln \left( P(t, t + \Delta) \frac{S(t + \Delta)}{S(t)} \right)$$

avec

- $\Delta$  : le pas de temps
- $P(t, t + m)$  : le prix au temps  $t$  d'une obligation zéro-coupon de maturité  $m$ , délivrant le nominal en  $t + m$  pour  $(t, m) \in \mathbb{R}^2$
- $S(t)$  : l'indice action à l'instant  $t$

Dans ce cadre, le log-rendement des actions suit une dynamique AR(1)-GARCH(1,1) qui prend en compte l'hétéroscédasticité du rendement des actions, on fixe  $t_0$  l'instant initial,  $\forall k \in [2, T]$  :

$$Y(t_0 + k\Delta) = \lambda + \gamma Y(t_0 + (k - 1)\Delta) + v(t_0 + k\Delta)$$

avec

- $v(t_0 + k\Delta) = \sqrt{h(t_0 + k\Delta)} \varepsilon(k)$
- $h(t_0 + k\Delta) = \omega + \alpha v(t_0 + (k - 1)\Delta)^2 + \beta h(t_0 + (k - 1)\Delta)$ ,  $k \in [3, T]$

et

- $(\lambda, \gamma) \in \mathbb{R}^2$  : les paramètres du modèle AR(1)
- $(\omega, \alpha, \beta) \in (\mathbb{R}^{*+})^3$  : les paramètres du modèle GARCH(1,1)

—  $(\varepsilon(k))_{k \in [0, T-1]}$  : les variables aléatoires Gaussiennes centrées réduites indépendantes  
Ce modèle est plus complexe et le calcul des corrélations repose sur la discrétisation du modèle. Ainsi, on considère le pas de temps  $\Delta$  et  $k \in \mathbb{N}$ . On définit  $(h_k)_{k \in \mathbb{N}}$  tel que :

$$h_k = h(t_0 + k\Delta) = \omega \sum_{m=0}^{k-1} \prod_{n=1}^m (\alpha \varepsilon(k-n)^2 + \beta) + h_0 \prod_{m=1}^k (\alpha \varepsilon(k-m)^2 + \beta) \quad (\text{B.5})$$

et la forme explicite de  $(Y_k)_{k \in \mathbb{N}}$  en est déduite  $\forall k \in \mathbb{N}$  :

$$Y_k = Y(t_0 + k\Delta) = \lambda \frac{(1 - \gamma^k)}{1 - \gamma} + \gamma^k Y_0 + \sum_{m=0}^{k-1} \gamma^m \sqrt{h_{k-m}} \varepsilon(k-m) \quad (\text{B.6})$$

dont la variance est :

$$V_k = \text{Var}[Y_k] = \theta \left( \frac{1 - \gamma^{2k}}{1 - \gamma^2} - \frac{(\alpha + \beta)^k - \gamma^{2k}}{1 - \frac{\gamma^2}{\alpha + \beta}} \right) + h_0 \frac{(\alpha + \beta)^k - \gamma^{2k}}{1 - \frac{\gamma^2}{\alpha + \beta}} \quad (\text{B.7})$$

En utilisant les expressions (B.6) et (B.5), la corrélation entre les quantités associées au modèle de Lee-Carter et au modèle AR-GARCH est déterminée :

- Lee-Carter :  $Y_{x,k}^1 = Y^1(x, k) = \eta_{x,k\Delta} - \eta_{x,(k-1)\Delta} = \beta_x(k\Delta\delta + \sqrt{k\Delta}\sigma\varepsilon_t^1)$  avec  $x$  l'âge et en reprenant les notations du modèle de Lee-Carter
- AR-AGRCH :  $Y_k = Y(k\Delta)$

$$\begin{aligned} \text{Cov}\left(Y_{x,k}^1, Y_k\right) &= \text{Cov}\left(\beta_x(\Delta\delta + \sqrt{\Delta}\sigma\varepsilon_t^1), \lambda \frac{(1 - \gamma^k)}{1 - \gamma} + \gamma^k Y_0 + \sum_{m=0}^{k-1} \gamma^m \sqrt{h_{k-m}} \varepsilon(k-m)\right) \\ &= \text{Cov}\left(\beta_x \sqrt{\Delta}\sigma\varepsilon_t^1, \sum_{m=0}^{k-1} \gamma^m \sqrt{h_{k-m}} \varepsilon(k-m)\right) \\ &= \beta_x \sqrt{\Delta}\sigma \sum_{m=0}^{k-1} \gamma^m \text{Cov}\left(\varepsilon_k^1, \sqrt{h_{k-m}} \varepsilon(k-m)\right) \end{aligned} \quad (\text{B.8})$$

Or les termes de la série  $(\varepsilon_k)_{k \in \mathbb{N}}$  sont indépendants et  $\mathbb{E}[\sqrt{h_k}\varepsilon(k)] = \mathbb{E}[\varepsilon_k^1] = 0$  donc

$$\text{Cov}(\varepsilon_k^1, \sqrt{h_{k-m}} \varepsilon(k-m)) = \mathbb{E}[\varepsilon_k^1 \sqrt{h_{k-m}} \varepsilon_{k-m}]$$

et  $\varepsilon_k^1$  est indépendant de  $\varepsilon_{k-m}^2$  et  $\varepsilon_k^1$  est indépendant de  $\sqrt{h_{k-m}} \varepsilon_{k-m}$  donc  $\mathbb{E}[\varepsilon_k^1 \sqrt{h_{k-m}} \varepsilon_{k-m}] = 0$

On en déduit la covariance et la corrélation à partir des équations (B.8) et (B.7) et de la propriété précédente :

$$\begin{aligned} \text{Cov}\left(Y_{x,k}^1, Y_k\right) &= \beta_x \sqrt{\Delta}\sigma \text{Cov}(\varepsilon_k^1, \sqrt{h_k} \varepsilon_k) \\ &= \beta_x \sqrt{\Delta}\sigma \rho \mathbb{E}\left[\sqrt{h_k}\right] \end{aligned}$$

$$= \beta_x \sqrt{\Delta} \sigma \text{Cov}(\varepsilon_k^1, \varepsilon_k) \left( \sqrt{\theta} + \frac{(h_0 - \theta)(\alpha + \beta)^k}{2\sqrt{\theta}} \right)$$

avec  $\theta = \frac{\omega}{1 - (\alpha + \beta)}$  et  $\sqrt{h_k}$  fonction déterministe de  $(\varepsilon_0, \dots, \varepsilon_{k-1})$ .

L'expression de la corrélation est donc :

$$\text{Corr}(Y_{x,k}^1, Y_k) = \frac{\text{Cov}(\varepsilon_k^1, \varepsilon_k) \left( \sqrt{\theta} + \frac{(h_0 - \theta)(\alpha + \beta)^k}{2\sqrt{\theta}} \right)}{\theta \left( \frac{1 - \gamma^{2k}}{1 - \gamma^2} - \frac{(\alpha + \beta)^k - \gamma^{2k}}{1 - \frac{\gamma^2}{\alpha + \beta}} \right) + h_0 \frac{(\alpha + \beta)^k - \gamma^{2k}}{1 - \frac{\gamma^2}{\alpha + \beta}}} \quad (\text{B.9})$$

L'expression de la corrélation dépend du temps, l'objectif étant d'obtenir une expression de la corrélation indépendante du temps pour le calcul des corrélations entre mouvements Browniens. Les simulations numériques permettent de montrer le résultat asymptotique suivant :

$$\text{Corr}(Y^1(x, t), Y(t)) \xrightarrow{t \rightarrow +\infty} \text{Cov}(\xi_k, \varepsilon_k) \sqrt{1 - \gamma^2} \quad (\text{B.10})$$

### Corrélation entre le modèle de Nelson-Siegel Vasicek et le modèle de Lee-Carter

Le modèle de Nelson-Siegel Vasicek (NSV) est un modèle de structure des taux, utilisé pour modéliser les taux nominaux et les taux réels ainsi que l'inflation. Il permet de simuler la structure par terme en spécifiant l'évolution de la fonction de Nelson-Siegel en fonction du temps. Les quantités modélisées par le modèle NSV sont les mêmes quantités que modélise le modèle de PCA-Vasicek, et les notations sont reprises. Les structures historiques de la quantité  $Y$  sont modélisées selon le cadre de Nelson-Siegel. Ainsi la définition du modèle est la suivante pour  $k = 1, \dots, T$  :

$$Y(t_0 + k\Delta, t_0 + k\Delta + m) = \beta_{1,t_0+k\Delta} + \beta_{2,t_0+k\Delta} \times \frac{1 - e^{-m/\lambda}}{m/\lambda} + \beta_{3,t_0+k\Delta} \times \left( \frac{1 - e^{-\frac{m}{\lambda}}}{\frac{m}{\lambda}} - e^{-\frac{m}{\lambda}} \right) \quad (\text{B.11})$$

L'évolution du paramètre  $(\beta_{i,t_0+k\Delta})_{i \in \{1,2,3\}, k \in \{1, \dots, T\}}$  en fonction du temps est décrite par un processus AR(1) homogène à la dynamique de Vasicek :

$$\beta_{i,t_0+(k+1)\Delta} = \beta_{i,t_0+k\Delta} - \alpha_i (\beta_{i,t_0+k\Delta} - \beta_i^\infty) \Delta + v_i \sqrt{\Delta} \varepsilon_{i,k}$$

avec  $(\varepsilon_{i,k})_{i \in \{1,2,3\}, k \in \{1, \dots, T\}}$  des variables Gaussiennes standards indépendantes et  $(\lambda, (\alpha_i, \beta_i^\infty, v_i)_{i \in \{1,2,3\}})$  les paramètres à estimer. On note, pour tout  $k \in \mathbb{N}^*$ ,  $Y_k = Y(k\Delta)$  et on pose  $(\tilde{Y}_{k,k+m})_{k \in \mathbb{N}^*}$  tel que  $\tilde{Y}_{k,k+m} = Y_{k,k+m} - Y_{k-1,k-1+m}$ . La variance du processus

$(\tilde{Y}_{k,k+m})_{k \in \mathbb{N}^*}$  est définie pour  $m = 1, \dots, M$  et pour tout  $k \in \mathbb{N}^*$  :

$$\widetilde{V}_k^m = \text{Var} [\tilde{Y}_{k,k+m}] = \Delta \left( v_1^2 + v_2^2 \left( \frac{1 - \exp(-\frac{m}{\lambda})}{\frac{m}{\lambda}} \right)^2 + v_3^2 \left( \frac{1 - \exp(-\frac{m}{\lambda})}{\frac{m}{\lambda}} - \exp(-\frac{m}{\lambda}) \right)^2 \right) \quad (\text{B.12})$$

La covariance entre le modèle NSV et le modèle de Lee-Carter est déduite en utilisant (B.11) et (B.12) :

$$\text{Corr}(Y_{x,k}^1, \tilde{Y}_{k,k+m}) = \frac{(\nu_1 \rho_1 + \nu_2 \rho_2 (\frac{1 - e^{-\frac{m}{\lambda}}}{\frac{m}{\lambda}}) + \nu_3 \rho_3 (\frac{1 - e^{-\frac{m}{\lambda}}}{\frac{m}{\lambda}} - e^{-\frac{m}{\lambda}}))}{\sqrt{\left( v_1^2 + v_2^2 \left( \frac{1 - \exp(-\frac{m}{\lambda})}{\frac{m}{\lambda}} \right)^2 + v_3^2 \left( \frac{1 - \exp(-\frac{m}{\lambda})}{\frac{m}{\lambda}} - \exp(-\frac{m}{\lambda}) \right)^2 \right)}} \quad (\text{B.13})$$

avec  $\rho_i = \text{Cov}(\varepsilon_k^1, \varepsilon_{i,k})$ ,  $i \in \{1, 2, 3\}$ , la covariance entre les mouvements Browniens des deux modèles.

### Corrélation entre le modèle de Black-Scholes avec changement de régime et le modèle de Lee-Carter

Le calcul des corrélations avec le modèle de Black-Scholes avec changement de régime est proche du calcul avec volatilité constante. L'équation aux dérivées partielles du modèle est  $dS_t = \sigma S_t dW_t$  avec  $\sigma$  variable aléatoire caractérisant le changement de régime.  $\sigma$  prend un nombre fini  $n$  de valeurs.

Calcul du moment d'ordre 2 de  $(S_t)$  :

$$\mathbb{E} [S_t^2] = \mathbb{E} [\mathbb{E} [S_t^2 | \sigma]] = \mathbb{E} \left[ \sum_{i=1}^n \mathbb{E} [S_t^2 | \sigma = \sigma_i] \mathbf{1}_{\sigma = \sigma_i} \right] = \mathbb{E} \left[ \sum_{i=1}^n \sigma_i \Delta \mathbf{1}_{\sigma = \sigma_i} \right] = \sum_{i=1}^n \sigma_i^2 \Delta \mathbb{P}(\sigma = \sigma_i) \quad (\text{B.14})$$

L'expression de la covariance est :

$$\text{Cov} \left( \eta_{x,t} - \eta_{x,t-\Delta}, \log \left( \frac{S_t}{S_{t-\Delta}} \right) \right) = \mathbb{E} \left[ \mathbb{E} \left[ \left( \eta_{x,t} - \eta_{x,t-\Delta} - \mathbb{E}[\eta_{x,t} - \eta_{x,t-\Delta}] \right) \right. \right. \quad (\text{B.15})$$

$$\begin{aligned} & \left. \left( \log \left( \frac{S_t}{S_{t-\Delta}} \right) - \mathbb{E} \left[ \log \left( \frac{S_t}{S_{t-\Delta}} \right) \right] \middle| \sigma \right) \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^n \mathbb{E} [(\beta_x \sqrt{\Delta} \sigma_1 \epsilon_1)(\sigma \sqrt{\Delta} \epsilon_2) | \sigma = \sigma_i] \mathbf{1}_{\sigma = \sigma_i} \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^n \beta_x \sigma_1 \Delta \sigma_i \mathbb{E} [\epsilon_1 \epsilon_2] \mathbf{1}_{\sigma = \sigma_i} \right] \\ &= \sum_{i=1}^n \beta_x \sigma_1 \Delta \sigma_i \text{Corr}(\epsilon_1, \epsilon_2) \mathbb{P}(\sigma = \sigma_i) \quad (\text{B.16}) \end{aligned}$$

La formulation de la corrélation est :

$$\begin{aligned} \text{Corr}\left(\eta_{x,t} - \eta_{x,t-\Delta}, \log\left(\frac{S_t}{S_{t-\Delta}}\right)\right) &= \frac{\text{Cov}\left(\eta_{x,t} - \eta_{x,t-\Delta}, \log\left(\frac{S_t}{S_{t-\Delta}}\right)\right)}{\sqrt{\text{Var}\left(\eta_{x,t} - \eta_{x,t-\Delta}\right)}\sqrt{\text{Var}\left(\log\left(\frac{S_t}{S_{t-\Delta}}\right)\right)}} \\ &= \frac{\beta_x \Delta \text{Corr}(\epsilon_1, \epsilon_2) \sum_{i=1}^n \sigma_i \mathbb{P}(\sigma = \sigma_i)}{\sqrt{\beta_x^2 \sigma_1^2 \Delta} \sqrt{\sum_{i=1}^n \sigma_i^2 \Delta \mathbb{P}(\sigma = \sigma_i)}} \end{aligned} \quad (\text{B.17})$$

### Calcul des corrélations entre le modèle de Cairns-Blake-Dowd et le modèle de Black-Scholes

Le modèle de Cairns-Blake-Dowd modélise la mortalité en intégrant deux composantes stochastiques, l'une pondérée par l'effet de l'âge. Le modèle est défini tel que :

$$\eta_{x,t} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} \quad (\text{B.18})$$

Où :

- $\kappa_t^{(1)} = \kappa_{t-1}^{(1)} + \delta_1 + \xi_t^{(1)}$  avec  $\xi_t^{(1)} \sim \mathcal{N}(0, \sigma_1^2)$  iid.
- $\kappa_t^{(2)} = \kappa_{t-1}^{(2)} + \delta_2 + \xi_t^{(2)}$  avec  $\xi_t^{(2)} \sim \mathcal{N}(0, \sigma_2^2)$  iid.
- $\text{Cov}(\xi_t^{(1)}, \xi_t^{(2)}) = \rho$

Par construction du modèles (B.18), la série temporelle  $\kappa_t^{(i)}$  peut s'écrire :  $\kappa_t^{(i)} = \kappa_0^{(i)} + t\delta_i + \sum_{j=0}^{t-1} \xi_{t-j}^{(i)}$ . L'expression des log-incréments des taux de mortalité pour un pas de temps  $\Delta$  est donc :

$$\begin{aligned} \eta_{x,t} - \eta_{x,t-\Delta} &= \left(\Delta\delta_1 + \sum_{i=0}^{t-1} \xi_{t-i}^{(1)}\right) + (x - \bar{x}) \left(\Delta\delta_2 + \sum_{i=0}^{t-1} \xi_{t-i}^{(2)}\right) \\ &= (\Delta\delta_1 + \sqrt{\Delta}\sigma_1\varepsilon_t^{(1)}) + (x - \bar{x})(\Delta\delta_2 + \sqrt{\Delta}\sigma_2\varepsilon_t^{(2)}) \end{aligned} \quad (\text{B.19})$$

Où :

- $\varepsilon_t^{(1)} \sim \mathcal{N}(0, 1)$  iid.
- $\varepsilon_t^{(2)} \sim \mathcal{N}(0, 1)$  iid.
- $\text{Cov}(\varepsilon_t^{(1)}, \varepsilon_t^{(2)}) = \Delta^2\rho$

La variance de  $\eta_{x,t} - \eta_{x,t-\Delta}$  en est déduite de (B.19) :

$$\begin{aligned} \text{Var}(\eta_{x,t} - \eta_{x,t-\Delta}) &= \text{Cov}(\eta_{x,t} - \eta_{x,t-\Delta}, \eta_{x,t} - \eta_{x,t-\Delta}) \\ &= \Delta\sigma_1^2 \text{Cov}(\varepsilon_t^{(1)}, \varepsilon_t^{(1)}) + 2\Delta\sigma_1\sigma_2(x - \bar{x})\text{Cov}(\varepsilon_t^{(1)}, \varepsilon_t^{(2)}) + \\ &\quad (x - \bar{x})^2 \Delta\sigma_2^2 \text{Cov}(\varepsilon_t^{(2)}, \varepsilon_t^{(2)}) \end{aligned} \quad (\text{B.20})$$

$$= \Delta\sigma_1^2 + 2\Delta\sigma_1\sigma_2(x - \bar{x}) + (x - \bar{x})^2 \Delta\sigma_2^2 \Delta^2\rho \quad (\text{B.21})$$

En reprenant l'expression du modèle de Black-Scholes (2.5), on en déduit covariance et la corrélation entre le modèle de CBD et le modèle de Black-Scholes est :

$$\text{Cov}\left(\eta_{x,t} - \eta_{x,t-\Delta}, \log\left(\frac{S_t}{S_{t-\Delta}}\right)\right) = \Delta\sigma_1\sigma\text{Cov}(\varepsilon_t^{(1)}, \varepsilon_t^{(3)}) + (x - \bar{x})\Delta\sigma_2\sigma\text{Cov}(\varepsilon_t^{(2)}, \varepsilon_t^{(3)}) \quad (\text{B.22})$$

$$\text{Corr}\left(\eta_{x,t} - \eta_{x,t-\Delta}, \log\left(\frac{S_t}{S_{t-\Delta}}\right)\right) = \frac{\Delta\sigma_1\sigma\text{Cov}(\varepsilon_t^{(1)}, \varepsilon_t^{(3)}) + (x - \bar{x})\Delta\sigma_2\sigma\text{Cov}(\varepsilon_t^{(2)}, \varepsilon_t^{(3)})}{\sqrt{\Delta\sigma_1^2 + 2\Delta\sigma_1\sigma_2(x - \bar{x}) + (x - \bar{x})^2\Delta\sigma_2^2\Delta^2\rho\sqrt{\sigma^2\Delta}}} \quad (\text{B.23})$$



## Annexe C

# Impact de la pandémie de COVID-19 sur la modélisation avec les modèles APC, CBD et M7

### Résultats du modèles APC

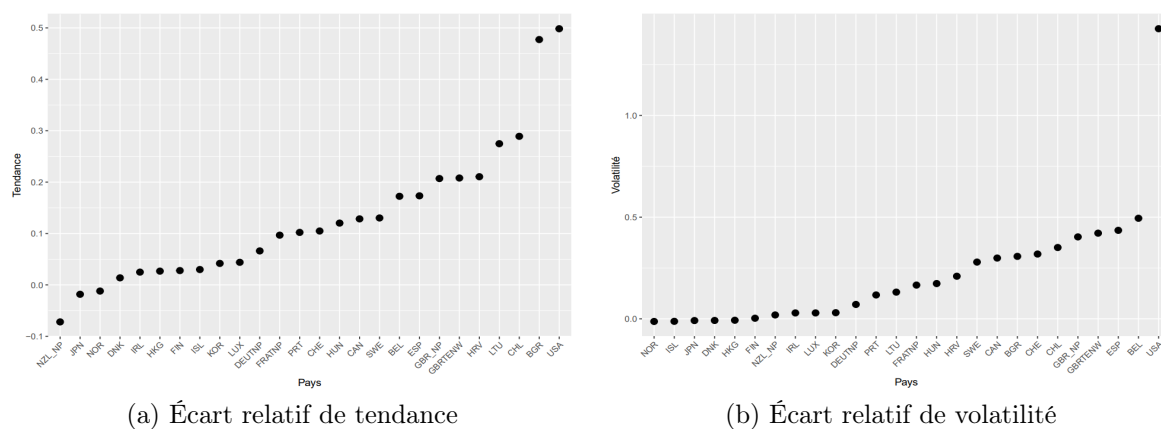


FIGURE C.1 – Écart relatif de tendance (a) et volatilité (b) de  $\kappa_t^{(1)}$  entre les modèles APC avec et sans l'année 2020 en fonction du pays

	Tendance 2019	Tendance 2020	Écart relatif
<b>Japan</b>	-0,0190	-0,0193	-2%
<b>France</b>	-0,0208	-0,0188	10%
<b>USA</b>	-0,0092	-0,0046	50%

TABLE C.1 – Tendances des modèles APC 2019 et 2020

	Volatilité 2019	Volatilité 2020	Écart relatif
<b>Japon</b>	0,0206	0,0204	-1%
<b>France</b>	0,0236	0,0275	17%
<b>USA</b>	0,0125	0,0305	145%

TABLE C.2 – Volatilités des modèles APC 2019 et 2020

### Résultats du modèle CBD

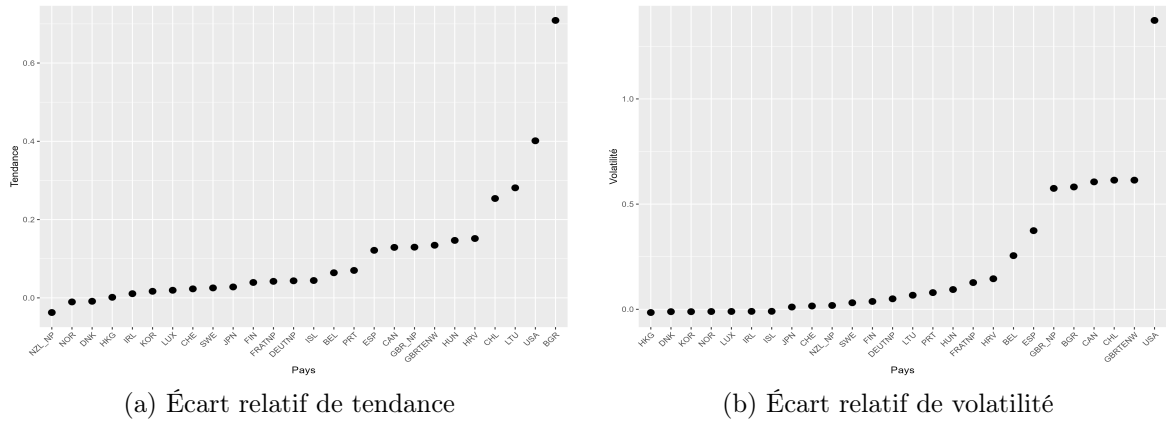


FIGURE C.2 – Écart relatif de tendance (a) et volatilité (b) de  $\kappa_t^{(1)}$  entre les modèles CBD avec et sans l'année 2020 en fonction du pays

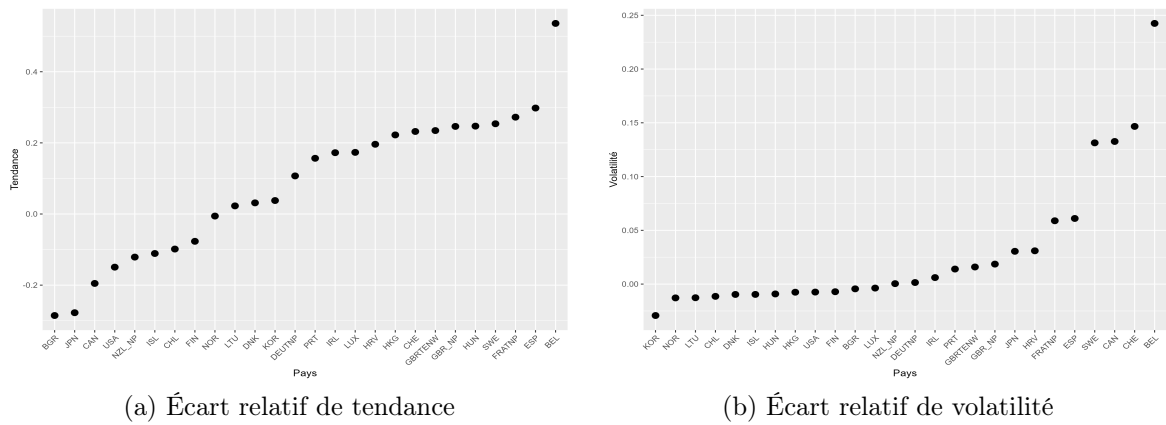


FIGURE C.3 – Écart relatif de tendance (a) et volatilité (b) de  $\kappa_t^{(2)}$  entre les modèles CBD avec et sans l'année 2020 en fonction du pays

	$\kappa_t^{(1)}$			$\kappa_t^{(2)}$		
	Tendance 2019	Tendance 2020	Écart re- latif	Tendance 2019	Tendance 2020	Écart re- latif
<b>Japon</b>	-0,0242	-0,0236	3%	0,00015	0,00011	-28%
<b>France</b>	-0,0225	-0,0215	4%	0,00023	0,00030	27%
<b>USA</b>	-0,0116	-0,0070	40%	0,00008	0,00007	-15%

TABLE C.3 – Tendances des séries temporelles  $\kappa_t^{(1)}$  et  $\kappa_t^{(2)}$  du modèle CBD

	$\kappa_t^{(1)}$			$\kappa_t^{(2)}$		
	Volatilité 2019	Volatilité 2020	Écart re- latif	Volatilité 2019	Volatilité 2020	Écart re- latif
<b>Japon</b>	0,020	0,020	1%	0,00091	0,00094	3%
<b>France</b>	0,011	0,012	13%	0,00105	0,00112	6%
<b>USA</b>	0,014	0,032	136%	0,00074	0,00073	-1%

TABLE C.4 – Volatilités des séries temporelles  $\kappa_t^{(1)}$  et  $\kappa_t^{(2)}$  du modèles CBD

### Comparaison des modèles

	<b>USA</b>			<b>FRANCE</b>			<b>JAPON</b>		
	2019	2020	Écart	2019	2020	Écart	2019	2020	Écart
<b>M1</b>	28,44	28,53	0,091	27,56	26,72	-0,84	24,36	22,38	-1,98
<b>M3</b>	27,85	28,00	0,15	27,70	26,93	-0,77	24,43	22,32	-2,11
<b>M5</b>	28,04	28,13	0,09	26,70	25,82	-0,88	23,79	21,66	-2,13

TABLE C.5 – Espérances de vie projetées à l'âge de 60 ans en 2030

	<b>USA</b>		<b>FRANCE</b>		<b>JAPON</b>	
	2019	2020	2019	2020	2019	2020
<b>M1</b>	99,85%	99,85%	99,95%	99,94%	99,88%	99,87%
<b>M3</b>	99,93%	99,93%	99,80%	99,79%	99,55%	99,53%
<b>M5</b>	95,60%	95,52%	93,78%	93,78%	96,25%	96,26%
<b>M7</b>	99,84%	99,85%	99,93%	99,93%	99,95%	99,95%

TABLE C.6 – Coefficients  $R^2$  des modèles

# Annexe D

## Application des méthodes de retraitement

### Surmortalité de la grippe saisonnière de 2014-2015

La grippe de 2014-2015 a causé 19000 décès.

Méthodes	Augmentation de la surmortalité entre 2014 et 2015 (pour 100 00 habitants)	Augmentation de la surmortalité entre 2014 et 2015 (nombre de décès)
Week specific average	100	63180
Week specific trend	130	82134
Week specific lower quartiles	106	66970
Yearly lower quartiles	105	66339
Alternative within-year	39	24640

TABLE D.1 – Différence de surmortalité de l'année 2015 de l'année 2014 en France

## Annexe E

# Application du modèle de Lee-Carter tenant compte du risque climatique

### Variables climatiques

Variable	Description
$tmax\_y$	Température maximale sur chaque mois $y$ de la période
$tmax$	Température maximale sur la période
$dx\_y$	Nombre de jours où la température maximale est supérieure au seuil $y$
$dx\_y\_suc$	Plus longue période où la température maximale est restée au-dessus du seuil $y$
$ep\_x\_d\_y$	Nombre de périodes de $x$ jours au-dessus du seuil $y$
$dt\_x\_seuil\_y\_ec$	Nombre de jours où l'écart de température est inférieur au seuil $x$ sachant que la température maximale est supérieure au seuil $y$ .

TABLE E.1 – Tableau répertoriant les variables climatiques des USA

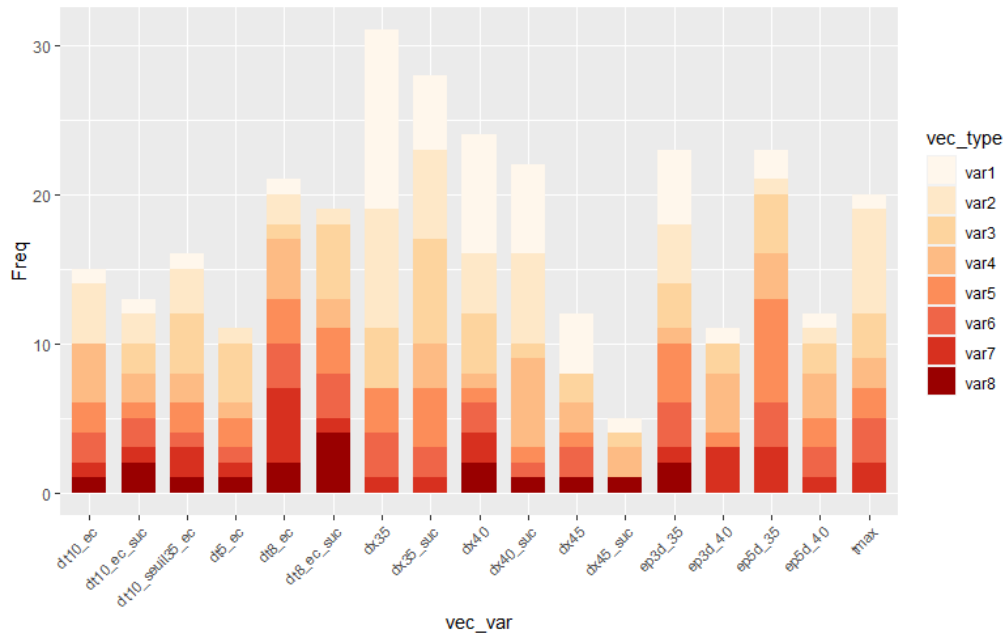


FIGURE E.1 – Variables retenues pour la construction de l'indicateur climatique

### Application du modèle de Lee-Carter climatique à l'Oklahoma

La mortalité climatique en fonction du temps en Oklahoma peut être expliquée à l'aide de huit variables climatiques sélectionnées :

- *ep3d\_35* : Nombre de périodes de 3 jours au-dessus du seuil de 35°C.
- *tmax* : Température maximale enregistrée sur l'année.
- *dx45* : Nombre de jours où la température maximale est au-dessus de 45°C.
- *ep5d\_40* : Nombre de périodes de 5 jours au-dessus du seuil de 40°C.
- *dx40\_suc* : Plus longue période où la température maximale est restée au-dessus du seuil de 40°C.
- *dt8\_ec* : Nombre de jours où l'écart entre la température maximale et minimale sur la journée est inférieur à un seuil de 8°C.
- *ep5d\_35* : Nombre de périodes de 5 jours au-dessus du seuil de 35°C.
- *dx40* : Nombre de jours où la température maximale est au-dessus de 40°C.

# Annexe F

## Résultats ALM

### Méthode sans retraitement

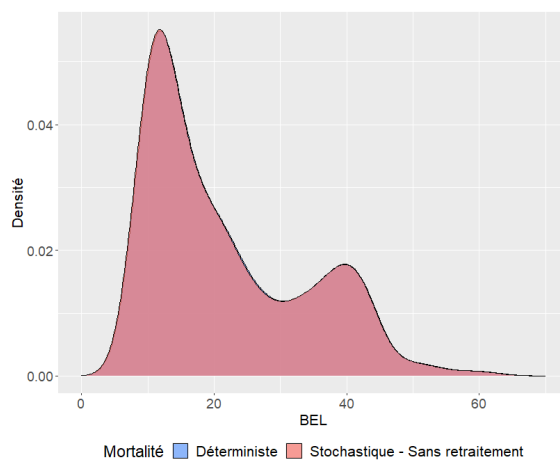


FIGURE F.1 – Impact de la mortalité stochastique sur le *Best Estimate* pour un groupe d'assurés d'âge 30 ans

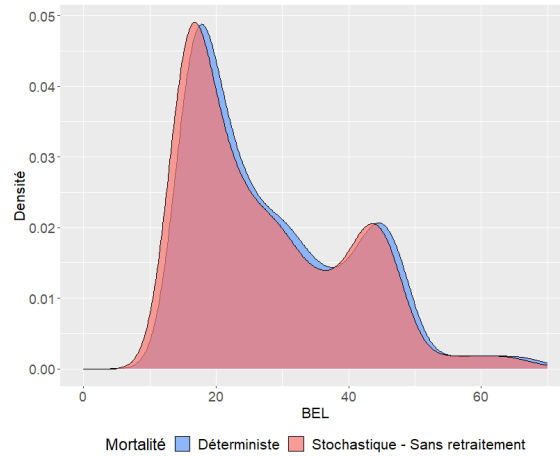


FIGURE F.2 – Impact de la mortalité stochastique sur le *Best Estimate* pour un groupe d'assurés d'âge 75 ans

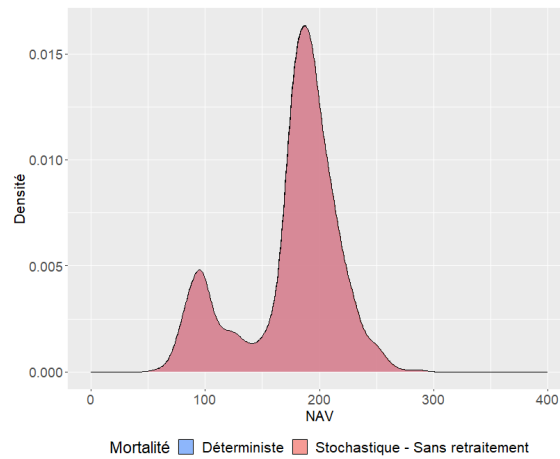


FIGURE F.3 – Impact de la mortalité stochastique sur la *Net Asset Value* pour un groupe d'assurés d'âge 30 ans



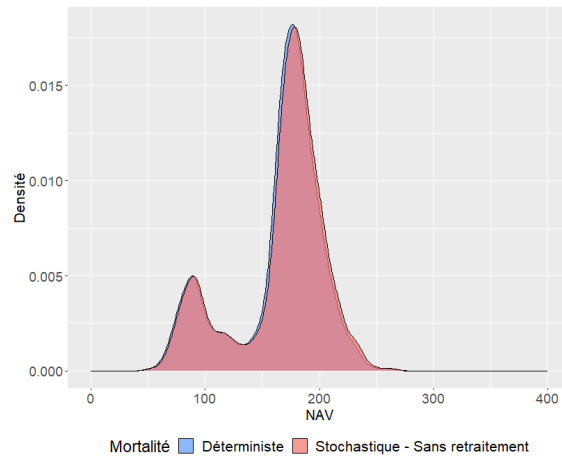


FIGURE F.4 – Impact de la mortalité stochastique sur la *Net Asset Value* pour un groupe d'assurés d'âge 75 ans

### Méthode week specific lower quartiles

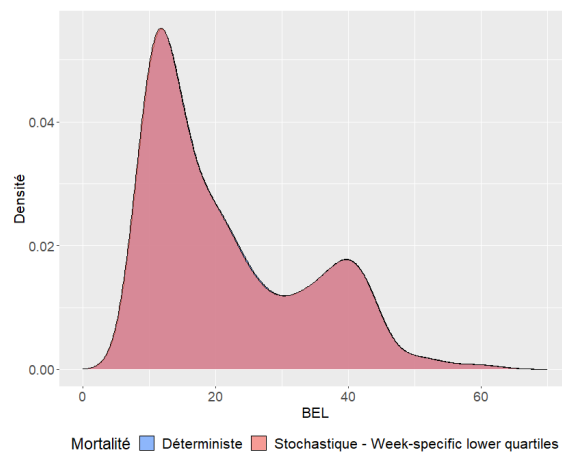


FIGURE F.5 – Impact de la méthode alternative intra-annuelle sur le *Best Estimate* pour un groupe d'assurés d'âge 30 ans

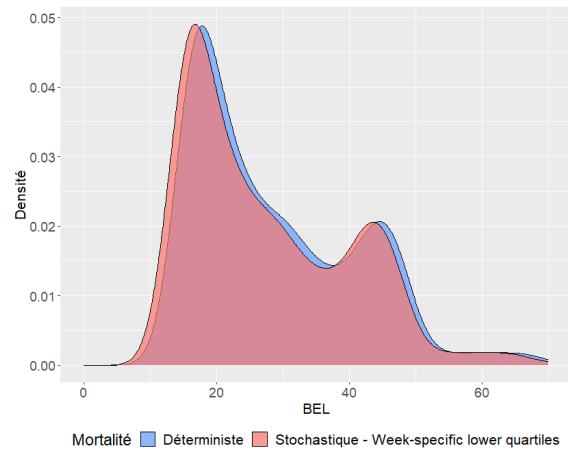


FIGURE F.6 – Impact de la méthode alternative intra-annuelle sur le *Best Estimate* pour un groupe d'assurés d'âge 75 ans

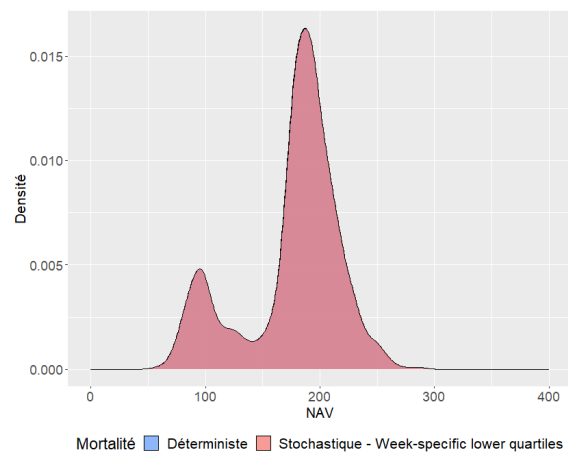


FIGURE F.7 – Impact de la méthode alternative intra-annuelle sur la *Net Asset Value* pour un groupe d'assurés d'âge 30 ans

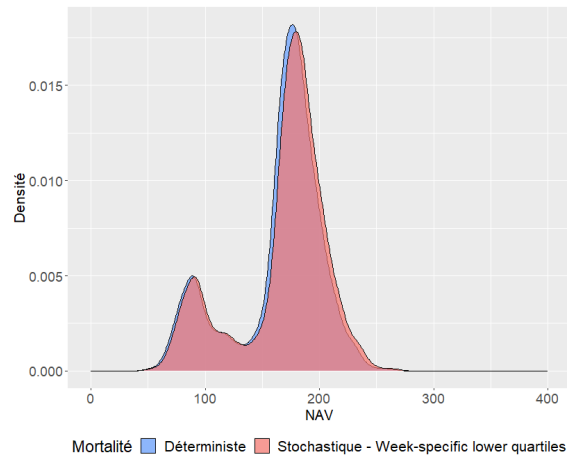


FIGURE F.8 – Impact de la méthode alternative intra-annuelle sur la *Net Asset Value* pour un groupe d'assurés d'âge 75 ans

### Synthèse des bilans économiques

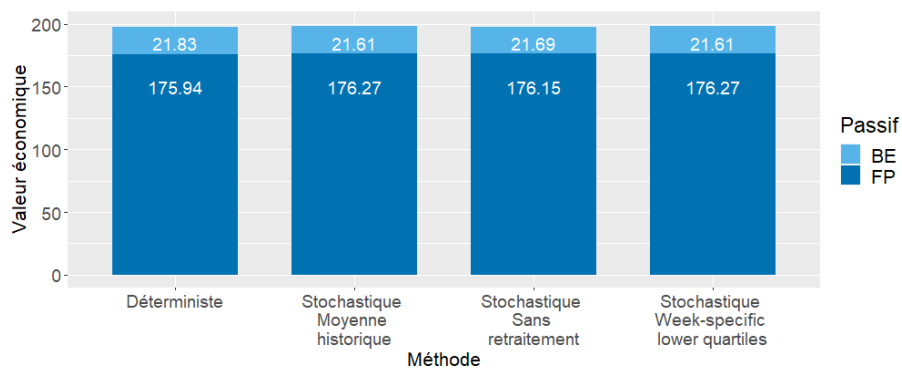


FIGURE F.9 – Impact des méthodes de projection de la mortalité sur le passif pour un groupe d'assurés d'âge 30 ans

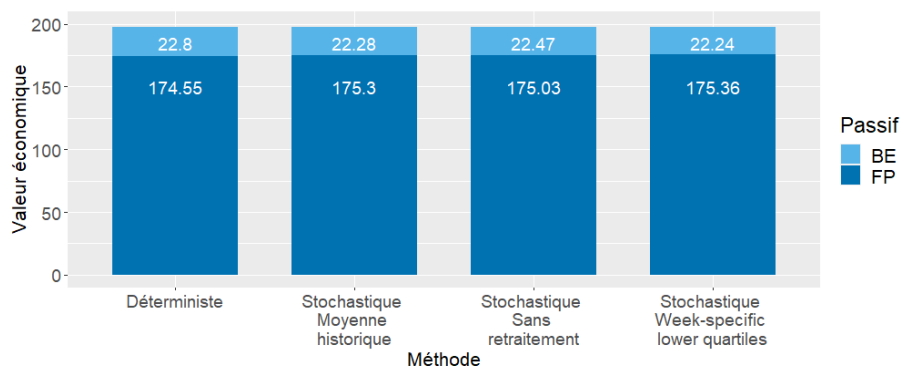


FIGURE F.10 – Impact des méthodes de projection de la mortalité sur le passif pour un groupe d'assuré d'âge 50 ans

# Bibliographie

- Bonam, D., et A. Smădu. 2021. “The long-run effects of pandemics on inflation : Will this time be different ?” *Economics letters*.
- Boumezoued, A., et al. 2019. “Modeling and forecasting cause-of-death mortality”. *Society of Actuaries Research Report*.
- . 2022. “Modelling the impact of climate risks on mortality.” *Milliman White Paper*.
- Buccioli, A., et L. Zarri. 2015. “The shadow of the past : Financial risk taking and negative life events.” *Journal of Economic Psychology*.
- Buckdhan, R. 2023. “Modèles de valorisation”. *Support de cours EURIA*.
- Byström, H. 2021. “Credit Risk in a Pandemic”. *The Journal of Fixed Income*.
- Cairns, A. J., et al. 2009. “A quantitative comparison of stochastic mortality models using data from England and Wales and the United States”. *North American Actuarial Journal*.
- Caporali, A., et al. 2022. “The demography of COVID-19 deaths database, a gateway to well-documented international data.” *Scientific data*.
- Chițu, L., et al. 2022. “The role of credit risk in recent global corporate bond valuations”. *Economic Bulletin Boxes*.
- CMI. 2023. “Continuous Mortality Investigation”. *Working Paper 169*.
- Code des assurances, Article R351-2. 2015.
- Costemalle, V., et al. 2021. “En quatre vagues, l'épidémie de Covid-19 a causé 116 000 décès et lourdement affecté le système de soins.” *INSEE*.
- Currie, I. D. 2006. “Smoothing and forecasting mortality rates with P-splines”. *Talk given at the Institute of Actuaries*.
- Dacorogna, M. M., et M. Cadena. 2015. “Exploring the dependence between mortality and market risks.” *Social Science Research Network*.
- Duffy, S. 2022. “When Inflation Rises, Health Outcomes Fall.” *Harvard Business Review*.
- EIOPA. 2018. “EIOPA’s second set of advice to the European Commission on specific items in the Solvency II Delegated Regulation”.

- Floryszczak, A., O. Le Courtois et M. Majri. 2016. “Inside the Solvency 2 black box : net asset values and solvency capital requirements with a least-squares Monte-Carlo approach.” *Insurance : Mathematics and Economics*.
- Germain, V. 2021. “Prise en compte du changement climatique dans la modélisation des risques biométriques.” *Mémoire d’actuariat*.
- Hanna, V., et P. Devolder. 2023. “Valuation of mixed life insurance contracts under stochastic correlated mortality and interest rates.” *European Actuarial Journal*.
- Hibbert, J., et al. 2001. “A stochastic asset model calibration for long-term financial planning purposes.” *Finance and investment conference*.
- Higham, N. J. 2002. “Computing the nearest correlation matrix—a problem from finance”. *IMA journal of Numerical Analysis*.
- HMD. 2018. “The Human Mortality Database.” *University of California, Berkeley (USA), and Max Planck institute for demographic research (Germany)*, <https://www.mortality.org>.
- Huang, P., et al. 2023. “Links between COVID-19 and Parkinson’s disease/Alzheimer’s disease : reciprocal impacts, medical care strategies and underlying mechanisms.” *Translational Neurodegeneration*.
- INSEE. 2023. “Taux de mortalité mensuels, annualisés (décès pour 1000 habitants) - France métropolitaine”, <https://www.insee.fr/en/statistiques/serie/000436395#Documentation>.
- Jdanov, D., et al. 2021. “Short-Term Mortality Fluctuations Dataseries Methods Protocol”. *Max Planck Institute for Demographic Research*.
- Klein, A., A. Boumezoued et A. Elfassihi. 2022. “Modelling the pandemic risk”. *Milliman White Paper*.
- Lee, R. D., et L. R. Carter. 1992. “Modeling and forecasting US mortality”. *Journal of the American statistical association*.
- Mathieu, E., et al. 2021. “A global database of COVID-19 vaccinations.” *Nat Hum Behav.*
- Maurer, T. A. 2011. “Asset pricing implications of demographic change.” *Australasian Finance and Banking Conference*.
- Németh, L., et al. 2021. “An open-sourced, web-based application to analyze weekly excess mortality based on the Short-term Mortality Fluctuations data series.” *PLoS One*.
- Papon, S. 2020. “Depuis plus d’un siècle, les décès sont les plus nombreux en hiver.” *Division Enquêtes et études démographiques, Insee*.
- Planchet, F. 2017. “Modèles financiers en assurance et analyse dynamiques”. *Support de cours ISFA*.

- Santé publique France. 2022. “Impact de la pandémie Covid-19 sur la santé mentale des Français”, <https://www.santepubliquefrance.fr/les-actualites/2022/impact-de-la-pandemie-covid-19-sur-la-sante-mentale-des-francais.-le-dossier-de-la-sante-en-action-n-461-septembre-2022>.
- Schnürch, S., et al. 2022. “The impact of mortality shocks on modelling and insurance valuation as exemplified by COVID-19”. *Annals of Actuarial Science*.
- Shkolnikov, V. M., et al. 2022. “What should be the baseline when calculating excess mortality? New approaches suggest that we have underestimated the impact of the COVID-19 pandemic and previous winter peaks.” *SSM-Population Health*.
- Soulas, V., et al. 2023. “What is the impact of wearable technologies on life expectancy?” *Milliman White Paper*.
- The United States Social Security Administration. 2023. “Actuarial life table”, <https://www.mortality.org>.
- Vermet, F. 2023. “Modèles de mortalité.” *Support de cours EURIA*.
- Villegas, A., et al. 2015. “StMoMo : An R package for stochastic mortality modelling.” *In 7th Australasian Actuarial Education and Research Symposium*.
- World Health Organization. 2008. “Improving public health responses to extreme weather”. *Copenhagen : WHO Regional Office for Europe*.

# Table des figures

1	Écart relatif des tendance (a) et volatilité (b) de $\kappa_t^{(1)}$ des modèles de Lee-Carter dont les données de calibrage s'arrêtent en 2019 ou 2020 . . . . .	v
2	Chocs de mortalité aux États-Unis avec ou sans méthode de retraitement . . . . .	vii
3	Relative gap of trend (a) and volatility (b) of $\kappa_t^{(1)}$ for Lee-Carter models with calibration data ending in 2019 or 2020 . . . . .	xi
4	Mortality shocks in the United States with and without adjustment method . . . . .	xiii
2.1	Corrélation des taux de mortalité français avec les variables financières . . . . .	18
2.2	Corrélation des taux de mortalité mensuels français avec les variables financières . . . . .	18
2.3	Corrélation des taux de mortalité mensuels français avec les spreads de crédit . . . . .	19
2.4	Comparaison des corrélations des taux de mortalité mensuels français avec les variables financières dont les données s'arrêtent en 2019 ou 2020 . . . . .	20
2.5	Corrélation des taux de mortalité américains avec les variables financières . . . . .	22
3.1	Paramètres dépendants de l'âge $\alpha_x$ et $\beta_x$ du modèle de Lee-Carter . . . . .	33
3.2	Paramètre $\kappa_t$ du modèle de Lee-Carter . . . . .	33
3.3	Carte de chaleur des résidus du modèle de Lee-Carter sur les États-Unis . . . . .	34
3.4	Coefficient $R^2$ du modèle de Lee-Carter en fonction de l'âge . . . . .	34
3.5	Simulation des taux de mortalité à l'âge de 65 ans . . . . .	35
3.6	Carte de chaleur de l'amélioration de la mortalité sur les taux historiques des États-Unis . . . . .	36
3.7	Carte de chaleur de l'amélioration de la mortalité sur les taux modélisés avec le modèle de Lee-Carter calibré sur des données des États-Unis . . . . .	36
4.1	Taux de mortalité (a) et leur logarithme (b) de la population totale des États-Unis en 2019 en fonction de l'âge . . . . .	42
4.2	Taux de mortalité (a) et leur logarithme (b) en 2019 aux États-Unis en fonction de l'âge et du genre . . . . .	43
4.3	Taux de mortalité (a) et leur logarithme (b) de la population totale en 2019 en fonction de l'âge et du pays (États-Unis ou Japon) . . . . .	43
4.4	Taux de mortalité (a) et leur logarithme (b) de la population totale aux États-Unis en 2019 et 2020 en fonction de l'âge . . . . .	44



4.5	Taux de mortalité (a) et leur logarithme (b) de la population totale au Japon en 2019 et 2020 en fonction de l'âge . . . . .	44
4.6	Carte de chaleur du logarithme des taux de mortalité aux Etats-Unis . . . . .	45
4.7	Modélisation et projection des taux de mortalité à 75 ans des États-Unis avec le modèle de Lee-Carter . . . . .	46
4.8	Modélisation et projection des taux de mortalité à 75 ans en France avec le modèle de Lee-Carter . . . . .	47
4.9	Modélisation et projection des taux de mortalité à 75 ans du Japon avec le modèle de Lee-Carter . . . . .	47
4.10	Écart relatif (en %) de tendance (a) et volatilité (b) de $\kappa_t^{(1)}$ entre les modèles de Lee-Carter avec et sans l'année 2020 en fonction du pays . . . . .	48
4.11	Carte de chaleur de la surmortalité (pour 100 000 habitants) déterminée avec la méthode <i>week-specific lower quartiles</i> . . . . .	57
4.12	Carte de chaleur de la surmortalité (pour 100 000 habitants) déterminée avec la méthode <i>week-specific trends</i> . . . . .	58
4.13	Carte de chaleur de la surmortalité (pour 100 000 habitants) déterminée avec la méthode <i>alternative within-year</i> . . . . .	58
4.14	Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction de l'année 2020 par <i>alternative within-year</i> . . . . .	62
4.15	Chocs de mortalité aux États-Unis avec et sans l'année 2020 . . . . .	64
4.16	Chocs de mortalité en France avec et sans l'année 2020 . . . . .	65
4.17	Chocs de mortalité au Japon avec et sans l'année 2020 . . . . .	65
4.18	Chocs de mortalité aux États-Unis (a) et en France (b) avec et sans correction de l'année 2020 . . . . .	66
4.19	Taux de mortalité de 2019 à 2021 au Japon et aux États-Unis . . . . .	68
4.20	Taux de mortalité hebdomadaires (STMF) de 2018 à 2021 en France et aux États-Unis . . . . .	69
4.21	Schéma de la méthodologie de retraitement et modélisation des taux de mortalité . . . . .	70
4.22	Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter sans correction . . . . .	71
4.23	Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction par la moyenne historique . . . . .	71
4.24	Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction par soustraction des décès . . . . .	72
4.25	Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction par la méthode <i>week-specific average</i> . . . . .	72
4.26	Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction par la méthode <i>week-specific lower quartiles</i> . . . . .	73
4.27	Modélisation et projection des taux de mortalité à 70 ans aux États-Unis avec le modèle de Lee-Carter avec correction par la méthode <i>alternative within-year</i> . . . . .	73

4.28	Choc de mortalité aux États-Unis sans retraitement des années 2020 et 2021	74
4.29	Choc de mortalité aux États-Unis avec retraitement des années 2020 et 2021 par la méthode <i>week-specific lower quartiles</i>	75
5.1	Score $R^2$ de la régression de l'indicateur climatique par état	83
5.2	Indicateur climatique et taux de mortalité climatique en Oklahoma	83
5.3	Mortalité climatique par classe d'âge en Oklahoma	84
5.4	Comparaison des modélisations à l'âge de 75 ans en Oklahoma	84
5.5	Coefficient d'ajustement $R^2$ des deux modèles sur les données d'Oklahoma	85
6.1	Schéma de l'organisation des flux financiers au sein d'une compagnie	91
6.2	Distribution de <i>Best Estimate</i> avec une table de mortalité déterministe pour différents âges	94
6.3	Impact de la mortalité stochastique sur le <i>Best Estimate</i> pour des groupes d'assurés d'âge 30 ans et 50 ans	94
6.4	Impact des méthodes de projection de la mortalité sur le passif pour un groupe d'assurés d'âge 75 ans	95
C.1	Écart relatif de tendance (a) et volatilité (b) de $\kappa_t^{(1)}$ entre les modèles APC avec et sans l'année 2020 en fonction du pays	109
C.2	Écart relatif de tendance (a) et volatilité (b) de $\kappa_t^{(1)}$ entre les modèles CBD avec et sans l'année 2020 en fonction du pays	110
C.3	Écart relatif de tendance (a) et volatilité (b) de $\kappa_t^{(2)}$ entre les modèles CBD avec et sans l'année 2020 en fonction du pays	110
E.1	Variables retenues pour la construction de l'indicateur climatique	114
F.1	Impact de la mortalité stochastique sur le <i>Best Estimate</i> pour un groupe d'assurés d'âge 30 ans	115
F.2	Impact de la mortalité stochastique sur le <i>Best Estimate</i> pour un groupe d'assurés d'âge 75 ans	116
F.3	Impact de la mortalité stochastique sur la <i>Net Asset Value</i> pour un groupe d'assurés d'âge 30 ans	116
F.4	Impact de la mortalité stochastique sur la <i>Net Asset Value</i> pour un groupe d'assurés d'âge 75 ans	117
F.5	Impact de la méthode alternative intra-annuelle sur le <i>Best Estimate</i> pour un groupe d'assurés d'âge 30 ans	117
F.6	Impact de la méthode alternative intra-annuelle sur le <i>Best Estimate</i> pour un groupe d'assurés d'âge 75 ans	118
F.7	Impact de la méthode alternative intra-annuelle sur la <i>Net Asset Value</i> pour un groupe d'assurés d'âge 30 ans	118
F.8	Impact de la méthode alternative intra-annuelle sur la <i>Net Asset Value</i> pour un groupe d'assurés d'âge 75 ans	119
F.9	Impact des méthodes de projection de la mortalité sur le passif pour un groupe d'assurés d'âge 30 ans	119

---

F.10 Impact des méthodes de projection de la mortalité sur le passif pour un groupe d'assuré d'âge 50 ans . . . . .	120
--	-----

# Liste des tableaux

1	Corrélations empiriques des facteurs de risque observables avec le risque de mortalité . . . . .	iv
2	Corrélations des aléas des modèles avec l'aléa du modèle de Lee-Carter . . .	iv
3	Espérances de vie (en années) projetées en 2021 avec le modèle de Lee-Carter	v
4	Surmortalité (pour 100 000 habitants) de l'année 2020 en fonction des méthodes de retraitement . . . . .	vii
5	Empirical correlations of observable risk factors with mortality risk . . . . .	x
6	Correlation between the random numbers driving the Lee-Carter model and the random numbers driving Black-Scholes and PCA-Vasicek models . . . . .	x
7	Life expectancy (in years) projected to 2021 using the Lee-Carter model . . .	xi
8	Excess mortality (per 100,000 inhabitants) in 2020 according to adjusting methods . . . . .	xii
1.1	Description des facteurs de risque . . . . .	15
2.1	Corrélations empiriques des facteurs de risque observables avec le risque de mortalité . . . . .	30
2.2	Corrélations des aléas des modèles avec l'aléa du modèle de Lee-Carter . . .	30
2.3	Corrélation des aléas des modèles avec l'aléa du modèle de Lee-Carter avant et après retraitement en matrice semi-définie positive . . . . .	31
3.1	Paramètres de la série temporelle $\kappa_t$ . . . . .	33
3.2	AIC et BIC du modèle . . . . .	34
4.1	Pays étudiés dans la base HMD au 03/03/2023 . . . . .	40
4.2	Nombre de décès COVID-19 par pays . . . . .	41
4.3	Tendances des modèles 2019 et 2020 de Lee-Carter . . . . .	48
4.4	Volatilités des modèles 2019 et 2020 de Lee-Carter . . . . .	48
4.5	Espérances de vie projetées en 2023 avec le modèle de Lee-Carter . . . . .	49
4.6	Pays étudiés dans la base STMF . . . . .	52
4.7	Surmortalité (pour 100 000 habitants) de l'année 2020 en fonction des méthodes . . . . .	59
4.8	Récapitulatif des méthodes de retraitement . . . . .	60
4.9	Surmortalité de l'été 2003 en France en fonction des méthodes . . . . .	61

---

4.10	Espérances de vie projetées en 2023 avec le modèle de Lee-Carter avec correction . . . . .	63
4.11	Moyenne des chocs de mortalité en % par tranche d'âge aux États-Unis . . .	66
4.12	Moyenne des chocs de mortalité en % par tranche d'âge en France . . . . .	67
4.13	Espérances de vie aux États-Unis en 2023 . . . . .	74
4.14	Moyenne des chocs de mortalité en % par tranche d'âge aux États-Unis . . .	75
5.1	Étapes de calibrage du modèle de Lee-Carter climatique . . . . .	82
5.2	Scores AIC et BIC des modèles de Lee-Carter classique et climatique . . . . .	85
5.3	Scores AIC et BIC des modèles de Lee-Carter classique et climatique . . . . .	85
C.1	Tendances des modèles APC 2019 et 2020 . . . . .	109
C.2	Volatilités des modèles APC 2019 et 2020 . . . . .	110
C.3	Tendances des séries temporelles $\kappa_t^{(1)}$ et $\kappa_t^{(2)}$ du modèle CBD . . . . .	111
C.4	Volatilités des séries temporelles $\kappa_t^{(1)}$ et $\kappa_t^{(2)}$ du modèles CBD . . . . .	111
C.5	Espérances de vie projetées à l'âge de 60 ans en 2030 . . . . .	111
C.6	Coefficients $R^2$ des modèles . . . . .	111
D.1	Différence de surmortalité de l'année 2015 de l'année 2014 en France . . . . .	112
E.1	Tableau répertoriant les variables climatiques des USA . . . . .	113