INSTITUT DE STATISTIQUE

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

DISCUSSION PAPER

0219

Importance de la période d'observation et des âges considérés dans la projection de la mortalité selon la méthode de Lee-Carter

A. DELWARDE & M. DENUIT

http://www.stat.ucl.ac.be

Importance de la période d'observation et des âges considérés dans la projection de la mortalité selon la méthode de Lee-Carter

Antoine Delwarde & Michel Denuit
Institut de Statistique & Institut des Sciences Actuarielles
Université Catholique de Louvain
Voie du Roman Pays, 20
Belgium

August 27, 2002

Résumé

Cet article a pour but d'examiner comment les projections de mortalité obtenues à l'aide de la méthode de Lee-Carter sont influencées par la période d'observation et par les âges considérés.

Summary

This paper aims to examine how the mortality projections obtained with the Lee-Carter method depend on the observation period and the age range considered.

1 Introduction

La mortalité en Belgique est un phénomène en pleine mutation, comme l'ont mis en évidence Brouhns & Denuit (2001a). Les calculs effectués sur base des tables du moment (ou transversales), bien qu'ils contiennent un chargement de sécurité implicite pour les opérations en cas de décès (du moins en période d'allongement de la durée de la vie humaine), conduisent quasi-systématiquement à des sous-estimations des primes relatives aux opérations en cas de vie. Si les assureurs désirent disposer de bases techniques leur permettant de tarifer des produits où le risque viager est important (notamment, des rentes viagères achetées au moment du départ à la retraite) et d'estimer correctement les provisions à constituer en vue de faire face à l'allongement de la vie humaine, il est essentiel d'établir des tables de mortalité prospectives.

Le point de départ de toute analyse de l'évolution de la mortalité consiste à récolter pour un ensemble d'années $t_{\min}, \ldots, t_{\max}$ et d'âges $x_{\min}, \ldots, x_{\max}$ les nombres annuels de décès à chaque âge et les expositions au risque correspondantes. Ceci permet alors d'estimer les taux de mortalité. Contrairement à d'autres auteurs, nous travaillons ici avec des données annuelles ventilées par sexe et par âge. Les tables triennales de l'INS ne sont en effet disponibles régulièrement que depuis récemment. Les précédentes couvraient des périodes de longueur souvent différente, et avaient une périodicité variable, ce qui risquait de biaiser les résultats de l'analyse.

La première étape de toute projection de mortalité consiste à réduire la dimensionalité des données. Il est en effet impossible de traiter simultanément autant de séries chronologiques décrivant l'évolution au cours du temps des quotients de mortalité ou des taux de mortalité aux différents âges.

Une première approche, très appréciée par les actuaires, consiste à ajuster les observations de chaque année à l'aide d'un modèle paramétrique (Makeham, par exemple); voyez par exemple Benjamin & Pollard (1993) ou Benjamin & Soliman (1993). Ceci permet de condenser l'information annuelle dans un petit nombre de paramètres (3 dans notre exemple). Ensuite, l'évolution au cours du temps de ces paramètres est à son tour modélisée, afin de fournir des projections de la mortalité dans le futur. Il ne faut néanmoins pas perdre de vue que, chaque année, ces paramètres seront vraisemblablement fort corrélés, ayant été estimés sur base des mêmes données. On ne peut donc pas se contenter d'une modélisation univariée mais on doit recourir à un modèle de série temporelle multivariée. Ce point de vue est malheureusement trop souvent ignoré dans la littérature. De plus, la pertinence de cette approche dépend bien évidemment de la bonne spécification du modèle paramétrique. Si ce dernier s'avère erroné, cela compromet gravement les projections qui en découlent.

Dans cet article, nous optons pour la méthode de Lee & Carter (1992), qui a fait ses preuves en démographie et a été présentée aux actuaires par Lee (2000). L'idée est de passer par une décomposition en valeurs singulières de la matrice des taux de mortalité (doublement indexés, par l'âge et le temps calendaire). La matrice initiale sera ainsi approximée au rang 1 par un produit de deux vecteurs propres: l'un d'entre eux traduira l'effet de l'âge, et l'autre l'effet du temps calendaire. Il suffira alors d'extrapoler dans le futur le vecteur décrivant l'évolution temporelle pour en déduire des tables de mortalité prospectives. Cette projection se fera à l'aide d'un modèle ARIMA, en suivant la méthodologie désormais classique de Box et Jenkins. Nous serons alors à même de calculer les valeurs actuelles de rentes viagères octroyées à un individu de la population belge. Ces valeurs permettent par exemple d'apprécier le coût réel des promesses de pension dans le cadre de la sécurité sociale.

Dans cet article, nous désirons examiner l'impact du choix de la période d'observation $[t_{\min}, t_{\max}]$ et de la fourchette d'âges $[x_{\min}, x_{\max}]$ sur les projections obtenues à l'aide de la méthode de Lee-Carter.

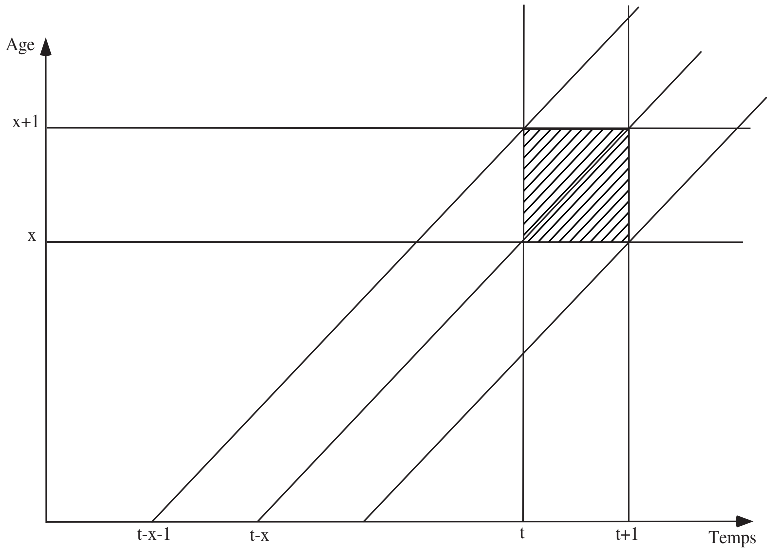

Figure 1: *Carré dans lequel en vertu de* (1) *le taux de mortalité est supposé constant.*

2 Tables de mortalité longitudinales

Les tables de mortalité longitudinales permettent de déterminer le temps restant à vivre pour un individu compte tenu, non pas des conditions du moment, mais de l'évolution future présumée des conditions de vie. Pour ce faire, nous introduirons des indicateurs démographiques doublement indicés, par l'âge et le temps calendaire. Plus précisément, nous allons travailler avec les quantités $q_x(t)$ et $\mu_x(t)$ définies comme suit:

$q_x(t)$ = probabilité de décéder à l'âge x durant l'année t;

$\mu_x(t)$ = taux de mortalité à l'âge x durant l'année t.

Dorénavant, nous noterons $T_x(t)$ la durée de vie restante d'un individu qui a atteint l'âge x au cours de l'année t. Ainsi, $q_x(t) = \Pr[T_x(t) \leq 1]$.

Tout naturellement, nous supposons ici que le taux instantané de mortalité $\mu_x(t)$ est constant dans chaque carré du diagramme de Lexis représenté à la Figure 1, i.e. quels que soient l'âge x et l'année t, x et t entiers,

$$\mu_{x+\xi}(t+\tau) = \mu_x(t) \text{ pour tout } 0 \leq \xi, \tau < 1. \tag{1}$$

Cette hypothèse entraîne notamment que

$$p_x(t) = 1 - q_x(t) = \exp\big(-\mu_x(t)\big).$$

Calculons l'espérance de vie d'un individu d'âge x l'année t sous l'hypothèse (1):

$$
\begin{aligned}
e_x(t) &= \int_{\xi\geq 0} \Pr[T_x(t)>\xi]d\xi \\
&= \int_{\xi=0}^{1} \exp\big(-\mu_x(t)\xi\big)d\xi \\
&\quad +\sum_{k\geq 1}\left\{\prod_{j=0}^{k-1}\exp\big(-\mu_{x+j}(t+j)\big)\right\}\int_{\xi=k}^{k+1}\exp\big(-\mu_{x+k}(t+k)(\xi-k)\big)d\xi \\
&= \frac{1-\exp\big(-\mu_x(t)\big)}{\mu_x(t)} \\
&\quad +\sum_{k\geq 1}\left\{\prod_{j=0}^{k-1}\exp\big(-\mu_{x+j}(t+j)\big)\right\}\frac{1-\exp\big(-\mu_{x+k}(t+k)\big)}{\mu_{x+k}(t+k)}. \qquad (2)
\end{aligned}
$$

On obtient également, toujours sous l'hypothèse (1), l'exposition au risque à l'âge x durant l'année t (à savoir, le temps total vécu par ces individus durant l'année t), notée $L_x(t)$, en fonction de l'effectif $l_x(t)$ d'âge x au 1er janvier de l'année t, donnée par

$$
L_x(t)=\int_{\xi=0}^{1} l_{x+\xi}(t)d\xi=\frac{-l_x(t)q_x(t)}{\ln(1-q_x(t))}
$$

et le prix d'une rente viagère vendue l'année t à un individu d'âge x

$$
a_x(t)=\sum_{k\geq 0}\left\{\prod_{j=0}^{k}p_{x+j}(t+j)\right\}\left(\frac{1}{1+i}\right)^{k+1} \qquad (3)
$$

où i est le taux d'intérêt annuel.

Il est évident que nous ne pouvons pas nous permettre de considérer toutes les séries chronologiques $\{q_x(t),\ t=t_{\min},\ldots,t_{\max}\}$, $x=x_{\min},\ldots,x_{\max}$, ou $\{\mu_x(t),\ t=t_{\min},\ldots,t_{\max}\}$, $x=x_{\min},\ldots,x_{\max}$, pour en déduire des prévisions pour $t\geq t_{\max}+1$. En effet, ces séries chronologiques ne peuvent être considérées isolément, en raison des fortes dépendances existant entre elles. Nous sommes donc amenés à réduire la dimension du problème, et à utiliser des méthodes garantissant la cohérence des projections.

Ne disposant pas d'observations pour les âges supérieurs à $x_{\max}$ (98 ans dans le cas des données INS), nous figerons les probabilités annuelles de décès aux âges supérieurs à $x_{\max}$ à la dernière valeur observée, i.e.

$$
q_{x_{\max}+k}(t)=q_{x_{\max}}(t),\ \ k=1,2,\ldots,
$$

quelle que soit l'année t option qui devrait s'avérer prudente dans le contexte des rentes viagères.

3 Méthode de Lee-Carter

3.1 Le modèle

Nous allons appliquer la méthode de Lee-Carter à ce problème. Cette méthode a déjà prouvé son efficacité à de maintes reprises, et sur diverses populations (USA cf. Lee & Carter (1992), Canada cf. Lee & Nault (1993), Chili cf. Lee & Rofman (1994), Japon cf. Wilmoth (1996), Belgique cf. Brouhns & Denuit (2001b)). Elle consiste à décomposer la mortalité en deux composantes, l'une propre à l'âge et l'autre tendancielle, et ensuite à extrapoler celle relative à l'âge. Il est bon de noter d'emblée que la méthode de Lee-Carter possède les avantages et les inconvénients de l'objectivité: elle n'incorpore pas d'avis d'expert sur l'évolution présumée de la mortalité, sur les progrès de la médecine, l'apparition de nouvelles maladies ou encore l'évolution du style de vie. Vu les avis parfois diamétralement opposés des experts, nous considérons cela comme un avantage. La méthode se borne donc à extrapoler dans le futur les tendances constatées dans le passé.

Nous utiliserons en fait une variante de la méthode originale de LEE & CARTER (1992) suggérée par LEE (2000, note 3 en bas de page 81), laquelle a été légèrement modifiée et appliquée avec succès par RENSHAW & HABERMAN (2001). L'idée est ici de décomposer l'estimation $\widehat{\mu}_x(t)$ de $\mu_x(t)$ comme suit sur l'échelle logarithmique:

$$\ln \widehat{\mu}_x(t) = \alpha_x + \beta_x \kappa_t + \epsilon_{xt} \tag{4}$$

où les erreurs ϵ_{xt} sont supposées centrées, indépendantes et de même variance σ^2 (hypothèse d'homoscédasticité) et où les paramètres β_x et κ_t satisfont aux contraintes

$$\sum_{t=t_{\min}}^{t_{\max}} \kappa_t = 0 \text{ et } \sum_{x=x_{\min}}^{x_{\max}} \beta_x = 1 \tag{5}$$

garantissant l'identifiabilité[1] du modèle.

En vertu de (4), le taux de mortalité à l'âge x pour l'année t est donc décomposé sur l'échelle logarithmique, à un terme d'erreur près, en la somme d'une composante spécifique à l'âge x et en un produit entre un paramètre temporel décrivant l'évolution générale de la mortalité et un paramètre propre à l'âge décrivant l'évolution du taux à l'âge x par rapport à ceux relatifs aux autres âges. On espère bien entendu que la variance des erreurs ϵ_{xt} sera aussi petite que possible. Ainsi, la majeure partie de la variance des $\mu_x(t)$ à chaque âge x sera expliquée par le paramètre κ_t, les ϵ_{xt} n'étant plus qu'un bruit blanc.

Le modèle (4) décompose donc l'évolution des taux instantanés de mortalité en différentes composantes, dont il est bon de saisir précisément la signification:

α_x: décrit le comportement moyen des $\mu_x(t)$ au cours du temps (sur l'échelle logarithmique);

β_x: décrit (toujours sur l'échelle logarithmique) l'écart des $\mu_x(t)$ par rapport au comportement moyen α_x puisque

$$\frac{d}{dt} \ln \mu_x(t) = \beta_x \frac{d}{dt} \kappa_t.$$

Les âges x pour lesquels β_x est grand seront donc très sensibles à l'évolution temporelle. Si les κ_t sont (approximativement) linéaires, $d\kappa_t/dt$ est constant et les taux de mortalité $\mu_x(t)$ présentent une décroissance exponentielle à un taux constant.

κ_t: décrit l'évolution de la mortalité au cours du temps.

L'avantage de cette méthode est de réduire le problème à l'étude de la série chronologique univariée $\{\kappa_t,\ t = 1, 2, \dots\}$.

[1] Rappelons qu'un modèle est identifiable lorsqu'il n'est pas possible de trouver deux jeux de paramètres distincts conduisant au même modèle. Le modèle (4) n'est pas identifiable si on ne formule pas de conditions supplémentaires (telles (5)), notamment car si on remplace β_x et κ_t par $c\beta_x$ et κ_t/c, quel que soit $c \in \mathbb{R}$, $c \neq 0$, on obtient le même modèle.

3.2 Estimation des paramètres

Il est clair que le modèle (4) ne peut pas être ajusté par une simple régression linéaire, puisqu'il n'y a pas de variables observables dans le membre de droite de (4). L'estimation des paramètres s'effectue par la méthode des moindres carrés ordinaire, c'est-à-dire en résolvant le problème d'optimisation

$$(\widehat{\alpha}_x, \widehat{\beta}_x, \widehat{\kappa}_t) = \arg\min_{\alpha_x, \beta_x, \kappa_t} \sum_{x,t} \Big(\ln \widehat{\mu}_x(t) - \alpha_x - \beta_x \kappa_t \Big)^2. \tag{6}$$

Une solution de (6) existe et est fournie par la méthode de décomposition aux valeurs singulières. Cette solution est unique grâce aux contraintes (5).

L'estimation des paramètres α_x, β_x et κ_t se décompose en les étapes suivantes:

Etape 1 estimation des $\mu_x(t)$ par

$$\begin{aligned} \widehat{\mu}_x(t) &= \frac{\#\text{ décès à l'âge } x \text{ durant l'année } t}{\frac{\#\text{ d'individus d'âge } x \text{ au } 1/1/t + \#\text{ d'individus d'âge } x \text{ au } 1/1/t+1}{2}} \\ &= \frac{d_x(t)}{\frac{l_x(t)+l_x(t+1)}{2}}. \end{aligned}$$

Les décès dont le nombre figure au numérateur sont localisés dans le carré hachuré de la Figure 1. Le dénominateur peut se voir comme une approximation de l'intégrale définissant $L_x(t)$ basée sur l'idée que les individus meurent en moyenne en milieu d'année.

Etape 2 estimation des α_x par les moyennes des $\ln \widehat{\mu}_x(t)$ au cours du temps, i.e.

$$\widehat{\alpha}_x = \frac{1}{x_{\max} - x_{\min} + 1} \ln \left\{ \prod_{t=t_{\min}}^{t_{\max}} \widehat{\mu}_x(t) \right\}.$$

Ceci se justifie par le fait que l'annulation de la dérivée de (6) par rapport à α_x donne

$$\sum_{t=t_{\min}}^{t_{\max}} \ln \widehat{\mu}_x(t) = (x_{\max} - x_{\min} + 1)\alpha_x + \beta_x \underbrace{\sum_{t=t_{\min}}^{t_{\max}} \kappa_t}_{=0 \text{ en vertu de la contrainte (5)}}.$$

Etape 3 calcul de la matrice $\boldsymbol{Z}$ de dimension $(x_{\max} - x_{\min} + 1) \times (t_{\max} - t_{\min} + 1)$ dont l'élément (x, t), noté z_{xt}, est donné par

$$z_{xt} = \ln \widehat{\mu}_x(t) - \widehat{\alpha}_x.$$

Il s'agit donc d'un centrage des $\ln \widehat{\mu}_x(t)$ par rapport à leur moyenne temporelle.

Etape 4 approximer $\boldsymbol{Z}$ de manière optimale (aux sens des moindres carrés ordinaires, i.e. en minimisant $\sum_{x,t}(z_{xt} - \beta_x \kappa_t)^2$) par le produit d'une matrice colonne et d'une matrice ligne

$$\boldsymbol{Z} \approx \widehat{\boldsymbol{\beta}} \widehat{\boldsymbol{\kappa}}^t \tag{7}$$

avec $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_{x_{\min}}, , ..., \widehat{\beta}_{x_{\max}})^t$ et $\widehat{\boldsymbol{\kappa}} = (\widehat{\kappa}_{t_{\min}}, ..., \widehat{\kappa}_{t_{\max}})^t$. La résolution de ce problème est directe en passant par la décomposition en valeurs singulières de $\boldsymbol{Z}$.

Notons $\boldsymbol{u}_i$ le $i$ème vecteur propre normé de la matrice carrée $\boldsymbol{Z}^t\boldsymbol{Z}$ de dimension $(t_{\max} - t_{\min} + 1) \times (t_{\max} - t_{\min} + 1)$ correspondant à la valeur propre λ_i; ceci signifie que $\boldsymbol{u}_i$ satisfait les conditions suivantes:

$$\boldsymbol{Z}^t \boldsymbol{Z} \boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i \text{ et } \boldsymbol{u}_i^t \boldsymbol{u}_i = 1.$$

En prémultipliant les deux membres de la première égalité par $\boldsymbol{Z}$, on obtient

$$(\boldsymbol{Z}\boldsymbol{Z}^t)\boldsymbol{Z}\boldsymbol{u}_i = \lambda_i(\boldsymbol{Z}\boldsymbol{u}_i).$$

Cette dernière relation montre qu'à tout vecteur propre $\boldsymbol{u}_i$ de $\boldsymbol{Z}^t\boldsymbol{Z}$ relatif à une valeur propre λ_i non nulle correspond un vecteur propre $\boldsymbol{Z}\boldsymbol{u}_i$ de $\boldsymbol{Z}\boldsymbol{Z}^t$, relatif à la même valeur propre λ_i. Ainsi, toutes les valeurs propres de $\boldsymbol{Z}^t\boldsymbol{Z}$ et $\boldsymbol{Z}\boldsymbol{Z}^t$ sont égales (avec le même ordre de multiplicité, le cas échéant). Si on note $\boldsymbol{v}_i$ le $i$ème vecteur propre de la matrice carrée $\boldsymbol{Z}\boldsymbol{Z}^t$ de dimension $(x_{\max} - x_{\min} + 1) \times (x_{\max} - x_{\min} + 1)$, correspondant à la valeur propre λ_i, on a alors pour $\lambda_i \neq 0$,

$$\boldsymbol{v}_i = \frac{1}{\sqrt{\lambda_i}}\boldsymbol{Z}\boldsymbol{u}_i$$

ou encore

$$\boldsymbol{u}_i = \frac{1}{\sqrt{\lambda_i}}\boldsymbol{Z}^t\boldsymbol{v}_i.$$

Partons à présent de la relation

$$\boldsymbol{Z}\boldsymbol{u}_i = \boldsymbol{v}_i\sqrt{\lambda_i}$$

et postmultiplions les deux membres de cette égalité par $\boldsymbol{u}_i^t$, avant de sommer sur toutes les valeurs propres de $\boldsymbol{Z}^t\boldsymbol{Z}$

$$\boldsymbol{Z}\left(\sum_{i\geq 1}\boldsymbol{u}_i\boldsymbol{u}_i^t\right) = \sum_{i\geq 1}\sqrt{\lambda_i}\boldsymbol{v}_i\boldsymbol{u}_i^t.$$

Les vecteurs propres $\boldsymbol{u}_i$ étant orthogonaux et de norme 1,

$$\sum_{i\geq 1}\boldsymbol{u}_i\boldsymbol{u}_i^t = \text{Identité}$$

de sorte qu'on aboutit à la décomposition suivante de la matrice $\boldsymbol{Z}$:

$$\boldsymbol{Z} = \sum_{i\geq 1}\sqrt{\lambda_i}\boldsymbol{v}_i\boldsymbol{u}_i^t.$$

Cette dernière formule est appelée décomposition aux valeurs singulières. Elle garantit que, sous des conditions assez générales, une matrice rectangulaire peut être écrite de façon unique comme une "somme optimale" de matrices de rang 1 (c'est-à-dire de produits d'une matrice ligne par une matrice colonne). L'optimalité dont il est question signifie que la première matrice de rang 1 constitue la meilleure approximation de rang 1 de la matrice initiale (au sens des moindres carrés), que la somme des deux premières constituent la meilleure approximation de rang 2, etc.

Si la valeur propre λ_1 domine nettement les autres, on obtient alors l'approximation

$$\boldsymbol{Z} \approx \sqrt{\lambda_1}\boldsymbol{v}_1\boldsymbol{u}_1^t. \tag{8}$$

On mesure la qualité de l'approximation (8) par le taux d'inertie τ_1 (ou pourcentage de variance expliqué), donné par

$$\tau_1 = \frac{\lambda_1}{\sum_{i\geq 1}\lambda_i}.$$

En confrontant (7) à (8), on voit qu'il suffit de prendre

$$\widehat{\boldsymbol{\beta}} = \frac{\boldsymbol{v}_1}{\sum_j v_{1j}} \text{ et } \widehat{\boldsymbol{\kappa}} = \lambda_1 \left(\sum_j v_{1j} \right) \boldsymbol{u}_1$$

en supposant $\sum_j v_{1j} \neq 0$. Nous constatons que la contrainte (5) est satisfaite par les $\widehat{\beta_x}$. Montrons que les $\widehat{\kappa_t}$ satisfont aux aussi la contrainte (5). Notons $\overline{\boldsymbol{u}}_1$ le vecteur colonne dont toutes les composantes sont égales à $\frac{1}{t_{\max} - t_{\min} + 1} \sum_j u_{1j}$ et réécrivons $\widehat{\boldsymbol{\kappa}}$ sous la forme

$$\begin{aligned} \widehat{\boldsymbol{\kappa}} &= \left(\frac{1 - \overline{\boldsymbol{u}}_1^t \boldsymbol{u}_1}{1 - \overline{\boldsymbol{u}}_1^t \boldsymbol{u}_1} \right) \cdot \lambda_1 \cdot \left(\sum_j v_{1j} \right) \boldsymbol{u}_1 \\ \text{(par orthogonalité)} &= \left(\frac{\lambda_1 \sum_j v_{1j}}{1 - \overline{\boldsymbol{u}}_1^t \boldsymbol{u}_1} \right) \cdot (\boldsymbol{u}_1 - \overline{\boldsymbol{u}}_1) \\ &= \text{constante} \cdot (\boldsymbol{u}_1 - \overline{\boldsymbol{u}}_1) \end{aligned}$$

en supposant que

$$\overline{\boldsymbol{u}}_1^t \boldsymbol{u}_1 \neq 1.$$

Nous en concluons donc que la contrainte (5) est également satisfaite par les $\widehat{\kappa_t}$.

Etape 5 pour terminer, nous procédons à un second ajustement des $\widehat{\kappa_t}$ afin que le nombre total de décès recensés chaque année corresponde à celui fourni par (4) étant donnés les $\widehat{\alpha_x}$ et $\widehat{\beta_x}$. Si on note $D_{x,t}$ le nombre de décès relevés à l'âge x durant l'année t et $L_{x,t}$ le nombre d'années vécues par les individus d'âge x durant l'année t, on exige donc que

$$\sum_x D_{x,t} = \sum_x \exp(\widehat{\alpha}_x + \widehat{\beta}_x \widehat{\kappa}_t) L_{x,t}. \tag{9}$$

On se référera à Lee & Carter (1992), Bell (1997), Lee (2000) et Lee & Miller (2000) pour une description complète de la méthode. Il est utile de souligner que cette méthode qui a été appliquée avec succès à des situations très variées, a aussi ses détracteurs parmi les actuaires; voyez par exemple Gutterman & Vanderhoof (1999) pour un regard critique à ce sujet.

4 Statistiques de mortalité pour la Belgique

4.1 Collecte des données

Avant de pouvoir appliquer la méthode de Lee-Carter à la population belge, nous devons construire les tableaux de données nécessaires aux techniques de projection que nous appliquerons par la suite. Notre objectif est de construire deux tables à deux entrées (l'âge x et l'année calendaire t) reprenant les effectifs au premier janvier $l_x(t)$ et le nombre de décès $d_x(t)$ survenus pendant l'année en Belgique, et ce pour les deux sexes séparément. Nous appliquons donc une approche transversale.

Tout naturellement, la première tâche est de considérer les données brutes mises à notre disposition par l'INS. L'absence totale de données durant la Première Guerre Mondiale nous contraint à démarrer nos tables en 1920, année de recensement. Au final, nous obtenons les résultats suivants :

1. Pour $l_x(t)$:

- Pour les années de recensement (1920, 1930, 1947), tout est disponible par sexe jusqu'à 84 ans, puis par sexe et catégories d'âge (5 ans) jusqu'à 98 ans.
- De 1921 à 1929, de 1931 à 1946 et de 1948 à 1949, nous ne disposons que de l'effectif total par sexe.
- De 1950 à 1959, tout est disponible jusqu'à 84 ans.
- De 1960 à 2000, tout est disponible jusqu'à 98 ans.

2. Pour $d_x(t)$:

- Pour 1920, tout est disponible par sexe jusqu'à 98 ans.
- De 1921 à 1923, nous disposons des décès par âge jusqu'à 98 ans mais les deux sexes sont confondus.
- De 1924 à 1940, tout est disponible jusqu'à 40 ans, puis par sexe et catégorie d'âge (5 ans) jusque 99 ans et plus.
- De 1941 à 2000, tout est disponible par sexe jusque 98 ans.

4.2 Traitement des données manquantes

La méthode de Lee et Carter que nous appliquerons par la suite nécessite une décomposition en valeurs singulières, et donc des tables rectangulaires. L'objectif de cette section est donc de compléter les données que nous possédons. Etant donné la quantité des retouches à effectuer et le peu d'informations complémentaires, nous optons pour les moyennes pondérées. Cette pratique revient à interpoler linéairement deux extrémités d'un intervalle pour en estimer les points intérieurs. Nous sommes bien conscients du fait que les manipulations décrites ci-après sont loin d'être exemptes de défaut. Principalement, cette manière de procéder donnera encore plus d'importance aux données observées puisqu'elles serviront à générer les données manquantes. Il n'y a cependant pas d'alternative si l'on désire incorporer la mortalité belge avant 1950 dans notre étude.

En guise d'exemple, estimons les effectifs féminins au premier janvier pour les années 1921 à 1929. En $t = 1920$ et 1930, nous pouvons représenter chaque âge x du sexe féminin par un taux $r_x(t) = \frac{l_x(t)}{\sum_{x=0}^{98} l_x(t)}$. Dès lors, par interpolation linéaire, il vient

$$r_x(t) = r_x(1920) + \left(\frac{t-1920}{1930-1920}\right)\Big(r_x(1930) - r_x(1920)\Big) \text{ pour } t \in [1920, 1930].$$

En multipliant ces taux par l'effectif total féminin de l'année t, nous trouvons les $l_x(t)$ pour les années spécifiées. Précisons que, quand c'est possible, nous préférons appliquer cette méthode à des taux (comme ci-dessus) plutôt qu'une moyenne pondérée directe car les résultats sont plus stables et nous profitons d'informations supplémentaires (ici : l'effectif total par sexe).

En procédant de la sorte autant de fois que nécessaire, nous parvenons à remplir les $d_x(t)$ pour $t \in \{1920, ..., 2000\}$ et $x \in \{0, ..., 98\}$ et les $l_x(t)$ pour $t \in \{1920, ..., 2000\}$ et $x \in \{0, ..., X_t\}$ avec

$$X_t = \begin{cases} 84 \text{ si } t \in \{1920, ..., 1929\} \\ 89 \text{ si } t \in \{1930, ..., 1946\} \\ 98 \text{ si } t = 1947 \\ 84 \text{ si } t \in \{1948, ..., 1960\} \\ 98 \text{ si } t \in \{1961, ..., 2000\} \\ 99 \text{ si } t = 2000 \end{cases}.$$

Nos données sont donc maintenant dépourvues de toute valeur manquante, mais restent néanmoins non-rectangulaires (X_t varie avec l'année calendaire t). Comme la méthode de Lee-Carter repose sur une décomposition de la matrice des logarithmes centrés des taux instantanés de mortalité estimés, il est nécessaire de compléter les données (jusqu'à l'âge de 98 ans, disons). Nous expliquerons dans la Section 4.3 les moyens mis en oeuvre pour y parvenir.

4.3 Complétion des tables jusqu'à 110 ans

L'objectif de cette section est de lisser les estimations $\widehat{q}_x(t)$ et $\widehat{m}_x(t)$ des quotients et taux de mortalité obtenues à partir de l'hypothèse (1) et de les extrapoler jusqu'à $x_{\max} = 110$ ans, afin d'obtenir des tableaux rectangulaires. Le lissage est utile pour atténuer la forte variabililité de ces indicateurs aux âges avancés, vu le peu d'observations disponibles[2].

Il convient néanmoins de souligner que les deux méthodes développées ci-après se servent des données disponibles pour estimer la mortalité aux grands âges. Elles ont donc pour effet de gonfler l'information contenue dans les statistiques mises à notre disposition. Cette pratique est cependant nécessaire pour les raisons évoquées au paragraphe précédent. Les méthodes développées ci-dessous traitent chaque année calendaire séparément. C'est pourquoi, dans un but de clarté, nous omettons le facteur t jusqu'à la fin de cette section.

4.3.1 Âge ultime

L'un des problèmes soulevés par les méthodes de fermeture des tables est la question d'un éventuel âge ultime, ou âge biologique maximal. La plupart des actuaires, s'appuyant sur des avis autorisés de médecins ou d'éminents biologistes, croient à la réalité de cette borne supérieure à la durée de vie humaine, qui serait inscrite dans nos gênes. Une étude relativement récente (Vaupel (1997)), construite sur une importante base de données[3] a fait apparaître les caractéristiques suivantes :

- passé l'âge de 80 ans, les probabilités annuelles de décès q_x augmentent à un taux décroissant (la fonction $x \longmapsto q_x$ est non-décroissante et concave);
- la mortalité décroît au cours du temps dans cette tranche d'âge.

Ces conclusions nous poussent à la prudence dans la détermination d'un éventuel âge ultime. Plusieurs auteurs considèrent $q_{120} = 1$ comme une approximation raisonnable de la réalité. Néanmoins, pour des raisons techniques liées à la méthode de Coale et Kisker (voir ci-après), nous préférons choisir

$$m_{110} = \begin{cases} 0.8 \text{ pour les femmes} \\ 1 \text{ pour les hommes} \end{cases}, \qquad (10)$$

qui semble une être une bonne approximation pour les pays occidentaux. Dans la section 4.3.5, nous analyserons l'impact de ce choix sur l'estimation de la mortalité aux grands âges.

4.3.2 La méthode de Coale et Kisker

La technique proposée par Coale & Kisker (1990) travaille sur les taux bruts de mortalité m_x. Elle propose de recalculer ces valeurs et de les extrapoler pour les grands âges (jusque $x = 110$ ans) en se basant sur la formule de Gompertz. Celle-ci postule

$$m_x = m_{65} \exp[(x-65)\overline{k}_x]$$

[2] La méthode de Coale et Kisker lisse les valeurs brutes à partir de $x = 70$ ans tandis que celle de Jaumain réalise cette opération dès $x = 3$ ans.

[3] 70 millions d'êtres humains de 14 pays ayant atteint l'âge de 80 ans, dont 200 000 centenaires.

pour $x \geqslant 65$, où $\overline{k}_x$ représente le taux moyen de croissance de m_x entre 65 et x ans. En faisant apparaître les taux annuels $k_{66}, k_{67}, ..., k_x$, il vient

$$m_x = m_{65} \exp \left[\sum_{y=66}^{x} k_y \right] \text{ avec } \overline{k}_x = \frac{1}{x-65} \sum_{y=66}^{x} k_y.$$

Le but de la méthode est donc de calculer les coefficients k_x jusqu'à un certain âge et de les extrapoler afin de pouvoir recomposer les taux m_x.

Les auteurs, après avoir examiné différents graphiques relatifs aux pays d'Europe Occidentale, ont remarqué que les courbes des k_x possèdent en général un pic aux alentours de 80 ans, avant de décroître linéairement. Nous posons donc

$$k_x = k_{80} + s(x-80) \tag{11}$$

pour $x \geqslant 80$. Afin de déterminer le coefficient s relatif à la pente, nous devons fixer une valeur-limite à l'extrémité de la table. Par l'approximation classique $q_{120} = 1$, et sous l'hypothèse de constance par morceaux des taux de mortalité, il vient

$$m_{120} = -\ln(1 - q_{120}) = +\infty.$$

Pour éviter ce problème, nous préférons recourir à (10). Nous devons donc résoudre l'équation

$$m_{110} = m_{79} \exp \left[\sum_{x=80}^{110} k_x \right] = m_{79} \exp \left[\sum_{x=80}^{110} \{k_{80} + s\,(x-80)\} \right]$$

dont la solution est donnée par

$$\begin{aligned} s &= -\frac{\left[\ln\left(\frac{m_{79}}{m_{110}}\right) + (110-80+1)k_{80}\right]}{\sum_{x=80}^{110}(x-80)} \\ &= -\frac{\left[\ln\left(\frac{m_{79}}{m_{110}}\right) + 31k_{80}\right]}{465}. \end{aligned} \tag{12}$$

En pratique, nous commençons par lisser les k_x disponibles (jusqu'à $x = 84$ ans dans notre situation). La première étape est le calcul des taux moyens de croissance pour une période de 5 ans centrée en x :

$$k_x^{'} = \frac{\ln\left(\frac{m_{x+2}}{m_{x-3}}\right)}{5}.$$

Ensuite, les $k_x^{''}$ lissés sont les moyennes de 5 valeurs de $k_x^{'}$ autour de x :

$$k_x^{''} = \frac{k_{x-2}^{'} + k_{x-1}^{'} + k_x^{'} + k_{x+1}^{'} + k_{x+2}^{'}}{5}.$$

Nous effectuons ce calcul pour $x \in \{70, 71, ..., 80\}$ et en déduisons les valeurs de $k_x^{''}$ pour les âges supérieurs par la formule (11). Les estimations des taux de mortalité pour $x \geqslant 70$ sont maintenant facilement accessibles :

$$m_x^* = m_{69} \exp \left[\sum_{y=70}^{x} k_y^{''} \right]. \tag{13}$$

Les taux bruts pour $x \geqslant 70$ sont donc remplacés par leurs valeurs lissées. Constatons que toute la construction des m_x^* est basée sur la valeur de m_{69}, laquelle n'a pas été lissée. C'est la raison pour laquelle il est préférable d'adopter la formule

$$m_x^* = m_{69}^{'} \exp\left[\sum_{y=70}^{x} k_y^{''}\right] \text{ où } m_{69}^{'} = \frac{m_{67}+m_{68}+m_{69}+m_{70}+m_{71}}{5}. \tag{14}$$

Il faut bien évidemment garder à l'esprit que, dans notre situation, les formules ci-dessus sont utilisées avec les estimations $\widehat{m}_x$ obtenues à partir de l'hypothèse de constance des taux instantanés de mortalité à chaque âge. Pour illustrer la méthode, appliquons-la aux hommes en $t = 2000$. Nous trouvons $s = -0.003$. De plus, nous disposons pour cette année de données jusque $x = 99$ ans et pouvons donc calculer les $k_x^{''}$ jusque 95 ans ; il nous est alors possible d'éprouver l'hypothèse de linéarité (11). Comme nous le montre la Figure 2, l'approximation est loin d'être parfaite. Néanmoins, la figure 3 qui compare les taux bruts $\widehat{m}_x(2000)$ et les valeurs estimées par l'équation (14) laisse apparaître une courbe raisonnable. Nous pouvons notamment y remarquer la partie concave en fin de vie.

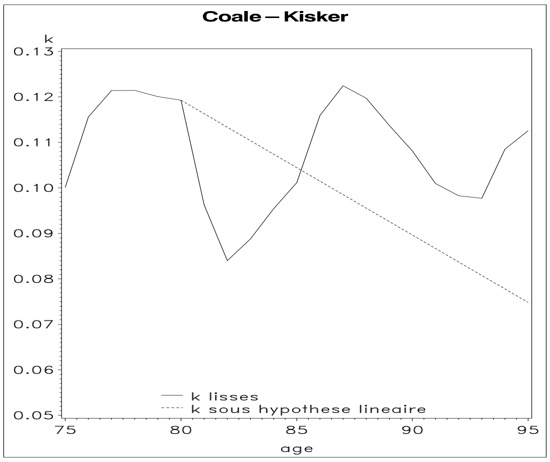

Figure 2: Comparaison des $k_x^{''}$ avec leur estimation par l'hypothèse linéaire (11) en $t = 2000$.

4.3.3 La méthode de Jaumain

Au contraire de celle de Coale et Kisker, la méthode de Jaumain (2001) pour la fermeture des tables travaille sur les quotients de mortalité. Là encore, la définition d'une valeur-limite est nécessaire. Dans son article, l'auteur choisit d'imposer $q_{120} = 1$. Néanmoins, dans le but d'obtenir des résultats comparables à ceux du chapitre précédent, nous préférons la contrainte

$$q_{110} = 1 - \exp(-m_{110}) \text{ avec } m_{110} = \begin{cases} 0.8 \text{ pour les femmes} \\ 1 \text{ pour les hommes} \end{cases}$$

toujours sous l'hypothèse de constance par morceaux des taux de mortalité. Ensuite, la technique consiste à ajuster un modèle linéaire en x et x^2 aux logarithmes des quotients de mortalité pour les grands âges, tout en respectant la valeur-limite convenue :

$$\ln q_x = a + bx + cx^2 + \varepsilon_x \text{ pour } x \geqslant 81 \tag{15}$$

avec ε_x les erreurs supposées centrées et homoscédastiques.

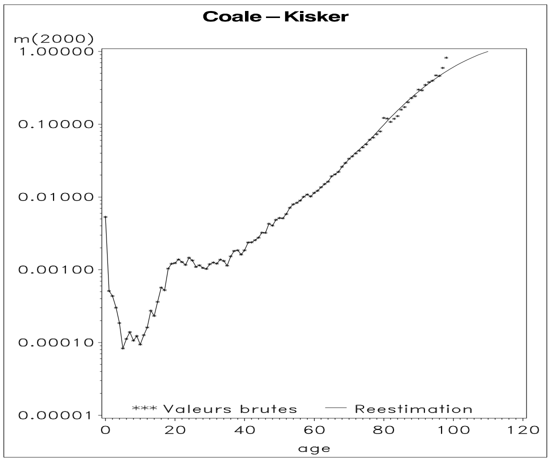

Figure 3: Comparaison des $\widehat{m}_x(2000)$ masculins bruts avec leur estimation $m^*_x(2000)$ par la méthode de Coale et Kisker.

La dernière étape permet alors de lisser la table, en effectuant des moyennes géométriques. Dans un premier temps, des quotients sont calculés de la manière suivante :

$$
\begin{aligned}
q'_x &= q_x \text{ pour } x \in \{0, 1, 2, 110\} \\
q'_x &= (q_{x-1} \cdot q_x \cdot q_{x+1})^{\frac{1}{3}} \text{ pour } x \in \{3, 109\} \\
q'_x &= (q_{x-2} \cdot q_{x-1} \cdot q_x \cdot q_{x+1} \cdot q_{x+2})^{\frac{1}{5}} \text{ pour } x \in \{4, 5, ..., 108\}
\end{aligned} \tag{16}
$$

Finalement, nous recommençons cette dernière opération à partir des q'_x pour obtenir les quotients de mortalité ajustés q''_x. Remarquons que $q'_0 = q_0$ et $q'_1 = q_1$ n'interviennent pas dans le processus de lissage car ces valeurs représentent la mortalité infantile. Elles risqueraient donc de provoquer une surestimation des q''_x pour les âges proches ($x = 2, 3, 4, ...$).

La Figure 4 représente le résultat graphique de la régression pour les hommes en $t = 2000$. La Figure 5 reprend quant à elle les courbes $\widehat{q}_x$ et $\widehat{q}''_x$ pour les hommes en $t = 2000$.

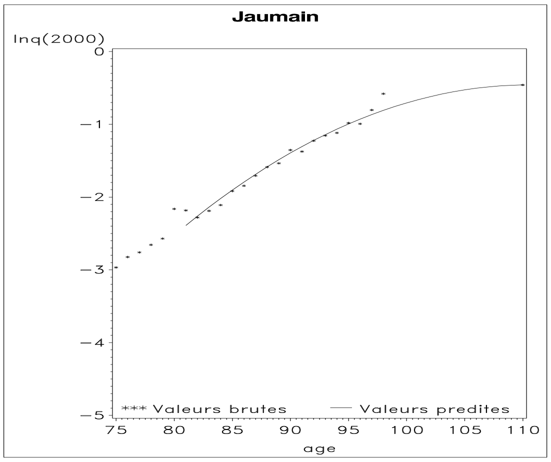

Figure 4: Résultat graphique de la régression des $\ln \widehat{q}_x(t)$ pour les hommes et $t = 2000$.

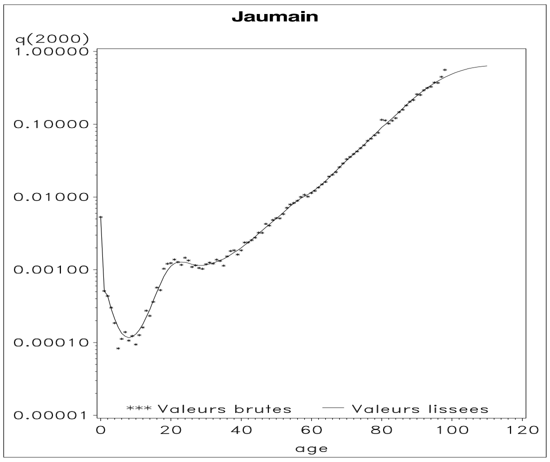

Figure 5: Comparaison des $\widehat{q}_x$ (valeurs brutes) et $\widehat{q}''_x$ (valeurs lissées) pour les hommes en $t = 2000$.

4.3.4 Comparaison des 2 méthodes

Nous pouvons à présent comparer les résultats obtenus à partir des deux techniques. Notons m_x^J et m_x^{CK} les taux de mortalité calculés respectivement par les méthodes de Jaumain et de Coale-Kisker.

La Figure 6 représente ces indicateurs pour les années 1920, 1960 et 2000. Nous constatons dans la majorité des cas que l'allure des deux courbes est comparable, voire similaire, avec une partie concave aux très grands âges. Néanmoins, nous remarquons que les m_x^J sont parfois décroissants aux âges avancés (comme chez les hommes en 1960), ce qui est improbable. Pour éviter ce genre de situation, nous optons pour la méthode de Coale et Kisker qui offre des courbes plus raisonnables.

Finalement, pour motiver ce choix, observons qu'originellement la méthode de Jaumain sert uniquement à lisser une table de mortalité à partir de la suite des quotients q_x bruts, qui est en général fortement brisée. C'est pourquoi elle effectue un lissage dès $x = 3$ ans. Au contraire, la méthode de Coale et Kisker a plutôt été conçue pour étendre les tables à des âges pour lesquels aucune donnée brute n'est disponible.

4.3.5 Impact de la valeur de m_{110}

Comme nous l'avons vu précédemment, les méthodes de fermeture des tables nécessitent une valeur-limite, souvent à $x = 110$ ou 120 ans. Nous avons opté pour $m_{110} = 0.8$ pour les femmes et 1 pour les hommes mais aimerions maintenant savoir si une modification de ces valeurs apporterait des changements conséquents pour l'estimation du taux de mortalité aux grands âges. La Figure 7 montre ces courbes en $t = 2000$ avec différentes valeurs-limites pour la méthode de Coale et Kisker. La table 1 quantifie les résultats pour $x = 100$ ans et pour l'espérance de vie à la naissance calculée en vision transversale (l'approche longitudinale, plus réaliste, n'est pas encore applicable à ce stade de l'étude).

Nous observons donc qu'une modification conséquente de la valeur-limite de l'ordre de 2 dixièmes affecterait le taux de mortalité à 100 ans de quelques 10%. Un tel impact est relativement faible et peut être considéré comme négligeable vu le petit nombre de centenaires. Cette remarque est confirmée par les valeurs de $e_0(2000)$. En effet, pour cet indicateur, les écarts relatifs par rapport à la valeur de référence sont inférieurs à 10^{-3}.

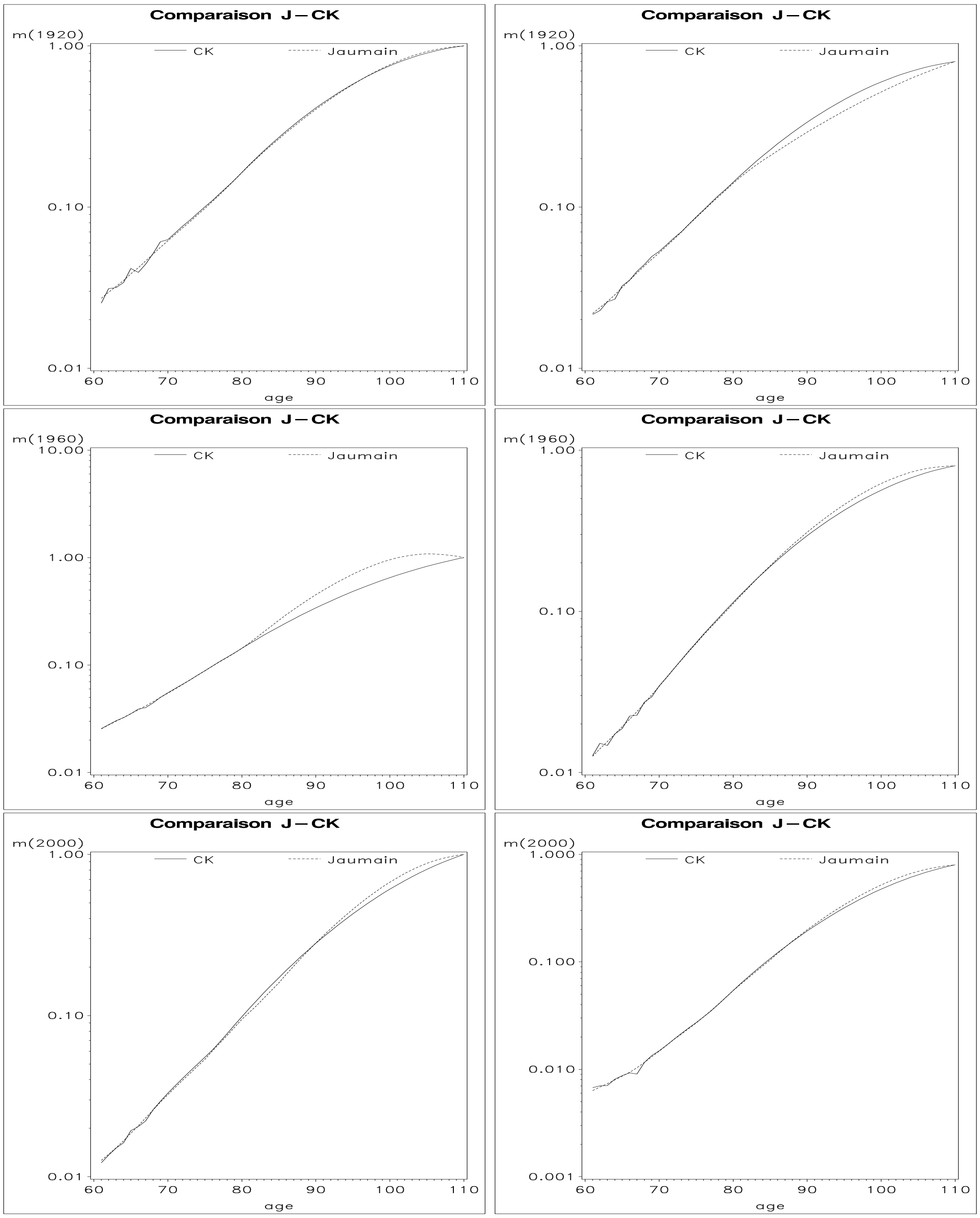

Figure 6: Comparaison des méthodes de Coale-Kisker et Jaumain pour les années 1920, 1960 et 2000 (les hommes sont à gauche et les femmes à droite).

	Hommes				
$m_{110}(2000)$	0.8	0.9	**1.0**	1.1	1.2
$\widehat{m}_{100}(2000)$	0.570	0.600	**0.624**	0.649	0.672
Écart relatif	−0.087	−0.042	**0**	0.040	0.077
$e_0(2000)$	74.056	74.048	**74.042**	74.036	74.030
Écart relatif	0.0002	0.0001	**0**	−0.0001	−0.0002
	Femmes				
$m_{110}(2000)$	0.6	0.7	**0.8**	0.9	1.0
$\widehat{m}_{100}(2000)$	0.426	0.456	**0.485**	0.511	0.536
Écart relatif	−0.122	−0.059	**0**	0.055	0.106
$e_0(2000)$	80.679	80.646	**80.618**	80.595	80.574
Écart relatif	0.0007	0.0003	**0**	−0.0003	−0.0006

Table 1: Impact du choix de la valeur-limite.

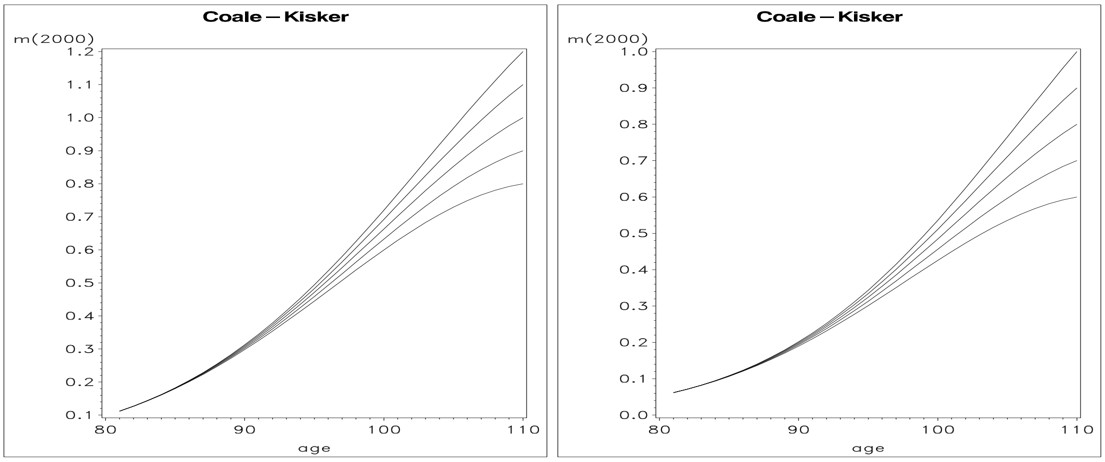

Figure 7: Taux de mortalité aux grands âges avec $m_{110}(2000) \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$ pour les hommes (à gauche) et $\{0.6, 0.7, 0.8, 0.9, 1.0\}$ pour les femmes (à droite).

5 Ajustement du modèle de Lee-Carter aux données belges

Nous sommes à présent en mesure d'appliquer la méthode de Lee-Carter aux données belges que nous avons complétées comme décrit précédemment. La période d'observation est donc composée des 81 années $\{1920, ..., 2000\}$.

Etape 1: L'estimation brute des taux instantanés de mortalité fournit le graphique présenté à la Figure 8.

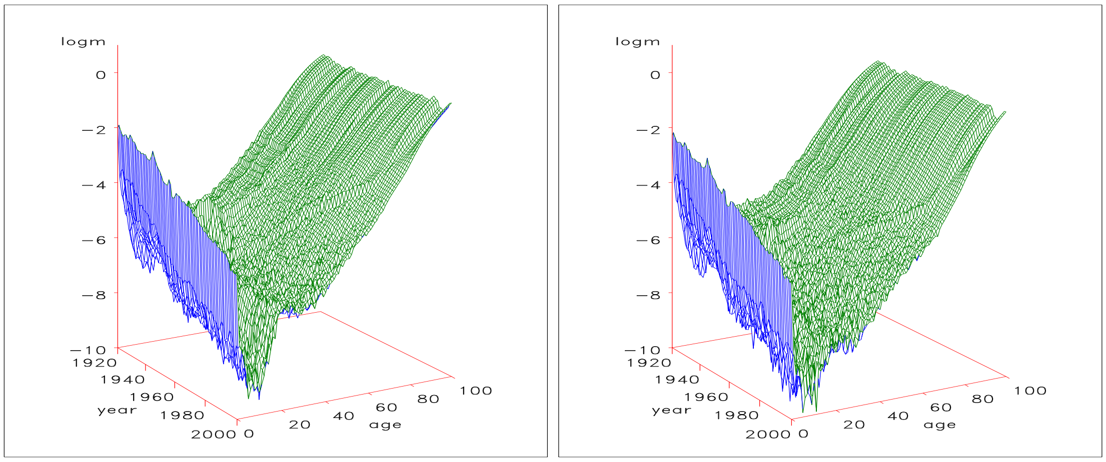

Figure 8: Estimation brute des taux instantanés de mortalité (les hommes sont à gauche et les femmes à droite).

Etape 2: La Figure 9 reprend l'estimation des α_x pour les hommes et pour les femmes. Nous pouvons y constater que la courbe des $\widehat{\alpha}_x$ est conforme à nos attentes : le taux de mortalité progresse avec l'âge excepté chez les nouveaux-nés tandis que la bosse accident aux alentours de $x = 20$-25 ans est présente et plus marquée chez les hommes. Nous retrouvons en outre la partie concave en fin de vie.

Etapes 3-4: Nous trouvons les taux d'inertie $\tau_1 = 0.9043$ pour les hommes et 0.9332 pour les femmes; l'approximation semble donc être de qualité. Ces valeurs sont par ailleurs meilleures que celles obtenues par Brouhns & Denuit (2001b) sur les données brutes relatives aux plus de 60 ans sur la période d'observation [1960,1999] (0.55 et 0.77 respectivement pour les hommes et les femmes). Ceci est sans doute la conséquence des divers traitements apportés aux données : d'une part la technique de complétion des tables utilisée pour les années inférieures à 1950, et d'autre part le lissage effectué aux grands âges par la méthode de Coale et Kisker.

La Figure 9 représente en outre les estimations des différents paramètres β_x et κ_t.

Les plus grandes variations temporelles du taux de mortalité ($\widehat{\beta}_x$) se situent chez les jeunes et sont probablement le résultat des progrès réalisés par la médecine en près d'un siècle pour freiner les mortalités infantile et juvénile. Remarquons que la méthode utilisée pour fermer les tables lissait les taux de mortalité à partir de $x = 70$ ans. Cet artifice se retrouve dans la partie correspondante des $\widehat{\beta}_x$. De plus, la tendance temporelle est annulée pour $x = 110$ ans ($\widehat{\beta}_{110} = 0$) puisque l'on a imposé à $\widehat{m}_{110}(t) = \widehat{\mu}_{110}(t)$ une valeur identique pour chaque année. $\widehat{\alpha}_{110}$ atteint d'ailleurs la valeur attendue, à savoir $\ln(1) = 0$ chez les hommes et $\ln(0.8)$ chez les femmes. Enfin, la courbe des $\widehat{\kappa}_t$ est en constante décroissance

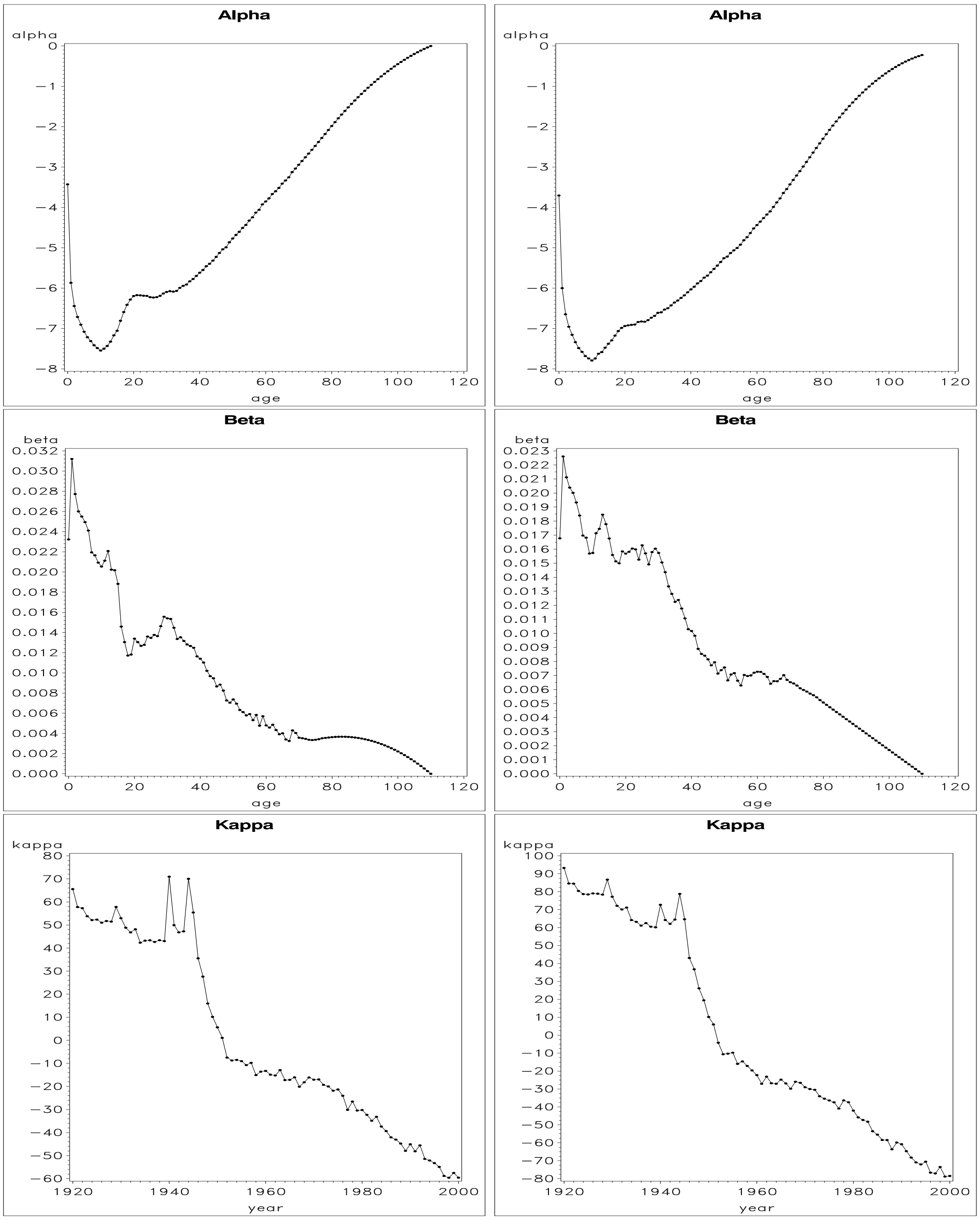

Figure 9: Estimation des paramètres de la méthode Lee et Carter pour la période d'observation $[1920, 2000]$ (les hommes sont à gauche et les femmes à droite).

hormis pendant la guerre (visible par les deux pics de mortalité en $t = 1940$ et 1944) et subit une fracture au sortir de celle-ci.

Etape 5: Les taux de mortalité ajustés $\widehat{\widehat{\mu}}_x(t) = \exp(\widehat{\alpha}_x + \widehat{\beta}_x\widehat{\kappa}_t)$, appliqués aux expositions au risque $L_x(t)$, ne reconstituent pas exactement le nombre total de décès observés pour l'année t en question, i.e.

$$\sum_x d_x(t) \neq \sum_x L_x(t)\widehat{\widehat{\mu}}_x(t).$$

Cette supposition est confirmée par la Figure 10.

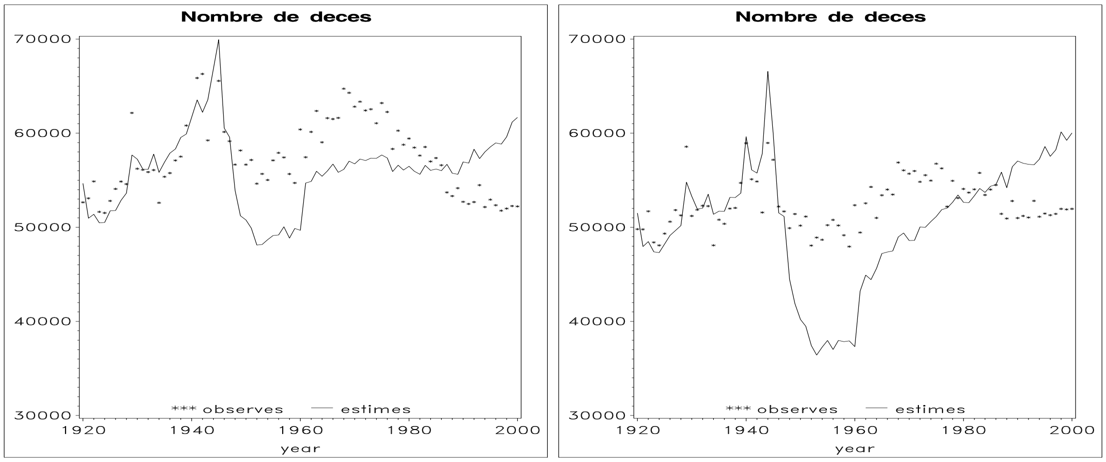

Figure 10: Comparaison des nombres de décès observés et estimés par la méthode de Lee-Carter (les hommes sont à gauche et les femmes à droite).

Pour palier à ce problème nous recalculons des $\widehat{\widehat{\kappa}}_t$ en résolvant

$$\sum_x d_x(t) = \sum_x L_x(t)\exp(\widehat{\alpha}_x + \widehat{\beta}_x\widehat{\widehat{\kappa}}_t) \tag{17}$$

c'est-à-dire

$$\sum_x \Big[L_x(t)\exp(\widehat{\alpha}_x)\exp(\widehat{\beta}_x\widehat{\widehat{\kappa}}_t) - d_x(t)\Big] = \sum_x \Big[K_{x,t}\exp(\widehat{\beta}_x\widehat{\widehat{\kappa}}_t)\Big] - K'_t = 0 \tag{18}$$

où $K_{x,t}$ et K'_t sont des constantes. Étant en présence de sommes d'exponentielles, donc des fonctions strictement croissantes, l'unicité de la racine de chaque fonction est donc évidente. Dès lors, une méthode itérative facile à implémenter telle que Newton-Raphson donnera de bons résultats en peu d'itérations (la convergence est quadratique)[4]. Pour rappel, cette technique se base sur l'itération suivante :

$$x_{i+1,t} = x_{i,t} - \frac{F(x_{i,t})}{F'(x_{i,t})}$$

où $x \longmapsto F(x)$ est la fonction à annuler. Dans notre situation, l'initialisation de la récurrence s'impose d'elle-même : nous prenons $x_{0,t} = \widehat{\kappa}_t$.

[4]Nous aurions pu choisir d'autres méthodes d'ordre de convergence égal (méthode d'interpolation de la sécante) ou supérieur à 2 (méthodes de Muller et Halley). Néanmoins, elles auraient été plus longues à implémenter et le nombre d'itérations gagnées aurait été faible.

Vu la régularité des fonctions à étudier, la précision exigée peut être relativement importante. Nous décidons donc d'arrêter les itérations quand

$$\frac{x_{i,t} - x_{i-1,t}}{x_{i-1,t}} \leqslant 0.0001.$$

En pratique, l'algorithme converge en 3 à 5 itérations. En guise d'exemple, la table 2 décrit le comportement de la méhode chez les hommes en $t = 2000$.

itération i	$x_{i,2000}$
0	-59.54792
1	-93.43137
2	-97.35765
3	-97.40053
4	-97.40054

Table 2: Réestimation des κ_t chez les hommes en $t = 2000$.

Les résultats de la réestimation sont visibles sur la Figure 11.

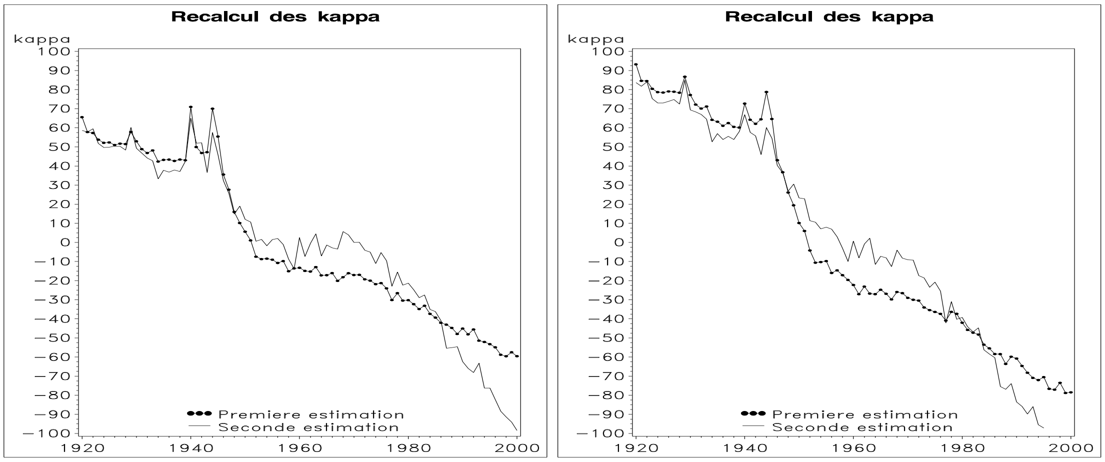

Figure 11: Réestimation des κ_t par la méthode Newton-Raphson (les hommes sont à gauche et les femmes à droite).

Bien entendu, la contrainte (5) sur les κ_t est alors violée par les nouveaux estimateurs $\widehat{\widehat{\kappa}}_t$; un nouveau réajustement s'impose donc. Posons

$$\beta_x^* = \widehat{\beta}_x \text{ et } \kappa_t^* = \widehat{\widehat{\kappa}}_t - \overline{\kappa_t}$$

où $\overline{\kappa_t}$ est la moyenne arithmétique des $\widehat{\widehat{\kappa}}_t$. Les contraintes (5) sont maintenant clairement respectées par les κ_t^* et β_x^*. Il ne nous reste qu'à déterminer les α_x^* qui satisferont $\alpha_x^* + \beta_x^*\kappa_t^* = \widehat{\alpha}_x + \widehat{\beta}_x\widehat{\widehat{\kappa}}_t$. Nous trouvons donc

$$\begin{aligned}
\alpha_x^* &= \widehat{\alpha}_x + \widehat{\beta}_x\widehat{\widehat{\kappa}}_t - \beta_x^*\kappa_t^* \\
&= \widehat{\alpha}_x + \widehat{\beta}_x\widehat{\widehat{\kappa}}_t - \widehat{\beta}_x(\widehat{\widehat{\kappa}}_t - \overline{\kappa_t}) \\
&= \widehat{\alpha}_x + \widehat{\beta}_x\overline{\kappa_t}
\end{aligned}$$

et les α_x^*, β_x^* et κ_t^* sont dès lors acceptables.

6 Comparaison de deux périodes d'observation : $[1920, 2000]$ et $[1950, 2000]$

L'intérêt de ce paragraphe est de déterminer s'il est nécessaire de travailler avec une large période d'observation ou s'il est raisonnable de se contenter de moins d'années. Dans ce but, nous comparons les taux de mortalité observés après 1950 avec leurs ajustements par la méthode de Lee et Carter sur base de deux périodes d'observation : $[1920, 2000]$ et $[1950, 2000]$. Les graphiques correspondants pour $x = 20, 40, 60, 80$ et 100 ans pour les hommes se trouvent à la Figure 12. Les courbes relatives aux femmes sont comparables. Enfin, les taux d'inertie obtenus à partir de la période d'observation $[1950, 2000]$ valent $\tau_1 = 0.7803$ et 0.7897 respectivement pour les hommes et les femmes.

Quand il existe une différence significative entre les deux estimations, nous pouvons constater que c'est systématiquement la période $[1950, 2000]$ qui semble la meilleure. Afin de quantifier ces résultats, il est bon de calculer un indice traduisant l'écart entre les estimations et les valeurs observées. Posons donc

$$R_j^2 = \sum_{\tau=1950}^{2000} \sum_{x=0}^{110} (\widehat{\mu}_x(\tau) - \widehat{\widehat{\mu}}_x^{j}(\tau))^2$$

où j caractérise la période d'observation avec

$$j = \begin{cases} 1920 \text{ si la période est } [1920, 2000] \\ 1950 \text{ si la période est } [1950, 2000] \end{cases}$$

Nous obtenons les valeurs

R_j^2	$j = 1920$	$j = 1950$
Hommes	1.6316	**1.4891**
Femmes	1.2113	**0.9435**

qui confirment l'impression laissée par la Figure 12 : la période $[1950, 2000]$ donne de meilleurs résultats. Ceci peut être la conséquence de différents phénomènes. Nous pensons naturellement à l'existence éventuelle d'une différence significative entre les deux parties du vingtième siècle dans l'évolution de la mortalité en Belgique. Cependant, il faut également garder à l'esprit l'importance des traitements apportés aux données manquantes avant 1950, qui peuvent avoir fortement influencé l'analyse.

Au vu des résultats obtenus ci-dessus, nous ne retiendrons que les années supérieures à $t = 1950$ pour la suite de cette étude. Par ce choix, nous évitons de traiter le délicat problème de la Seconde Guerre Mondiale.

7 Extrapolation de la tendance temporelle

Le modèle classique de Lee et Carter présenté ci-dessus possède la propriété de synthétiser dans la série $\{\kappa_t\}_t$ toute l'information relative à l'évolution de la mortalité dans le temps. L'objectif de cette section est donc de modéliser cette série temporelle afin d'être en mesure d'effectuer des prévisions sur les indices de mortalité des années futures. Cette manière de procéder possède les avantages et les inconvénients de l'objectivité : aucune attention n'est accordée à l'avis d'expert sur l'évolution présumée de la mortalité, sur les progrès de la médecine, l'apparition de nouvelles maladies ou encore l'évolution du style de vie. Le but est donc de se borner à extrapoler dans le futur les tendances observées dans le passé. Cette démarche a été critiquée,

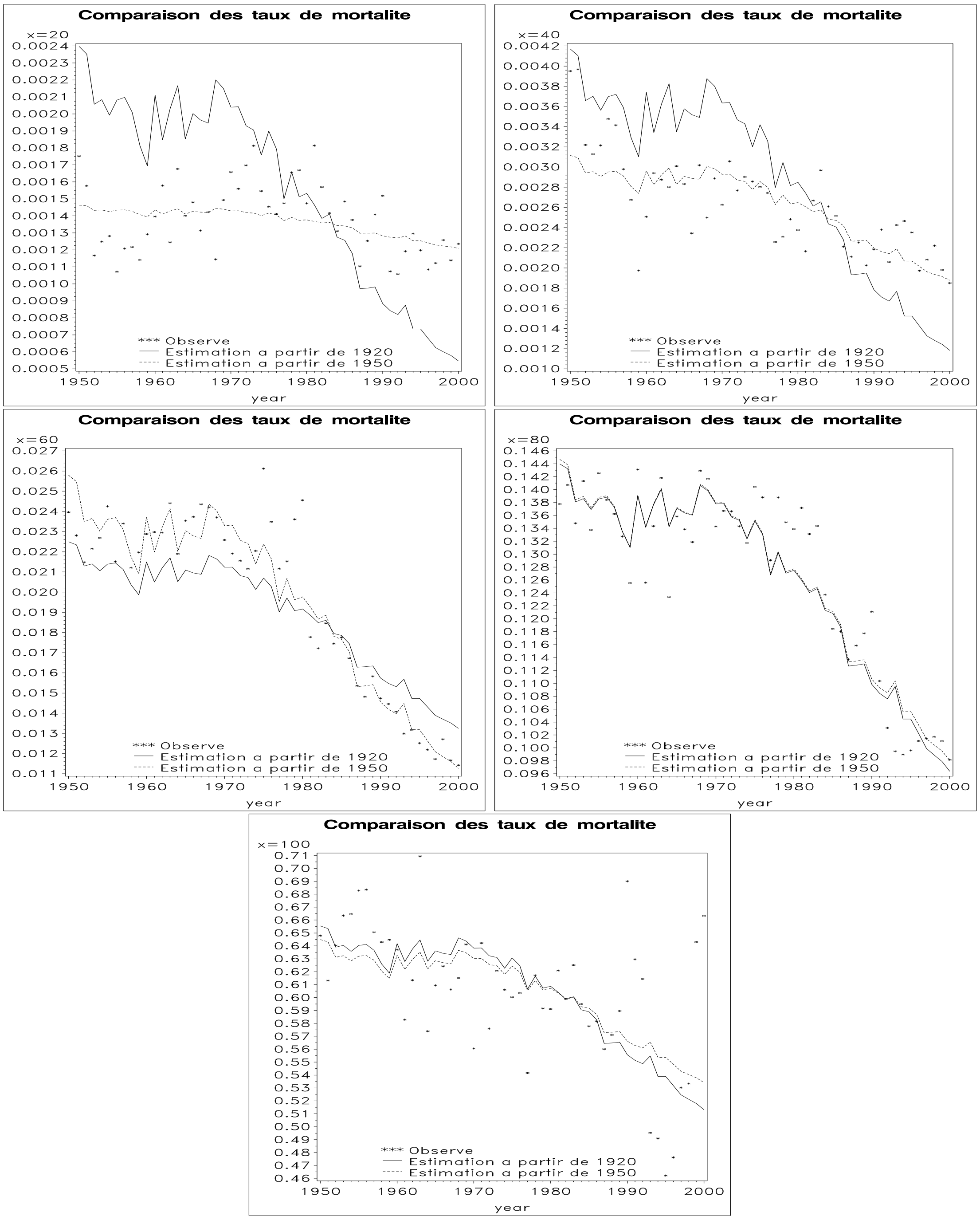

Figure 12: Comparaison des deux périodes d'observation pour l'estimation des taux de mortalité masculins pour les âges $x = 20, 40, 60, 80$ et 100.

car elle est incapable de prévoir des variations subites de la mortalité (par exemple la découverte d'un nouveau traitement médical). Néanmoins, la mortalité étant influencée par l'interaction complexe de facteurs sociaux et biologiques difficiles à modéliser, la prédiction par extrapolation est en fait la *technique par défaut*.

Nous disposons donc de deux séries temporelles que nous notons $\{\kappa_t^w\}_t$ et $\{\kappa_t^m\}_t$. Les exposants w et m font respectivement référence aux femmes et aux hommes. Nous appliquerons la désormais célèbre méthodologie de Box et Jenkins à ces deux séries temporelles. Nous serons alors en mesure d'effectuer des projections de nos séries temporelles jusqu'en 2110.

Enfin, nous effectuerons des prévisions sur la période $[1996, 2000]$ à partir des modèles choisis appliqués à la période d'observation $[1950, 1995]$. Leur pertinence pourra ainsi être testée en comparant les prévisions et les valeurs observées.

7.1 Modélisation ARIMA

Nous voyons à présent $\{\kappa_t^*,\ t = t_{\min}, \ldots, t_{\max}\}$ comme une réalisation de dimension finie de la série chronologique $\mathcal{K} = \{\kappa_t,\ t \in \mathbb{N}\}$. Cette dernière sera modélisée à l'aide des processus ARIMA (pour AutoRegressive Integrated Moving Average) en suivant la méthodologie proposée par Box et Jenkins (dont nous rappelons très brièvement les étapes principales ci-dessous). La modélisation paramétrique d'une série temporelle se construit essentiellement à partir de la structure de dépendance existant entre les éléments de la série. Celle-ci sera étudiée au moyen de la fonction d'autocovariance et de la fonction d'autocorrélation de la série. Pour une introduction aux modèles ARIMA, on pourra utilement se référer à Brockwell & Davis (1996); l'ouvrage de Brockwell & Davis (1991) offre un traitement plus théorique de la question.

Nous suivons ici pas à pas les étapes de la méthodologie de Box & Jenkins:

1. Stationnarisation de la série. La méthode proposée par Box et Jenkins s'applique à des processus stochastiques faiblement stationnaires d'ordre 2 (ce qui signifie l'existence des deux premiers moments, la constance de la moyenne et de la variance au cours du temps et des fonctions d'autocovariance et d'autocorrélation qui ne dépendent que du délai et pas de l'instant). La première étape de la modélisation consiste donc à s'assurer que $\mathcal{K}$ satisfait bien ces conditions; dans le cas contraire, il s'agira de transformer $\mathcal{K}$ adéquatement afin d'atteindre la stationnarité faible d'ordre 2.

1.(a) Stationnarisation en variance. La première condition nécessaire à l'ajustement d'un modèle ARIMA est la stabilité en variance des séries $\{\kappa_t\}_t$ considérées. Il faut donc que la variabilité de chacune des séries soit homogène sur toute la période d'observation.

Nous remarquons sur la Figure 13 que la variabilité ne croît ni ne decroît avec le temps. Aucune transformation ne doit donc être appliquée à nos séries.

1.(b) Stationnarisation en moyenne. La méthodologie de Box et Jenkins exige également des séries stationnaires en moyenne, c'est-à-dire dépourvues de tendance temporelle.

Adoptons les notations suivantes :

- B est l'opérateur de retard, c'est-à-dire $B\kappa_t = \kappa_{t-1}$
- ∇_d est l'opérateur de différence d'ordre d, c'est-à-dire $\nabla_d = 1 - B^d$.

La méthode des différences itérées consiste à appliquer l'opérateur ∇_1 à la série initiale $\{\kappa_t\}_t$ autant de fois que nécessaire jusqu'à l'annulation de sa tendance. La Figure 13 fait apparaître des séries clairement décroissantes, donc non-stationnaires, mais dépourvues de saisonnalité (nous sommes en présence de données annuelles). Nous devons donc appliquer une première différentiation, tant chez les hommes que chez les

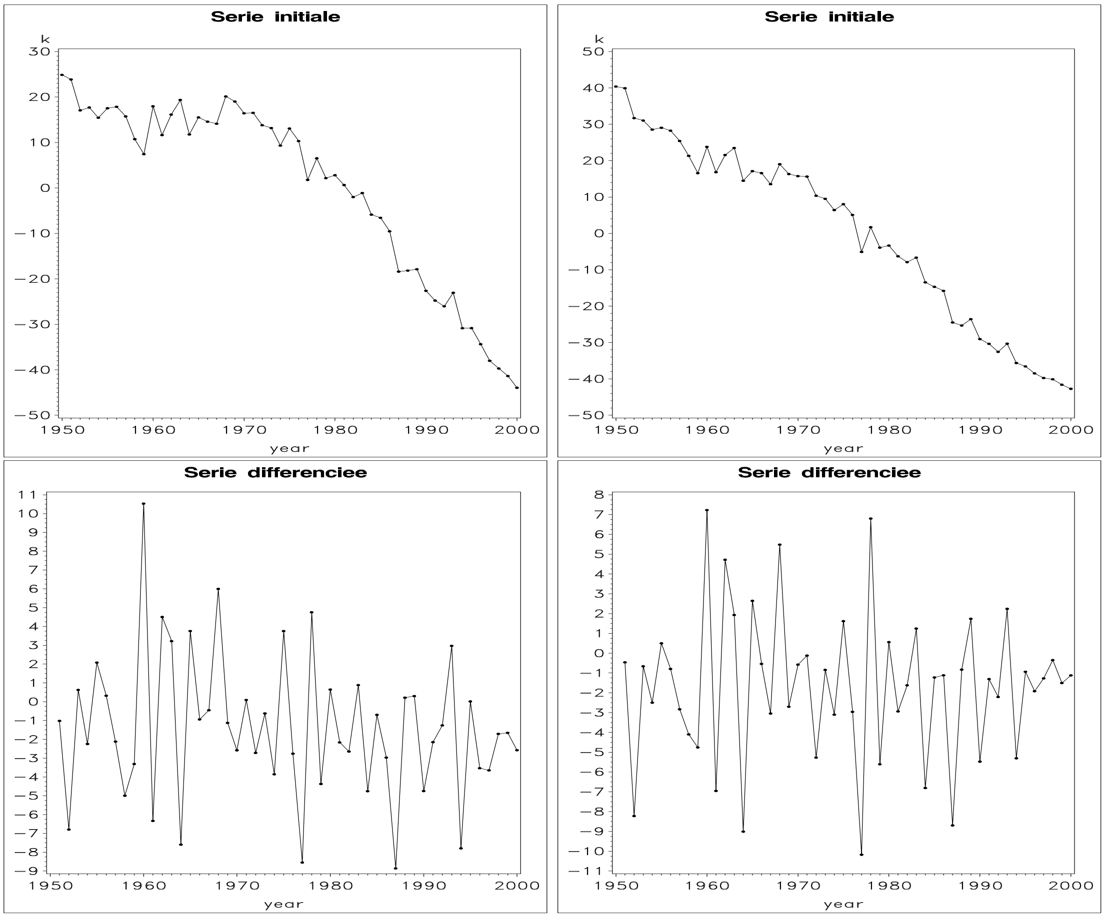

Figure 13: Séries initiales et différences premières (les hommes sont à gauche et les femmes à droite).

femmes. Remarquons au passage que cette technique nous fait perdre une observation. Les séries obtenues $\mathcal{K}_d$ sont apparemment stationnaires en moyenne (voir Figure 13). Afin de confirmer ces impressions, il est pratique d'utiliser des tests formels, comme ceux de Dickey-Fuller ou de Phillips-Perron. Ceux-ci ont pour but d'éprouver l'hypothèse nulle de non-stationnarité selon plusieurs aspects (existence d'une tendance,...). Nous disposons donc d'une batterie de tests dont les p-valeurs sont toujours inférieures à 0.0149 pour les hommes et 0.0102 pour les femmes. La stationnarité en moyenne de nos deux séries différenciées n'est donc pas contestée.

2. **Identification du modèle.** Les graphiques des fonctions d'autocorrélation et d'autocorrélation partielle de la série différenciée sont repris à la Figure 14. Si la première donne le coefficient de corrélation entre deux éléments de la série en fonction du délai qui les sépare, la seconde présente l'avantage d'éliminer l'influence des variables intermédiaires. Ces fonctions traduisent l'influence entre deux observations en fonction du délai; elles jouent donc un rôle majeur dans l'analyse de la dépendance temporelle de la série. Ces graphiques nous permettent d'apprécier la stationnarité de la série (une décroissance lente des autocorrélations partielles indique une non-stationnarité). L'amortissement rapide du corrélogramme et du corrélogramme partiel conforte l'hypothèse de stationarité de $\mathcal{K}_d$.

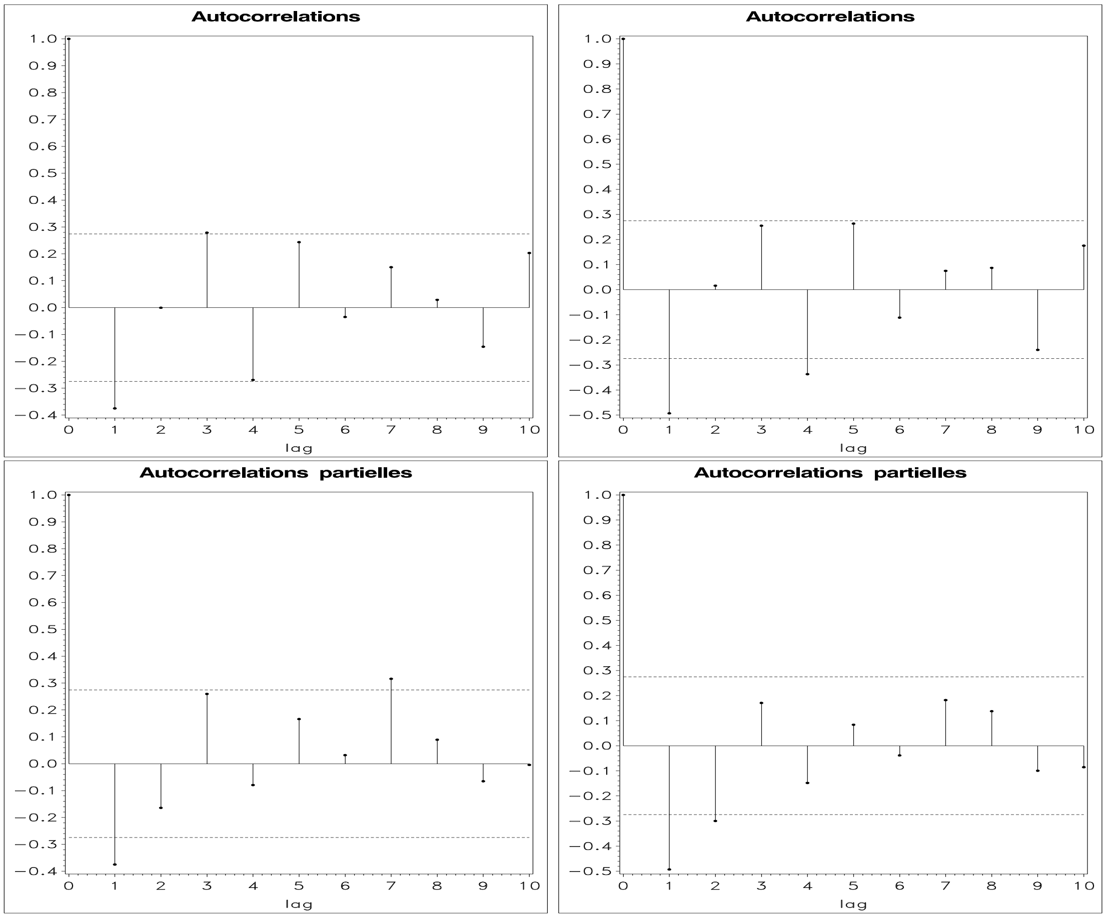

Figure 14: Autocorrélations et autocorrélations partielles des séries différenciées (les hommes sont à gauche et les femmes à droite).

Si nous avions affaire à un processus ARIMA(0,1,0) pour les séries initiales comme c'était le cas dans les applications de la méthode de Lee et Carter aux populations américaine, canadienne, chilienne et japonaise, nous devrions reconnaître dans les corrélogrammes la structure d'un bruit blanc. Remarquons qu'il n'en est pas ainsi : les corrélations et corrélations partielles au délai 1 (lag 1) sortent systématiquement des bandes de confiance. De plus, les p-valeurs des tests éprouvant la nullité des 6 premières autocorrélations valent 0.0036 pour les hommes et 0.0001 pour les femmes, et amènent donc à rejeter cette hypothèse, confirmant donc nos impressions.

Aucun modèle ne semble se dégager explicitement de la Figure 14. Avec un peu d'imagination, nous pourrions sélectionner quelques-uns, comme un MA(1) ou un AR(1). Par souci de parcimonie, nous nous restreignons à un nombre de paramètres inférieur ou égal à 4. Enfin, nous départageons les modèles candidats par le critère BIC (Bayesian information criterion). Les résultats sont les suivants :

BIC	Hommes	Femmes
$ARMA(0,0)$	279.17	277.65
AR(1)	275.49	267.64
MA(1)	275.59	**266.08**
ARMA(1,1)	278.82	268.73
AR(2)	278.02	266.83
MA(2)	277.95	268.20
ARMA(2,1)	**273.97**	267.57
ARMA(1,2)	278.51	271.86
AR(3)	278.45	269.27
MA(3)	280.90	272.05

C'est donc le modèle MA(1) qui est finalement choisi chez les femmes tandis que le ARMA(2,1) est préféré chez les hommes. Précisons que le modèle féminin avait également été retenu par les méthodes SCAN et ESACF de SAS. Au contraire, c'est le AR(1) qui avait été sélectionné pour les hommes.

3. Estimation des paramètres. Nous pouvons maintenant passer à la phase d'estimation. Nous utiliserons pour ce faire la méthode du maximum de vraisemblance (ou des moindres carrés) conditionnelle.

Considérons donc les modèles ARMA(2,1) pour les hommes

$$\begin{aligned}\nabla_1\kappa_t^m &= \rho^m + \phi_1^m\nabla_1\kappa_{t-1}^m + \phi_2^m\nabla_1\kappa_{t-2}^m \\ &+ \theta^m\varepsilon_{t-1}^m + \varepsilon_t^m \end{aligned} \tag{19}$$

et MA(1) pour les femmes

$$\nabla_1\kappa_t^w = \rho^w + \theta^w\varepsilon_{t-1}^w + \varepsilon_t^w \tag{20}$$

avec ε_t^m et ε_t^w des bruits blancs de variances σ_m^2 et σ_w^2. L'estimation par maximum de vraisemblance des paramètres nous donne :

Paramètre	Estimation
ρ^m	-1.3603
ϕ_1^m	-1.2687
ϕ_2^m	-0.5045
σ_m	3.3399
ρ^w	-1.6596
θ^w	-0.5182
σ_w	3.2675

Tous les paramètres sont jugés significativement différents de 0 au niveau $\alpha = 0.05$.

4. **Analyse des résidus.** Pour que nos deux modèles soient corrects, il nous reste à vérifier que les résidus ainsi obtenus sont la réalisation d'un bruit blanc. Nous disposons à cet effet d'un test global éprouvant la nullité des corrélations des résidus. Les résultats en sont repris dans le tableau ci-dessous.

To lag	p-valeur hommes	p-valeur femmes
6	0.2092	0.1488
12	0.5992	0.2733
18	0.5040	0.0720

Nous ne sommes donc pas en mesure de rejeter l'hypothèse de non-corrélation des résidus au niveau $\alpha = 0.05$.

Observons à la Figure 15 les QQ-plots normaux des résidus de nos deux modèles. Les quantiles empiriques et théoriques sont visiblement proches, ce qui conforte encore le modèle retenu.

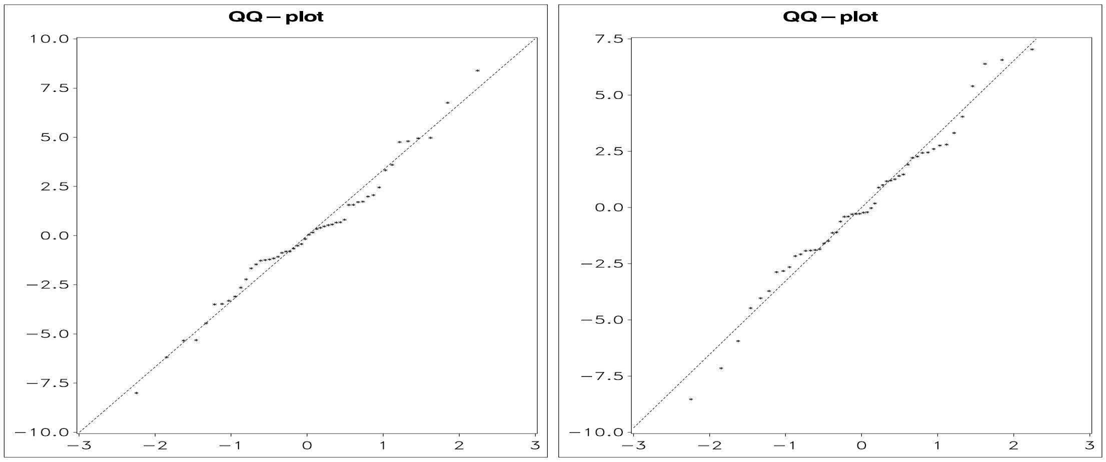

Figure 15: QQ-plots normaux des résidus des modèles ARIMA (les hommes sont à gauche et les femmes à droite).

5. **Prévision.** Finalement, nous pouvons projeter les valeurs de $\mathcal{K}$ dans le futur. On peut voir à la Figure 16 la projection des κ_t à l'horizon 2050, de même qu'un intervalle de confiance pour chacune des valeurs prédites. Il convient ici d'insister sur le fait que nos 51 années d'observation ne nous autorisent normalement pas à effectuer des projections à un horizon lointain. Notons que les intervalles de confiance sur les prédictions des κ_t fournissent directement des intervalles de confiance sur les indicateurs démographiques (du moins si on néglige les erreurs d'échantillonage sur les α_x et les β_x).

   Il suffit alors d'utiliser les projections des κ_t pour obtenir les tables de mortalité prospectives. Une fois la projection $\{\widehat{\kappa}_{2000+s},\ \ s = 1, 2, \dots\}$ obtenue, on en déduit

$$\widehat{\mu}_x(2000 + s) = \exp(\widehat{\alpha}_x + \widehat{\beta}_x \widehat{\kappa}_{2000+s}),\ \ s = 1, 2, \dots .$$

7.2 Validation des modèles et prédictions

Pour éprouver nos modèles, nous comparons les valeurs observées et prédites sur la période [1996,2000] à partir de la période d'observation [1950, 1995]. Les résultats se trouvent à la Table

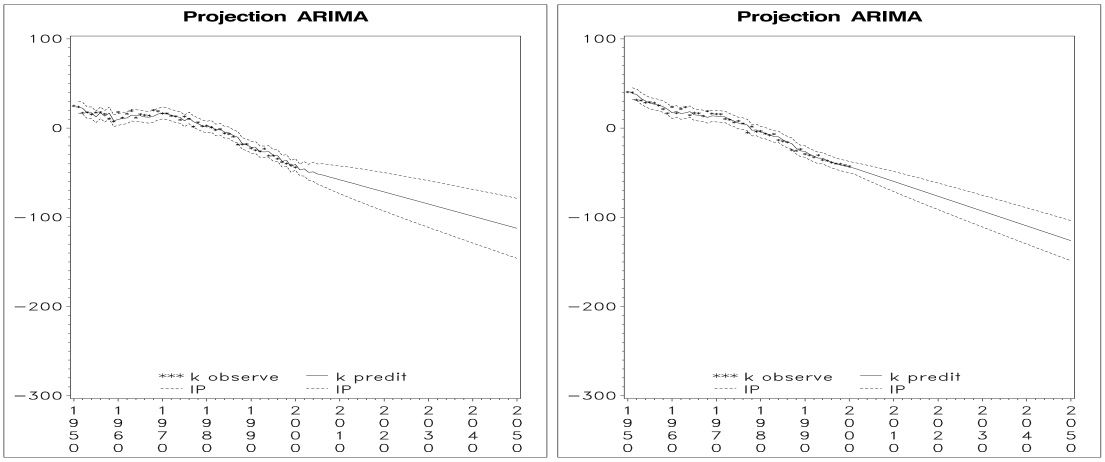

Figure 16: Prévisions à l'horizon 2050 par les différents modèles et intervalles de prédiction à 95% (les hommes sont à gauche et les femmes à droite).

3. Il en ressort que le modèle de prédiction féminin semble efficace, avec des erreurs relatives à un horizon de 5 ans inférieures à 5%. Les erreurs masculines, moins bonnes, restent toutefois raisonnables.

Hommes			
t	κ_t	$\widehat{\kappa}_t$	Erreur relative
1996	-34.36	-30.87	0.102
1997	-38.01	-34.60	0.090
1998	-39.72	-33.62	0.154
1999	-41.37	-36.75	0.112
2000	-43.95	-37.04	0.157
Femmes			
t	κ_t	$\widehat{\kappa}_t$	Erreur relative
1996	-38.48	-38.08	0.010
1997	-39.75	-39.74	0
1998	-40.10	-41.40	-0.032
1999	-41.60	-43.06	-0.035
2000	-42.72	-44.72	-0.047

Table 3: Résultats des prédictions sur la période [1996,2000].

8 Estimation de $a_x(2000)$ et $e_x(2000)$

Nous sommes maintenant capables d'estimer la valeur d'indices de mortalité à un horizon plus ou moins lointain et donc de calculer l'espérance de vie et la prime pure unique d'une rente viagère souscrite en $t = 2000$ par un individu d'âge x grâce aux formules (2) et (3) . La Table 4 reprend les résultats pour un taux d'intérêt annuel $i = 0.04$.

Notons que les espérances de vie restante à 65 ans calculées sur base des statistiques INS 1997-99 donnent 15.62 pour les hommes et 19.85 pour les femmes, respectivement. La projection $e_{65}(2000)$ pour les hommes est donc inférieure à celle obtenue sur base des statistiques triennales, alors que le gain de longévité pour les femmes n'est que de un an. D'autre part, Brouhns &

Hommes		
x	$e_x(2000)$	$a_x(2000)$
0	81.09	23.59
20	59.85	22.05
40	39.02	18.77
65	15.55	10.45
Femmes		
0	90.23	23.98
20	68.44	22.91
40	46.36	20.23
65	20.84	13.80

Table 4: Valeurs de $e_x(2000)$ et $a_x(2000)$ calculées à partir des différents modèles.

Denuit (2001b) avaient obtenu en appliquant la méthode de Lee-Carter sur les plus de 60 ans 16.09 pour les hommes et 21.33 pour les femmes. Ceci met clairement en évidence un des défauts de la méthode de Lee-Carter: l'évolution temporelle est synthétisée dans le seul indice κ_t dont la détermination dépend de la fourchette d'âges considérés. On constate en effet au cours du 20ème siècle un renforcement de la bosse accident, donc une agravation de la mortalité parmi les 20-30 ans. Ceci perturbe l'estimation des κ_t et, partant, leur extrapolation. S'agissant de rentes viagères, il est donc préférable de modéliser les taux de mortalité du public cible, soit les plus de 50 ans (afin d'être à même de traiter les préretraites).

9 En guise de conclusion

Comme l'a très justement fait remarquer Alho (2000), le critère des moindres carrés (qui mène à la décomposition aux valeurs singulières), voire sa version pondérée suggérée par Wilmoth (1993), est raisonnable mais certainement pas optimal. L'approche du maximum de vraisemblance dans le modèle de Poisson est certainement de nature à fournir des résultats intellectuellement plus acceptables.

Spécifiquement, il s'agira ici de considérer que

$$D_{x,t} \sim \text{Poisson}(\mu_x(t)L_{x,t}) \tag{21}$$

où $D_{x,t}$ est le nombre de décès à l'âge x durant l'année t et $L_{x,t}$ est le nombre de personnes/années d'âge x durant l'année t (i.e. l'exposition au risque). La spécification (21) présente plusieurs avantages, au nombre desquels

1. le modèle reconnaît le caractère entier du nombre de décès $D_{x,t}$
2. la loi de Poisson permet de lever l'hypothèse d'homoscédasticité et reconnaît la variabilité plus grande aux âges élevés, due aux effectifs réduits
3. Le recours au principe du maximum de vraisemblance pour estimer les paramètres ne nécessite plus des statistiques de mortalité sous forme rectangulaire. Ceci nous évite de devoir compléter les données, opération qui, comme nous l'avons relevé plus haut, n'est pas sans impact sur les résultats.

La formulation (21) est semblable à celle de Renshaw & Haberman (1996). La différence se situe au niveau de la modélisation des taux instantanés de décès $\mu_x(t)$. Alho (2000) suggère de retenir la forme

$$\ln \mu_x(t) = \alpha_x + \beta_x \kappa_t.$$

Le taux de mortalité relatif à l'âge x et à l'année t est donc de forme log-bilinéaire. Il s'agit en fait d'une version des modèles d'association (voir par exemple GOODMAN (1991)). Cette approche a été suivie dans BROUHNS, DENUIT & VERMUNT (2002).

Remerciements

Cet article reprend une partie des résultats de DELWARDE (2002).

Les auteurs tiennent à remercier chaleureusement l'Association Royale des Actuaires Belges pour son support financier. Ils expriment leur gratitude envers les membres du Groupe de Travail "Mortalité" de l'ARAB, présidé par Monsieur Philippe Delfosse. Les intéressantes discussions avec les professionnels de l'assurance lors des réunions de ce groupe ont motivé la présente étude.

Les auteurs ont bénéficié du soutien précieux des Professeurs Josianne Duchêne et Christian Jaumain de l'Université catholique de Louvain. Ils remercient également le Professeur Montserrat Guillén du Département d'Econométrie de l'Université de Barcelone, le Professeur Ermanno Pitacco du Département de Mathématique Appliquée et d'Actuariat de l'Université Degli Studi de Trieste, le Professeur Annamaria Olivieri du Département d'Economie de l'Université de Parme, et les Professeurs Steven Haberman et Richard Verrall, de la City University de Londres. Enfin, le concours de Mademoiselle Amélie Louis pour la récolte et la préparation des données a été précieux.

Les auteurs tiennent à remercier tous ceux qui ont bien voulu commenter leur travail, tout particulièrement Pierre Devolder et Bernadette Govaerts, pour leur lecture critique et constructive.

Les auteurs assument bien sûr seuls l'entière responsabilité des erreurs éventuelles subsistant dans leurs écrits.

Références

ALHO, J.M. (2000). Discussion of *Lee* (2000). *North American Actuarial Journal* **4**, 91-93.

BELL, W.R. (1997). Comparing and assessing time series methods for forecasting age-specific fertility and mortality rates. *Journal of Official Statistics* **13**, 279-303.

BENJAMIN, B., & POLLARD, J.H. (1993). *The Analysis of Mortality and Other Actuarial Statistics (3d Edition)*. Heinemann, London.

BENJAMIN, B., & SOLIMAN, A.S. (1993). *Mortality on the Move*. Institute of Actuaries, Oxford.

BROCKWELL, P.J., & DAVIS, R.A. (1991). *Time Series: Theory and Methods*. Springer Verlag. New York.

BROCKWELL, P.J., & DAVIS, R.A. (1996). *Introduction to Time Series and Forecasting*. Springer Verlag. New York.

BROUHNS, N., & DENUIT, M. (2001a). Risque de longévité et rentes viagères I. Evolution de la mortalité en Belgique de 1880 à nos jours. Discussion Paper, Institut de Statistique, Université catholique de Louvain, Louvain-la-Neuve, Belgique.

BROUHNS, N., & DENUIT, M. (2001b). Risque de longévité et rentes viagères II. Tables de mortalité prospectives pour la population belge. Discussion Paper, Institut de Statistique, Université catholique de Louvain, Louvain-la-Neuve, Belgique.

BROUHNS, N., DENUIT, M., & VERMUNT, J. (2002). A Poisson regression approach to the construction of projected lifetables for life annuitants. Discussion Paper, Institut de Statistique, Université catholique de Louvain, Louvain-la-Neuve, Belgique.

Coale, A., & Kisker, E.E. (1990). Defects in data on old age mortality in the United States: New procedures for calculating approximately accurate mortality schedules and life tables at the highest ages. *Asian and Pacific Population Forum* **4**, 1-31.

Delwarde, A. (2002). *Analyse prospective de la charge des pensions pour l'Etat belge.* Mémoire présenté en vue de l'obtention du grade d'Ingénieur Civil en Mathématiques Appliquées à l'Université catholique de Louvain, Louvain-la-Neuve.

Goodman, L. (1991). Measures, models and graphical displays in the analysis of cross-classified data. *Journal of the American Statistical Association* **86**, 1085-1111.

Gutterman, S., & Vanderhoof, I.T. (1999). Forecasting changes in mortality: a search for a law of causes and effects. *North American Actuarial Journal* **2**, 135-138.

Jaumain, Ch. (2001). Tables de mortalité en Belgique: questions d'actualité. Actu-L **1**, 67-108.

Lee, R.D. (2000). The Lee-Carter method of forecasting mortality, with various extensions and applications. *North American Actuarial Journal* **4**, 80-91.

Lee, R.D., & Carter, L. (1992). Modelling and forecasting the time series of US mortality. *Journal of the American Statistical Association* **87**, 659-671.

Lee, R.D., & Miller, T. (2000). Assessing the performance of the Lee-Carter approach to modelling and forecasting mortality. see www.demog.berkeley.edu/ rlee

Lee, R.D., & Nault, F. (1993). Modeling and forecasting provincial mortality in Canada. Paper presented at the World Congress of International Union for the Scientific Study of Population, Montréal.

Lee, R.D., & Rofman, R. (1994). Modelacion y proyeccion de la mortalidad in Chile. *Natas* **22(59)**, 182-313. (available in English from the authors, titled "Modeling and forecasting mortality in Chile")

Renshaw, A.E., & Haberman, S. (1996). The modelling of recent mortality trends in the United Kingdom male assured lives. *British Actuarial Journal* **2**, 449-477.

Renshaw, A.E., & Haberman, S. (2001). On the forecasting of mortality reduction factors. Actuarial Research Paper **135**, Department of Actuarial Science and Statistics, City University, London.

Vaupel, J.W. (1997). Trajectory of mortality at advances ages. In *Between Zeus and the Salmon: The Biodemography of Longevity*, pp. 17-37, National Academy of Science.

Wilmoth, J.R. (1993). Computational methods for fitting and extrapolating the Lee-Carter model of mortality change. Technical Report, Department of Demography, University of California, Berkeley.

Wilmoth, J.R. (1996). Mortality projections for Japan: A comparison of four methods. In *Health and Mortality among Elderly Populations*, Caselli, G., and Lopez, A. editors, pp. 266-287. Oxford University Press, New York.