

Document de recherche

Application du modèle de crédibilité hiérarchique à la modélisation des taux de mortalité de plusieurs populations

**Sous-commission sur la recherche universitaire –
Commission exécutive de recherche**

Avril 2018

Document 218042

This document is available in English

© 2018 Institut canadien des actuaires

Les documents de recherche ne représentent pas nécessairement l'opinion de l'Institut canadien des actuaires. Les membres devraient connaître les documents de recherche. Les documents de recherche ne constituent pas des normes de pratique et sont donc de caractère non exécutoire. Il n'est pas obligatoire que les documents de recherche soient conformes aux normes de pratique. Le mode d'application de normes de pratique dans un contexte particulier demeure la responsabilité des membres.

Application du modèle de crédibilité hiérarchique à la modélisation des taux de mortalité de plusieurs populations

Cary Chi-Liang Tsai et Adelaide Di Wu

Département de statistiques et d'actuariat, Université Simon Fraser,
Burnaby, C.-B.

Résumé

Le modèle de crédibilité hiérarchique est une généralisation des modèles de crédibilité de Bühlmann et de Bühlmann-Straub avec une structure arborescente d'au moins quatre niveaux. Dans le présent document, nous proposons d'intégrer la théorie de la crédibilité hiérarchique, utilisée en assurances IARD, à la modélisation de la dépendance des taux de mortalité de plusieurs populations. La qualité prévisionnelle des modèles de crédibilité hiérarchique à cinq, à quatre ou à trois niveaux est comparée à celle du modèle traditionnel de Lee-Carter et de ses trois prolongements pour plusieurs populations (k-conjoint, cointégré et augmenté de facteurs communs). À l'aide d'exemples numériques basés sur des données de mortalité des États-Unis, du Royaume-Uni et du Japon pour les deux sexes, et avec une série d'intervalles d'années d'ajustement et trois périodes de prévision, nous montrons que le modèle de crédibilité hiérarchique améliore la qualité des prévisions telle qu'elle est mesurée par la moyenne des erreurs absolues moyennes en pourcentage (AMAPE, en anglais). Le modèle proposé est facile à mettre en œuvre et les taux de mortalité prévus de plusieurs populations peuvent servir à construire un indice de mortalité qui améliorera la tarification des titres indexés sur la mortalité.

Mots-clés : théorie de la crédibilité hiérarchique; théorie de la crédibilité de Bühlmann; modèle de Lee-Carter; modèle de mortalité de plusieurs populations.

1. Introduction

La mortalité est l'un des éléments clés du calcul des primes et des réserves des produits de rentes et d'assurance-vie ainsi que de la tarification des titres indexés sur la mortalité (p. ex., *q-forwards*, obligations de longévité, swaps de longévité, contrats à terme standardisés sur rentes, options sur mortalité et *survivor caps*). Le risque de longévité est devenu un enjeu important dans la société humaine. Au cours des dernières décennies, les taux de mortalité ont connu une amélioration spectaculaire. Étant donné que les contrats de rente viagère et les régimes de retraite s'étendent sur plusieurs décennies, l'augmentation graduelle de l'espérance de vie dans la plupart des pays développés fait en sorte que les fournisseurs de rentes, les programmes de retraite et les systèmes de régime de retraite et de soins de longue durée risquent fort de devoir verser plus de prestations de survivants que prévu aux retraités et aux rentiers. C'est ce qu'on appelle le risque de longévité. Un tel risque pourrait être source de difficultés financières ou d'insolvabilité pour les fournisseurs de rentes, les programmes de retraite et les régimes de sécurité sociale. D'autre part, les assureurs-vie font face à un risque de mortalité en raison de la détérioration de la mortalité due aux catastrophes (p. ex., la pandémie de grippe espagnole de 1918 et le tremblement de terre et le tsunami de 2004 dans l'océan Indien). Pour réduire le risque de mortalité, Swiss Re a émis en décembre 2003 une obligation de mortalité de 400 millions de dollars sur trois ans en cas de catastrophe. Le capital payé à l'investisseur, à l'échéance, dépend d'un indice de mortalité – un indice spécialement

construit de taux de mortalité reflétant les deux sexes et cinq pays (États-Unis, Royaume-Uni, France, Italie et Suisse). On voit donc toute l'urgence de construire des modèles permettant de calculer avec exactitude les taux de mortalité et ainsi améliorer la tarification des polices d'assurance-vie, des produits de rentes, des régimes de retraite, des régimes de sécurité sociale et des titres indexés sur la mortalité.

L'utilisation d'un modèle efficace de mortalité portant sur plusieurs populations facilite la construction d'un indice de mortalité précis sur lequel peut être basée la tarification des dérivés et des titres indexés sur la mortalité pour couvrir les risques de mortalité et de longévité des assureurs-vie et des fournisseurs de rentes. Pour modéliser les taux de mortalité de plusieurs populations, bien que les modèles traditionnels puissent projeter les taux de mortalité séparément pour chacune des populations, ces modèles ne tiennent pas compte de la corrélation entre les populations. Les taux de mortalité des femmes et des hommes dans un pays donné sont fortement corrélés puisque ces deux groupes sont exposés aux mêmes conditions de vie et bénéficient des mêmes services de santé publique et d'éducation ainsi que des mêmes services médicaux. Les taux de mortalité des pays développés ou des provinces/États d'un pays présentent également un différent degré de corrélation en raison d'expositions similaires.

La projection des taux de mortalité et la modélisation de leur caractère aléatoire ont suscité beaucoup d'attention ces dernières décennies. Le modèle de Lee et Carter (1992) est le modèle le plus souvent cité dans les prévisions et les applications de mortalité. Le modèle CBD, proposé par Cairns et coll. (2006), est un autre modèle bien connu servant à modéliser les taux de mortalité des personnes âgées. Tsai et Yang (2015) ont lié chacune des séries de séquences de taux de mortalité par période pendant des années consécutives à celles d'une année de base au moyen d'une régression linéaire, et modélisé chacune des deux séquences résultantes de paramètres d'ordonnée à l'origine et de pente avec une marche aléatoire avec dérive pour prédire les taux de mortalité. Lin et coll. (2015) ont proposé des modèles AR-GARCH pour prévoir les taux de mortalité à un âge donné et ont utilisé une copule pour prendre en compte la dépendance de la mortalité entre les âges. De nombreux prolongements des modèles CBD et de Lee-Carter ont été élaborés. Par exemple, Renshaw et Haberman (2006) ont généralisé le modèle Lee-Carter pour en faire un modèle non linéaire qui prévoit des effets de cohorte et des effets de période spécifiques à l'âge; Li et coll. (2009) ont proposé un prolongement du modèle de Lee-Carter qui produit des prévisions par intervalles plus prudentes du taux central de mortalité en tenant compte des différences individuelles dans chaque cellule âge-période; Plat (2009) a construit un modèle qui combine les belles caractéristiques du modèle CBD et des modèles de Lee-Carter et de Renshaw-Haberman, tout en éliminant leurs inconvénients; Mitchell et coll. (2013) ont produit un modèle en se basant sur l'idée d'une modélisation bilinéaire de l'âge et du temps comme dans le modèle original de Lee-Carter, mais qui modélise la variation du logarithme du taux central de mortalité plutôt que le niveau du taux central de mortalité.

Il existe des prolongements du modèle de Lee-Carter qui tiennent compte de la dépendance entre les taux de mortalité de plusieurs populations. Carter et Lee (1992) ont introduit le modèle k-conjoint de Lee-Carter, qui suppose que les taux de mortalité de deux populations sont régis conjointement par un indice commun variable dans le temps. Li et Lee (2005) ont proposé le modèle de Lee-Carter augmenté de facteurs communs qui comprend des facteurs communs et des facteurs propres à la population. Li et Hardy (2011) ont démontré l'existence d'une cointégration dans les tendances temporelles des taux de mortalité de deux populations

et ils ont modélisé ces tendances par une relation linéaire appelée modèle de Lee-Carter cointégré.

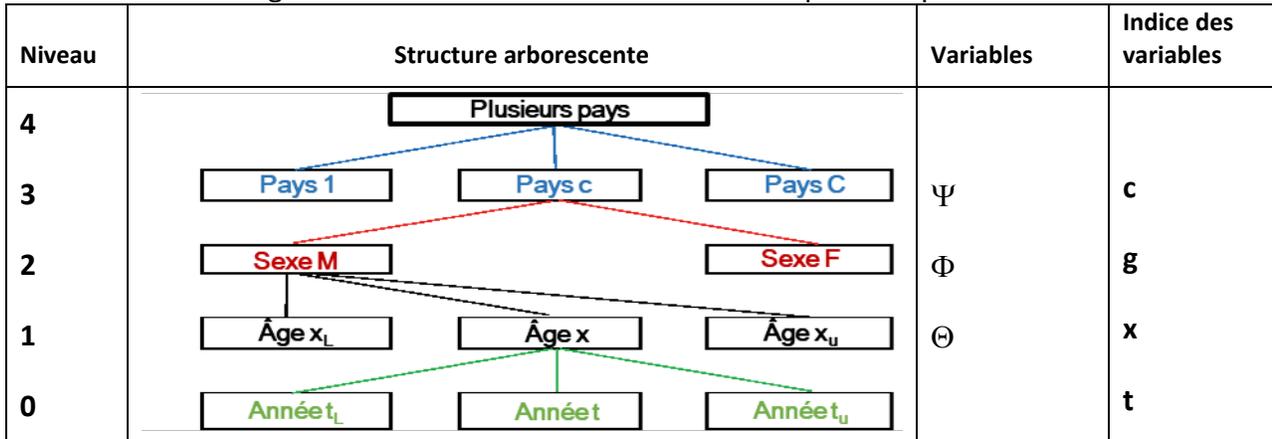
De plus, il y a des articles, non fondés sur le modèle de Lee-Carter, qui étudient la corrélation entre les taux de mortalité de plusieurs populations. Par exemple, Cairns et coll. (2011) ont introduit un nouveau cadre pour modéliser le développement conjoint dans le temps des taux de mortalité d'une paire de populations connexes, avec pour objectif principal de produire des prévisions de mortalité cohérentes pour deux populations. Wang et coll. (2015) ont utilisé des modèles de copule variables dans le temps pour prendre en compte la structure de dépendance de la mortalité à l'échelle des pays, en examinant les structures de dépendance symétrique et asymétrique. Chen et coll. (2015) ont d'abord filtré la dynamique de mortalité de chaque population à l'aide d'un processus ARMA-GARCH avec innovations à queue lourde, puis ils ont modélisé le risque résiduel au moyen d'un modèle de copule à un facteur qui est souvent appliqué aux données à plusieurs dimensions et qui est très flexible en termes de spécification. Li et coll. (2015) ont proposé un processus systématique pour construire un modèle de mortalité à deux populations et ont développé des généralisations à deux populations de chacun des sept modèles à une population étudiés dans Cairns et coll. (2009). Li et coll. (2016) ont introduit un nouveau concept appelé semi-cohérence pour produire des prévisions de mortalité semi-cohérentes à l'aide d'un seuil vecteur autorégressif. Enchev et coll. (2016) ont examiné un certain nombre de modèles de mortalité de plusieurs populations et ont développé des modèles de prévision qui produisent des scénarios de taux de mortalité non divergents et conjoints.

La théorie de la crédibilité est souvent appliquée aux assurances IARD. Bühlmann (1967) a proposé une formule de crédibilité pour déterminer l'estimation de la crédibilité au moyen de la moyenne pondérée de la moyenne de l'échantillon des données de réclamations d'un titulaire de police et de la moyenne réelle des réclamations. Bühlmann et Straub (1970) ont prolongé le modèle de Bühlmann en permettant un nombre inégal d'unités d'exposition pour chaque risque. Les données de réclamations peuvent représenter des montants ou des fréquences et peuvent être traitées pour tous les titulaires d'un groupe comme une structure arborescente à trois niveaux. Tsai et Lin (2017a) ont intégré l'approche de crédibilité de Bühlmann aux modèles CBD, de Lee-Carter et de Tsai et Yang (2015) pour améliorer la qualité des prévisions de la mortalité. Tsai et Lin (2017b) ont proposé une approche de crédibilité non paramétrique de Bühlmann pour modéliser les taux de mortalité et ont démontré que cette méthode surpassait le modèle de Lee-Carter du point de vue de la qualité des prévisions.

Bühlmann et Gisler (2005) ont intégré l'idée de structure hiérarchique à la théorie de la crédibilité pour obtenir le modèle de crédibilité hiérarchique. On constate que la crédibilité hiérarchique est une généralisation du modèle de Bühlmann et du modèle de Bühlmann-Straub avec une structure arborescente comportant plus de niveaux. Les modèles de mortalité traditionnels pour une seule population regroupent les données de mortalité en trois niveaux (population, âge et année). Dans le présent document, nous proposons une approche de crédibilité hiérarchique pour modéliser les taux de mortalité de plusieurs populations. Nous regroupons les données de mortalité des deux sexes, de trois pays développés, en quatre niveaux (pays, sexe, âge et année) ou en cinq niveaux (plusieurs pays, pays, sexe, âge et année), et nous appliquons la crédibilité hiérarchique à chaque niveau pour mieux refléter la corrélation des taux de mortalité entre les populations ou les sexes/pays, puis nous comparons la qualité prévisionnelle des modèles de mortalité sous-jacents.

Le reste du document se divise comme suit. La section 2 présente les approches de crédibilité hiérarchique à quatre et à cinq niveaux pour modéliser les taux de mortalité de plusieurs populations, avant de les appliquer à la projection des taux de mortalité de six populations (les deux sexes aux États-Unis, au Royaume-Uni et au Japon). La section 3 compare la qualité prévisionnelle des modèles hiérarchiques proposés à trois, à quatre et à cinq niveaux et du modèle de Lee-Carter et de ses trois variantes (k-conjoint, cointégré, et augmenté de facteurs communs; voir l'annexe A) pour plusieurs populations. La section 4 présente les conclusions.

Figure 1 : Structure arborescente hiérarchique à cinq niveaux



2. Modèle de mortalité selon la crédibilité hiérarchique de Bühlmann

Nous appliquons ici l'idée de crédibilité hiérarchique de Bühlmann et de Gisler (2005) à la proposition d'un modèle de mortalité selon la crédibilité hiérarchique pour plusieurs populations, qui est une généralisation du modèle non paramétrique de Bühlmann pour une seule population, proposée par Tsai et Lin (2017b).

La figure 1 présente les données de mortalité suivant un arbre hiérarchique à cinq niveaux, où le niveau le plus élevé représente plusieurs pays, C pays précisément. À chaque pays correspondent G sexes ($G = 2$) et à chaque sexe correspondent des âges consécutifs x_L, \dots, x_U . Enfin, à chaque âge correspondent des données annuelles allant de l'année t_L à l'année t_U , qui est le niveau inférieur. Soit $m_{c,g,x,t}$ le taux central de mortalité correspondant au pays c , au sexe g , à l'âge x et à l'année t . Le modèle de Lee-Carter et ses trois variantes pour plusieurs populations utilisent $\ln(m_{c,g,x,t})$ pour modéliser les taux de mortalité. La figure 2 montre que les $\ln(m_{c,g,x,t})$ historiques des É.-U., du R.-U. et du Japon, tirés de la Human Mortality Base, affichent une tendance à la baisse sur les années $t = 1950, \dots, 2010$ pour $x = 25, 50, 75$ et Avg, où la courbe Avg représente la moyenne des $\ln(m_{c,g,x,t})$ sur $x = 20, \dots, 84$. À l'exemple du modèle de crédibilité de Bühlmann proposé par Tsai et Lin (2017a, b), nous appliquons l'approche de crédibilité hiérarchique à la modélisation de $Y_{c,g,x,t} \triangleq \ln(m_{c,g,x,t}) - \ln(m_{c,g,x,t})$, le décrément du logarithme du taux central de mortalité correspondant au pays c , au sexe g et à l'âge x sur $[t-1, t]$, afin d'éliminer la tendance à la baisse (figure 3). Vu que nous utiliserons les indices $x = 1, \dots, X$ et $t = 1, \dots, T$ pour simplifier les notations, les $\ln(m_{c,g,x,t})$ donnés dans un rectangle âge-année $[x_L, x_U] \times [t_L, t_U]$ pour une population du pays c et de sexe g , l'intervalle d'âges $[x_L, x_U]$ et l'intervalle d'années $[t_L, t_U]$ pour $\ln(m_{c,g,x,t})$ correspondent à $[1, X]$ et $[1, T]$ pour $Y_{c,g,x,t}$ respectivement; c'est-à-dire que $Y_{c,g,x,t} = \ln(m_{c,g,x_L+x-1, t_L+t}) - \ln(m_{c,g,x_L+x-1, t_L+t-1})$ pour $x = 1, \dots, X$ et $t = 1, \dots, T$, où $X - 1 = x_U - x_L$ et $T = t_U - t_L$.

Figure 2 : $\ln(m_{x,t,i})$ en fonction de t

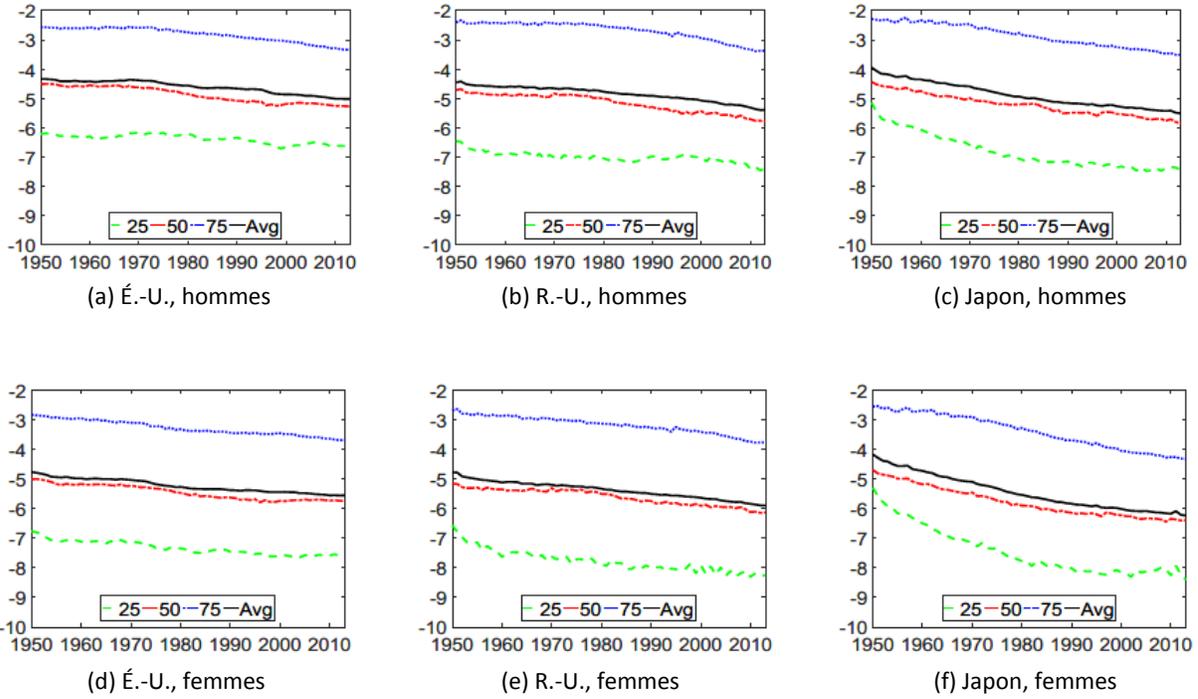
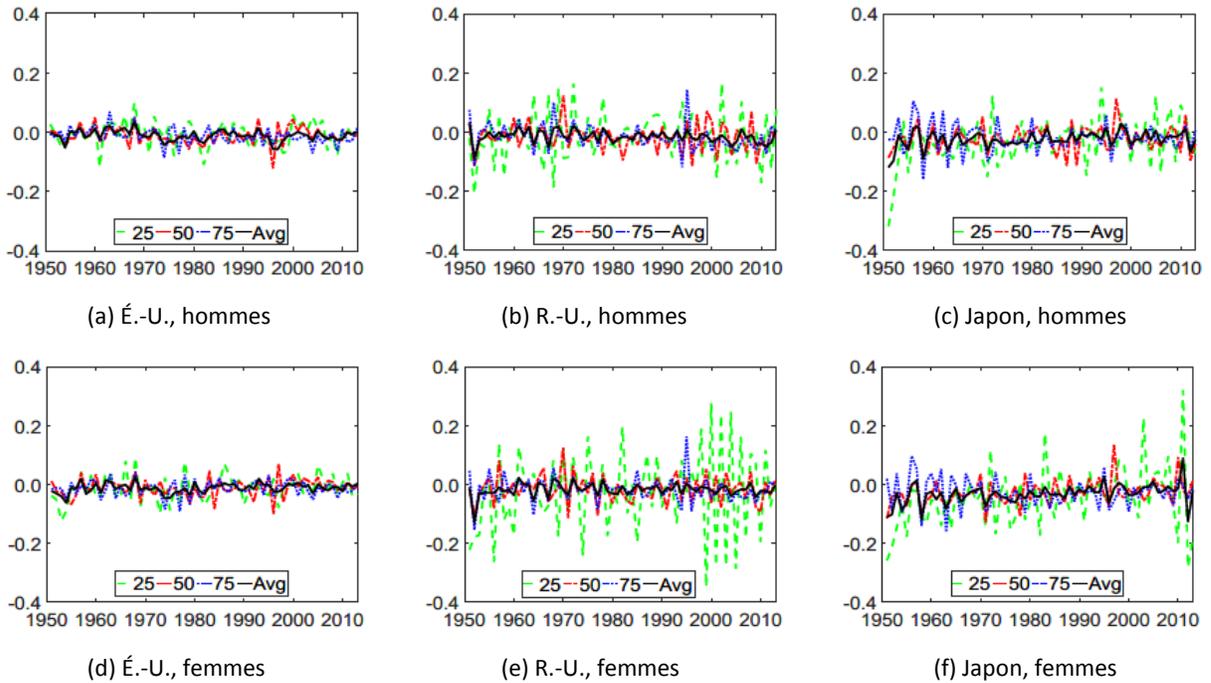


Figure 3 : $Y_{x,t,i} = \ln(m_{x,t,i}) - \ln(m_{x,t-1,i})$ en fonction de t



2.1 Hypothèses et notations

Nous indiquons d'abord les hypothèses, puis les notations pour les moyennes conditionnelles, la moyenne, les variances conditionnelles et les variances pour un arbre hiérarchique. Le modèle hiérarchique repose sur les hypothèses suivantes :

- Niveau 0 : $Y_{c,g,x,t} | \Theta_{c,g,x}, t, = 1, \dots, T$, sont indépendantes et identiquement distribuées pour des valeurs fixes de c, g et x ;
- Niveau 1 : $\Theta_{c,g,x} | \Phi_{c,g}, x = 1, \dots, X$, sont indépendantes et identiquement distribuées pour des valeurs fixes de c et g ;
- Niveau 2 : $\Phi_{c,g} | \Psi_c, g = 1, \dots, G$, sont indépendantes et identiquement distribuées pour une valeur fixe de c ;
- Niveau 3 : $\Psi_c, c = 1, \dots, C$, sont indépendantes et identiquement distribuées.

Tout d'abord, nous dénotons comme suit les moyennes conditionnelles et la moyenne pour les niveaux 0 à 3 :

- Niveau 0 : $\mu_1(\Theta_{c,g,x}) \triangleq E(Y_{c,g,x,t} | \Theta_{c,g,x})$, l'espérance conditionnelle de $Y_{c,g,x,t}$ sachant $\Theta_{c,g,x}$;
- Niveau 1 : $\mu_2(\Phi_{c,g}) \triangleq E[\mu_1(\Theta_{c,g,x}) | \Phi_{c,g}] = E[E(Y_{c,g,x,t} | \Theta_{c,g,x}) | \Phi_{c,g}]$, l'espérance conditionnelle de $\mu_1(\Theta_{c,g,x})$ sachant $\Phi_{c,g}$;
- Niveau 2 : $\mu_3(\Psi_c) \triangleq E[\mu_2(\Phi_{c,g}) | \Psi_c] = E\{E[\mu_1(\Theta_{c,g,x}) | \Phi_{c,g}] | \Psi_c\} = E\{E[E(Y_{c,g,x,t} | \Theta_{c,g,x}) | \Phi_{c,g}] | \Psi_c\}$, l'espérance conditionnelle de $\mu_2(\Phi_{c,g})$ sachant Ψ_c ;
- Niveau 3 : $\mu_4 \triangleq E[\mu_3(\Psi_c)] = E\{E[\mu_2(\Phi_{c,g}) | \Psi_c]\} = E\{E\{E[\mu_1(\Theta_{c,g,x}) | \Phi_{c,g}] | \Psi_c\}\} = E\{E\{E[E(Y_{c,g,x,t} | \Theta_{c,g,x}) | \Phi_{c,g}] | \Psi_c\}\}$, l'espérance de $\mu_3(\Psi_c)$.

Suivant le théorème de l'espérance totale, il est facile de démontrer que $\mu_4 = E[Y_{c,g,x,t}]$, $\mu_3(\Psi_c) = E[Y_{c,g,x,t} | \Psi_c]$, et $\mu_2(\Phi_{c,g}) = E[Y_{c,g,x,t} | \Phi_{c,g}]$.

Ensuite, nous dénotons les variances conditionnelles suivantes pour les niveaux 0 à 2 :

- Niveau 0 : $\frac{\sigma_1^2(\Theta_{c,g,x})}{w_{c,g,x,t}} \triangleq Var[Y_{c,g,x,t} | \Theta_{c,g,x}] = E\{[Y_{c,g,x,t} - \mu_1(\Theta_{c,g,x})]^2 | \Theta_{c,g,x}\}$, la variance conditionnelle de $Y_{c,g,x,t}$ sachant $\Theta_{c,g,x}$, où $w_{c,g,x,t}$ est une unité d'exposition connue;
- Niveau 1 : $\sigma_2^2(\Phi_{c,g}) \triangleq Var[\mu_1(\Theta_{c,g,x}) | \Phi_{c,g}] = E\{[\mu_1(\Theta_{c,g,x}) - \mu_2(\Phi_{c,g})]^2 | \Phi_{c,g}\}$, la variance conditionnelle de $\mu_1(\Theta_{c,g,x})$ sachant $\Phi_{c,g}$;
- Niveau 2 : $\sigma_3^2(\Psi_c) \triangleq Var[\mu_2(\Phi_{c,g}) | \Psi_c] = E\{[\mu_2(\Phi_{c,g}) - \mu_3(\Psi_c)]^2 | \Psi_c\}$, la variance conditionnelle de $\mu_2(\Phi_{c,g})$ sachant Ψ_c .

Enfin, nous dénotons comme suit les variances conditionnelles attendues pour les niveaux 0 à 2 et la variance pour le niveau 3 :

- Niveau 0 : $\sigma_1^2 \triangleq E[\sigma_1^2(\Theta_{c,g,x})] = E\{w_{c,g,x,t} \cdot Var[Y_{c,g,x,t} | \Theta_{c,g,x}]\}$, l'espérance de $w_{c,g,x,t}$ multipliée par la variance conditionnelle de $Y_{c,g,x,t}$ sachant $\Theta_{c,g,x}$;
- Niveau 1 : $\sigma_2^2 \triangleq E[\sigma_2^2(\Phi_{c,g})] = E\{Var[\mu_1(\Theta_{c,g,x}) | \Phi_{c,g}]\}$, l'espérance de la variance conditionnelle de $\mu_1(\Theta_{c,g,x})$ sachant $\Phi_{c,g}$;

- Niveau 2 : $\sigma_3^2 \triangleq E[\sigma_3^2(\Psi_c)] = E\{Var[\mu_2(\Phi_{c,g})|\Psi_c]\}$, l'espérance de la variance conditionnelle de $\mu_2(\Phi_{c,g})$ sachant Ψ_c ;
- Niveau 3 : $\sigma_4^2 \triangleq Var[\mu_3(\Psi_c)] = E\{[\mu_3(\Psi_c) - \mu_4]^2\}$, la variance de $\mu_3(\Psi_c)$.

À noter que, suivant le théorème de la variance totale, nous avons

$$\begin{aligned} Var[\mu_2(\Phi_{c,g})] &= E\{Var[\mu_2(\Phi_{c,g})|\Psi_c]\} + Var\{E[\mu_2(\Phi_{c,g})|\Psi_c]\} \\ &= \sigma_3^2 + Var[\mu_3(\Psi_c)] = \sigma_3^2 + \sigma_4^2, \end{aligned}$$

et

$$\begin{aligned} Var[\mu_1(\Theta_{c,g,x})] &= E\{Var[\mu_1(\Theta_{c,g,x})|\Phi_{c,g}]\} + Var\{E[\mu_1(\Theta_{c,g,x})|\Phi_{c,g}]\} \\ &= \sigma_2^2 + Var[\mu_2(\Phi_{c,g})] = \sigma_2^2 + \sigma_3^2 + \sigma_4^2. \end{aligned}$$

Il est évident que la structure arborescente du modèle de crédibilité hiérarchique recouvre celle du modèle de crédibilité de Bühlmann-Straub, qui peut être obtenue en appliquant une structure arborescente à trois niveaux avec seulement les niveaux zéro, un et deux de la figure 1 pour une seule population du pays c et de sexe g . Pour une population du pays c et de sexe g , si par surcroît $w_{c,g,x,t} = 1$ pour $x = 1, \dots, X$ et $t = 1, \dots, T$, le modèle de crédibilité de Bühlmann-Straub se réduit au modèle de crédibilité de Bühlmann.

2.2 Estimations des paramètres

D'après la section 6.4 de Bühlmann et Gisler (2005), les données comprimées B et les facteurs de crédibilité α sont calculés du niveau inférieur au niveau supérieur en minimisant les fonctions de projection associées de haut en bas.

Estimateurs de crédibilité

Pour trouver l'estimateur de crédibilité $\hat{Y}_{c,g,x,T+1}$, correspondant à l'âge x , au sexe g et au pays c et à l'année $T + 1$, nous devons tout d'abord trouver les estimateurs de crédibilité $\hat{Y}_{c,g,x}$ et $\hat{Y}_{c,g}$. Tous les estimateurs de crédibilité peuvent s'écrire comme des combinaisons linéaires de la moyenne globale $\hat{\mu}_4$ et des données comprimées B , comme suit : $\hat{Y}_{c,g} = \alpha_c^{(3)} \cdot B_c^{(3)} + (1 - \alpha_c^{(3)}) \cdot \hat{\mu}_4$,

$$\begin{aligned} \hat{Y}_{c,g,x} &= \alpha_{c,g}^{(2)} \cdot B_{c,g}^{(2)} + (1 - \alpha_{c,g}^{(2)}) \cdot \hat{Y}_{c,g} \\ &= \alpha_{c,g}^{(2)} \cdot B_{c,g}^{(2)} + [(1 - \alpha_{c,g}^{(2)}) \cdot \alpha_c^{(3)}] \cdot B_c^{(3)} + [(1 - \alpha_{c,g}^{(2)}) \cdot (1 - \alpha_c^{(3)})] \cdot \hat{\mu}_4, \end{aligned}$$

et

$$\begin{aligned} &\hat{Y}_{c,g,x,T+1} \\ &= \alpha_{c,g,x}^{(1)} \cdot B_{c,g,x}^{(1)} + (1 - \alpha_{c,g,x}^{(1)}) \cdot \hat{Y}_{c,g,x} \\ &= \alpha_{c,g,x}^{(1)} \cdot B_{c,g,x}^{(1)} + [(1 - \alpha_{c,g,x}^{(1)}) \cdot \alpha_{c,g}^{(2)}] \cdot B_{c,g}^{(2)} + [(1 - \alpha_{c,g,x}^{(1)}) \cdot (1 - \alpha_{c,g}^{(2)})] \cdot \hat{Y}_{c,g} \\ &= \alpha_{c,g,x}^{(1)} \cdot B_{c,g,x}^{(1)} + [(1 - \alpha_{c,g,x}^{(1)}) \cdot \alpha_{c,g}^{(2)}] \cdot B_{c,g}^{(2)} + [(1 - \alpha_{c,g,x}^{(1)}) \cdot (1 - \alpha_{c,g}^{(2)}) \cdot \alpha_c^{(3)}] \cdot B_c^{(3)} \\ &\quad + [(1 - \alpha_{c,g,x}^{(1)}) \cdot (1 - \alpha_{c,g}^{(2)}) \cdot (1 - \alpha_c^{(3)})] \cdot \hat{\mu}_4, \end{aligned} \tag{2.1}$$

où $B_{c,g,x}^{(1)}$, $B_{c,g}^{(2)}$ et $B_c^{(3)}$ représentent des données comprimées; $\alpha_{c,g,x}^{(1)}$, $\alpha_{c,g}^{(2)}$ et $\alpha_c^{(3)}$ sont les facteurs de crédibilité correspondants pour les niveaux un, deux et trois, respectivement. Les

expressions pour les données comprimées ($B_{c,g,x}^{(1)}$, $B_{c,g}^{(2)}$ et $B_c^{(3)}$), les facteurs de crédibilité ($\alpha_{c,g,x}^{(1)}$, $\alpha_{c,g}^{(2)}$ et $\alpha_c^{(3)}$) et $\hat{\mu}_4$ selon le théorème 6.4 de Bühlmann et Gisler (2005) sont indiqués dans la partie A du tableau 1.

Estimation des paramètres structurels

Tableau 1 Estimation de crédibilité hiérarchique de σ_1^2 , σ_2^2 , σ_3^2 et σ_4^2

Partie A : Formules pour $B_{c,g,x}^{(1)}$, $B_{c,g}^{(2)}$, $B_c^{(3)}$, $\alpha_{c,g,x}^{(1)}$, $\alpha_{c,g}^{(2)}$, $\alpha_c^{(3)}$ et $\hat{\mu}_4$		
$B_{c,g,x}^{(1)} = \sum_{t=1}^T \frac{w_{c,g,x,t}}{w_{c,g,x}^{(1)}} \cdot Y_{c,g,x,t}$,	$w_{c,g,x}^{(1)} = \sum_{t=1}^T w_{c,g,x,t}$,	$\alpha_{c,g,x}^{(1)} = \frac{w_{c,g,x}^{(1)} \cdot \sigma_2^2}{w_{c,g,x}^{(1)} \cdot \sigma_2^2 + \sigma_1^2}$
$B_{c,g}^{(2)} = \sum_{x=1}^X \frac{\alpha_{c,g,x}^{(1)}}{w_{c,g}^{(2)}} \cdot B_{c,g,x}^{(1)}$,	$w_{c,g}^{(2)} = \sum_{x=1}^X \alpha_{c,g,x}^{(1)}$,	$\alpha_{c,g}^{(2)} = \frac{w_{c,g}^{(2)} \cdot \sigma_3^2}{w_{c,g}^{(2)} \cdot \sigma_3^2 + \sigma_2^2}$
$B_c^{(3)} = \sum_{g=1}^G \frac{\alpha_{c,g}^{(2)}}{w_c^{(3)}} \cdot B_{c,g}^{(2)}$,	$w_c^{(3)} = \sum_{g=1}^G \alpha_{c,g}^{(2)}$,	$\alpha_c^{(3)} = \frac{w_c^{(3)} \cdot \sigma_4^2}{w_c^{(3)} \cdot \sigma_4^2 + \sigma_3^2}$
$\hat{\mu}_4 = \sum_{c=1}^C \frac{\alpha_c^{(3)}}{w^{(4)}} \cdot B_c^{(3)}$,	$w^{(4)} = \sum_{c=1}^C \alpha_c^{(3)}$.	
Partie B : Estimation de crédibilité hiérarchique de σ_1^2 , σ_2^2 , σ_3^2 et σ_4^2		
$\sigma_1^2 = E[\sigma_1^2(\Theta_{c,g,x})]$	$\sigma_1^2(\Theta_{c,g,x}) = w_{c,g,x,t} \cdot Var[Y_{c,g,x,t} \Theta_{c,g,x}]$	
$\hat{\sigma}_1^2 = \frac{1}{C \cdot G \cdot X} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \hat{\sigma}_1^2(\Theta_{c,g,x})$	$\hat{\sigma}_1^2(\Theta_{c,g,x}) \triangleq \frac{1}{T-1} \sum_{t=1}^T w_{c,g,x,t} \cdot [Y_{c,g,x,t} - B_{c,g,x}^{(1)}]^2$	
$\sigma_2^2 = E\{Var[\mu_1(\Theta_{c,g,x}) \Phi_{c,g}]\}$	$z_{c,g}^{(1)} = \sum_{x=1}^X w_{c,g,x}^{(1)}$, $\bar{B}_{c,g,\bullet}^{(1)} = \sum_{x=1}^X \frac{w_{c,g,x}^{(1)}}{z_{c,g}^{(1)}} \cdot B_{c,g,x}^{(1)}$ $c_{c,g}^{(1)} = \frac{X-1}{X} \cdot \left\{ \sum_{x=1}^X \frac{w_{c,g,x}^{(1)}}{z_{c,g}^{(1)}} \cdot \left[1 - \frac{w_{c,g,x}^{(1)}}{z_{c,g}^{(1)}} \right] \right\}^{-1}$	
$\hat{\sigma}_2^2 = \frac{1}{C \cdot G} \sum_{c=1}^C \sum_{g=1}^G \max[\hat{T}_{c,g}^{(1)}, 0]$	$\hat{T}_{c,g}^{(1)} = c_{c,g}^{(1)} \cdot \left\{ \frac{X}{X-1} \sum_{x=1}^X \frac{w_{c,g,x}^{(1)}}{z_{c,g}^{(1)}} \left[B_{c,g,x}^{(1)} - \bar{B}_{c,g,\bullet}^{(1)} \right]^2 - \frac{X}{z_{c,g}^{(1)}} \cdot \hat{\sigma}_1^2 \right\}$	
$\sigma_3^2 = E\{Var[\mu_2(\Phi_{c,g}) \Psi_c]\}$	$z_c^{(2)} = \sum_{g=1}^G w_{c,g}^{(2)}$, $\bar{B}_{c,\bullet}^{(2)} = \sum_{g=1}^G \frac{w_{c,g}^{(2)}}{z_c^{(2)}} \cdot B_{c,g}^{(2)}$ $c_c^{(2)} = \frac{G-1}{G} \cdot \left\{ \sum_{g=1}^G \frac{w_{c,g}^{(2)}}{z_c^{(2)}} \cdot \left[1 - \frac{w_{c,g}^{(2)}}{z_c^{(2)}} \right] \right\}^{-1}$	
$\hat{\sigma}_3^2 = \frac{1}{C} \sum_{c=1}^C \max[\hat{T}_c^{(2)}, 0]$	$\hat{T}_c^{(2)} = c_c^{(2)} \cdot \left\{ \frac{G}{G-1} \sum_{g=1}^G \frac{w_{c,g}^{(2)}}{z_c^{(2)}} \left[B_{c,g}^{(2)} - \bar{B}_{c,\bullet}^{(2)} \right]^2 - \frac{G}{z_c^{(2)}} \cdot \hat{\sigma}_2^2 \right\}$	
$\sigma_4^2 = Var[\mu_3(\Psi_c)]$	$z_c^{(3)} = \sum_{g=1}^G w_c^{(3)}$, $\bar{B}_{c,\bullet}^{(3)} = \sum_{g=1}^G \frac{w_c^{(3)}}{z_c^{(3)}} \cdot B_c^{(3)}$ $c_c^{(3)} = \frac{C-1}{C} \cdot \left\{ \sum_{g=1}^G \frac{w_c^{(3)}}{z_c^{(3)}} \cdot \left[1 - \frac{w_c^{(3)}}{z_c^{(3)}} \right] \right\}^{-1}$	
$\hat{\sigma}_4^2 = \max[\hat{T}_c^{(3)}, 0]$	$\hat{T}_c^{(3)} = c_c^{(3)} \cdot \left\{ \frac{C}{C-1} \sum_{g=1}^G \frac{w_c^{(3)}}{z_c^{(3)}} \left[B_c^{(3)} - \bar{B}_{c,\bullet}^{(3)} \right]^2 - \frac{C}{z_c^{(3)}} \cdot \hat{\sigma}_3^2 \right\}$	

À noter que, dans le cas d'une structure arborescente à quatre niveaux, nous pouvons poser $\alpha_c^{(3)} = 1$ et n'avons pas besoin de calculer $\hat{\mu}_4$ et $w^{(4)}$ dans la partie A du tableau 1. De même, pour obtenir la formule correspondante (2.1) de la structure arborescente à trois niveaux, qui est le modèle de crédibilité de Bühlmann-Straub, nous posons $\alpha_c^{(3)} = \alpha_{c,g}^{(2)} = 1$ et il n'est pas nécessaire de calculer $B_c^{(3)}$, $w_c^{(3)}$, $\hat{\mu}_4$ et $w^{(4)}$.

Pour obtenir l'estimation de crédibilité $\hat{Y}_{c,g,x,T+1}$, en (2.1), il nous faut les estimations $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, $\hat{\sigma}_3^2$ et $\hat{\sigma}_4^2$. L'estimation de σ_1^2 , σ_2^2 , σ_3^2 et σ_4^2 est indiquée dans la partie B du tableau 1. Pour une estimation détaillée, consultez la section 6.6 de Bühlmann et Gisler (2005). En regardant de près les formules du tableau 1, nous constatons que les valeurs de $\hat{\sigma}_2^2$, $\hat{\sigma}_3^2$ et $\hat{\sigma}_4^2$ peuvent être nulles, ce qui rendrait nulles les valeurs de $\hat{\alpha}_{c,g,x}^{(1)}$, $\hat{\alpha}_{c,g}^{(2)}$ et $\hat{\alpha}_c^{(3)}$, respectivement. Puisque les paramètres structurels sont estimés du bas vers le haut de la structure arborescente, il est facile d'étendre l'analyse à une structure comportant des niveaux supérieurs. Toutefois, une structure hiérarchique arborescente comportant des niveaux supérieurs possède davantage de paramètres structurels. Comme l'a suggéré Bühlmann et Gisler (2005), il faut choisir avec soin le nombre de niveaux dans le modèle de crédibilité hiérarchique.

Ensuite, nous présentons un cas particulier où toutes les unités d'exposition connues, $w_{c,g,x,t}$ sont égales à 1. Nous utiliserons ce cas particulier pour notre modèle de crédibilité hiérarchique appliqué à la mortalité.

Un cas particulier

Si $w_{c,g,x,t} = 1$ pour tous $c = 1, \dots, C$, $g = 1, \dots, G$, $x = 1, \dots, X$, et $t = 1, \dots, T$, les quantités de la partie A du tableau 1 se réduisent à

$$w_{c,g,x}^{(1)} = T \triangleq w^{(1)}, \quad B_{c,g,x}^{(1)} = \frac{1}{T} \sum_{t=1}^T Y_{c,g,x,t} \triangleq \bar{Y}_{c,g,x,\bullet},$$

$$\hat{\alpha}_{c,g,x}^{(1)} = \frac{T \cdot \hat{\sigma}_2^2}{T \cdot \hat{\sigma}_2^2 + \hat{\sigma}_1^2} \triangleq \hat{\alpha}^{(1)}; \quad (2.2)$$

$$w_{c,g}^{(2)} = X \cdot \hat{\alpha}^{(1)}, \quad B_{c,g}^{(2)} = \frac{1}{X} \sum_{x=1}^X B_{c,g,x}^{(1)} = \frac{1}{X \cdot T} \sum_{x=1}^X \sum_{t=1}^T Y_{c,g,x,t} \triangleq \bar{Y}_{c,g,\bullet,\bullet},$$

$$\hat{\alpha}_{c,g}^{(2)} = \frac{X \cdot \hat{\alpha}^{(1)} \cdot \hat{\sigma}_3^2}{X \cdot \hat{\alpha}^{(1)} \cdot \hat{\sigma}_3^2 + \hat{\sigma}_2^2} \triangleq \hat{\alpha}^{(2)}; \quad (2.3)$$

$$w_c^{(3)} = G \cdot \hat{\alpha}^{(2)}, \quad B_c^{(3)} = \frac{1}{G} \sum_{g=1}^G B_{c,g}^{(2)} = \frac{1}{G \cdot X \cdot T} \sum_{g=1}^G \sum_{x=1}^X \sum_{t=1}^T Y_{c,g,x,t} \triangleq \bar{Y}_{c,\bullet,\bullet,\bullet},$$

$$\hat{\alpha}_c^{(3)} = \frac{G \cdot \hat{\alpha}^{(2)} \cdot \hat{\sigma}_4^2}{G \cdot \hat{\alpha}^{(2)} \cdot \hat{\sigma}_4^2 + \hat{\sigma}_3^2} \triangleq \hat{\alpha}^{(3)}, \quad (2.4)$$

$$w^{(4)} = C \cdot \hat{\alpha}^{(3)}, \quad \hat{\mu}_4 = \frac{1}{C} \sum_{c=1}^C B_c^{(3)} = \frac{1}{C \cdot G \cdot X \cdot T} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \sum_{t=1}^T Y_{c,g,x,t} \triangleq \bar{Y}_{\bullet,\bullet,\bullet,\bullet}.$$

De plus, les quantités se réduisent à, pour l'estimation de

- σ_1^2 : $B_{c,g,x} = Y_{c,g,x,\bullet}$, et

$$\hat{\sigma}_1^2(\Theta_{c,g,x}) = \frac{1}{T-1} \sum_{t=1}^T (Y_{c,g,x,t} - \bar{Y}_{c,g,x,\bullet})^2;$$

- σ_2^2 : $w_{c,g,x} = T$, $z_{c,g} = X \cdot T$, $\bar{B}_{c,g,\bullet} = \bar{Y}_{c,g,\bullet,\bullet}$, $c_{c,g}^{(1)} = 1$, et

$$\hat{T}_{c,g}^{(1)} = \frac{1}{X-1} \sum_{x=1}^X (\bar{Y}_{c,g,x,\bullet} - \bar{Y}_{c,g,\bullet,\bullet})^2 - \frac{\hat{\sigma}_1^2}{T};$$

- σ_3^2 : $w_{c,g} = X \cdot \alpha^{(1)}$, $z_c^{(2)} = G \cdot X \cdot \alpha^{(1)}$, $B_{c,g}^{(2)} = \bar{Y}_{c,g,\bullet,\bullet}$, $\bar{B}_{c,\bullet}^{(2)} = \bar{Y}_{c,\bullet,\bullet,\bullet}$, $c_c^{(2)} = 1$, et

$$\hat{T}_c^{(2)} = \frac{1}{G-1} \sum_{g=1}^G (\bar{Y}_{c,g,\bullet,\bullet} - \bar{Y}_{c,\bullet,\bullet,\bullet})^2 - \left[\frac{\hat{\sigma}_2^2}{X} + \frac{\hat{\sigma}_1^2}{X \cdot T} \right];$$

- σ_4^2 : $w_c^{(3)} = G \cdot \alpha^{(2)}$, $z^{(3)} = C \cdot G \cdot \alpha^{(2)}$, $B_c^{(3)} = \bar{Y}_{c,\bullet,\bullet,\bullet}$, $\bar{B}_{\bullet}^{(3)} = \bar{Y}_{\bullet,\bullet,\bullet,\bullet}$, $c^{(3)} = 1$, et

$$\hat{T}^{(3)} = \frac{1}{C-1} \sum_{c=1}^C (\bar{Y}_{c,\bullet,\bullet,\bullet} - \bar{Y}_{\bullet,\bullet,\bullet,\bullet})^2 - \left[\frac{\hat{\sigma}_3^2}{G} + \frac{\hat{\sigma}_2^2}{G \cdot X} + \frac{\hat{\sigma}_1^2}{G \cdot X \cdot T} \right]$$

L'estimation de crédibilité hiérarchique du décrement logarithmique du taux central de mortalité pour le pays c , le sexe g et l'âge x sur $[T, T+1]$ dans (2.1), pour ce cas particulier, devient

$$\begin{aligned} & \hat{Y}_{c,g,x,T+1} \\ = & \hat{\alpha}^{(1)} \cdot \bar{Y}_{c,g,x,\bullet} + [(1 - \hat{\alpha}^{(1)}) \cdot \hat{\alpha}^{(2)}] \cdot \bar{Y}_{c,g,\bullet,\bullet} + [(1 - \hat{\alpha}^{(1)}) \cdot (1 - \hat{\alpha}^{(2)}) \cdot \hat{\alpha}^{(3)}] \cdot \bar{Y}_{c,\bullet,\bullet,\bullet} \\ & + [(1 - \hat{\alpha}^{(1)}) \cdot (1 - \hat{\alpha}^{(2)}) \cdot (1 - \hat{\alpha}^{(3)})] \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}. \end{aligned} \quad (2.5)$$

À noter que $\hat{Y}_{c,g,x,T+1}$ représente la moyenne, pondérée des facteurs de crédibilité, de

- $\bar{Y}_{c,g,x,\bullet} = \frac{1}{T} \sum_{t=1}^T Y_{c,g,x,t} = \frac{1}{T} \sum_{t=1}^T [\ln(m_{c,g,x,t-1})]$ (le décrement annuel moyen de la tendance temporelle **individuelle** $\{\ln(m_{c,g,x,t}) : t = 0, \dots, T\}$ sur $[0, T]$ pour une personne âgée de x et de sexe g dans le pays c);
- $\bar{Y}_{c,g,\bullet,\bullet} = \frac{1}{T} \sum_{t=1}^T \bar{Y}_{c,g,\bullet,t} = \frac{1}{T} \sum_{t=1}^T [\ln(m_{c,g,\bullet,t}) - \ln(m_{c,g,\bullet,t-1})]$ (le décrement annuel moyen de la tendance temporelle de la **population** $\{\ln(m_{c,g,\bullet,t}) = \frac{1}{X} \sum_{x=1}^X \ln(m_{c,g,x,t}) : t = 0, \dots, T\}$ sur $[0, T]$ pour une population de sexe g dans le pays c);
- $\bar{Y}_{c,\bullet,\bullet,\bullet} = \frac{1}{T} \sum_{t=1}^T \bar{Y}_{c,\bullet,\bullet,t} = \frac{1}{T} \sum_{t=1}^T [\ln(m_{c,\bullet,\bullet,t}) - \ln(m_{c,\bullet,\bullet,t-1})]$ (le décrement annuel moyen de la tendance temporelle du **pays** $\{\ln(m_{c,\bullet,\bullet,t}) = \frac{1}{G} \sum_{g=1}^G \ln(m_{c,g,\bullet,t}) : t = 0, \dots, T\}$ sur $[0, T]$ dans le pays c);
- $\bar{Y}_{\bullet,\bullet,\bullet,\bullet} = \frac{1}{T} \sum_{t=1}^T \bar{Y}_{\bullet,\bullet,\bullet,t} = \frac{1}{T} \sum_{t=1}^T [\ln(m_{\bullet,\bullet,\bullet,t}) - \ln(m_{\bullet,\bullet,\bullet,t-1})]$ (le décrement annuel moyen de la tendance temporelle de **plusieurs pays** $\{\ln(m_{\bullet,\bullet,\bullet,t}) = \frac{1}{C} \sum_{c=1}^C [\ln(m_{c,\bullet,\bullet,t})] : t = 0, \dots, T\}$ sur $[0, T]$ pour la totalité des C pays).

Nous utilisons (g, x, t) pour désigner le sexe g et l'âge x dans l'année t pour une structure hiérarchique à quatre niveaux appliquée à un pays. Les paramètres structurels d'un arbre hiérarchique à quatre niveaux peuvent aussi être estimés avec ceux d'un arbre hiérarchique à cinq niveaux en posant $C = 1$ et $\hat{\alpha}^{(3)} = 1$, et nous n'avons pas besoin de calculer $\hat{T}^{(3)}$ et $\hat{\sigma}_3^2$. L'estimation de crédibilité hiérarchique du décretement logarithmique du taux central de mortalité pour le sexe g et l'âge x sur $[T, T + 1]$, selon le cas particulier, est ensuite donnée par

$$\hat{Y}_{g,x,T+1} = \hat{\alpha}^{(1)} \cdot \bar{Y}_{g,x,\bullet} + [(1 - \hat{\alpha}^{(1)}) \cdot \hat{\alpha}^{(2)}] \cdot \bar{Y}_{g,\bullet,\bullet} + [(1 - \hat{\alpha}^{(1)}) \cdot (1 - \hat{\alpha}^{(2)})] \cdot \bar{Y}_{\bullet,\bullet,\bullet}, \quad (2.6)$$

où $\bar{Y}_{g,x,\bullet} = \frac{1}{T} \sum_{t=1}^T Y_{g,x,t}$, $\bar{Y}_{g,\bullet,\bullet} = \frac{1}{X} \sum_{x=1}^X \bar{Y}_{g,x,\bullet} = \frac{1}{X \cdot T} \sum_{x=1}^X \sum_{t=1}^T Y_{g,x,t}$ et

$$\bar{Y}_{\bullet,\bullet,\bullet} = \frac{1}{G} \sum_{g=1}^G \bar{Y}_{g,\bullet,\bullet} = \frac{1}{G \cdot X \cdot T} \sum_{g=1}^G \sum_{x=1}^X \sum_{t=1}^T Y_{g,x,t}.$$

À noter que les expressions pour $\alpha^{(1)}$ et $\alpha^{(2)}$ dans (2.6) sont les mêmes que (2.2) et (2.3), respectivement, et $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$ et $\hat{\sigma}_3^2$ dans $\alpha^{(1)}$ et $\alpha^{(2)}$ deviennent

$$\hat{\sigma}_1^2 = \frac{1}{G \cdot X} \sum_{g=1}^G \sum_{x=1}^X \left[\frac{1}{T-1} \sum_{t=1}^T (Y_{g,x,t} - \bar{Y}_{g,x,\bullet})^2 \right],$$

$$\hat{\sigma}_2^2 = \frac{1}{G} \sum_{g=1}^G \max[\hat{T}_g^{(1)}, 0] \quad \text{et} \quad \hat{\sigma}_3^2 = \max[\hat{T}^{(2)}, 0], \quad \text{où}$$

$$\hat{T}_g^{(1)} = \frac{1}{X-1} \sum_{x=1}^X (\bar{Y}_{g,x,\bullet} - \bar{Y}_{g,\bullet,\bullet})^2 - \frac{\hat{\sigma}_1^2}{T},$$

et

$$\hat{T}^{(2)} = \frac{1}{G-1} \sum_{g=1}^G (\bar{Y}_{g,\bullet,\bullet} - \bar{Y}_{\bullet,\bullet,\bullet})^2 - \left[\frac{\hat{\sigma}_2^2}{X} + \frac{\hat{\sigma}_1^2}{X \cdot T} \right].$$

De même, les paramètres structurels d'un arbre hiérarchique à trois niveaux peuvent être estimés en posant $C = G = 1$ et $\alpha^{(3)} = \alpha^{(2)} = 1$. Nous n'avons pas besoin ensuite de calculer $\hat{\sigma}_3^2$ et $\hat{\sigma}_4^2$. Par conséquent, le modèle de crédibilité hiérarchique à trois niveaux est exactement le même que le modèle de crédibilité non paramétrique de Bühlmann, proposé par Tsai et Lin (2017b).

L'expression en (2.5) donne l'estimation de crédibilité hiérarchique $\hat{Y}_{c,g,x,T+1}$ correspondant au pays c , au sexe g et à l'âge x dans l'année $T + 1$. Pour obtenir l'estimation de crédibilité hiérarchique $\hat{Y}_{c,g,x,T+\tau}$ pour l'année $T + \tau$ ($\tau \geq 2$), que nous représentons par

$$\begin{aligned}\hat{Y}_{c,g,x,T+\tau} &= \hat{\alpha}_\tau^{(1)} \cdot \bar{Y}_{c,g,x,\bullet}^{T+\tau} + [(1 - \hat{\alpha}_\tau^{(1)}) \cdot \hat{\alpha}_\tau^{(2)}] \cdot \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} \\ &\quad + [(1 - \hat{\alpha}_\tau^{(1)}) \cdot (1 - \hat{\alpha}_\tau^{(2)}) \cdot \hat{\alpha}_\tau^{(3)}] \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} \\ &\quad + [(1 - \hat{\alpha}_\tau^{(1)}) \cdot (1 - \hat{\alpha}_\tau^{(2)}) \cdot (1 - \hat{\alpha}_\tau^{(3)})] \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau},\end{aligned}\quad (2.7)$$

nous adoptons les deux mêmes stratégies que celles du modèle de crédibilité de Bühlmann proposé par Tsai et Lin (2017b), à savoir la fenêtre extensible (EW, en anglais) et la fenêtre mobile (MW, en anglais).

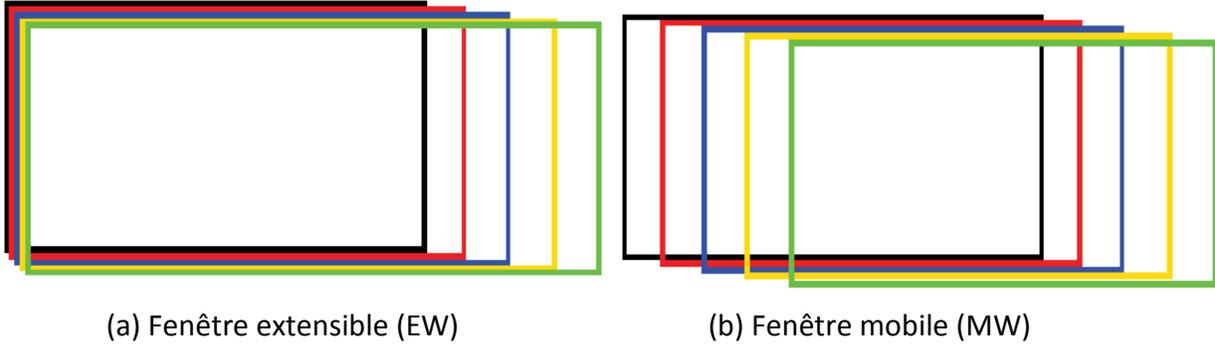
Stratégie EW : Extension de la fenêtre d'un an

La stratégie EW élargit l'intervalle initial d'années d'ajustement de $\tau - 1$ années pour le transformer en $[1, T + \tau - 1]$, et ce, en ajoutant $\{\hat{Y}_{c,g,x,T+1}, \dots, \hat{Y}_{c,g,x,T+\tau-1}\}$ à $\{Y_{c,g,x,1}, \dots, Y_{c,g,x,T}\}$ pour tous c, g et x . Voir la figure 4 (a).

Tout d'abord, nous calculons le décrétement annuel moyen sur l'intervalle d'années $[1, T + \tau - 1]$ pour le pays c , le sexe g et l'âge x , $\bar{Y}_{c,g,x,\bullet}^{T+\tau}$, $\tau \geq 2$, comme suit :

$$\bar{Y}_{c,g,x,\bullet}^{T+\tau} = \frac{1}{T + \tau - 1} \left[\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau-1} \hat{Y}_{c,g,x,t} \right]. \quad (2.8)$$

Figure 4 : Fenêtre extensible (EW) et fenêtre mobile (MW)



Les quantités $\bar{Y}_{c,g,\bullet,\bullet}^{T+\tau}$, $\bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau}$ et $\bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau}$ s'obtiennent comme suit, au moyen de la même formule puisque $\tau = 1$:

$$\bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} = \frac{1}{X} \sum_{x=1}^X \bar{Y}_{c,g,x,\bullet}^{T+\tau} \quad (2.9)$$

$$\bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} = \frac{1}{G} \sum_{g=1}^G \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} \quad (2.10)$$

et

$$\bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} = \frac{1}{C} \sum_{c=1}^C \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} \quad (2.11)$$

Ensuite, nous calculons $\hat{\alpha}_\tau^{(1)}$, le facteur de crédibilité affecté à $\bar{Y}_{c,g,x,\bullet}^{T+\tau}$, comme suit :

$$\hat{\alpha}_\tau^{(1)} = \frac{(T + \tau - 1) \cdot \hat{\sigma}_2^2}{(T + \tau - 1) \cdot \hat{\sigma}_2^2 + \hat{\sigma}_1^2}. \quad (2.12)$$

Les formules pour les facteurs de crédibilité $\hat{\alpha}_\tau^{(2)}$ et $\hat{\alpha}_\tau^{(3)}$ sont les mêmes que celles pour $\tau = 1$ indiquées en (2.3) et (2.4), soit

$$\hat{\alpha}_\tau^{(2)} = \frac{X \cdot \hat{\alpha}_\tau^{(1)} \cdot \hat{\sigma}_3^2}{X \cdot \hat{\alpha}_\tau^{(1)} \cdot \hat{\sigma}_3^2 + \hat{\sigma}_2^2} = \frac{X (T + \tau - 1) \cdot \hat{\sigma}_3^2}{X (T + \tau - 1) \cdot \hat{\sigma}_3^2 + (T + \tau - 1) \cdot \hat{\sigma}_2^2 + \hat{\sigma}_1^2}, \quad (2.13)$$

et

$$\begin{aligned} \hat{\alpha}_\tau^{(3)} &= \frac{G \cdot \hat{\alpha}_\tau^{(2)} \cdot \hat{\sigma}_4^2}{G \cdot \hat{\alpha}_\tau^{(2)} \cdot \hat{\sigma}_4^2 + \hat{\sigma}_3^2} \\ &= \frac{G X (T + \tau - 1) \cdot \hat{\sigma}_4^2}{G X (T + \tau - 1) \cdot \hat{\sigma}_4^2 + X (T + \tau - 1) \cdot \hat{\sigma}_3^2 + (T + \tau - 1) \cdot \hat{\sigma}_2^2 + \hat{\sigma}_1^2}. \end{aligned} \quad (2.14)$$

Notons par ailleurs que les valeurs de $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, $\hat{\sigma}_3^2$ et $\hat{\sigma}_4^2$ ne changent pas à mesure que τ augmente.

Enfin, l'estimation de crédibilité hiérarchique $\hat{Y}_{c,g,x,T+\tau}$ pour le pays c , le sexe g et l'âge x dans l'année $T + \tau$ s'obtient par (2.7).

Stratégie MW : Déplacement de la fenêtre d'un an

La stratégie MW déplace l'intervalle initial d'années d'ajustement de $\tau - 1$ années pour le transformer en $[\tau, T + \tau - 1]$, et ce, en ajoutant les estimations de crédibilité hiérarchique $\{\hat{Y}_{c,g,x,T+1}, \dots, \hat{Y}_{c,g,x,T+\tau-1}\}$ à $\{\hat{Y}_{c,g,x,1}, \dots, \hat{Y}_{c,g,x,T}\}$ et en supprimant de celle-ci $\{\hat{Y}_{c,g,x,1}, \dots, \hat{Y}_{c,g,x,\tau-1}\}$ pour tous c, g et x où $\hat{Y}_{c,g,x,t} = Y_{c,g,x,t}$ pour $t \leq T$. Voir figure 4 (b).

Tout d'abord, nous calculons le décrétement annuel moyen sur l'intervalle d'années $[\tau, T + \tau - 1]$ pour le pays c , le sexe g et l'âge x , $\bar{Y}_{c,g,x,\bullet}^{T+\tau}$, comme suit :

$$\bar{Y}_{c,g,x,\bullet}^{T+\tau} = \frac{1}{T} \sum_{t=\tau}^{T+\tau-1} \hat{Y}_{c,g,x,t}, \quad (2.15)$$

et $\bar{Y}_{c,g,\bullet,\bullet}^{T+\tau}$, $\bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau}$ et $\bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau}$ s'obtiennent par (2.9), (2.10) et (2.11), respectivement.

Ensuite, le facteur de crédibilité à attribuer à $\bar{Y}_{c,g,x,\bullet}^{T+\tau}$ est obtenu par $\hat{\alpha}_\tau^{(1)} = \frac{T \hat{\sigma}_2^2}{T \hat{\sigma}_2^2 + \hat{\sigma}_1^2}$. Vu que $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, $\hat{\sigma}_3^2$, $\hat{\sigma}_4^2$ restent constants pour tout τ , nous avons $\hat{\alpha}_\tau^{(1)} = \hat{\alpha}_1^{(1)}$, $\hat{\alpha}_\tau^{(2)} = \hat{\alpha}_1^{(2)}$ et $\hat{\alpha}_\tau^{(3)} = \hat{\alpha}_1^{(3)}$. Par conséquent, $\hat{\alpha}_\tau^{(1)}$, $\hat{\alpha}_\tau^{(2)}$ et $\hat{\alpha}_\tau^{(3)}$ ne varient pas en fonction de τ selon la stratégie MW.

Enfin, nous pouvons calculer $\hat{Y}_{c,g,x,T+\tau}$ le décrétement logarithmique du taux central de mortalité pour le pays c , le sexe g et l'âge x sur $[T + \tau - 1, T + \tau]$, en utilisant (2.7).

Les stratégies EW et MW obéissent à certaines propriétés qui sont énoncées dans les trois propositions suivantes. Les démonstrations figurent à l'annexe B.

Proposition 1 Selon les stratégies EW et MW, la moyenne des estimations de crédibilité hiérarchique $\hat{Y}_{c,g,x,T+\tau}$ sur les âges $1, \dots, X$, les sexes $1, \dots, G$ et les pays $1, \dots, C$ pour l'année $T + \tau$ est égale à la moyenne de $\bar{Y}_{c,g,x,\bullet}^{T+\tau}$ sur les mêmes âges, sexes et pays pour l'année $T + \tau$. Plus précisément,

$$\frac{1}{C \cdot G \cdot X} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \hat{Y}_{c,g,x,T+\tau} = \frac{1}{C \cdot G \cdot X} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \bar{Y}_{c,g,x,\bullet}^{T+\tau} = \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau}, \tau = 1, 2, \dots$$

Proposition 2 Selon la stratégie EW, la moyenne globale des estimations de crédibilité hiérarchique pour l'année $T + \tau$, $\bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau}$, $\tau = 1, 2, \dots$, ne varie pas en fonction de τ , c'est-à-dire que,

$$\bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} = \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+1}, \quad \tau = 2, 3, \dots$$

Proposition 3 Selon la stratégie EW, l'estimation de crédibilité hiérarchique $\hat{Y}_{c,g,x,T+\tau}$ est constante pour $\tau = 1, 2, \dots$, c'est-à-dire que, $\hat{Y}_{c,g,x,T+\tau} = \hat{Y}_{c,g,x,T+1}$, $\tau = 2, 3, \dots$.

Nous déduisons de la proposition 3 que

$$\ln(\hat{m}_{c,g,x,T+\tau}) = \ln(m_{c,g,x,T}) + \sum_{t=1}^{\tau} \hat{Y}_{c,g,x,T+t} = \ln(m_{c,g,x,T}) + (\hat{Y}_{c,g,x,T+1}) \cdot \tau.$$

Par conséquent, $\ln(\hat{m}_{c,g,x,T+\tau})$ est une fonction linéaire de τ avec une pente de $\hat{Y}_{c,g,x,T+1}$ et une ordonnée à l'origine de $\ln(m_{c,g,x,T})$.

3. Exemples numériques

Nous appliquons ici les modèles hiérarchiques à trois, à quatre ou à cinq niveaux et les paramètres structurels estimés de la section 2 à la prévision des taux de mortalité de six populations (les deux sexes aux États-Unis, au R.-U. et au Japon) afin de produire des exemples numériques. Les données de mortalité sont tirées de la Human Mortality Database. Nous ajustons des modèles de crédibilité hiérarchique à cinq, à quatre ou à trois niveaux à un grand intervalle d'âges et à une série d'intervalles d'années d'ajustement et nous faisons des prévisions hors échantillon pour des années consécutives futures. Le même ensemble de données est aussi ajusté au modèle traditionnel de Lee-Carter et aux modèles k-conjoint, cointégré et augmenté de facteurs communs de Lee-Carter avec six populations et deux populations de chacun des trois pays, respectivement. La qualité prévisionnelle est mesurée par la moyenne des erreurs absolues moyennes en pourcentage (AMAPE), qui montre que tous les modèles de mortalité hiérarchique à trois, à quatre et à cinq niveaux surpassent le modèle traditionnel et les trois modèles pour plusieurs populations de Lee-Carter.

Pour chacun des modèles de crédibilité hiérarchique à trois, à quatre ou à cinq niveaux, nous adoptons les stratégies EW et MW pour prévoir les taux de mortalité pour trois intervalles d'années. Nous désignons par EW- l et MW- l les stratégies EW et MW, respectivement, selon le modèle de crédibilité hiérarchique à l niveaux, où $l = 3, 4, 5$. Le modèle de crédibilité hiérarchique à trois niveaux avec une structure arborescente suivant l'année, l'âge et la population (homme ou femme d'un pays) est appliqué à chacune des six populations. Selon le modèle de crédibilité hiérarchique à quatre (cinq) niveaux, la structure arborescente de bas en

haut est spécifiée suivant l'année, l'âge, le sexe et le pays (année, âge, sexe, pays et plusieurs pays), et le modèle est appliqué aux deux sexes de chacun des trois pays (les six populations). Afin de comparer la qualité prévisionnelle du modèle de crédibilité hiérarchique à trois niveaux, les données de mortalité de chacune des six populations sont respectivement ajustées au modèle traditionnel de Lee-Carter, que nous représentons par LC1-Ind. Également, afin de comparer la qualité prévisionnelle du modèle de crédibilité hiérarchique à quatre (cinq) niveaux, les données de mortalité des deux sexes de chacun des trois pays (six populations) sont respectivement ajustées aux modèles k-conjoint, cointégré et augmenté de facteurs communs de Lee-Carter avec la population des hommes d'un pays (les hommes aux É.-U.) comme population de base pour le modèle cointégré de Lee-Carter, que nous représentons par LC2-JoK, LC2-Col et LC2-ACF (LC6-JoK, LC6-Col et LC6-ACF).

Soit $[T_1, T_2]$ la période à l'étude pour laquelle les taux de mortalité sont disponibles. Nous supposons que nous sommes à la fin de l'année t_U et que nous aimerions ajuster les modèles aux données de mortalité dans le rectangle $[x_L, x_U] \times [t_L, t_U]$, projeter les taux de mortalité sur les années $[t_U + 1, T_2]$ et évaluer la qualité prévisionnelle des modèles de mortalité sous-jacents. Voici les hypothèses en détail :

- Pour l'intervalle d'âges $[x_L, x_U]$, nous choisissons $x_L = 20$ et $x_U = 84$ et la longueur de l'intervalle est $m = 65$.
- Pour la période à l'étude $[T_1, T_2]$, nous utilisons la période de 63 ans 1951-2013, c.-à-d., $T_1 = 1951$ et $T_2 = 2013$, soit l'année la plus récente pour laquelle les taux de mortalité sont disponibles pour les deux sexes aux É.-U., au R.-U. et au Japon.
- Pour les intervalles d'années d'ajustement $[t_L, t_U]$, une série de périodes, $[1951, t_U], \dots, [t_U - 4, t_U]$, sont choisies, où t_U prend les valeurs 1983, 1993 et 2003, et la période la plus courte est de cinq ans.
- Pour l'intervalle d'années de prévision $[t_U + 1, T_2]$, nous choisissons $[2004, 2013]$ (étendue de 10 ans), $[1994, 2013]$ (étendue de 20 ans) et $[1984, 2013]$ (étendue de 30 ans).

Le tableau 2 fait le sommaire des intervalles d'années d'ajustement et de prévision.

Tableau 2 Sommaire des intervalles d'années d'ajustement et de prévision

<i>Intervalles d'années d'ajustement</i>	$[t_L, t_U]$	[1951, 2003] [1952, 2003] ⋮ [1999, 2003]	[1951, 1993] [1952, 1993] ⋮ [1989, 1993]	[1951, 1983] [1952, 1983] ⋮ [1979, 1983]
<i>Année terminale des intervalles d'années d'ajustement</i>	t_U	2003	1993	1983
<i>Nombre d'intervalles d'années d'ajustement</i>	J	49	39	29
<i>Intervalles d'années de prévision</i>	$[t_U + 1, T_2]$	[2004, 2013]	[1994, 2013]	[1984, 2013]
<i>Étendue des intervalles d'années de prévision</i>	$T_2 - t_U$	10	20	30

Nous comparons la qualité prévisionnelle des modèles de crédibilité hiérarchique avec les modèles traditionnels, k-conjoint, cointégré et augmenté de facteurs communs de Lee-Carter, au moyen de la mesure de la MAPE, une mesure couramment utilisée par Lin et coll. (2015) et Tsai et Lin (2017a, b). La MAPE s'obtient en divisant la valeur absolue de la différence entre le taux de mortalité prévu $\hat{q}_{c,g,x,t_U+\tau}$ et le taux de mortalité réel $q_{c,g,x,t_U+\tau}$, par le taux de mortalité réel $q_{c,g,x,t_U+\tau}$. Plus précisément, la $MAPE_{c,g,x,t_U+\tau}^{t_L,t_U}$ correspondant à l'âge x , au sexe g et au pays c dans l'année de prévision $t_U + \tau$ et basée sur l'intervalle d'années d'ajustement $[t_L, t_U]$ est définie par

$$MAPE_{c,g,x,t_U+\tau}^{[t_L,t_U]} = \left| \frac{\hat{q}_{c,g,x,t_U+\tau} - q_{c,g,x,t_U+\tau}}{q_{c,g,x,t_U+\tau}} \right|, \quad (3.1)$$

où $\hat{q}_{c,g,x,t_U+\tau} = 1 - e^{-\hat{m}_{c,g,x,t_U+\tau}}$ selon l'hypothèse d'un taux de mortalité instantané constant sur $[x, x + 1] \times [t_U + \tau, t_U + \tau + 1]$. La $MAPE_{c,g,x,t_U+\tau}^{t_L,t_U}$ ci-dessus est calculée pour une seule prévision. Ici, nous faisons intervenir l' $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ pour évaluer la qualité prévisionnelle à l'égard des taux de mortalité dans la fenêtre $[x_U, x_U] \times [t_U + 1, T_2]$, que nous obtenons en calculant la moyenne des $MAPE_{c,g,x,t_U+\tau}^{t_L,t_U}$ sur la fenêtre, comme suit :

$$AMAPE_{c,g,[t_U+1,T_2]}^{[t_L,t_U]} = \frac{1}{T_2 - t_U} \cdot \frac{1}{m} \sum_{\tau=1}^{T_2-t_U} \sum_{x=x_L}^{x_U} MAPE_{c,g,x,t_U+\tau}^{[t_L,t_U]}.$$

La valeur de l' $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ dépend dans une large mesure de l'intervalle d'années d'ajustement $[t_L, t_U]$; en conséquence, nous évaluons la qualité prévisionnelle globale d'un modèle de mortalité en calculant la moyenne des $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ sur les intervalles d'années d'ajustement $[t_L, t_U]$ pour $t_L = T_1, T_1 + 1, \dots, t_U - 4$, pour obtenir l' $AAMAPE_{c,g,[t_U+1,T_2]}$, qui s'obtient comme suit :

$$AAMAPE_{c,g,[t_U+1,T_2]} = \frac{1}{t_U - 4 - T_1 + 1} \sum_{t_L=T_1}^{t_U-4} AMAPE_{c,g,[t_U+1,T_2]}^{[t_L,t_U]}.$$

Plus l' $AAMAPE_{c,g,[t_U+1,T_2]}$ produite par un modèle est basse, plus la prévision pour la période $[t_U + 1, T_2]$ est exacte dans l'ensemble. Les modèles de mortalité qui sous-tendent le présent article seront classés en fonction de l' $AAMAPE_{c,g,[t_U+1,T_2]}$.

Nous produisons sept figures et construisons trois tableaux correspondant à trois intervalles d'années de prévision, soit [2004, 2013], [1994, 2013] et [1984, 2013]. Les figures 5 à 10 tracent l' $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ en fonction de t_L , où $t_U = 2003, 1993, 1983$ pour chacune des six populations (homme É.-U., femme É.-U., homme R.-U., femme R.-U., homme Japon et femme Japon), respectivement, tandis que la figure 11 affiche la moyenne des $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ sur les six populations en fonction de t_L , où $t_U = 2003, 1993, 1983$. Dans chaque figure, les graphiques des $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ qui correspondent aux trois différents intervalles d'années de prévision sont présentés sur trois lignes, tandis que les graphiques relatifs à une seule population, à deux populations et à six populations auxquelles les modèles sont respectivement ajustés sont présentés sur trois colonnes. Plus précisément, la première colonne indique l' $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ en fonction de t_L , la première année de l'intervalle d'années d'ajustement, pour les modèle traditionnel de Lee-Carter et le modèle de crédibilité

hiérarchique à trois niveaux; la deuxième colonne indique l' $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ en fonction de t_L pour le modèle de crédibilité hiérarchique à quatre niveaux et les trois modèles de Lee-Carter pour deux populations; et la troisième colonne présente l' $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ en fonction de t_L pour le modèle de crédibilité hiérarchique à cinq niveaux et les trois modèles de Lee-Carter pour six populations. À noter que les $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ selon les trois modèles de Lee-Carter pour deux populations et six populations sont différentes. En ce qui concerne les graphiques à deux populations, les modèles k-conjoint, cointégré et augmenté de facteurs communs de Lee-Carter sont ajustés aux données des deux sexes d'un pays; par exemple, les trois modèles de Lee-Carter sont appliqués aux deux sexes aux É.-U., et les $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ correspondantes sont indiquées en (b), (e) et (h) des figures 5 et 6, respectivement. Pour ce qui est des graphiques à six populations, les modèles k-conjoint, cointégré et augmenté de facteurs communs de Lee-Carter sont ajustés aux données des six populations, et les $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ correspondantes pour les hommes et les femmes aux É.-U. sont indiquées en (c), (f) et (i) des figures 5 et 6, respectivement.

Les figures appellent plusieurs observations :

- Les valeurs $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ des modèles de crédibilité hiérarchique et des modèles de Lee-Carter décroissent généralement en fonction de t_U , ce qui veut dire que plus l'étendue de la période de prévision est grande, plus la valeur de l' $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ est élevée.
- Les valeurs $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ de tous les modèles et des deux stratégies ne diminuent pas de façon monotone ni n'augmentent avec t_L (la première année de l'intervalle d'années d'ajustement), c.-à-d. que les valeurs $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ dépendent de la longueur et du lieu de l'intervalle d'années d'ajustement. Le schéma des courbes $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ dépend au sens large de l'ensemble de données. Par exemple, la courbe $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ relative à la stratégie MW, à la figure 6 (h), augmente pour les vingt premières valeurs t_L , puis diminue pour les cinq valeurs t_L suivantes, et enfin augmente pour les dernières valeurs t_L . Toutefois, la figure 7 (h) montre que la courbe $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ relative à la stratégie MW diminue en fonction de t_L , sauf pour quelques valeurs de t_L situées aux deux extrémités du domaine. Puisque nous ne savons pas quel intervalle d'années d'ajustement produira la plus faible $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$, nous calculons l' $AAMAPE_{c,g,[t_U+1,T_2]}$ (la moyenne des $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ sur toutes les valeurs de t_L) et nous nous en servons pour classer les modèles de mortalité sous-jacents.
- Les valeurs $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ relatives aux deux stratégies EW et MW des modèles de crédibilité hiérarchique sont, dans l'ensemble, plus basses que celles des modèles de Lee-Carter sauf dans quelques cas, par exemple, pour les hommes aux É.-U. (figure 5 (i)), pour les femmes aux É.-U. (figure 6 (i)) et pour les hommes au Japon (figure 9 (h)); les valeurs $AAMAPE_{c,g,[t_U+1,T_2]}$ correspondantes sont indiquées au tableau 3. Par ailleurs, les courbes $AMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ en fonction de t_L qui se rapportent à la stratégie MW paraissent plus lisses que celles relatives à la stratégie EW.

Nous tirons les observations suivantes des tableaux 1 à 3, qui montrent l' $AAMAPE_{c,g,[t_U+1,2013]}$, l'Avg 2 (la moyenne des $AAMAPE_{c,g,[t_U+1,2013]}$ sur les deux sexes d'un pays) et l'Avg 6 (la moyenne des $AAMAPE_{c,g,[t_U+1,2013]}$ sur les six populations) pour $t_U = 2003$, 1993 et 1983, respectivement :

Tableau 1 : $AAMAPE_{c,g,[2000,2013]}$ (%)

AAMAPE Modèle	Pays Avg 6	É.-U.			R.-U.			Japon		
		H	F	Avg 2	H	F	Avg 2	H	F	Avg 2
Modèle de crédibilité hiérarchique										
EW-5	6,63	5,58	4,92	5,25	9,48	7,87	8,68	5,74	6,20	5,97
MW-5	6,66	5,60	4,93	5,26	9,57	7,95	8,76	5,77	6,16	5,97
EW-4	7,23	6,36	5,53	5,95	9,86	8,24	9,05	5,86	7,52	6,69
MW-4	7,16	6,41	5,56	5,99	9,79	8,17	8,98	5,76	7,27	6,52
EW-3	7,47	6,00	6,04	6,02	9,78	8,33	9,06	5,83	8,85	7,34
MW-3	7,41	5,96	6,14	6,05	9,61	8,36	8,98	5,85	8,54	7,19
Modèle de Lee-Carter appliqué à six populations (LC6)										
LC6-JoK	10,61	10,27	8,54	9,40	15,10	10,23	12,66	7,98	11,56	9,77
LC6-Col	9,69	9,23	8,57	8,90	12,64	9,79	11,22	7,93	9,95	8,94
LC6-ACF	9,22	7,46	8,45	7,96	13,94	8,24	11,09	6,26	10,96	8,61
Modèle de Lee-Carter appliqué aux deux sexes d'un pays (LC2)										
LC2-JoK	9,77	10,06	8,75	9,41	12,95	8,89	10,92	7,61	10,39	9,00
LC2-Col	9,59	9,23	8,57	8,90	12,28	9,32	10,80	7,63	10,50	9,06
LC2-ACF	9,60	9,58	8,25	8,92	12,16	9,26	10,71	7,57	10,77	9,17
Modèle de Lee-Carter appliqué à une seule population (LC1)										
LC1-Ind	9,64	9,23	8,57	8,90	12,28	9,36	10,82	7,63	10,77	9,20

Tableau 2 : $AAMAPE_{c,g,[1994,2013]}$ (%)

AAMAPE Modèle	Pays Avg 6	É.-U.			R.-U.			Japon		
		H	F	Avg 2	H	F	Avg 2	H	F	Avg 2
Modèle de crédibilité hiérarchique										
EW-5	10,41	10,12	7,56	8,84	14,00	9,05	11,53	8,52	13,18	10,85
MW-5	10,55	10,41	7,13	8,77	14,39	9,45	11,92	8,66	13,24	10,95
EW-4	11,85	13,77	6,05	9,91	14,94	10,03	12,48	12,19	14,10	13,15
MW-4	11,74	14,39	6,12	10,26	15,07	10,14	12,61	11,17	13,56	12,37
EW-3	11,98	14,97	6,14	10,55	15,57	9,50	12,53	10,05	15,63	12,84
MW-3	11,81	15,04	6,32	10,68	15,48	9,79	12,63	9,49	14,75	12,12
Modèle de Lee-Carter appliqué à six populations (LC6)										
LC6-JoK	14,71	16,64	9,59	13,11	19,11	13,63	16,37	12,93	16,37	14,65
LC6-Col	13,98	16,48	8,32	12,40	18,71	13,81	16,26	11,31	15,27	13,29
LC6-ACF	14,25	18,19	8,95	13,57	18,92	10,75	14,83	11,30	17,39	14,34
Modèle de Lee-Carter appliqué aux deux sexes d'un pays (LC2)										
LC2-JoK	14,31	16,75	8,55	12,65	17,64	13,05	15,34	13,19	16,71	14,95
LC2-Col	14,14	16,48	8,32	12,40	17,43	13,12	15,28	12,74	16,73	14,73
LC2-ACF	14,02	16,36	7,61	11,98	17,34	13,01	15,17	12,68	17,14	14,91
Modèle de Lee-Carter appliqué à une seule population (LC1)										
LC1-Ind	14,23	16,48	8,50	12,49	17,43	13,07	15,25	12,74	17,14	14,94

Tableau 3 : $AAMAPE_{c,g,[1984,2013]}$ (%)

AAMAPE	Pays	É.-U.			R.-U.			Japon		
Modèle	Avg 6	H	F	Avg 2	H	F	Avg 2	H	F	Avg 2
Modèle de crédibilité hiérarchique										
EW-5	14,01	13,57	17,33	15,45	16,94	10,63	13,78	10,69	14,90	12,79
MW-5	14,02	13,48	18,26	15,87	16,44	10,60	13,52	10,85	14,50	12,67
EW-4	14,60	11,45	13,42	12,44	18,00	10,03	14,01	17,63	17,09	17,36
MW-4	14,28	11,63	15,38	13,50	17,03	9,53	13,28	16,12	15,98	16,05
EW-3	15,03	11,71	15,41	13,56	19,26	9,56	14,41	13,88	20,39	17,13
MW-3	14,55	11,83	16,37	14,10	17,98	9,58	13,78	12,62	18,94	15,78
Modèle de Lee-Carter appliqué à six populations (LC6)										
LC6-JoK	19,57	14,43	17,01	15,72	24,29	16,90	20,60	19,99	24,79	22,39
LC6-Col	17,97	13,43	15,60	14,52	23,86	17,15	20,51	16,39	21,37	18,88
LC6-ACF	17,26	12,32	16,15	14,24	23,02	13,30	18,16	15,36	23,43	19,40
Modèle de Lee-Carter appliqué aux deux sexes d'un pays (LC2)										
LC2-JoK	18,40	13,42	17,11	15,27	22,63	16,59	19,61	17,99	22,61	20,30
LC2-Col	17,92	13,43	15,60	14,52	22,53	16,62	19,58	17,05	22,26	19,65
LC2-ACF	18,41	14,32	16,49	15,41	22,43	16,89	19,66	16,90	23,46	20,18
Modèle de Lee-Carter appliqué à une seule population (LC1)										
LC1-Ind	18,25	13,43	16,26	14,85	22,53	16,78	19,66	17,05	23,44	20,24

- Au vu des valeurs des tableaux 1 à 3, il est évident que les valeurs $AAMAPE_{c,g,[t_U+1,2013]}$ croissent lorsque la longueur de l'intervalle d'années de prévision passe de 10 à 30 ans, ce qui concorde avec les observations des figures 5 à 10, selon lesquelles les valeurs $AAMAPE_{c,g,[t_U+1,T_2]}^{t_L,t_U}$ augmentent en fonction de l'étendue de l'intervalle d'années de prévision.
- Selon la moyenne des $AAMAPE_{c,g,[t_U+1,2013]}$ sur six populations (« Avg 6 »), nous constatons que plus le modèle de crédibilité hiérarchique a de niveaux, plus sa qualité prévisionnelle est grande. Toutefois, cette conclusion ne tient pas dans le cas de l'« Avg 2 » (la moyenne des $AAMAPE_{c,g,[t_U+1,2013]}$ sur les deux sexes d'un pays). Le classement des qualités prévisionnelles selon « Avg 2 » dépend du pays. Par exemple, selon l'« Avg 2 », dans le tableau 3, pour la période de prévision [1984, 2013], c'est le modèle de crédibilité hiérarchique à cinq niveaux qui a le moins de précision et le modèle à quatre niveaux qui en a le plus aux É.-U., alors qu'au Japon, c'est le modèle de crédibilité hiérarchique à cinq niveaux qui a le plus de précision et le modèle à quatre niveaux qui en a le moins.
- Les valeurs $AAMAPE_{c,g,[t_U+1,2013]}$ et leurs moyennes « Avg 6 » et « Avg 2 », selon les deux stratégies EW et MW et un modèle de crédibilité hiérarchique comportant le même nombre de niveaux, sont près l'une de l'autre.
- Dans le cas des modèles de Lee-Carter appliqués à six populations, les moyennes des $AAMAPE_{c,g,[t_U+1,2013]}$ sur six populations (« Avg 6 ») montrent que c'est le modèle augmenté de facteurs communs, LC6-ACF, qui est le plus précis pour les périodes de prévision [2004, 2013] et [1984, 2013], que le modèle cointégré, LC6-Col, produit les meilleures prévisions pour la période [1994, 2013], et que le modèle k-conjoint, LC6-JoK, est le moins précis pour les trois périodes de prévision. Par ailleurs, le modèle LC2-JoK surpasse le modèle LC6-JoK pour les trois périodes, le modèle LC2-Col est meilleur que le modèle LC6-Col pour [2004, 2013] et [1984, 2013], et le modèle LC2-

ACF donne de pires résultats que le modèle LC6-ACF pour [2004, 2013] et [1984, 2013].

- Les tableaux 1 à 3 nous montrent que la plupart des valeurs $AAMAPE_{c,g,[t_U+1,2013]}$ des modèles de crédibilité hiérarchique pour les trois périodes de prévision et les six populations sont plus basses que celles des modèles de Lee-Carter. Il s'ensuit que les moyennes des $AAMAPE_{c,g,[t_U+1,2013]}$ sur les deux sexes d'un pays et sur six populations sont, pour les trois périodes de prévision des modèles de crédibilité hiérarchique, nettement inférieures à celles des modèles de Lee-Carter. Par exemple, pour la période prévisionnelle de 30 ans [1984, 2013], les moyennes des $AAMAPE_{c,g,[t_U+1,2013]}$ sur six populations pour les modèles k-conjoint, cointégré et augmenté de facteurs communs de Lee-Carter appliqués à six populations et à deux populations et pour le modèle traditionnel de Lee-Carter sont respectivement égales à 19,57 %, 17,97 %, 17,26 %, 18,40 %, 17,92 %, 18,41 % et 18,25 %, tandis que, dans le cas des stratégies EW et MW selon les modèles de crédibilité hiérarchique à cinq, quatre et trois niveaux, elles sont respectivement égales à 14,01 %, 14,02 %, 14,60 %, 14,28 %, 15,03 % et 14,55 %. Les exemples numériques confirment donc fortement la conclusion que les modèles de crédibilité hiérarchique surpassent les modèles de Lee-Carter.

En résumé, plus un modèle de crédibilité hiérarchique comporte de niveaux, plus il donne de meilleures prévisions, et les stratégies EW et MW produisent des prévisions de qualité similaire. Quelle que soit la longueur de l'intervalle d'années d'ajustement et de l'intervalle d'années de prévision, les modèles de crédibilité hiérarchique produisent, dans l'ensemble, des prévisions plus exactes que celles des modèles de Lee-Carter. Par conséquent, nous concluons que le modèle de crédibilité hiérarchique est un moyen efficace de modéliser les taux de mortalité de plusieurs populations.

4. Conclusions

Le risque de longévité représente un défi majeur pour les fournisseurs de rentes, les régimes de sécurité sociale et les régimes de retraite à prestations déterminées. Nous avons appliqué ici la théorie de la crédibilité hiérarchique avec des structures arborescentes à trois, quatre et cinq niveaux à la modélisation des taux de mortalité de plusieurs populations. La structure à cinq niveaux est représentée, de bas en haut, par l'« année t », l'« âge x », le « sexe g », le « pays c » et « plusieurs pays », tandis que la structure à quatre niveaux l'est par l'« année t », l'« âge x », le « sexe g », le « pays c ». Les modèles de crédibilité hiérarchique sont ajustés aux données de mortalité des hommes et des femmes de trois pays développés (les É.-U., le R.-U. et le Japon) tirées de la Human Mortality Database Bas pour l'intervalle d'âges de 25 à 84 ans et une série d'intervalles d'années d'ajustement.

Figure 5 : $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ en fonction de t_L , hommes âgés de 20 à 84 ans, É.-U.

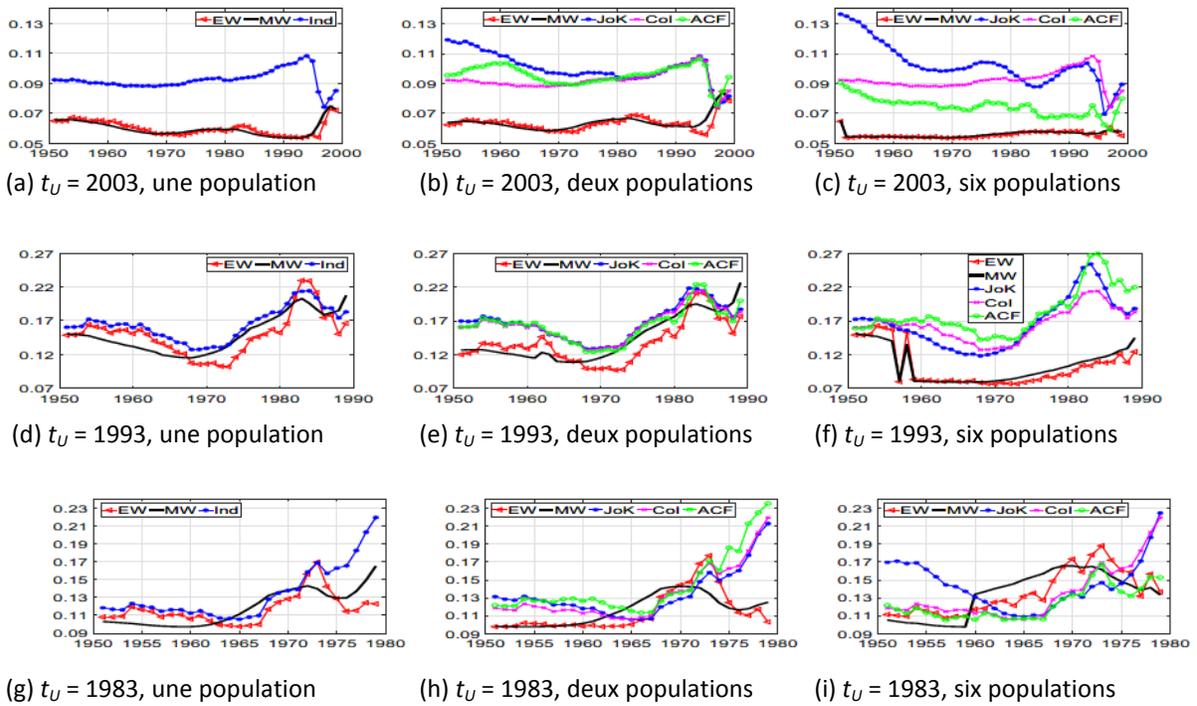


Figure 6 : $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ en fonction de t_L , femmes âgées de 20 à 84 ans, É.-U.

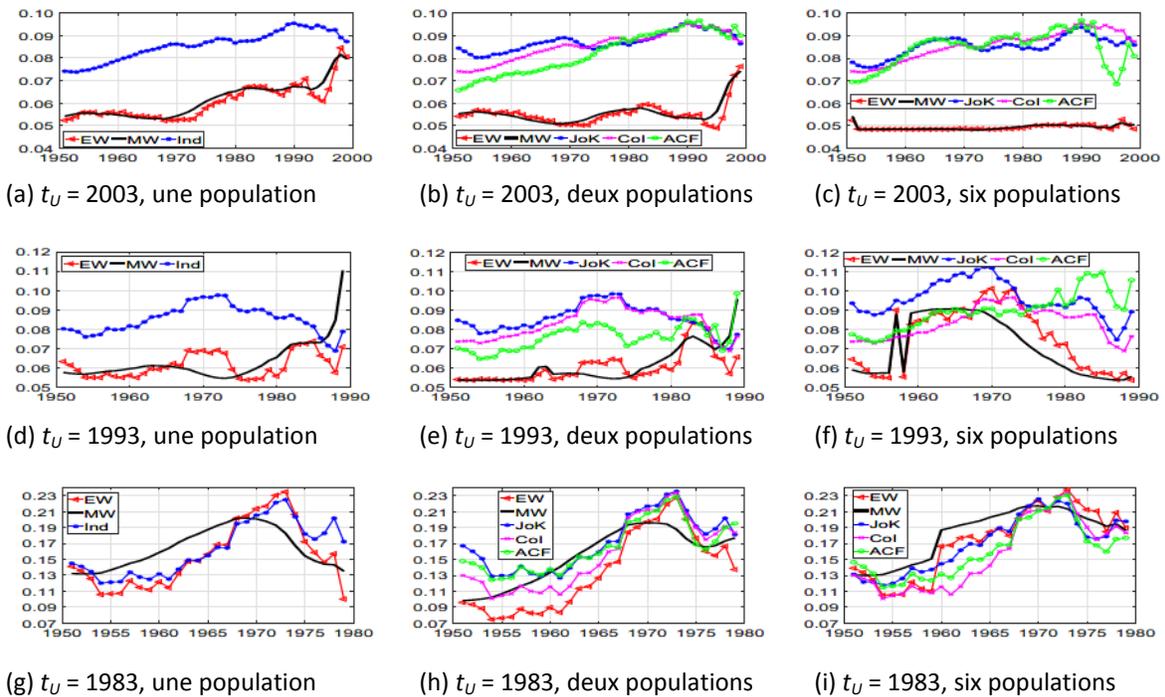


Figure 7 : $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ en fonction de t_L , hommes âgés de 20 à 84 ans, R.-U.

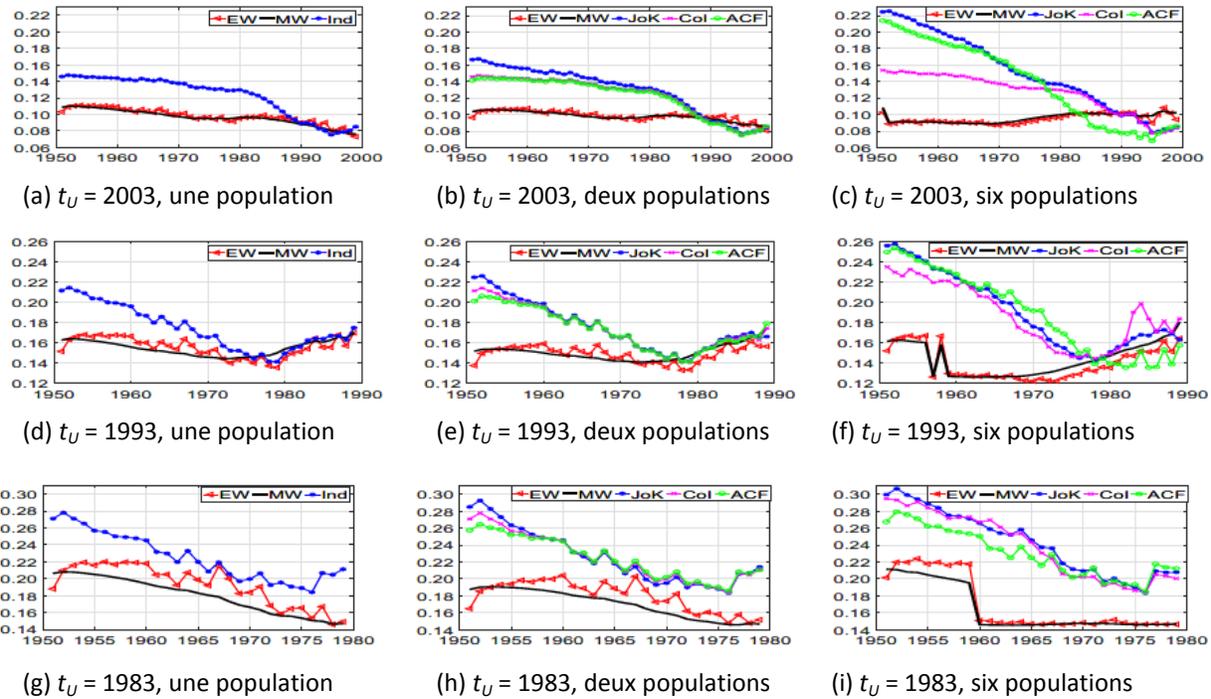


Figure 8 : $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ en fonction de t_L , femmes âgées de 20 à 84 ans, R.-U.

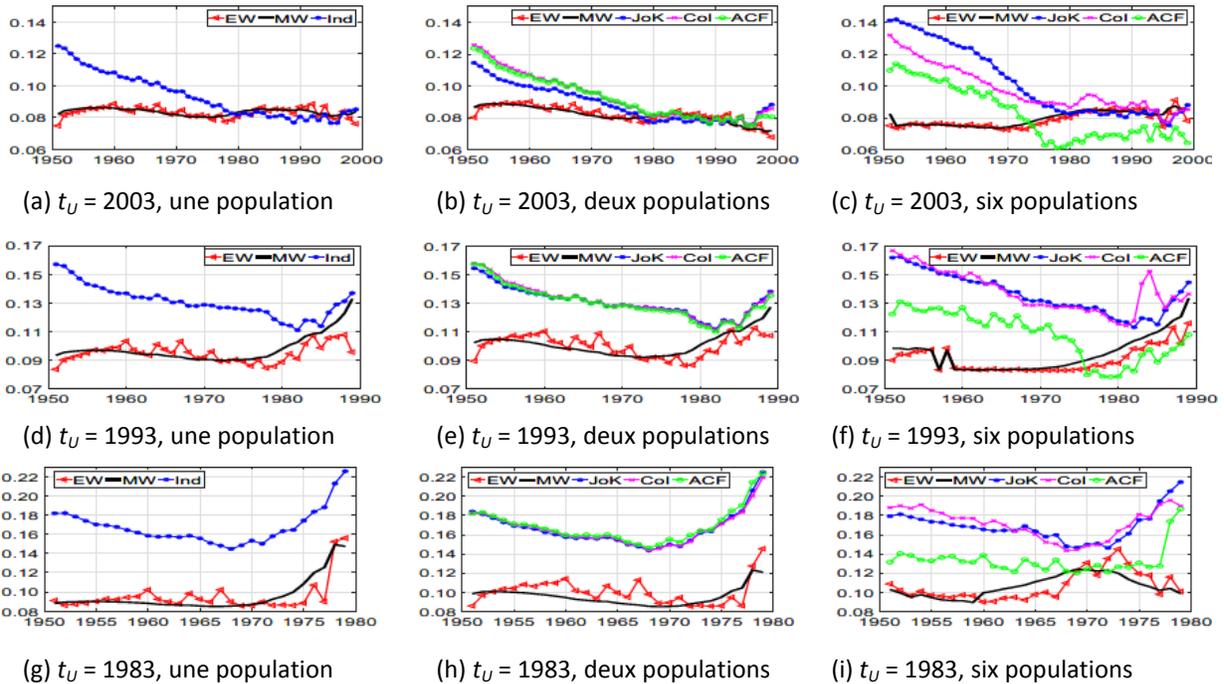


Figure 9 : $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ en fonction de t_L , hommes âgés de 20 à 84 ans, Japon

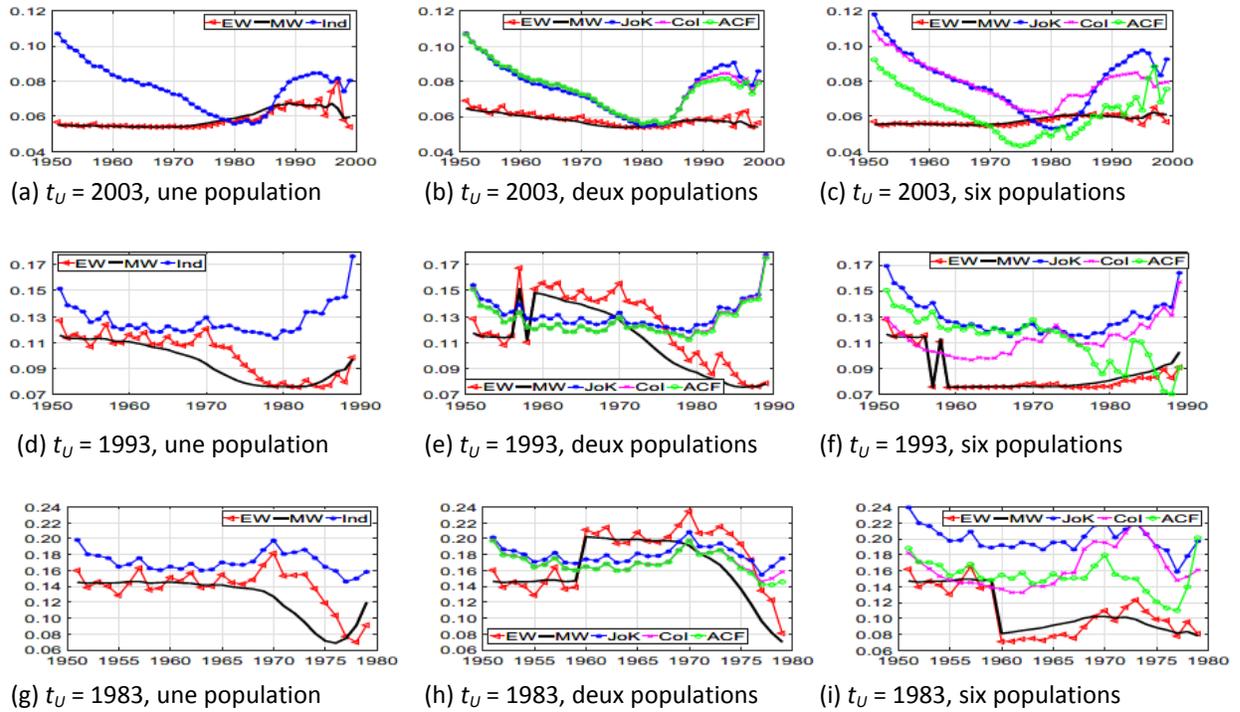


Figure 10 : $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ en fonction de t_L , femmes âgées de 20 à 84 ans, Japon

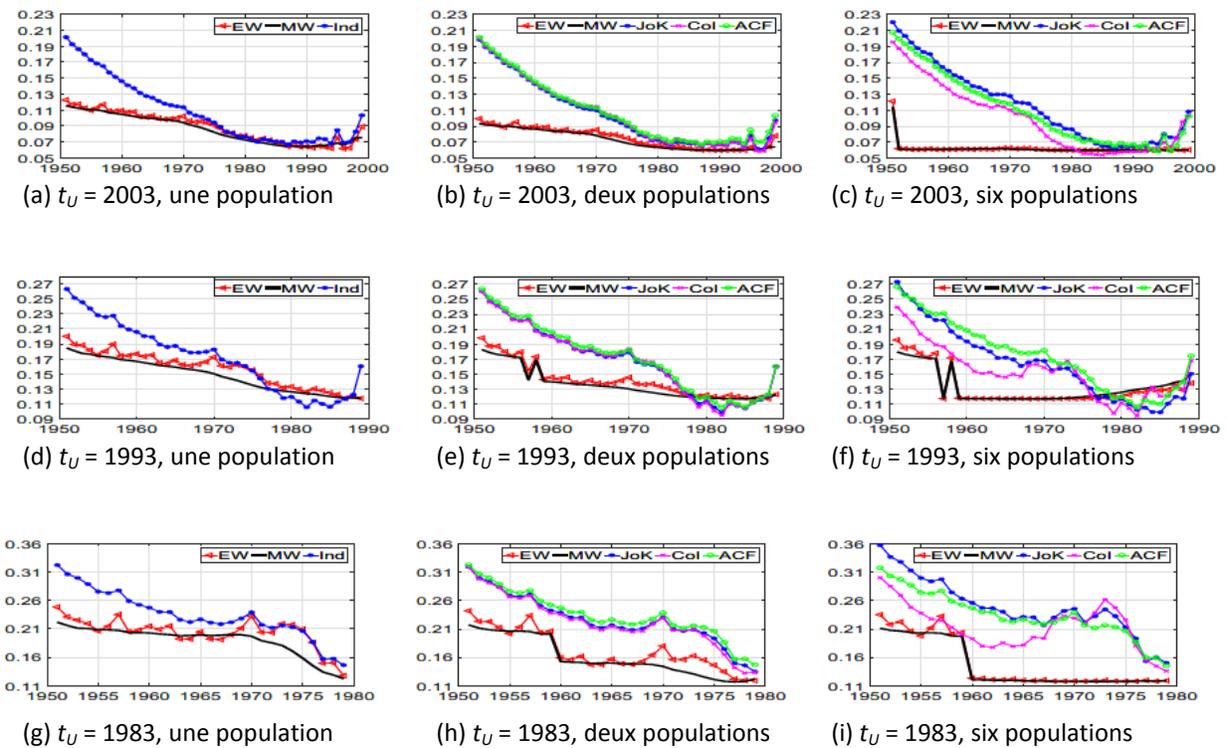
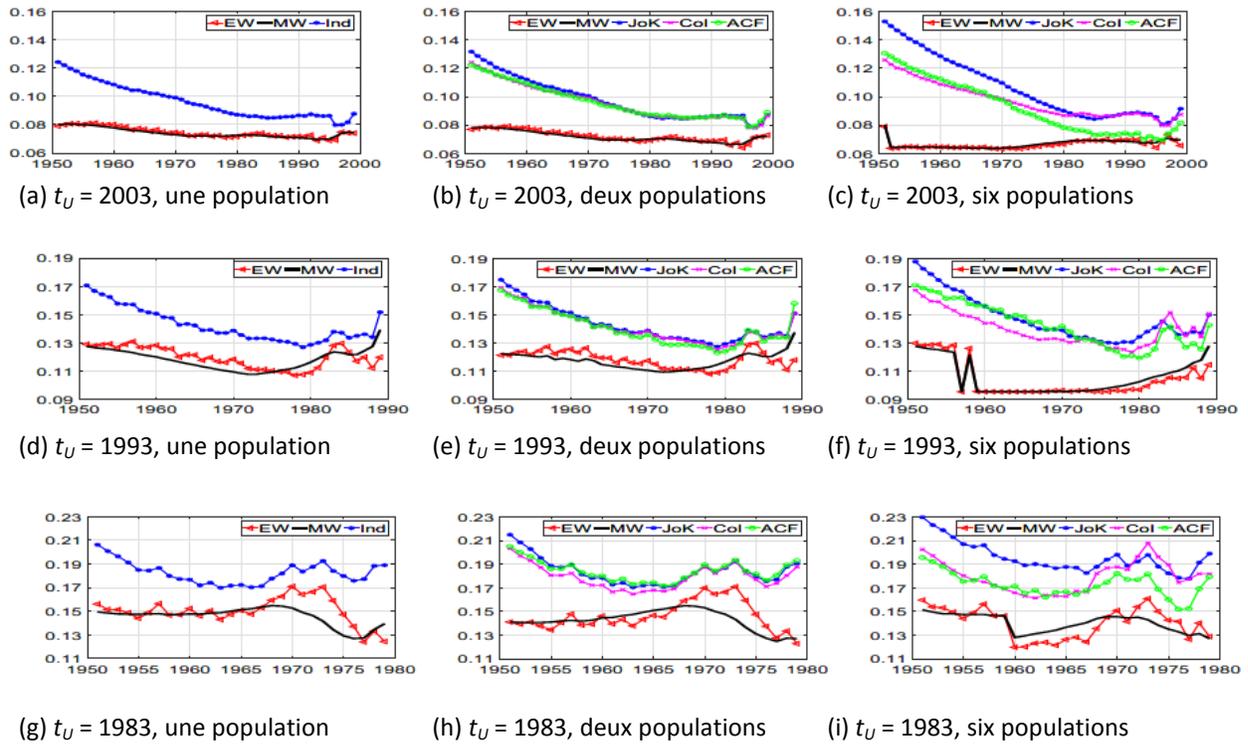


Figure 11 : Moyenne des $AMAPE_{c,g,[t_U+1,2013]}^{t_L,t_U}$ sur 6 populations en fonction de t_L , âges 20 à 84



Le modèle traditionnel de Lee-Carter et ses trois prolongements pour plusieurs populations (k-conjoint, cointégré, augmenté de facteurs communs) ont aussi été ajustés aux mêmes données à des fins de comparaison.

La formule pour $\hat{Y}_{c,g,x,T+1}$, l'estimation de crédibilité hiérarchique du décretement logarithmique du taux central de mortalité pour le pays c , le sexe g et l'âge x sur $[T, T+1]$, selon une structure hiérarchique à cinq niveaux pour le cas particulier ($w_{c,g,x,t} = 1$ pour tous $c = 1, \dots, C$, $g = 1, \dots, G$, $x = 1, \dots, X$ et $t = 1, \dots, T$) est la moyenne, pondérée des facteurs de crédibilité, des décretements annuels moyens des quatre tendances temporelles $\bar{Y}_{c,g,x,t}$, $\bar{Y}_{c,g,\cdot,t}$, $\bar{Y}_{c,\cdot,\cdot,t}$ et $\bar{Y}_{\cdot,\cdot,\cdot,t}$. Pour le modèle de mortalité hiérarchique, nous adoptons aussi les stratégies de la fenêtre extensible (EW) et de la fenêtre mobile (MW), proposées par Tsai et Lin (2017a, b), pour prévoir les taux de mortalité sur deux années ou plus. Selon la stratégie de la fenêtre extensible, l'estimation de crédibilité hiérarchique $\hat{Y}_{c,g,x,T+\tau}$ est constante pour $\tau = 1, 2, \dots$. En conséquence, $\ln(\hat{m}_{c,g,x,T+\tau})$ est une fonction linéaire de τ avec une pente de $\hat{Y}_{c,g,x,T+1}$ et une ordonnée à l'origine de $\ln(m_{c,g,x,T})$.

La qualité prévisionnelle est mesurée par l'AMAPE pour un seul intervalle d'années d'ajustement. D'après les chiffres présentés à la section 3, nous concluons que plus l'étendue de l'intervalle d'années de prévision est grande, plus l'AMAPE des modèles est élevée. Puisque nous ne savons pas quel intervalle d'années d'ajustement produira l'AMAPE la plus faible, et qu'en fait l'AMAPE varie largement dans un intervalle d'années d'ajustement, nous classons les modèles selon l'AMAPE, la moyenne des AMAPE sur l'ensemble des intervalles d'années d'ajustement. Au vu des tableaux 1 à 3, nous concluons que plus le modèle de crédibilité hiérarchique comporte de niveaux, plus il donne de meilleures prévisions et que quelle que soit

l'étendue de l'intervalle d'années de prévision (10, 20 ou 30 ans), les modèles de crédibilité hiérarchique donnent des prévisions plus précises que celles des modèles de Lee-Carter.

Le modèle proposé fait progresser la littérature sur la modélisation de la mortalité de plusieurs populations en utilisant la théorie de la crédibilité hiérarchique, qui est couramment utilisée en assurances IARD, pour modéliser la dépendance des taux de mortalité de plusieurs populations. Le modèle est facile à mettre en place et peut être appliqué à un arbre hiérarchique d'un nombre arbitraire de niveaux pour ajuster un ensemble de données, ou réduit en une structure arborescente à quatre ou trois niveaux. Il s'agit d'une généralisation du modèle de crédibilité proposé par Tsai et Lin (2017b), qui peut être considéré comme ayant une structure arborescente à trois niveaux (population, année, âge). Les taux de mortalité prévus avec un modèle de crédibilité hiérarchique pour plusieurs populations avec dépendance peuvent servir à construire un indice de mortalité permettant de mieux tarifier les titres indexés sur la mortalité.

Annexe A : Modèles de Lee-Carter pour plusieurs populations

A.1 Modèle k-conjoint de Lee-Carter

Carter et Lee (1992) ont proposé le modèle k-conjoint pour régir les covariations des taux de mortalité de plusieurs populations. Ce modèle est construit de la même façon que le modèle indépendant de Lee-Carter, sauf que l'indice variable dans le temps, $k_{t,i} = K_t$, $i = 1, \dots, r$. Le logarithme des taux centraux de mortalité, $\ln(m_{x,t,i})$, des personnes âgées de x dans l'année t et appartenant à la population i peut s'exprimer comme suit :

$$\ln(m_{x,t,i}) = \alpha_{x,i} + \beta_{x,i} \times K_t + \varepsilon_{x,t,i}, \quad x = x_L, \dots, x_U, t = t_L, \dots, t_U, i = 1, \dots, r, \quad (\text{A.1})$$

où $\alpha_{x,i}$ est le facteur de mortalité moyen spécifique à l'âge x pour la population i , $\beta_{x,i}$ est la réaction spécifique à l'âge x à K_t pour la population i , K_t est l'indice commun du niveau de mortalité dans l'année t , et les erreurs du modèle, $\varepsilon_{x,t,i}$, $t = t_L, \dots, t_U$, qui tiennent compte des effets spécifiques à l'âge ignorés par le modèle, sont supposées indépendantes et identiquement distribuées.

Le modèle k-conjoint de Lee-Carter est soumis à deux contraintes : $\sum_{t=t_L}^{t_U} K_t = 0$ et $\sum_{i=1}^r \sum_{x=x_L}^{x_U} \beta_{x,i} = 1$. La première contrainte, $\sum_{t=t_L}^{t_U} K_t = 0$, donne les estimations de α_x ,

$$\hat{\alpha}_{x,i} = \frac{\sum_{t=t_L}^{t_U} \ln(m_{x,t,i})}{t_U - t_L + 1}, \quad x = x_L, \dots, x_U,$$

et la seconde, $\sum_{i=1}^r \sum_{x=x_L}^{x_U} \beta_{x,i} = 1$, implique l'estimation de K_t ,

$$\hat{K}_t = \sum_{i=1}^r \sum_{x=x_L}^{x_U} [\ln(m_{x,t,i}) - \hat{\alpha}_{x,i}], \quad t = t_L, \dots, t_U.$$

Enfin, nous régressons $[\ln(m_{x,t,i}) - \hat{\alpha}_{x,i}]$ sur \hat{K}_t sans le terme constant pour chaque âge x afin d'obtenir $\hat{\beta}_{x,i}$.

L'indice commun variable dans le temps, \hat{K}_t , est supposé suivre une marche aléatoire avec une dérive de ϑ pour la prévision de la mortalité : $\hat{K}_t = \hat{K}_{t+1} + \vartheta + \varepsilon_t$, où les erreurs de tendance

temporelle, ϵ_t , $t = t_L+1, \dots, t_U$, sont supposées être indépendantes et identiquement distribuées, et les erreurs de tendance temporelle, $\{\epsilon_t\}$, sont supposées être indépendantes des erreurs du modèle, $\{\epsilon_{x,t,i}\}$. Ensuite, nous pouvons estimer ϑ , le paramètre de dérive, avec

$$\hat{\theta} = \frac{1}{n-1} \sum_{t=t_L+1}^{t_U} (\hat{K}_t - \hat{K}_{t-1}) = \frac{\hat{K}_{t_U} - \hat{K}_{t_L}}{n-1},$$

où $\eta = t_U - t_L + 1$. Le logarithme du taux central projeté de mortalité pour l'âge x dans l'année $t_U + \tau$ et la population i est donné par

$$\ln(\hat{m}_{x,t_U+\tau,i}) = \hat{\alpha}_{x,i} + \hat{\beta}_{x,i} \times (\hat{K}_{t_U} + \tau \times \hat{\theta}) = \ln(\hat{m}_{x,t_U,i}) + (\hat{\beta}_{x,i} \times \hat{\theta}) \times \tau, \quad \tau = 1, \dots, \quad (\text{A.2})$$

une fonction linéaire de τ avec ordonnée à l'origine $\ln(\hat{m}_{x,t_U,i})$ et pente $(\hat{\beta}_{x,i} \times \hat{\theta})$, où $\ln(\hat{m}_{x,t_U,i}) = \hat{\alpha}_{x,i} + \hat{\beta}_{x,i} \times \hat{K}_{t_U}$.

A.2 Modèle cointégré de Lee-Carter

À la différence du modèle k-conjoint, qui suppose que toutes les populations ont l'indice commun K qui varie dans le temps, le modèle cointégré proposé par Li et Hardy (2011) suppose que l'indice variable dans le temps pour la population i ($i \geq 2$) est en relation linéaire avec l'indice variable dans le temps pour la population 1, la population de base. Par conséquent, dans ce modèle, l'indice variable dans le temps pour la population i ($i \geq 2$) doit faire l'objet d'une nouvelle estimation.

Nous supposons que le taux de mortalité des personnes âgées de x dans l'année t et la population i suit le modèle traditionnel de Lee-Carter :

$$\ln(m_{x,t,i}) = \alpha_{x,i} + \beta_{x,i} \times k_{t,i} + \epsilon_{x,t,i}, \quad x = x_L, \dots, x_U, t = t_L, \dots, t_U, i = 1, \dots, r. \quad (\text{A.3})$$

Le modèle cointégré de Lee-Carter est soumis à deux contraintes : $\sum_{t=t_L}^{t_U} k_{t,i} = 0$ et $\sum_{x=x_L}^{x_U} \hat{\beta}_{x,i} = 1$, $i = 1, \dots, r$. L'estimation de α_x peut être obtenue comme suit par la contrainte $\sum_{t=t_L}^{t_U} k_{t,i} = 0$

$$\hat{\alpha}_{x,i} = \frac{\sum_{t=t_L}^{t_U} \ln(m_{x,t,i})}{t_U - t_L + 1}, \quad x = x_L, \dots, x_U,$$

et l'estimation de $k_{t,i}$ s'obtient comme suit avec la contrainte restante $\sum_{x=x_L}^{x_U} \hat{\beta}_{x,i} = 1$

$$\hat{k}_{t,i} = \sum_{x=x_L}^{x_U} [\ln(m_{x,t,i}) - \hat{\alpha}_{x,i}], \quad x = x_L, \dots, x_U.$$

Ici encore, nous régressons $[\ln(m_{x,t,i}) - \hat{\alpha}_{x,i}]$ sur $\hat{k}_{t,i}$ sans le terme constant pour chaque âge x afin d'obtenir $\hat{\beta}_{x,i}$.

La tendance temporelle $\hat{k}_{t,i}$ est supposée suivre une marche aléatoire avec une dérive de ϑ pour la prévision de la mortalité : $\hat{k}_{t,i} = \hat{k}_{t-1,i} + \vartheta + \epsilon_{t,i}$, où les erreurs de tendance temporelle $\epsilon_{t,i}$, $t = t_L + 1, \dots, t_U$, sont supposées être indépendantes et identiquement distribuées, et les erreurs de tendance temporelle, $\{\epsilon_{t,i}\}$, sont supposées être indépendantes des erreurs du

modèle, $\{\varepsilon_{x,t,i}\}$. Ensuite, nous pouvons estimer le paramètre de dérive ϑ_i pour la population i par $\hat{\theta}_i = (\hat{k}_{t_U,i} - \hat{k}_{t_L,i})/(\eta - 1)$, $i = 1, \dots, \tau$.

Le modèle cointégré de Lee-Carter suppose l'existence d'une relation linéaire et d'un terme d'erreur $e_{t,i}$ entre $\hat{k}_{t,i}$ pour $i = 2, \dots, \tau$ et $\hat{k}_{t,1}$. Plus précisément, $\hat{k}_{t,i} = a_i + b_i \times \hat{k}_{t,1} + e_{t,i}$, $i = 2, \dots, \tau$. Ensuite, nous estimons à nouveau $\hat{k}_{t,i}$ par une régression linéaire simple de la forme $\hat{k}_{t,i} = \hat{a}_i + \hat{b}_i \times \hat{k}_{t,1}$ pour $i = 2, \dots, \tau$, ce qui implique que l'estimation de la dérive de l'indice variable dans le temps, pour la population i , $\hat{\theta}_i$, est donnée par

$$\hat{\theta}_i = \begin{cases} \frac{1}{n-1} \sum_{t=t_L+1}^{t_U} (\hat{k}_{t,1} - \hat{k}_{t-1,1}) = \frac{\hat{k}_{t_U,1} - \hat{k}_{t_L,1}}{n-1} = \hat{\theta}_1, & i = 1, \\ \frac{\hat{k}_{t_U,i} - \hat{k}_{t_L,i}}{n-1} = \hat{b}_i \times \frac{\hat{k}_{t_U,1} - \hat{k}_{t_L,1}}{n-1} = \hat{b}_i \times \hat{\theta}_1, & i = 2, \dots, r. \end{cases}$$

De même, le logarithme des taux centraux prévus de mortalité des personnes âgées de x dans l'année $t_U + \tau$ et la population i est donné par

$$\ln(\hat{m}_{x,t_U+\tau,i}) = \hat{\alpha}_{x,i} + \hat{\beta}_{x,i} \times (\hat{k}_{t_U,i} + \tau \times \hat{\theta}_i) = \ln(\hat{m}_{x,t_U,i}) + (\hat{\beta}_{x,i} \times \hat{\theta}_i) \times \tau, \quad \tau = 1, \dots, \quad (\text{A.4})$$

une fonction linéaire de τ avec $\ln(\hat{m}_{x,t_U,i})$ pour ordonnée à l'origine et $(\hat{\beta}_{x,i} \cdot \hat{\theta}_i)$ pour pente, où $\ln(\hat{m}_{x,t_U,i}) = \hat{\alpha}_{x,i} + \hat{\beta}_{x,i} \times \hat{k}_{t_U,i}$.

A.3 Modèle de Lee-Carter augmenté de facteurs communs

Afin de tenir compte de la divergence dans la prévision à long terme des taux de mortalité de plusieurs populations, Li et Lee (2005) ont proposé un modèle augmenté de facteurs communs qui non seulement prend en compte les points communs dans l'expérience historique, mais aussi les différences individuelles dans les tendances.

Tout d'abord, nous transformons le modèle indépendant de Lee-Carter en un modèle à facteurs communs en introduisant comme suit un terme commun spécifique à l'âge, B_x , et un indice uniforme variable dans le temps, K_t , pour toutes les populations :

$\ln(m_{x,t,i}) = \alpha_{x,i} + B_x \times K_t + \varepsilon_{x,t,i}$, $x = x_L, \dots, x_U$, $t = t_L, \dots, t_U$, $i = 1, \dots, r$, soumis à deux contraintes, $\sum_{t=t_L}^{t_U} K_t = 0$ et $\sum_{i=1}^{\tau} \sum_{x=x_L}^{x_U} w_i B_x = 1$, où w_i , que nous avons défini dans cet article comme étant égal à $1/\tau$, est le coefficient de pondération affecté à la population i et $\sum_{i=1}^{\tau} w_i = 1$. Nous pouvons de même estimer α_x par

$$\hat{\alpha}_{x,i} = \frac{\sum_{t=t_L}^{t_U} \ln(m_{x,t,i})}{t_U - t_L + 1}, \quad x = x_L, \dots, x_U,$$

et K_t par

$$\hat{K}_t = \sum_{i=1}^{\tau} \sum_{x=x_L}^{x_U} w_i \times [\ln(m_{x,t,i}) - \hat{\alpha}_{x,i}], \quad t = t_L, \dots, t_U.$$

Ensuite, \hat{B}_x peut de même être obtenu par la régression de $\sum_{i=1}^{\tau} w_i \times [\ln(m_{x,t,i}) - \hat{\alpha}_{x,i}]$ sur \hat{K}_t sans le terme constant pour chaque âge x .

Afin de prendre en compte les différences individuelles dans les tendances, Li et Lee (2005) ont ajouté le facteur $\beta'_{x,i} \times k'_{t,i}$ au modèle à facteurs communs, pour obtenir

$$\ln(m_{x,t,i}) = \alpha_{x,i} + B_x \times K_t + \beta'_{x,i} \times k'_{t,i} + \varepsilon_{x,t,i}, \quad (\text{A.5})$$

avec la contrainte supplémentaire $\sum_{x=x_L}^{x_U} \beta'_{x,i} = 1$, et le résultat est appelé modèle augmenté de facteurs communs. La contrainte supplémentaire implique $\hat{k}'_{t,i} = \sum_{x=x_L}^{x_U} [\ln(m_{x,t,i}) - \hat{\alpha}_{x,i} - \hat{B}_x \times \hat{K}_t]$ et $\hat{\beta}'_{x,i}$ peut être obtenu par la régression de $[\ln(m_{x,t,i}) - \hat{\alpha}_{x,i} - \hat{B}_x \times \hat{K}_t]$ sur $\hat{k}'_{t,i}$ sans le terme constant pour chaque âge x .

De même, nous supposons que les deux tendances temporelles \hat{K}_t et $\hat{k}'_{t,i}$ suivent une marche aléatoire avec dérivés ϑ et tendances $\hat{\theta}_i$, respectivement. Plus précisément, $\hat{K}_t = \hat{K}_{t-1} + \vartheta + \varepsilon_t$, et $\hat{k}'_{t,i} = \hat{k}'_{t-1,i} + \hat{\theta}'_i + \varepsilon_{t,i}$, où les erreurs de tendance temporelle ε_t et $\varepsilon_{t,i}$ $t = t_L + 1, \dots, t_U$, sont supposées être indépendantes et identiquement distribuées, et les trois termes d'erreur, $\{\varepsilon_{x,t,i}\}$, $\{\varepsilon_t\}$ et $\{\varepsilon_{t,i}\}$, sont supposés être indépendants. Ici encore, les paramètres de dérive sont $\hat{\theta} = (\hat{K}_{t_U} - \hat{K}_{t_L})/(\eta - 1)$ et $\hat{\theta}'_i = (\hat{k}'_{t_U,i} - \hat{k}'_{t_L,i})/(\eta - 1)$.

Enfin, le logarithme des taux centraux prévus de mortalité des personnes âgées de x dans l'année $t_U + \tau$ et la population i peut s'exprimer par

$$\begin{aligned} \ln(\hat{m}_{x,t_U+\tau,i}) &= \hat{\alpha}_{x,i} + \hat{B}_x \times (\hat{K}_{t_U} + \tau \times \hat{\theta}) + \hat{\beta}'_{x,i} \times (\hat{k}'_{t_U,i} + \tau \times \hat{\theta}'_i) \\ &= \ln(\hat{m}_{x,t_U,i}) + (\hat{B}_x \times \hat{\theta} + \hat{\beta}'_{x,i} \times \hat{\theta}'_i) \times \tau, \quad \tau = 1, \dots, \end{aligned} \quad (\text{A.6})$$

une fonction linéaire de τ avec $\ln(\hat{m}_{x,t_U,i})$ pour ordonnée à l'origine et $(\hat{\beta}_x \cdot \hat{\theta} + \beta'_{x,i} \hat{\theta}'_i)$ pour pente, où $\ln(\hat{m}_{x,t_U,i}) = \hat{\alpha}_{x,i} + \hat{\beta}_x \times \hat{K}_{t_U} + \hat{\beta}'_{x,i} \times \hat{k}'_{t_U,i}$.

Annexe B : Démonstrations des propositions 1 à 3

B.1 Démonstration de la proposition 1

Selon (2.7),

$$\begin{aligned}
 \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \hat{Y}_{c,g,x,T+\tau} &= \hat{\alpha}_\tau^{(1)} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \bar{Y}_{c,g,x,\bullet}^{T+\tau} + [(1 - \hat{\alpha}_\tau^{(1)}) \cdot \hat{\alpha}_\tau^{(2)}] \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} \\
 &+ [(1 - \hat{\alpha}_\tau^{(1)}) \cdot (1 - \hat{\alpha}_\tau^{(2)}) \cdot \hat{\alpha}_\tau^{(3)}] \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} \\
 &+ [(1 - \hat{\alpha}_\tau^{(1)}) \cdot (1 - \hat{\alpha}_\tau^{(2)}) \cdot (1 - \hat{\alpha}_\tau^{(3)})] \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} \\
 &= \hat{\alpha}_\tau^{(1)} \cdot C \cdot G \cdot X \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} + [(1 - \hat{\alpha}_\tau^{(1)}) \cdot \hat{\alpha}_\tau^{(2)}] \cdot C \cdot G \cdot X \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} \\
 &+ [(1 - \hat{\alpha}_\tau^{(1)}) \cdot (1 - \hat{\alpha}_\tau^{(2)}) \cdot \hat{\alpha}_\tau^{(3)}] \cdot C \cdot G \cdot X \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} \\
 &+ [(1 - \hat{\alpha}_\tau^{(1)}) \cdot (1 - \hat{\alpha}_\tau^{(2)}) \cdot (1 - \hat{\alpha}_\tau^{(3)})] \cdot C \cdot G \cdot X \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} \\
 &= C \cdot G \cdot X \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau}.
 \end{aligned}$$

En divisant par (C, G, X) des deux côtés, nous avons pour $\tau = 1, 2, \dots$,

$$\frac{1}{C \cdot G \cdot X} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \hat{Y}_{c,g,x,T+\tau} = \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} = \frac{1}{C \cdot G \cdot X} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \bar{Y}_{c,g,x,\bullet}^{T+\tau}.$$

B.2 Démonstration de la proposition 2

Cette proposition se démontre par induction mathématique sur τ . Tout d'abord, pour $\tau = 2$, par définition, (2.8) et la proposition 1,

$$\begin{aligned}
 \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+2} &= \frac{1}{C \cdot G \cdot X} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \bar{Y}_{c,g,x,\bullet}^{T+2} \\
 &= \frac{1}{C \cdot G \cdot X \cdot (T+1)} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \left[\sum_{t=1}^T Y_{c,g,x,t} + \hat{Y}_{c,g,x,T+1} \right] \\
 &= \frac{T}{T+1} \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+1} + \frac{1}{T+1} \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+1} = \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+1}.
 \end{aligned}$$

Ensuite, nous supposons la validité de l'égalité $\bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} = \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+1}$. L'équation (2.8) et la proposition 1 entraînent

$$\begin{aligned}
\bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau+1} &= \frac{1}{C \cdot G \cdot X} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \bar{Y}_{c,g,x,\bullet}^{T+\tau+1} \\
&= \frac{1}{C \cdot G \cdot X \cdot (T + \tau)} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \left[\left(\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau-1} \hat{Y}_{c,g,x,t} \right) + \hat{Y}_{c,g,x,T+\tau} \right] \\
&= \frac{1}{C \cdot G \cdot X \cdot (T + \tau)} \sum_{c=1}^C \sum_{g=1}^G \sum_{x=1}^X \left[(T + \tau - 1) \cdot \bar{Y}_{c,g,x,\bullet}^{T+\tau} + \hat{Y}_{c,g,x,T+\tau} \right] \\
&= \frac{T + \tau - 1}{T + \tau} \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} + \frac{1}{T + \tau} \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} = \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+1}.
\end{aligned}$$

Nous avons donc démontré que $\bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} = \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+1}$ pour $\tau = 2, 3, \dots$ selon la stratégie EW.

B.3 Démonstration de la proposition 3

Posons $f_\tau = (T + \tau) \hat{\sigma}_2^2 + \hat{\sigma}_1^2$, $g_\tau = X (T + \tau) \hat{\sigma}_3^2 + f_\tau$ et $h_\tau = G X (T + \tau) \hat{\sigma}_4^2 + g_\tau$. Ensuite, selon (2.12), (2.13) et (2.14), nous avons

$$\begin{aligned}
\hat{\alpha}_\tau^{(1)} &= \frac{f_{\tau-1} - \hat{\sigma}_1^2}{f_{\tau-1}}, & 1 - \hat{\alpha}_\tau^{(1)} &= \frac{\hat{\sigma}_1^2}{f_{\tau-1}}, & (1 - \hat{\alpha}_\tau^{(1)}) \hat{\alpha}_\tau^{(2)} &= \frac{\hat{\sigma}_1^2 (g_{\tau-1} - f_{\tau-1})}{f_{\tau-1} g_{\tau-1}}, \\
\hat{\alpha}_\tau^{(2)} &= \frac{g_{\tau-1} - f_{\tau-1}}{g_{\tau-1}}, & 1 - \hat{\alpha}_\tau^{(2)} &= \frac{f_{\tau-1}}{g_{\tau-1}}, & (1 - \hat{\alpha}_\tau^{(1)}) (1 - \hat{\alpha}_\tau^{(2)}) \hat{\alpha}_\tau^{(3)} &= \frac{\hat{\sigma}_1^2 (h_{\tau-1} - g_{\tau-1})}{g_{\tau-1} h_{\tau-1}}, \\
\hat{\alpha}_\tau^{(3)} &= \frac{h_{\tau-1} - g_{\tau-1}}{h_{\tau-1}}, & 1 - \hat{\alpha}_\tau^{(3)} &= \frac{g_{\tau-1}}{h_{\tau-1}}, & (1 - \hat{\alpha}_\tau^{(1)}) (1 - \hat{\alpha}_\tau^{(2)}) (1 - \hat{\alpha}_\tau^{(3)}) &= \frac{\hat{\sigma}_1^2}{h_{\tau-1}}.
\end{aligned}$$

D'après (2.7), nous pouvons exprimer $\hat{Y}_{c,g,x,T+\tau}$ par

$$\begin{aligned}
&\hat{Y}_{c,g,x,T+\tau} \\
&= \frac{1}{f_{\tau-1} g_{\tau-1} h_{\tau-1}} \left[g_{\tau-1} h_{\tau-1} (f_{\tau-1} - \hat{\sigma}_1^2) \cdot \bar{Y}_{c,g,x,\bullet}^{T+\tau} + \hat{\sigma}_1^2 (g_{\tau-1} - f_{\tau-1}) h_{\tau-1} \cdot \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} \right. \\
&\quad \left. + \hat{\sigma}_1^2 (h_{\tau-1} - g_{\tau-1}) f_{\tau-1} \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} + \hat{\sigma}_1^2 f_{\tau-1} g_{\tau-1} \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} \right]. \tag{B.1}
\end{aligned}$$

Notre objectif est de montrer que $f_\tau g_\tau h_\tau \hat{Y}_{c,g,x,T+\tau+1} = f_\tau g_\tau h_\tau \hat{Y}_{c,g,x,T+\tau}$, ce qui implique que $\hat{Y}_{c,g,x,T+\tau+1} = \hat{Y}_{c,g,x,T+\tau}$, $\tau = 1, 2, \dots$, et donc que $\hat{Y}_{c,g,x,T+\tau} = \hat{Y}_{c,g,x,T+1}$, $\tau = 2, 3, \dots$. Nous posons d'abord

$$DIFF = f_\tau g_\tau h_\tau \cdot \hat{Y}_{c,g,x,T+\tau+1} - f_{\tau-1} g_{\tau-1} h_{\tau-1} \cdot \hat{Y}_{c,g,x,T+\tau}.$$

Notre objectif consiste maintenant à démontrer que

$$f_\tau g_\tau h_\tau \cdot \hat{Y}_{c,g,x,T+\tau} = f_\tau g_\tau h_\tau \cdot \hat{Y}_{c,g,x,T+\tau+1} = f_{\tau-1} g_{\tau-1} h_{\tau-1} \cdot \hat{Y}_{c,g,x,T+\tau} + DIFF,$$

ou, de façon équivalente, que $DIFF = (f_\tau g_\tau h_\tau - f_{\tau-1} g_{\tau-1} h_{\tau-1}) \hat{Y}_{c,g,x,T+\tau}$.

Selon (B.1), $DIFF$ est la somme des quatre expressions (B.2) à (B.5) suivantes :

$$\begin{aligned}
& g_\tau h_\tau (f_\tau - \hat{\sigma}_1^2) \cdot \bar{Y}_{c,g,x,\bullet}^{T+\tau+1} - g_{\tau-1} h_{\tau-1} (f_{\tau-1} - \hat{\sigma}_1^2) \cdot \bar{Y}_{c,g,x,\bullet}^{T+\tau} \\
&= \hat{\sigma}_2^2 \left[g_\tau h_\tau \left(\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau} \hat{Y}_{c,g,x,t} \right) - g_{\tau-1} h_{\tau-1} \left(\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau-1} \hat{Y}_{c,g,x,t} \right) \right] \\
&= \hat{\sigma}_2^2 \left[(g_\tau h_\tau - g_{\tau-1} h_{\tau-1}) \left(\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau-1} \hat{Y}_{c,g,x,t} \right) + g_\tau h_\tau \cdot \hat{Y}_{c,g,x,T+\tau} \right] \\
&= (g_\tau h_\tau - g_{\tau-1} h_{\tau-1}) (f_{\tau-1} - \hat{\sigma}_1^2) \cdot \bar{Y}_{c,g,x,\bullet}^{T+\tau} + \hat{\sigma}_2^2 g_\tau h_\tau \cdot \hat{Y}_{c,g,x,T+\tau}, \tag{B.2}
\end{aligned}$$

$$\begin{aligned}
& \hat{\sigma}_1^2 (g_\tau - f_\tau) h_\tau \cdot \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau+1} - \hat{\sigma}_1^2 (g_{\tau-1} - f_{\tau-1}) h_{\tau-1} \cdot \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} \\
&= \hat{\sigma}_1^2 \hat{\sigma}_3^2 \left[h_\tau \sum_{x=1}^X \left(\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau} \hat{Y}_{c,g,x,t} \right) - h_{\tau-1} \sum_{x=1}^X \left(\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau-1} \hat{Y}_{c,g,x,t} \right) \right] \\
&= \hat{\sigma}_1^2 \hat{\sigma}_3^2 \left[(h_\tau - h_{\tau-1}) \sum_{x=1}^X \left(\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau-1} \hat{Y}_{c,g,x,t} \right) + h_\tau \sum_{x=1}^X \hat{Y}_{c,g,x,T+\tau} \right] \\
&= \hat{\sigma}_1^2 (h_\tau - h_{\tau-1}) (g_{\tau-1} - f_{\tau-1}) \cdot \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} + \hat{\sigma}_1^2 \hat{\sigma}_3^2 h_\tau \sum_{x=1}^X \hat{Y}_{c,g,x,T+\tau}, \tag{B.3}
\end{aligned}$$

$$\begin{aligned}
& \hat{\sigma}_1^2 (h_\tau - g_\tau) f_\tau \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau+1} - \hat{\sigma}_1^2 (h_{\tau-1} - g_{\tau-1}) f_{\tau-1} \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} \\
&= \hat{\sigma}_1^2 \hat{\sigma}_4^2 \left[f_\tau \sum_{g=1}^G \sum_{x=1}^X \left(\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau} \hat{Y}_{c,g,x,t} \right) \right. \\
&\quad \left. - f_{\tau-1} \sum_{g=1}^G \sum_{x=1}^X \left(\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau-1} \hat{Y}_{c,g,x,t} \right) \right] \\
&= \hat{\sigma}_1^2 \hat{\sigma}_4^2 \left[(f_\tau - f_{\tau-1}) \sum_{g=1}^G \sum_{x=1}^X \left(\sum_{t=1}^T Y_{c,g,x,t} + \sum_{t=T+1}^{T+\tau-1} \hat{Y}_{c,g,x,t} \right) + f_\tau \sum_{g=1}^G \sum_{x=1}^X \hat{Y}_{c,g,x,T+\tau} \right] \\
&= \hat{\sigma}_1^2 (f_\tau - f_{\tau-1}) (h_{\tau-1} - g_{\tau-1}) \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} + \hat{\sigma}_1^2 \hat{\sigma}_4^2 f_\tau \sum_{g=1}^G \sum_{x=1}^X \hat{Y}_{c,g,x,T+\tau}, \tag{B.4}
\end{aligned}$$

$$\hat{\sigma}_1^2 f_\tau g_\tau \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau+1} - \hat{\sigma}_1^2 f_{\tau-1} g_{\tau-1} \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} = \hat{\sigma}_1^2 (f_\tau g_\tau - f_{\tau-1} g_{\tau-1}) \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} \tag{B.5}$$

Ensuite,

$$f_\tau g_\tau h_\tau - f_{\tau-1} g_{\tau-1} h_{\tau-1} = (f_{\tau-1} + \hat{\sigma}_2^2) g_\tau h_\tau - f_{\tau-1} g_{\tau-1} h_{\tau-1} = \hat{\sigma}_2^2 g_\tau h_\tau + f_{\tau-1} (g_\tau h_\tau - g_{\tau-1} h_{\tau-1}),$$

et

$$(f_{\tau}g_{\tau}h_{\tau} - f_{\tau-1}g_{\tau-1}h_{\tau-1}) \cdot \hat{Y}_{c,g,x,T+\tau} = \hat{\sigma}_2^2 g_{\tau} h_{\tau} \cdot \hat{Y}_{c,g,x,T+\tau} + f_{\tau-1}(g_{\tau}h_{\tau} - g_{\tau-1}h_{\tau-1}) \cdot \hat{Y}_{c,g,x,T+\tau}. \quad (\text{B.6})$$

Le premier terme de (B.6) et le deuxième terme de (B.2) s'annulent, et le deuxième terme de (B.6) selon (B.1) donne

$$\begin{aligned} & (g_{\tau}h_{\tau} - g_{\tau-1}h_{\tau-1})(f_{\tau-1} - \hat{\sigma}_1^2) \cdot \bar{Y}_{c,g,x,\bullet}^{T+\tau} + \frac{\hat{\sigma}_1^2}{g_{\tau-1}}(g_{\tau}h_{\tau} - g_{\tau-1}h_{\tau-1})(g_{\tau-1} - f_{\tau-1}) \cdot \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} \\ & + \frac{\hat{\sigma}_1^2}{g_{\tau-1}h_{\tau-1}}(g_{\tau}h_{\tau} - g_{\tau-1}h_{\tau-1})(h_{\tau-1} - g_{\tau-1})f_{\tau-1} \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} \\ & + \frac{\hat{\sigma}_1^2}{h_{\tau-1}}(g_{\tau}h_{\tau} - g_{\tau-1}h_{\tau-1})f_{\tau-1} \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau}. \end{aligned} \quad (\text{B.7})$$

Le premier terme de (B.7) et le premier terme de (B.2) s'annulent, si bien qu'il nous reste à démontrer que la somme des trois derniers termes de (B.7) est égale à la somme de (B.3), (B.4) et (B.5). Puisque

$$\begin{aligned} \sum_{x=1}^X \hat{Y}_{c,g,x,T+\tau} &= \frac{X}{f_{\tau-1}g_{\tau-1}h_{\tau-1}} \left\{ \left[g_{\tau-1}h_{\tau-1}(f_{\tau-1} - \hat{\sigma}_1^2) + \hat{\sigma}_1^2(g_{\tau-1} - f_{\tau-1})h_{\tau-1} \right] \cdot \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} \right. \\ & \quad \left. + \hat{\sigma}_1^2(h_{\tau-1} - g_{\tau-1})f_{\tau-1} \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} + \hat{\sigma}_1^2 f_{\tau-1}g_{\tau-1} \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} \right\} \\ &= \frac{X}{f_{\tau-1}g_{\tau-1}h_{\tau-1}} \left[f_{\tau-1}h_{\tau-1}(g_{\tau-1} - \hat{\sigma}_1^2) \cdot \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} \right. \\ & \quad \left. + \hat{\sigma}_1^2(h_{\tau-1} - g_{\tau-1})f_{\tau-1} \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} + \hat{\sigma}_1^2 f_{\tau-1}g_{\tau-1} \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} \right] \\ &= \frac{X(g_{\tau-1} - \hat{\sigma}_1^2)}{g_{\tau-1}} \cdot \bar{Y}_{c,g,\bullet,\bullet}^{T+\tau} + \frac{X\hat{\sigma}_1^2(h_{\tau-1} - g_{\tau-1})}{g_{\tau-1}h_{\tau-1}} \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} + \frac{X\hat{\sigma}_1^2}{h_{\tau-1}} \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} \end{aligned} \quad (\text{B.8})$$

et

$$\begin{aligned} \sum_{g=1}^G \sum_{x=1}^X \hat{Y}_{c,g,x,T+\tau} &= \frac{GX}{f_{\tau-1}g_{\tau-1}h_{\tau-1}} \left\{ \left[f_{\tau-1}h_{\tau-1}(g_{\tau-1} - \hat{\sigma}_1^2) + \hat{\sigma}_1^2(h_{\tau-1} - g_{\tau-1})f_{\tau-1} \right] \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} \right. \\ & \quad \left. + \hat{\sigma}_1^2 f_{\tau-1}g_{\tau-1} \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau} \right\} \\ &= \frac{GX(h_{\tau-1} - \hat{\sigma}_1^2)}{h_{\tau-1}} \cdot \bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau} + \frac{GX\hat{\sigma}_1^2}{h_{\tau-1}} \cdot \bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau}, \end{aligned} \quad (\text{B.9})$$

en comparant les coefficients de $\bar{Y}_{c,g,\bullet,\bullet}^{T+\tau}$, $\bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau}$ et $\bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau}$ dans les trois derniers termes de (B.7) avec ceux dans (B.3), (B.4) et (B.5) associés avec (B.8) et (B.9), il suffit de montrer que

1. coefficient de $\bar{Y}_{c,g,\bullet,\bullet}^{T+\tau}$:

$$\frac{\hat{\sigma}_1^2}{g_{\tau-1}}(g_{\tau}h_{\tau} - g_{\tau-1}h_{\tau-1})(g_{\tau-1} - f_{\tau-1}) = \hat{\sigma}_1^2(h_{\tau} - h_{\tau-1})(g_{\tau-1} - f_{\tau-1}) + \frac{X\hat{\sigma}_1^2\hat{\sigma}_3^2h_{\tau}(g_{\tau-1} - \hat{\sigma}_1^2)}{g_{\tau-1}}; \quad (\text{B.10})$$

2. coefficient de $\bar{Y}_{c,\bullet,\bullet,\bullet}^{T+\tau}$:

$$\begin{aligned} & \frac{\hat{\sigma}_1^2}{g_{\tau-1}h_{\tau-1}}(g_{\tau}h_{\tau} - g_{\tau-1}h_{\tau-1})(h_{\tau-1} - g_{\tau-1})f_{\tau-1} \\ = & \frac{X\hat{\sigma}_1^4\hat{\sigma}_3^2h_{\tau}(h_{\tau-1} - g_{\tau-1})}{g_{\tau-1}h_{\tau-1}} + \hat{\sigma}_1^2(f_{\tau} - f_{\tau-1})(h_{\tau-1} - g_{\tau-1}) + \frac{GX\hat{\sigma}_1^2\hat{\sigma}_4^2f_{\tau}(h_{\tau-1} - \hat{\sigma}_1^2)}{h_{\tau-1}}; \end{aligned} \quad (\text{B.11})$$

3. coefficient de $\bar{Y}_{\bullet,\bullet,\bullet,\bullet}^{T+\tau}$:

$$\frac{\hat{\sigma}_1^2}{h_{\tau-1}}(g_{\tau}h_{\tau} - g_{\tau-1}h_{\tau-1})f_{\tau-1} = \frac{X\hat{\sigma}_1^4\hat{\sigma}_3^2h_{\tau}}{h_{\tau-1}} + \frac{GX\hat{\sigma}_1^4\hat{\sigma}_4^2f_{\tau}}{h_{\tau-1}} + \hat{\sigma}_1^2(f_{\tau}g_{\tau} - f_{\tau-1}g_{\tau-1}). \quad (\text{B.12})$$

Après réarrangement et simplification, les trois équations (B.10), (B.11) et (B.12) que nous devons montrer deviennent

$$(g_{\tau-1} - f_{\tau-1})(g_{\tau} - g_{\tau-1}) = X\hat{\sigma}_3^2(g_{\tau-1} - \hat{\sigma}_1^2),$$

$$(h_{\tau-1} - g_{\tau-1})(f_{\tau-1}g_{\tau}h_{\tau} - f_{\tau}g_{\tau-1}h_{\tau-1} - X\hat{\sigma}_1^2\hat{\sigma}_3^2h_{\tau}) = GX\hat{\sigma}_4^2f_{\tau}g_{\tau-1}(h_{\tau-1} - \hat{\sigma}_1^2),$$

$$g_{\tau}(f_{\tau-1}h_{\tau} - f_{\tau}h_{\tau-1}) = X\hat{\sigma}_1^2(\hat{\sigma}_3^2h_{\tau} + G\hat{\sigma}_4^2f_{\tau}),$$

respectivement, ce qui peut être vérifié directement par les définitions de f_{τ} , g_{τ} et h_{τ} .

Bibliographie

- Bühlmann, H. « Experience rating and credibility », *Astin Bulletin*, vol. 4, 1967, p. 199-207.
- Bühlmann, H. et Gisler, A. *A Course in credibility theory and its applications*, Springer, Pays-Bas, 2005.
- Bühlmann, H. et Straub, E. « Glaubwürdigkeit frschadenstze », *Bulletin of the Swiss Association of Actuaries* 70, 1970, p. 111-133.
- Cairns, A.J.G., Blake, D. et Dowd, K. « A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration », *Journal of Risk and Insurance*, 73, 2006, p. 687-718.
- Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A. et Balevich, I. « A quantitative comparison of stochastic mortality models using data from England and Wales and the United States », *North American Actuarial Journal*, vol. 13, n° 1, 2009, p. 1-35.
- Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D. et Khalaf-Allah, M. « Bayesian stochastic mortality modelling for two populations », *Astin Bulletin*, vol. 41, n° 1, 2011, p. 29-59.
- Carter, L.R. and Lee, R.D. « Forecasting demographic components: Modeling and forecasting U.S. sex differentials in mortality », *International Journal of Forecasting*, vol. 8, n° 3, 1992, p. 393-411.
- Chen, H., MacMinn, R. et Sun, T. « Multi-population mortality models: A factor copula approach », *Insurance: Mathematics and Economics*, vol. 63, 2015, p. 135-146.
- Enchev, V., Kleinow, T. et Cairns, A.J.G. « Multi-population mortality models: Fitting, forecasting and comparisons », *Scandinavian Actuarial Journal*, 2016, p. 1-24.
- Lee, R.D. et Carter, L.R. « Modelling and forecasting U.S. mortality », *Journal of the American Statistical Association*, vol. 87, 1992, p. 659-671.
- Li, J.S.H. et Hardy, M.R. « Measuring basis risk in longevity hedges », *North American Actuarial Journal*, vol. 15, n° 2, 2011, p. 177-200.
- Li, J.S.H., Hardy, M.R. et Tan, K.S. « Uncertainty in mortality forecasting: An extension to the classical Lee-Carter approach », *Astin Bulletin*, vol. 39, n° 1, 2009, p. 137-164.
- Li, J.S.H., Chan, W.S. et Zhou, R. « Semicohherent multipopulation mortality modeling: The impact on longevity risk securitization », *Journal of Risk and Insurance*, 2016.
- Li, J.S.H., Zhou, R. et Hardy, M. « A step-by-step guide to building two-population stochastic mortality models », *Insurance: Mathematics and Economics*, vol. 63, 2015, p. 121-134.
- Li, N. et Lee, R.D. « Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method », *Demography*, vol. 42, n° 3, 2005, p. 575-594.
- Lin, T., Wang, C.W. et Tsai, C.C.L. « Age-specific copula-AR-GARCH mortality models », *Insurance: Mathematics and Economics*, vol. 61, 2015, p. 110 à 124.
- Mitchell, D., Brockett, P., Mendoza-Arriaga, R. et Muthuraman, K. « Modeling and forecasting mortality rates », *Insurance: Mathematics and Economics*, vol. 52, n° 2, 2013, p. 275-285.
- Plat, R. « On stochastic mortality modeling », *Insurance: Mathematics and Economics*, vol. 45, 2009, p. 393-404.
- Renshaw, A.E. et Haberman, S. « A cohort-based extension to the Lee-Carter model for mortality reduction factors », *Insurance: Mathematics and Economics*, vol. 38, 2006, p. 556-570.

- Tsai, C.C.L. et Lin, T. « Incorporating Bühlmann credibility approach to improving mortality forecasting », *Scandinavian Actuarial Journal*, 2017a, p. 419-440.
- Tsai, C.C.L. et Lin, T. « A Bühlmann credibility approach to modeling mortality rates », *North American Actuarial Journal*, vol. 21, n° 2, 2017b, p. 204-227.
- Tsai, C.C.L. et Yang, S. « A linear regression approach to modeling mortality rates of different forms », *North American Actuarial Journal*, vol. 19, n° 1, 2015, p. 1-23.
- Wang, C.W., Yang, S.S., et Huang, H.C. « Modeling multi-country mortality dependence and its application in pricing survivor index swaps - A dynamic copula approach », *Insurance: Mathematics and Economics*, vol. 63, 2015, p. 30-39.