

HEC MONTRÉAL

Les modèles factoriels du risque de longévité

Par

Ibtihel Sassi

**Sciences de la gestion
(Finance)**

*Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences (M.Sc.)*

Août 2014

© Ibtihel Sassi, 2014

REMERCIEMENTS

« Voilà, j'y suis »; il y a eu des moments où j'ai cru ne pas y arriver mais « enfin, j'y suis ».

Ce projet de mémoire représente l'aboutissement d'un nombre incalculable d'heures de travail. Néanmoins, ce fut une aventure de tous les instants dans laquelle j'ai appris à redéfinir mes perceptions, capacités et orientations quotidiennement.

Si on m'a annoncé au début de cette expérience que la rédaction d'un mémoire était avant tout un voyage avant d'être une destination, je dois avouer que je n'ai réellement pu saisir le sens profond de cette analogie qu'au moment où il m'a fallu me préparer à la clore. Et quel voyage !

Une telle traversée ne pouvant se faire seule, j'aimerais sincèrement remercier et rendre hommage aux personnes qui, par leur soutien et leurs précieux conseils, m'ont permis de tenir la barre et aller de l'avant, les jours de tempête comme de beau temps. Un grand merci donc à :

Monsieur Christian Dorion et Monsieur Martin Boyer pour avoir accepté de me diriger tout en me laissant une certaine autonomie. Vos conseils judicieux et votre patience durant l'exercice parfois ardu menant à la rédaction de ce mémoire sont énormément appréciés. J'ai beaucoup appris avec vous, et pour cela je vous remercie profondément.

La direction administrative de la M.Sc. Un grand merci à Madame Louise Guay et à Madame Louisa D'Amato pour leur compréhension et leur soutien inconditionnel. Vous contribuez extraordinairement à la réussite de HEC Montréal et m'avez orientée pour en faire partie. Pour cela je vous serais toujours reconnaissante.

Ma famille, pour leur appui incommensurable et leur support moral, à travers l'ensemble des péripéties entourant mes études universitaires. Papa, Maman, mes adorables sœurs et les très chers Saida, Nour El Houda et Khémaïès : Merci pour TOUT, je ne serais pas allée si loin sans vous.

Mes amis, qui ont soutenu mes humeurs et m'ont encouragé à poursuivre dans mes efforts de recherche.

A tous, encore une fois, un grand merci.

Ibtihel Sassi

SOMMAIRE

Aujourd'hui plus que jamais, les assureurs et les fonds de pension sont tenus d'évaluer leurs provisions et les risques associés à leurs engagements de la manière la plus réaliste possible. Il est donc indispensable de développer des méthodes fiables de prévision de l'espérance de vie humaine.

Ce mémoire, intitulé « les modèles factoriels du risque de longévité », a pour objectif de proposer une alternative de prédiction des taux de survie. L'approche écarte les techniques de conception des tables de mortalité au profit d'une analyse « financière » du *LifeMetrics Index* de *JP Morgan* où les données de longévité sont traitées comme des indices boursiers ou des prix d'actions.

Nous nous inspirons des modèles standards de la mortalité et proposons une méthode qui combine l'analyse descriptive des données avec l'analyse des séries temporelles. L'idée est de proposer une approche récapitulative permettant de détecter les dimensions globales d'une surface donnée de longévité. La prédiction des taux de survie à des âges différents passe ensuite par la prédiction de l'évolution des dimensions globales en question. Le modèle exploite ainsi l'historique d'évolution de la durée de vie humaine et prend en compte son augmentation incessante.

En somme, cette recherche offre des apports théoriques et pratiques importants. Non seulement elle permet une meilleure compréhension du risque de longévité mais elle constitue une base fondamentale pour des travaux futurs se rapportant notamment au transfert du risque de longévité.

Mots clés : Risque, longévité, mortalité, solvabilité, assurance, fonds de pension, Analyse en Composantes Principales, analyse des séries temporelles, prévision, projection, *LifeMetrics*, *LifeMetrics Index*, indice de longévité, JP Morgan

TABLES DES MATIÈRES

REMERCIEMENTS	i
SOMMAIRE	ii
TABLE DES MATIÈRES	iv
LISTE DES TABLEAUX	vi
LISTE DES FIGURES	vii
CHAPITRE 1 : INTRODUCTION	1
CHAPITRE 2 : REVUE DE LITTÉRATURE	10
2.1. Introduction	10
2.2. L’approche expectative et ses modèles	10
2.3. L’approche explicative et ses modèles	11
2.4. L’approche extrapolative et ses modèles	12
2.4.1. Lee & Carter, le modèle référence, et ses extensions	13
2.4.2. Les modèles multifactoriels de la longévité	16
CHAPITRE 3 : FONDEMENTS THÉORIQUES DES CONCEPTS	20
3.1. Les Données : « <i>LifeMetrics Index</i> » ou Indice de longévité	21
3.1.1. Description	21
3.1.2. Les caractéristiques statistiques de l’indice de longévité.....	22
3.2. L’Analyse en Composantes Principales : La théorie au cœur de notre modélisation	22
3.2.1. L’ACP classique	23
3.2.2. L’ACP adaptée à l’analyse des séries temporelles	24
CHAPITRE 4 : APPLICATION : RÉSULTATS ET ÉVALUATION DU MODÈLE	27
4.1. Vue d’ensemble des données	27
4.2. Application de l’approche	30
4.2.1. Modèle sous-jacent aux données de longévité des Femmes aux Pays-Bas	32
4.2.2. Modèle de projection	38

A- Analyse de la première composante principale	39
B- Analyse de la deuxième composante principale	40
C- Analyse de la troisième composante principale	42
4.2.3. Prévission de l'espérance de vie résiduelle par tranche d'âge	43
CHAPITRE 5 : IDENTIFICATION DES DIMENSIONS GLOBALES DE LA LONGÉVITÉ	49
CHAPITRE 6 : CONCLUSION	59
BIBLIOGRAPHIE	61
ARTICLES	61
LIVRES	61
SITES INTERNET	62

LISTE DES TABLEAUX

Tableau 1 : Description statistique des données de longévité aux Pays-Bas.....	29
Tableau 2 : ACP classique : Variances expliquées et variances cumulées	32
Tableau 3 : ACP après M-SSA : Variances expliquées et variances cumulées	33
Tableau 4 : Nombre optimal de CP à retenir – critère de décision : « Akaike »	36
Tableau 5 : Paramètres de la relation établie entre la longévité et ses composantes principales	37
Tableau 6 : Estimation du processus sous-jacent à la CP(1)	39
Tableau 7 : Estimation du processus sous-jacent à la CP(2).....	40
Tableau 8 : Estimation du processus sous-jacent à la CP(3).....	42
Tableau 9 : « Substituts explicatifs potentiels » de la CP(1).....	55
Tableau 10 : « Substituts explicatifs potentiels » de la CP(2).....	56
Tableau 11 : « Substituts explicatifs potentiels » de la CP(3).....	56

LISTE DES FIGURES

Figure 1 : Evolution de la longévité aux Pays-Bas 27

Figure 2 : Espérance de vie résiduelle chez les femmes de 60 ans aux Pays-Bas 30

Figure 3 : Analyse de la fonction d'autocorrélation de l'indice de longévité féminine à 60 ans
aux Pays-Bas 31

Figure 4 : Série logarithmique de la longévité féminine à 60 ans aux Pays Bas 31

Figure 5 : Nombre optimal de CP à retenir – Critère de décision : « *Scree-test* » 34

Figure 6 : Nombre optimal de CP à retenir – Critère de décision : « *% cumulé de la variance
expliquée* » 34

Figure 7 : Évaluation graphique de L'ACP après M-SSA 35

Figure 8 : Qualité d'ajustement de la CP(1) par un ARMA(2,1) 39

Figure 9 : Validation du modèle spécifié pour la CP(1) : Analyse des résidus du
ARMA(2,1) 40

Figure 10 : Qualité d'ajustement de la CP(2) par un ARMA(2,1) 41

Figure 11 : Validation du modèle spécifié pour la CP(2) : Analyse des résidus du
ARMA(2,1) 41

Figure 12 : Qualité d'ajustement de la CP(3) par un ARMA(1,0) 42

Figure 13 : Validation du modèle spécifié pour la CP(3) : Analyse des résidus du
ARMA(1,0) 43

Figure 14 : Prévisions versus réalisations & qualité du modèle, par tranche d'âge de 40 à 80
ans 44

Figure 15 : Corrélation entre les CP et les « substituts explicatifs potentiels » 54

Chapitre 1 : Introduction :

Le vieillissement de la population est aujourd'hui une réalité que l'on se trouve tous appelé à confronter. Bien que le phénomène soit déclenché depuis un siècle et qu'il concerne les quatre coins de la planète, il évolue à un rythme différent d'une zone géographique à une autre.

La cartographie¹ des faits stipule qu'à ce jour, la longévité concerne principalement les pays de l'hémisphère nord. En effet, l'évolution de la démographie dans des pays tel que le Canada, les États-Unis ou encore la France et la Grande Bretagne reflète bien une tendance caractérisée par un allongement continu et généralisé de la durée de vie humaine. Demain, à un horizon de 20 à 30 ans ce sera au tour de la Chine et de l'Inde d'être confrontées au vieillissement accéléré de leurs populations sous l'effet combiné de l'effondrement du taux de mortalité et de l'allongement de la durée de vie. Après demain, tous les pays du monde connaîtront cette révolution démographique.

La longévité, notamment dans les pays développés, est donc désormais perçue comme un processus naturel. L'espérance de vie dépasse déjà 80 ans dans 11 pays. Aujourd'hui, les enfants naissent et grandissent avec la quasi-certitude d'atteindre « le grand âge ». Le nombre de personnes âgées de 65 ans risque de tripler d'ici 40 ans pour atteindre près d'un milliard et demi.

Outre les bouleversements sociaux et familiaux qu'elle engendre, cette amélioration de l'espérance de vie a des conséquences économiques importantes pour les organismes d'assurance-vie et de retraites. Pour un assureur, le risque de longévité est le risque que la durée de vie de sa population assurée soit plus longue que prévue. Que ce soit pour tarifier la prime unique d'une rente viagère ou pour établir le montant des cotisations à la caisse de retraite, la solvabilité d'une compagnie repose en partie sur sa capacité à pouvoir estimer la longévité future avec précision. Une

¹ Henri de Castries -Président-directeur général d'AXA. Global forum for longevity – Mars 2011

sous-estimation de celle-ci implique des provisions insuffisantes pour couvrir les futurs versements d'arrérages.

Alors techniquement, le risque de longévité trouve son origine dans l'inadéquation possible entre l'estimation établie de la mortalité et la survie réelle des rentiers. Un décalage potentiel entre la réalité et les niveaux de longévité que l'on s'est engagé à couvrir entraînerait des pertes conséquentes. Par exemple, l'actif sous-gestion des fonds de pension des pays membres de l'Organisation de Coopération et de Développement Économiques (OCDE) a récemment été estimé à 15 € trillion, ce qui constitue un risque en soi. En règle générale, chaque année supplémentaire d'espérance de vie, ajouterait 3-5% aux dus des régimes des fonds de pension. Ainsi, une sous-estimation de l'espérance de vie à l'échelle mondiale par seulement 1 an, exigerait une augmentation des actifs des régimes de retraite de plus de 500 milliards € pour pouvoir maintenir le niveau de financement actuel. Ce chiffre n'est qu'une illustration parmi tant d'autres de ce qui pourrait être engendré par une longévité future mal-estimée.

Si les régimes publics de retraite sont en mesure de réagir assez efficacement aux scénarios de longévité imprévue, les assureurs et fonds de pension privés ne disposent pas du même avantage. En effet, la logique de retenue à la source permettrait aux premiers d'augmenter les cotisations et / ou réduire les prestations en guise de réaction. Le risque de longévité se trouve ainsi transféré aux rentiers qui seraient obligés de payer plus et / ou recevoir moins. Cependant, l'architecture du contrat d'épargne et des rentes viagères rend ce type de réaction impossible chez les deuxièmes. De tels contrats prévoient généralement une phase de constitution et une phase de restitution ce qui rend presque impossible la correction de la tarification après le passage à la liquidation. Par ailleurs, les engagements des institutions privés de retraite sont souvent établis selon des termes fixes définis au moment de la conclusion du contrat d'assurance et non-rectifiables par la suite. Modifier ces caractéristiques afin de reporter le risque de longévité sur les assurés reviendrait à vider les rentes de leur intérêt social. Rappelons que la durée de l'engagement de

l'assureur pour ce type de contrat est considérable, augmentant l'incertitude liée aux taux de mortalité / survie utilisés pour la tarification et le provisionnement.

Il est donc évident qu'une mise en place d'un processus de gestion adéquate et efficace de ce phénomène est une question vitale pour le système non seulement d'un point de vue micro-économique, mais surtout d'un point de vue macro-économique. En effet, le scénario des gouvernements qui se lancent, sans préavis, dans des augmentations des cotisations et / ou réductions des prestations de retraite s'est produit dans de nombreux pays industrialisés au cours des dernières décennies. D'autres gouvernements se trouvent aujourd'hui obligés d'augmenter les âges légaux de départ à la retraite pour être en mesure de prendre en compte l'amélioration de la longévité et les effets du vieillissement de la population sur le financement des retraites.

Les enjeux sont donc considérables et de nombreux défis risquent d'être provoqués à divers niveaux de la société. Adresser le phénomène est désormais une urgence pour tous. Les fonds de pension et régimes de retraite sont effectivement dans l'obligation d'établir les bonnes stratégies de couverture et de transfert du risque de longévité. Pour ce faire, deux approches sont envisageables : L'approche assurance et l'approche marché des capitaux.

La première fait appel aux réassureurs qui, de par leur connaissance du marché, peuvent aider les compagnies d'assurance à mieux appréhender leur risque de longévité. Plusieurs contrats de réassurances peuvent alors être conçus. Entre autres, un *Excess of Loss* en durée est une couverture qui permettrait de transférer la responsabilité des engagements au réassureur après une certaine période fixée isolément pour chaque police d'un portefeuille d'assurance : Dès lors que l'assuré dépasse un certain âge, le réassureur se substitue à l'assureur pour le paiement de la rente. Deux options d'intervention sont donc possibles. La première fait en sorte que le réassureur paye les arrérages à la place de l'assureur jusqu'au décès de l'assuré. Dans ce cas, le réassureur assume seul l'incertitude de l'évolution ultérieure de la longévité et les conséquences qui lui seraient associées. Avec la seconde option, le

réassureur choisit de verser à l'assureur la provision mathématique de la rente évaluée à la date de conclusion du contrat. Ils se partagent ainsi le risque de dérive de la mortalité dans le temps. *L'Excess of Loss* en durée est un contrat qui s'adapte bien aux portefeuilles composés de têtes très âgées.

Donc en théorie, la réassurance est une solution éventuelle au risque de longévité. Cependant, son potentiel limité de diversification en fait une solution inefficace, nécessaire mais insuffisante. En effet, l'augmentation simultanée et fortement corrélée de la longévité dans presque tous les pays du monde et le fait que la majorité des rentes soient vendues dans les pays industrialisés se traduit par un risque peu diversifiable et très difficile à mutualiser géographiquement. La réalité se traduit donc par un secteur de réassurance incapable d'absorber l'ampleur globale du phénomène de la longévité où les réassureurs ne peuvent ni exercer pleinement leur capacité de diversification ni faire jouer convenablement le phénomène de mutualisation.

La deuxième approche de couverture, soit celle impliquant le transfert du risque de longévité aux marchés des capitaux vient donc soutenir la réassurance et propose les alternatives difficiles à développer ailleurs. Des instruments financiers liés à la mortalité / longévité y sont conçus et permettent aux fonds de pensions et fournisseurs de rentes de couvrir leurs expositions respectives. Les fournisseurs de couverture, quant à eux, sont compensés pour la prise en charge du risque de longévité par une prime de risque leur permettant un retour sur investissement supérieur aux taux d'intérêt sans risque si la mortalité / longévité évolue comme prévu.

Les tentatives de titrisation du risque de longévité sont assez nombreuses mais les principaux instruments de couverture à ce jour sont incontestablement : les q-forwards, les s-forwards, les obligations de longévité et les swaps de longévité.

Introduits par JP Morgan, les "*forwards*" de longévité sont les plus simples instruments permettant le transfert du risque de longévité aux marchés financiers. Il

s'agit de produits structurés impliquant l'échange du taux de mortalité d'une population réalisé à une date future contre un taux fixe convenu initialement.

Le "*q-forward*" est un contrat à terme de couverture contre la mortalité. Il s'agit d'une entente entre deux parties qui se traduit par l'échange, à une date future (échéance du contrat), d'un montant proportionnel au taux de mortalité réalisé contre un montant fixe convenu au moment de conclusion de l'entente². Le taux de mortalité utilisé pour un *q-forward* correspond à la probabilité de décès dans l'année. Le taux fixe se détermine au moment de conclusion du contrat en ayant recourt à des moyens sophistiqués de calcul actuariel et statistique. Le taux réalisé, quant à lui, dépend de l'issue d'un indice de longévité à l'échéance (le *LifeMetrics* de JP Morgan par exemple).

Le règlement net à l'échéance d'un *q-forward* dépend de la différence entre les deux taux. Si le taux de mortalité réalisé est au-dessous du taux initialement fixé, le fournisseur de couverture paye au fond de pension un montant proportionnel à la différence et vice versa. C'est ainsi qu'un *q-forward* reflète une couverture contre le risque de longévité. En effet, des taux de mortalité plus bas signifient un allongement de la durée de vie humaine. Donc si la mortalité observée est moins élevée que le taux initialement estimé, le fond de pension serait en mesure de payer des prestations supplémentaires de retraite car il reçoit de l'argent du fournisseur de couverture, l'acquéreur et détenteur de *q-forward*.

Un *s-forward* ou "*survivor*" *forward* est un produit assez similaire au *q-forward*. La seule différence est que le sous-jacent correspond au taux de survie d'une population donnée et non-pas à son taux de mortalité. Le taux fixe de survie que l'on retient au moment de conclusion du contrat correspond à la meilleure estimation du taux de survie de la population en question majorée d'une prime de risque payable par le régime de retraite. Le règlement net à l'échéance dépend de l'issue de l'indice de longévité sous-jacent au contrat. Le *s-forward* est tel que, si le taux de survie réalisé

² q-Forwards : Derivatives for transferring longevity and mortality risk, 2007, G. Coughlan, D. Epstein, A Sinha and P.Honig

dépasse le taux de survie initialement convenu, le fond de pension reçoit de l'argent de son fournisseur de couverture puisqu'on a plus de survivants que prévus à la conclusion du contrat.

Notons que, pour ce qui est de la gestion du risque de longévité, l'utilisation du taux de survie dans les *s-forwards* constitue un avantage comparativement au recours à la probabilité de décès dans l'année par les *q-forwards*. En effet, les fonds de pensions sont souvent plus intéressés par le taux de survie d'une population donnée que de son taux de mortalité. Un produit dont le sous-jacent est la probabilité de survivre N années leur seraient donc plus convenable. En outre, contrairement au *q-forward* où le régime de retraite est payeur du taux réalisé, on est payeur du taux fixe dans une *s-forward* et donc moins affecté par l'incertitude associée à l'évolution de la longévité.

Les obligations de longévité sont des obligations indexées sur la durée de vie humaine. Les réassureurs s'en servent pour émettre de la dette sur les marchés financiers. L'opération est généralement effectuée via une SPC (*Special Purpose Company*). En cas d'allongement excessif de la durée de vie humaine, les flux des coupons et / ou le principal qui auraient dû être remboursés aux détenteurs des obligations sont suspendus ou réduits. Ces flux sont alors reversés à l'assureur qui les utilise pour régler les rentes. Pour les investisseurs, l'avantage est de détenir une nouvelle classe d'actifs dont les rendements sont peu corrélés aux autres produits traditionnels et d'obtenir ainsi un bénéfice de diversification.

Swiss Re fut la première à s'aventurer dans la titrisation associée à la durée de vie humaine. En 2003, Vita I, une obligation basée sur un indice de mortalité fut émise sur les marchés. L'instrument à 3 ans de maturité, est une '*mortality bond*' dont le principal se trouvait réduit (respectivement annulé) si l'indice de mortalité excédait 130% (respectivement 150%) du niveau de 2002. Le succès de ce produit auprès des investisseurs motiva la reprise de l'expérience et l'émission de Vita II(2005) et Vita III(2007) par Swiss Re, Osiris(2006) par Axa et Tartan(2006) par Scottish Re.

Une “*mortality bond*” est donc clairement un instrument de couverture contre une éventuelle hausse de mortalité. Mais un raisonnement symétrique appliqué à un indice de longévité nous permettrait de concevoir des outils similaires assurant la couverture contre l’allongement de la durée de vie. C’est ainsi que les “*survivor bonds*” ont vu le jour en étant des obligations “*coupon-based*” dont le paiement est proportionnel au taux de survie d’une population de référence spécifiée au moment d’émission. Ces obligations peuvent avoir une durée limitée fixée à l’avance sinon elles arrivent à échéance lorsque le dernier survivant de la population référence meurt. Une obligation de longévité réduit le risque d’allongement de la durée de vie puisqu’un accroissement de la longévité conduit à des versements de coupons plus élevés permettant aux responsables des fonds de pension de mieux gérer leur exposition au risque de longévité.

En 2004, la “*survivor bond*” de PNB Paribas fut la première véritable obligation de longévité à être émise sur les marchés. Indexée au taux de survie des hommes anglais et gallois de 65 ans en 2003, l’obligation avait une maturité de 25 ans. La structure assez complexe du produit et la faible demande qu’il a rencontré, notamment car le risque couvert était trop grand et que le capital nécessaire était trop important pour la protection offerte, ont fait de l’opération soit un quasi échec. L’obligation fut abandonnée l’année suivante.

Autres que les *q-forwards*, les *s-forwards* et les obligations de longévité, les swaps de longévité sont l’une des formules les plus attractives pour traiter le risque d’allongement de la durée de vie humaine. L’instrument permet l’échange de flux monétaires fixes payé par l’assureur contre des paiements variables versés à intervalles réguliers par la partie contractante. Le revenu flottant est d’autant plus élevé lorsque la longévité est supérieure aux estimations et que l’écart est plus grand. Un swap de longévité n’est donc rien d’autre qu’un ensemble de *s-forwards* avec des échéances différentes. Le risque de longévité se trouve ainsi annulé puisque l’assureur paye un montant prévu par un contrat indépendamment des taux de survie observés ultérieurement. En comparaison avec les obligations de longévité, les swaps de longévité permettent d’éviter un intermédiaire (le réassureur via la *SPC*). En

revanche, l'absence de collatéral (tel que le principal de l'obligation) induit un risque supplémentaire de défaut de la contrepartie.

En 2009, Canada Life était parmi les premiers assureurs à avoir signé un tel contrat. Le Swap structuré par JPMorgan, permettait une couverture du risque de longévité pour une durée de 40 ans. Structuré par Crédit Swiss pour un nominal d'environ 500 millions de livres, le swap contracté par Babcock International est aussi une des transactions les plus célèbres.

Grace à un volume de transactions dépassant les 10 milliards d'euros en Europe, l'année 2011 s'est marqué comme année record pour ce marché qui continue à se développer lentement. Cependant, autres les questions éthiques qui interviennent remettant en question le fait qu'un swap pourrait être considéré comme un produit de spéculation sur la mort, la difficulté du tarification d'un tel instrument en complique la commercialisation.

En résumé, nous pouvons affirmer qu'en comparaison aux premières tentatives des années 2000, le marché de longévité s'est énormément développé. Cependant, il est encore assez opaque et loin d'être un marché liquide. Le nombre réduit d'intervenants se traduit par des transactions de gré à gré impliquant souvent le même groupe d'acteurs du marché. En conséquence, on ne reçoit que peu d'information concernant les prix des instruments de couverture de la longévité ce qui ne facilite guère le développement de nouvelles opérations.

Pour réussir l'objectif d'un marché mondial liquide de longévité, plusieurs questions doivent être adressées. La législation et une réglementation appropriée sont certainement au cœur du débat mais ne sont pas suffisantes. L'efficacité d'un tel marché et des différents mécanismes qui le constituent n'est garantie que par la convergence entre "acheteurs" et "vendeurs" vers des prix d'équilibre fiables. Pour ce faire, une évaluation réaliste des engagements de retraite, la mise à jour continuelle des hypothèses d'évolution de la mortalité, l'analyse profonde et la modélisation efficace du risque de longévité sont indispensables.

Nous consacrons donc ce mémoire à la modélisation du risque de longévité. Nous nous plaçons dans un cadre financier plutôt qu'actuariel et analysons des indices de longévité plutôt que des tables de mortalité. La quête d'une approche dynamique mais surtout adaptable (et favorable) au développement des moyens et techniques de couverture est la piste de réflexion ayant « inspiré » ce choix de sujet de recherche. Nous présentons d'abord une revue de littérature résumant les principales études antérieures portant sur la modélisation de la longévité. Nous développons ensuite notre méthodologie. Nous évaluons enfin les résultats et présentons les principales conclusions de notre approche.

Chapitre 2 : Revue de littérature :

2.1. Introduction :

La modélisation de la durée de vie humaine a une longue histoire. De nombreux modèles ont été proposés depuis la loi de mortalité de Gompertz en 1825. Des décennies d'évolution ont permis de substituer les méthodes préliminaires relativement simples par des méthodes plus sophistiquées. Ainsi, des plus primitives aux plus récentes et complexes, les méthodes de projection de la longévité peuvent être classées sous trois grandes approches : les méthodes de projection par avis (expectation), par cause (explication) et par extrapolation. La première est fondée sur les opinions subjectives d'experts portant sur différents degrés de formalité. La deuxième fait souvent appel à des modèles structurels ou épidémiologiques fondés sur des causes de mortalité connues et mesurables. La troisième, quant à elle, fait usage de la régularité caractérisant la structure selon l'âge de la mortalité et l'ensemble des tendances temporelles qu'elle contient. Les modèles extrapolatifs reposent ainsi sur des hypothèses d'évolution stochastiques de l'espérance de vie humaine et ignorent souvent toute influence potentielle des avancées médicales ou des évolutions sociales et environnementales. Nous passons en revue les trois approches mais nous nous concentrons sur l'approche extrapolative puisque notre proposition de modélisation de la longévité s'y inscrit.

2.2. L'approche expectative et ses modèles :

Les « expectations / attentes » ont souvent été utilisées dans la prévision des taux de mortalité sous la forme d'avis d'experts. La prévision en question correspond à la conception d'un scénario bien spécifié, souvent accompagné par des scénarios alternatifs à la hausse et à la baisse.

La méthode la plus couramment utilisée dans l'approche expectative est « le ciblage de l'espérance de vie » où une valeur estimée est prévue pour une date future avec un chemin de réalisation spécifique. Un exemple illustrant cette technique de ciblage est donné par son application à l'étude des facteurs de réduction de la mortalité réalisée

par le bureau de l'investigation continue de la mortalité (CMIB)³ de l'institut anglais des actuaires.

La méthode estime la proportion q_x des personnes d'âge « x » qui meurent avant d'atteindre l'âge « $x+1$ » en calculant un facteur de réduction $RF(x, t)$, où :

$$RF(x, t) = \alpha(x) + [1 - \alpha(x)] * [1 - f_n(x)]^{\frac{1}{n}}$$

et $\alpha(x)$ et $f_n(x)$ représentent respectivement le facteur de réduction ultime et la proportion de $[1 - \alpha(x)]$, le déclin total à observer dans n années. Cette méthode suppose un déclin exponentiel de la mortalité vers la valeur asymptotique $\alpha(x)$ et se fie à l'opinion des experts pour définir les valeurs de $\alpha(x)$ et $f_n(x)$.

L'opinion des experts accorde à l'approche expectative l'avantage d'incorporation des connaissances démographiques et épidémiologiques et permet la prise en compte des avancées médicales et de leurs mises à jour continues. Cependant, la subjectivité de cette opinion et le potentiel de biais qu'elle amène se traduisent inévitablement par une faiblesse. Cet inconvénient est d'autant plus prononcé lorsqu'on apprend que le conservatisme de l'opinion des experts résulte souvent en des taux de survie prévisionnels largement au-dessous des réalisations ex-post. Ce conservatisme est dû à l'incapacité et aussi à la réticence des spécialistes à poursuivre les tendances à long terme de la mortalité ce qui cause l'impopularité de l'approche expectative

2.3. L'approche explicative et ses modèles :

L'approche « explicative » de la mortalité est basée sur des modèles structurels et des modèles épidémiologiques retraçant des liens de causalités entre un certain nombre de causes de décès et les niveaux de mortalité / longévité observés. Les modèles explicatifs intègrent les avancées les plus à jour des processus médicaux et les connaissances se rapportant aux facteurs de risque. Cependant, l'ambiguïté

³ CMIB (2005)

persistante autour des liens de causalités en question freine le développement de cette approche.

La plupart des modèles explicatifs correspondent à des modèles de régression linéaire. Les variables explicatives ou les facteurs de risque intégrés dans ces modèles le sont soit comme des variables « retardées » ou comme des variables projetées dans le futur. Lorsque les modèles structurels sont basés exclusivement sur les facteurs de risque exogènes retardés, les horizons de prévision se trouvent limités à l'étendue du plus petit retard. De cette manière, l'approche explicative se trouve généralement limitée à des prévisions à court terme.

Manton et al (1991 et 1992) ont eu recours aux modèles structurels impliquant les autorégressions multivariées pour relier le mode de vie et autres facteurs de risque à la mortalité chez les personnes âgées. L'approche performe mieux à court terme comparée à certaines méthodes impliquant l'opinion des experts et l'analyse classique des séries temporelles. Cependant, l'efficacité des prévisions est réduite à plus long terme non seulement à cause de la brièveté des séries chronologiques des facteurs de risque mais aussi à cause du grand nombre de paramètres à estimer et la forte hypothèse de liens linéaires entre les différentes variables.

2.4. L'approche extrapolative et ses modèles :

Après l'expectation et l'explication, l'extrapolation est aujourd'hui à la base de la plupart des méthodes de prévision de la durée de vie humaine. L'approche suppose que les tendances futures sont essentiellement une continuation du passé. On calcule des estimations de la mortalité future en utilisant son niveau actuel ainsi qu'une estimation du taux de variation futur. En raison des régularités historiques caractérisant les données de la mortalité et de l'espérance de vie humaine, cette hypothèse clé s'annonce assez justifiée. Cependant et pour deux principales raisons, cet aspect objectif tant apprécié est parfois remis en question. D'abord, réussir les prévisions extrapolatives exige un minimum de subjectivité, le choix de la période d'observation entre autres. En outre, l'objectivité empêche la prise en compte des

avancées médicales et l'impact d'évènements exceptionnels (tel que les crises démographiques dues à des guerres ou des épidémies) ce qui constitue une faiblesse pour les modèles extrapolatifs.

Ces derniers varient de la simple prolongation linéaire des taux de survie tel que proposé par White (2002) jusqu'aux modèles factoriels de la longévité, souvent plus sophistiqué. Ces méthodes permettent d'éviter la lourde tâche d'un traitement simultané de toutes les séries chronologiques décrivant l'évolution des taux de mortalité ou de survie aux différents âges. La projection de la mortalité / longévité y consiste à une réduction de la dimension des données suivie par l'extrapolation statistique des composantes génératrices « affectées par le temps » à l'aide des techniques d'analyse des séries temporelles. Notons que l'introduction de telles techniques a permis de faire évoluer les analyses prédictives de longévité en substituant les prévisions ponctuelles déterministes par des distributions probabilistes prévisionnelles. La modélisation ARIMA univariée ou multivariée a souvent été utilisée dans ce contexte écartant le recours aux méthodes basées sur la régression linéaire et donc tenant compte de la non-linéarité de la mortalité dans le temps. Remarquons aussi que bien qu'elles aient été initialement développées pour les prévisions à court terme, les méthodes de séries chronologiques sont de plus en plus adaptées pour la prévision à long terme de la longévité. C'est donc en traitant les faiblesses majeures des deux approches précédentes que les modèles extrapolatifs réussissent à être les plus populaires.

Nous présentons ci-dessus une revue des modèles factoriels les plus importants de prévision de la mortalité / longévité.

2.4.1. Lee & Carter, le modèle référence, et ses extensions :

Lee & Carter (1992) est incontestablement le modèle référence de l'approche extrapolative en particulier et de la modélisation de la durée de vie humaine en générale. L'idée est de passer par une décomposition en valeurs singulières (SVD) de la matrice des taux de mortalité (doublement indexés, par l'âge et le temps

calendaire). Les données de mortalité sont ainsi réduites et approximées par un produit de deux vecteurs propres, l'un traduisant l'effet de l'âge et l'autre l'effet du temps calendaire. L'extrapolation, par la suite, du vecteur tendanciel décrivant l'évolution temporelle permet la projection de la mortalité future et donc la production de ses tables prospectives.

Les auteurs modélisent le taux instantané de mortalité tel que :

$$\ln\left(\hat{\mu}_{xt}\right) = \alpha_x + \beta_x k_t + \varepsilon_{x,t}$$

avec $t_m \leq t \leq t_M$ et en supposant les variables aléatoires $\varepsilon_{x,t}$ indépendantes, identiquement distribués selon une loi $N(0, \sigma)$, α_x décrit le comportement moyen des taux instantanés de mortalité au cours du temps, β_x traduit l'écart des taux par rapport au taux de mortalité moyen à l'âge x et k_t représente l'évolution de la mortalité au cours du temps. En d'autres termes, le taux de mortalité à l'âge x pour l'année t est décomposé par l'équation Lee & Carter (sur une échelle logarithmique) en la somme d'une composante spécifique à l'âge x et d'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). En outre, on fait l'hypothèse d'une faible variance des erreurs.

Appliqué aux données historiques de 1933 à 1987 aux États-Unis, le modèle LC performe bien et produit des résultats satisfaisants pour les différentes tranches d'âges, même pour celles dont le nombre d'observations est limité. Cependant et même qu'il ait fait ses preuves à maintes reprises, ce modèle référence présente quelques limites. D'abord, la méthode se borne à l'extrapolation dans le futur des tendances constatées dans le passé et subit ainsi les inconvénients de l'objectivité empêchant la prise en compte des avis des experts sur l'évolution présumée de la mortalité. Les progrès de la médecine, l'apparition de nouvelles maladies, l'évolution du style de vie ou encore l'occurrence d'événements exceptionnels sont tout simplement des aspects écartés du cadre Lee & Carter. Ensuite, les tests

empiriques mettant en évidence la forte corrélation entre les erreurs du modèle contredise l'hypothèse d'homoscédasticité et démontre que le biais de modélisation qui en résulte est d'autant plus important pour les tranches d'âges les plus élevés, notamment en raison du faible nombre d'observations. Enfin, l'hypothèse de la sensibilité constante de la mortalité instantanée à l'âge x avec une β_x fixe au cours du temps est un autre aspect de la méthode du Lee & Carter qui se trouve souvent remis en question.

Plusieurs travaux ont été consacré à l'amélioration de la version originale de ce modèle référence. Certains concernent uniquement le traitement préalable des données de mortalité, d'autres s'attardent sur l'application même de la méthode et rediscutent les aspects techniques de sa conception.

Contrairement à la version originale, Lee & Nault (1993) distinguent les femmes des hommes en appliquant le modèle LC aux données canadiennes. Leur modèle performe légèrement mieux que lorsqu'on retienne la même série de k_t pour estimer la structure par âge des taux de mortalité pour les deux sexes.

Lee & Miller (2001) limitent la période d'estimation à la deuxième moitié du vingtième siècle pour réduire l'impact des changements structurels et retiennent une mesure d'espérance de vie résiduelle plutôt que le taux de décès. Une évaluation de 15 ans de prévision appliquée à une population d'âge 20 ans montre que la version de Lee & Miller apporte des améliorations significatives par rapport au modèle Lee & Carter.

Bongarts (2004) remet en question l'hypothèse de la vitesse constante d'évolution de la mortalité. L'auteur applique le modèle LC à la population suédoise féminine et constate une plus forte hausse pour les décès à 75 ans pour la période 1950-2000 comparativement aux périodes 1850-1900 et 1900-1950. Il profite de ce constat pour rappeler, qu'en général, l'évolution de la mortalité au cours du temps est tout sauf constante. Elle diminue à une vitesse décroissante pour les plus jeunes contre une vitesse croissante pour les âges les plus avancés.

Plusieurs tentatives d'amélioration ont également été proposées par apport à l'inconvénient majeur du modèle LC, soit l'hypothèse d'homoscédasticité des erreurs. Planchet & Lelieur (2007) prouvent, qu'en pratique, cette condition est rarement vérifiée. Ils décident alors de consacrer leur étude à la prévision de la mortalité des tranches d'âges les plus élevés. Ils commencent par proposer une méthode d'extrapolation permettant d'augmenter l'échantillon relatif aux âges avancés. Ensuite, ils présentent une adaptation consistant à modéliser « D_t » le nombre de décès à l'âge x durant l'année t à l'aide d'une loi de poisson. Cette nouvelle spécification, qui a recours au principe de maximum de vraisemblance pour estimer les paramètres, permet de lever l'hypothèse d'homoscédasticité et reconnaît une plus grande variabilité des taux de mortalité aux âges avancés. Le *modèle poissonien* performe tout aussi bien que celui de Lee & Carter. Toutefois, il tend toujours à sous-estimer les taux de mortalité aux âges très élevés.

2.4.2. Les modèles multifactoriels de la longévité :

Les développements les plus récents d'analyse et de prévision de la longévité continuent à privilégier l'approche extrapolative. Les modèles unifactoriels semblent cependant être délaissés au profit des méthodes multifactorielles.

Dans « *Lee-Carter mortality forecasting with age-specific enhancement* », Renshaw & Haberman (2003) critiquent la dimension unifactorielle de l'approche classique de Lee & Carter qui fait qu'on ne retienne qu'une seule composante temporelle, la même pour tous les âges. Les auteurs démontrent que cet aspect amène une certaine rigidité du modèle lors de son application à certains types de données. Par exemple, ils remarquent qu'en Angleterre et au Pays de Galles, l'impact des maladies de société (le suicide essentiellement) ainsi que l'effet du VIH/SIDA sur la mortalité des âges 20-39 (notamment au cours de la deuxième moitié du 20^e siècle) ne peuvent pas être captés par le modèle de base de Lee & Carter.

Pour pallier à cette limite, Renshaw & Haberman proposent une extension consistant à rajouter une deuxième composante temporelle choisie parmi celles de la SVD. Leur

objectif principal est de permettre à d'éventuels effets spécifiques à l'âge d'être intégrés dans le processus de prévision de la mortalité. Les auteurs évitent aussi la critique relative à l'homoscédasticité en imposant une loi de poisson aux erreurs du modèle. Les prévisions sont générées en modélisant les deux composantes temporelles avec deux processus ARIMA différents. Les résultats obtenus amènent les auteurs à conclure que le choix d'une modélisation distincte et uni-variée pour chacune des composantes temporelles de leur modèle en constitue une vraie faiblesse. Ils recommandent une approche de modélisation plus unifiée et suggèrent le recours aux techniques d'analyse multi-variées des séries temporelles.

La remise en question de la dimension unifactorielle du modèle LC a motivé un autre travail tout aussi intéressant que celui de Renshaw & Haberman. Cairns, Blake & Dowd (CBD 2006) publient « *A two factor model for stochastic mortality with parameter uncertainty: theory and calibration* ». L'article présente un modèle bifactoriel et analyse l'évolution de la mortalité à 60 ans et plus, en Grande Bretagne.

Cairns, Blake & Dowd considèrent que le taux de mortalité de la tranche d'âge x pour l'année t est dicté par une sensibilité à une première composante temporelle qui vient affecter toutes les tranches d'âge d'une façon identique. Cependant, ils estiment qu'un modèle ainsi conçu est incomplet. Ils l'améliorent avec l'ajout d'un deuxième facteur venant affecter différemment la dynamique de la mortalité selon qu'il s'agisse d'âges jeunes ou avancés. L'effet de cette nouvelle composante est plus fort sur les âges les plus élevés. Ce qui permet au modèle d'échapper à la critique relative à l'homoscédasticité des erreurs.

Les auteurs démontrent que, pour un horizon de moins de 10 ans, le modèle CBD performe bien avec une faible erreur de prévision et un risque systématique non significatif. Pourtant, ils concluent qu'au-delà de 10 ans, ce risque devient de plus en plus important et que passé les 20 ans, l'effort d'estimation est dominé, à la fois, par le risque du modèle et le risque des paramètres.

Face à la difficulté de l'élimination du risque de longévité notamment à moyen et long terme, Cairns, Blake et Dowd suggèrent d'améliorer les moyens de son transfert au marché. Ils profitent de l'aspect stochastique de leur modèle et proposent une alternative de *pricing* des « *longevity bonds* » émis plus tôt par la Banque Européenne d'Investissement.

Depuis sa conception, le modèle CBD est adopté comme référence notamment pour les approches multifactorielles de la longévité. Ses avantages lui valent cette distinction puisqu'il réussit à combiner l'indicateur temporel classique avec un nouvel indicateur de dynamique de la mortalité. L'approche, à la fois intuitive et robuste, permet de modéliser des effets relatifs à l'âge et réussit à mieux capter les tendances analysées.

Les faiblesses du modèle CBD sont établies en comparant la performance de ce dernier à la performance du modèle LC. Le CBD présente une meilleure vue d'ensemble pour l'analyse de la mortalité. En revanche, le LC performe mieux pour les âges extrêmes, les plus faibles et les plus élevés.

Plusieurs études ont permis de produire des extensions du modèle de Cairns, Blake et Dowd. La proposition la plus intéressante vient des auteurs eux-mêmes. Cairns, Blake et Dowd (2006b) suggèrent une première généralisation du modèle initial : un modèle à trois facteurs permettant de capter l'effet de cohorte. L'étude s'intitulant « *A quantitative comparison of stochastic mortality models using data from England & Wales and the United States* » démontre que la version du modèle CBD intégrant l'effet de cohorte est un des modèles les plus efficaces lorsque les données de l'Angleterre et du Pays de Galles sont retenues. Les auteurs affinent davantage leur approche en proposant un modèle à quatre facteurs captant l'effet de la génération.

Au terme de cette présentation générale du sujet de notre recherche et de la revue de littérature s'y rapportant, nous prenons davantage conscience de l'ampleur du risque de longévité. Nous constatons notamment le caractère versatile de la longévité et apprenons que les méthodes les plus efficaces pour sa prévision seraient celles intégrant l'historique d'évolution de la durée de vie humaine. D'autant plus que le caractère stochastique des méthodes extrapolatives s'annonce favorable aux développements des outils de couverture et de transfert du risque de longévité.

Il nous paraît donc raisonnable de proposer une alternative de modélisation énoncée sous la forme d'un modèle factoriel extrapolatif. Nous tenons cependant à profiter des avancées des autres approches pour corriger l'aspect purement objectif de la méthode. Nous essayons donc d'interpréter les dimensions globales de la longévité, après les avoir définies et projetées statistiquement.

Nous formulons les fondements théoriques de notre modèle et les détails de sa conception dans la section suivante. Son application et sa performance sont évalués par la suite. Le rapprochement avec les méthodes non-extrapolative est effectué dans un dernier temps.

Chapitre 3 : Fondements théoriques des concepts exploités :

Le choix de la méthode de prévision à retenir est une étape cruciale dans la modélisation du risque de longévité. Ce phénomène, en perpétuelle évolution, pourrait avoir de lourdes conséquences sur la solvabilité des assureurs et des fonds de pensions. Les modèles extrapolatifs performant bien à long terme d'où notre motivation à retenir cette approche. En effet, une optimisation des calculs des primes relatives aux rentes viagères nécessite une amélioration des prévisions de la longévité, notamment à long terme. Cependant, l'objectif et l'horizon de projection ne sont que deux critères parmi tant d'autres pour décider de la méthode à retenir. La réussite d'un concept au détriment d'un autre est largement tributaire de la disponibilité des données. L'application de la plupart des modèles extrapolatifs n'est possible que lorsqu'on est en mesure de fournir des séries historiques assez longues. Dans notre cas, le *LifeMetrics Index* de JP Morgan nous permet de confirmer l'orientation de modélisation que nous choisissons.

Notre approche combine l'analyse descriptive des données avec l'analyse statistiques des séries temporelles. La méthode implique la détection des dimensions globales de la longévité et la spécification d'un modèle sous-jacent aux données puis la formulation d'un modèle de projection. Nous nous inspirons du modèle standard de Lee & Carter qui retient une approche de décomposition en valeurs singulières pour modéliser la longévité. Dans une logique identique, nous retenons une approche d'Analyse en Composantes Principales (ACP). Nous développons donc un modèle factoriel extrapolatif et ce en traitant des indices de longévité plutôt que des tables de mortalité.

Avant de présenter les fondements théoriques des principaux concepts utilisés pour élaborer notre modèle, nous présentons les avantages du *LifeMetrics Index* et les raisons qui en font un choix approprié pour la conception générale de notre méthode.

3.1. Les Données : « *LifeMetrics Index* » ou Indice de longévité :

3.1.1. Description :

Le *LifeMetrics Index* est le seul indice international de longévité. Il fait partie de la plate-forme *LifeMetrics* créée, en 2007, par la banque JP Morgan. Son objectif est d'établir un profil international, à la fois, pour la longévité et la mortalité.

« *En comparaison à d'autres approches, nous avons voulu offrir aux fonds de retraite, aux compagnies d'assurance et aux investisseurs, un ensemble complet d'outils et de possibilités visant à gérer le risque de longévité stratégiquement* » souligne Guy Coughlan⁴. Ce dernier précise que « *la plateforme LifeMetrics et le LifeMetrics Index sont conçus pour fournir aux clients les outils pratiques pour mesurer le risque de longévité et le hedger avec des solutions efficaces impliquant des dérivés et des produits structurés* »⁵.

Les données sur les taux de mortalité et sur l'espérance de vie y sont regroupées selon le sexe, l'âge et la nationalité. Au départ, seuls des indices pour les États-Unis, l'Angleterre et le Pays de Galles ont été développés. Aujourd'hui, le *LifeMetrics Index* inclut les Pays-Bas et l'Allemagne également. Les données se présentent comme suit :

- *Les États-Unis* : de 1968 à 2007, des données pour les tranches d'âges allant de 20 à 80 ans,
- *L'Angleterre & le Pays de Galles* : de 1961 à 2009, des données pour les tranches d'âges allant de 20 à 80 ans,
- *Les Pays-Bas* : de 1951 à 2009, des données pour les tranches d'âges allant de 20 à 80 ans,
- *L'Allemagne* : de 1952 à 2009, des données pour les tranches d'âges allant de 20 à 80 ans.

⁴ Managing Director and Global Head of Pension Asset Liability Management

⁵ JP Morgan – Site officiel, LifeMetrics - Index Description

Cette configuration des données convient parfaitement au concept de modélisation que nous nous proposons de développer. Nous retenons les Pays Bas puisque ses séries chronologiques sont les plus longues.

3.1.2. Les caractéristiques statistiques de l'indice de longévité :

Autre l'avantage de séries à long historique adaptées à la prévision de longévité à long terme, les caractéristiques statistiques du *LifeMetrics Index* constituent un aspect supplémentaire motivant le recours à ces données.

L'indice correspond à un processus stochastique multidimensionnel. Les séries des taux de mortalité et d'espérance de vie y sont présentés pour n tranches d'âge illustrant une famille de vecteurs stochastiques indexés par le temps et à valeurs dans R^n . Autrement dit, l'indice peut être schématisé par une matrice Z de dimension $(n \times T)$ telle que : Notons $Z = (z_{i,j})$, avec $1 \leq i \leq n$ et $1 \leq j \leq T$. Une telle disposition des données permet à la fois de retracer l'évolution dans le temps de longévité et de capter les éventuels liens entre l'historique des différentes tranches d'âge.

Ces caractéristiques nous permettent de confirmer que le *LifeMetrics Index* est bel est bien approprié pour réussir le modèle que nous espérons développer.

3.2. L'Analyse en Composantes Principales : La théorie au cœur de notre modélisation :

Lee & Carter retiennent un procédé de factorisation des matrices et ajustent la surface de mortalité par la méthode de décomposition en valeurs singulières (SVD). Dans une logique identique, nous retenons l'Analyse en Composante Principale (ACP) pour définir un modèle d'ajustement pour les indices de longévité de JP Morgan.

Les principes d'une ACP classique sont énoncés ci-dessous. Son adaptation à la modélisation de la longévité est expliquée par la suite.

La technique d'Analyse en Composantes Principales peut être présentée de différentes façons. C'est du point de vue des analystes de données que nous nous plaçons pour expliciter ses principes généraux. Dans un premier temps, nous présentons les bases d'une ACP dite classique, soit celle appliquée à n variables mesurées sur un ensemble de p individus. L'application de cette technique à des séries temporelles nécessite une certaine adaptation que nous expliquons par la suite.

3.2.1. L'ACP classique :

Un ensemble de n variables $(Z^{(1)}, Z^{(2)}, \dots, Z^{(n)})$ mesurées sur p individus peut être présenté sous la forme d'une matrice $(p \times n)$, soit la matrice Z définie comme suit :

$$Z = \begin{pmatrix} z_{1,1} & \dots & z_{1,j} & \dots & z_{1,n} \\ z_{i,1} & \dots & z_{i,j} & \dots & z_{i,n} \\ z_{p,1} & \dots & z_{p,j} & \dots & z_{p,n} \end{pmatrix} \text{ où } z_{i,j} \text{ est la valeur prise par la variable } Z^{(j)} \text{ pour}$$

l'individu $Z_{(i)}$.

L'idée de base derrière une ACP est de restituer le maximum de l'information contenue dans le nuage de points décrit par la matrice Z et ce en réduisant les n variables initiales en q facteurs composés appelés composantes principales. L'ACP exploite la matrice variance-covariance de Z et en déduit une matrice réduite définie comme suit :

$$F = \begin{pmatrix} f_{1,1} & \dots & z_{1,j} & \dots & f_{1,q} \\ f_{p,1} & \dots & z_{p,j} & \dots & f_{p,q} \end{pmatrix} \text{ où l'élément général } f_{i,k} \text{ est la valeur du facteur } k \text{ pour}$$

l'individu i .

Le pourcentage d'information restituée décroît à mesure que l'on retienne la première, la deuxième ou la $q^{\text{ème}}$ composante principale. Ces dernières respectent indispensablement les conditions de linéarité et d'indépendance tel que :

$$F^{(k)} = \sum_{j=1}^n a_j^{(k)} Z^{(j)} \text{ et } \text{corr}(F^{(k)}, F^{(m)}) = 0, (k \neq m)$$

3.2.2. L'ACP adaptée à l'analyse des séries temporelles :

Lorsqu'elles s'appliquent à des séries temporelles, les matrices de variance-covariance « classiques ou habituelles » ne tiennent pas compte des dépendances potentielles entre les variables chronologiques à différents retards. On risque qu'une ACP classique ne mette pas à profit tout le potentiel informatif du *LifeMetrics Index*. Nous jugeons alors qu'un traitement préalable des données est indispensable pour contourner cet inconvénient.

Pour ce faire, nous retenons la *Multivariate SSA* (M-SSA pour la suite) : une extension de la *Singular Spectrum Analysis* (SSA pour la suite) qui est une méthode d'analyse de données s'appliquant à une seule série temporelle.

Nous présentons d'abord les principes généraux de la SSA. Ensuite, nous expliquons la manière de déduire l'extension M-SSA applicable à plusieurs séries temporelles à la fois.

A- La technique SSA adapte l'ACP pour une meilleure application à une seule série temporelle :

La *Singular Spectrum Analysis* est une technique d'analyse de données qui s'applique à une seule série chronologique à la fois et permet de lui opérer une analyse multivariée. En pratique, la méthode SSA propose de transformer une variable unique de longueur T en une matrice de dimension $(m \times r)$. Elle a pour objectif de ressortir les dépendances / corrélations que pourrait avoir une partie des observations d'une variable unique avec d'autres parties plus décalées dans le temps.

Soit $Z_T = (z_1, z_2, \dots, z_T)$ le vecteur unique analysé, la SSA nous permet d'obtenir

$$Z = \begin{pmatrix} z_1 & z_2 & z_3 & \dots & z_r \\ z_2 & z_3 & z_4 & \dots & z_{r+1} \\ z_3 & z_4 & z_5 & \dots & z_{r+2} \\ \vdots & & & & \vdots \\ z_m & z_{m+1} & z_{m+2} & \dots & z_T \end{pmatrix}$$

Où : $m = T - r + 1$, r est appelé « largeur de la fenêtre » et $Z^{(j)}$, ($j = 1, \dots, r$) correspondent à des versions décalées de la série temporelle Z_T .

Les composantes principales de la série originale sont déduites par application de l'ACP à la matrice Z .

B- La technique M-SSA adapte l'ACP pour une meilleure application à l'indice de longévité :

Comme nous l'avons déjà expliqué, l'indice de longévité est décrit par la matrice $Z = (z_{i,j})$, avec $1 \leq i \leq T$ et $1 \leq j \leq n$, une matrice de dimension $(T \times n)$.

La technique SSA transforme **une seule** série de données de longueur T en une matrice de dimension $(m \times r)$. La technique M-SSA applique cette même transformation, d'une façon simultanée, aux n séries temporelles de l'indice de longévité. Elle propose ainsi de créer une matrice plus grande : \tilde{Z} de dimension $(T' \times n')$ dont la $i^{\text{ème}}$ ligne s'écrit pour $i = 1, 2, \dots, T'$:

$$\tilde{Z}_{(i)} = (z_{1,i}, z_{1,i+1}, \dots, z_{1,i+r-1}, z_{2,i}, z_{2,i+1}, \dots, z_{2,i+r-1}, \dots, z_{n,i}, z_{n,i+1}, \dots, z_{n,i+r-1})$$

avec : $T' = T - r + 1$ et $n' = n \times r$

Le vecteur $\tilde{Z}_{(i)}$ $(n \times r)$ -dimensionnel est alors constitué de n variables (les n tranches d'âge) r -dimensionnelles. L'ACP exige, à ce stade, le calcul de la matrice de covariance. Cette dernière, notée $\tilde{\Gamma}_{nr}$, a la forme suivante :

$$\tilde{\Gamma}_{nq} = \begin{pmatrix} \Gamma_r^{(11)} & \Gamma_r^{(12)} & \dots & \Gamma_r^{(1n)} \\ \Gamma_r^{(21)} & \Gamma_r^{(22)} & \dots & \Gamma_r^{(2n)} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_r^{(n1)} & \Gamma_r^{(n2)} & \dots & \Gamma_r^{(nn)} \end{pmatrix}$$

où les regroupements se font par « croisement » de séries, sur une largeur de fenêtre égale à r .

$\Gamma_r^{(kk)}$ est la matrice carrée d'ordre r , de covariance à différents retards de la $k^{\text{ième}}$ série. Cette matrice $\Gamma_r^{(kk)}$ a la même structure que la matrice de covariance de la $k^{\text{ième}}$ série dans une SSA.

Les composantes principales des « n » tranches d'âge de l'indice de longévité sont alors déduites grâce à l'application de l'ACP à $\tilde{\Gamma}_{nr}$. Cette dernière est une matrice de variance-covariance qui tient compte des dépendances entre les n séries temporelles du *LifeMetrics Index* en général en plus d'exploiter les dépendances entre les différents retards de ces variables.

Nous appliquons les théories énoncées ci-dessus au *LifeMetrics Index* et définissons donc le modèle sous-jacent à nos données. La détection des dimensions globales de la longévité est suivie par la formulation d'un modèle de projection.

Les résultats de notre méthode et l'évaluation de sa performance sont présentés dans le chapitre suivant.

Chapitre 4 : Application : Résultats et évaluation du modèle :

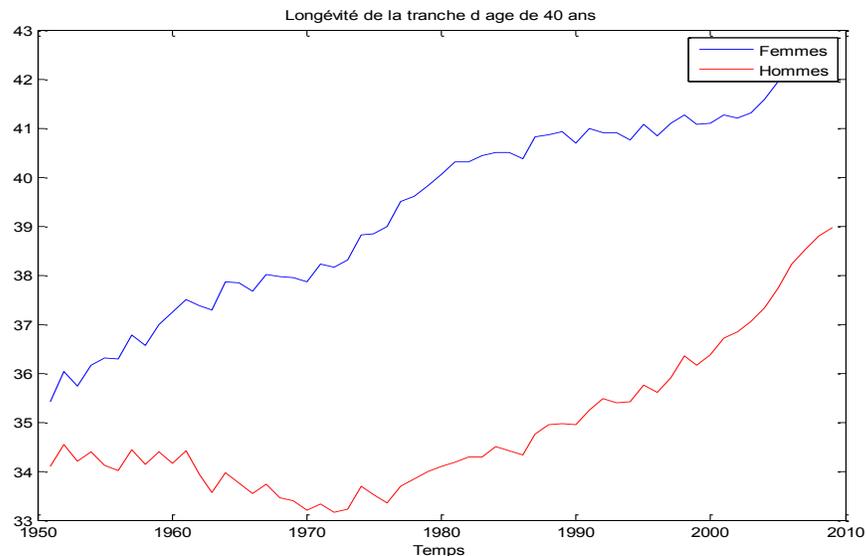
Nous appliquons notre concept de projection de la longévité au *LifeMetrics Index*. Les principaux résultats ainsi qu'une évaluation de l'approche sont présentés dans ce chapitre.

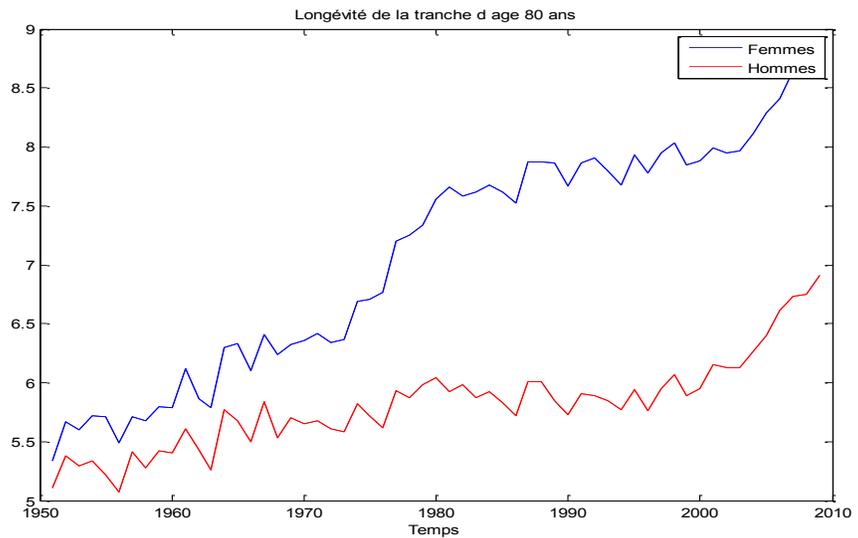
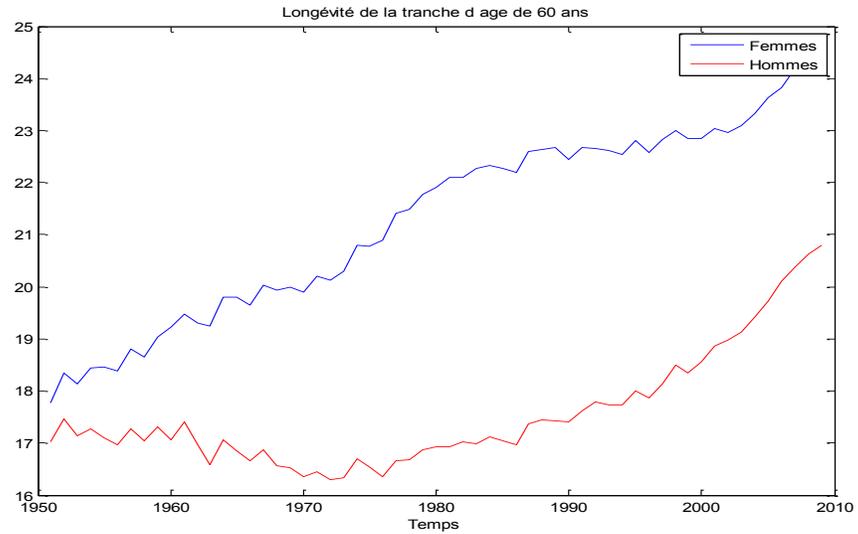
Nous devons tout d'abord décider si nous modélisons conjointement l'espérance de vie des femmes et des hommes ou si, tout comme Lee & Nault (1993), nous analysons séparément les données des deux sexes.

4.1. Vue d'ensemble des données :

Nous retenons les données historiques d'espérance de vie des Pays-Bas : 13 indices de longévité pour l'âge 20 ans, 25 ans, ..., 80 ans. Chaque indice regroupe des observations annuelles pour la période 1951-2009, soit 59 observations par tranche d'âge. La même structure matricielle des données est disponible séparément pour les « femmes » et pour les « hommes ».

Figure 1 : Évolution de la longévité aux Pays-Bas





Comme prévu, nous observons une tendance à la hausse de l'espérance de vie résiduelle aussi bien aux âges jeunes qu'aux âges les plus avancés. Ce constat est cohérent avec l'amélioration de la longévité au cours du temps documentée plus tôt dans ce mémoire. Nous remarquons cependant que, pour les femmes, cette tendance ascendante s'est maintenue depuis les années 50 alors que pour certains âges chez les hommes, nous observons une tendance mixte : descendante jusqu'à la moitié des années 70, ascendante par la suite.

Tableau 1 : Description statistique des données de longévité aux Pays-Bas

Pays-Bas	Femmes de 20 ans	Hommes de 20 ans
Max	62.35	58.42
Min	54.50	52.16
Moyenne	58.83	54.03
écart type	2.11	1.63

Pays-Bas	Femmes de 40 ans	Hommes de 40 ans
Max	42.70	38.97
Min	35.42	33.15
Moyenne	39.39	34.90
écart type	1.99	1.50

Pays-Bas	Femmes de 60 ans	Hommes de 60 ans
Max	24.28	20.79
Min	17.78	16.30
Moyenne	21.27	17.55
écart type	1.82	1.12

Pays-Bas	Femmes de 80 ans	Hommes de 80 ans
Max	8.65	6.91
Min	5.34	5.07
Moyenne	7.07	5.81
écart type	0.98	0.38

L'espérance de vie résiduelle moyenne, pour presque tous les âges, est plus élevée chez les femmes. La longévité féminine semble évoluer à un rythme plus élevé que celle des hommes. En revanche, les forts écarts types caractérisant les indices féminins par rapport aux indices masculins témoignent d'une évolution plus volatile chez les femmes.

Nous décidons, suite à ces constatations, d'écarter l'hypothèse d'une dynamique commune entre les évolutions respectives de la longévité des deux sexes aux pays bas. Nous appliquons notre approche aux données féminines.

Nous constatons en outre une divergence des dynamiques de progression des indices du *LifeMetrics* aux âges jeunes comparés aux âges plus élevés. L'efficacité d'une ACP est tributaire du degré de corrélation des variables analysées. Nous devons donc

nous contenter de l'un ou l'autre des groupes d'indices. Les fonds de pensions et compagnies d'assurance étant concernés principalement par les problèmes de prévisions à des âges plutôt avancés, nous décidons de retenir les séries d'espérance de vie résiduelle de la tranche 40 ans et plus.

4.2. Application de l'approche :

Rappelons que notre concept de modélisation de la longévité s'établit en deux étapes : Nous commençons par spécifier un modèle sous-jacent aux données en ayant recours à l'ACP. Nous identifions ensuite un modèle de projection en faisant usage des techniques d'analyse des séries temporelles.

L'identification du modèle sous-jacent nécessite une analyse préalable de la stationnarité des chroniques de longévité analysées. L'ACP se résumerait sinon en un seul axe représentant la tendance temporelle générale. Rappelons que, par définition, un processus n'est dit stationnaire que lorsque l'ensemble de ses moments d'ordre un et d'ordre deux sont indépendants du temps. Par opposition, un processus non stationnaire est un processus qui ne satisfait que l'une ou l'autre des deux conditions.

Retenons, pour exemple, la série d'espérance de vie résiduelle à 60 ans chez les femmes.

Figure 2 : Espérance de vie résiduelle chez les femmes de 60 ans aux Pays-Bas

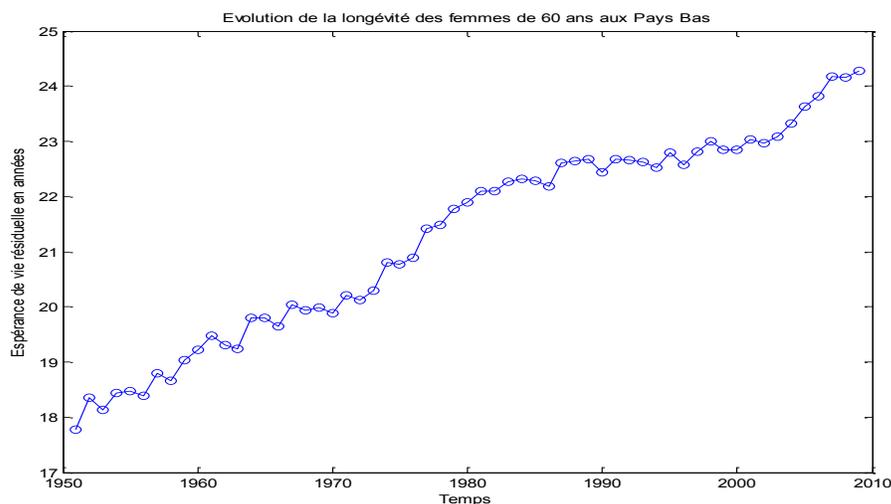
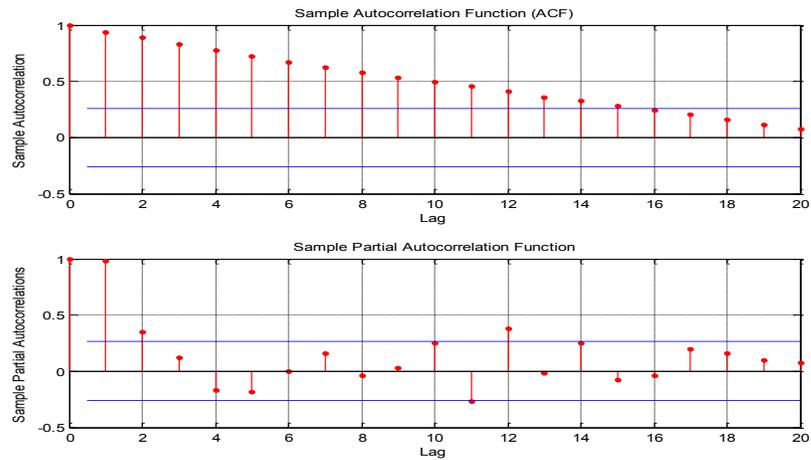


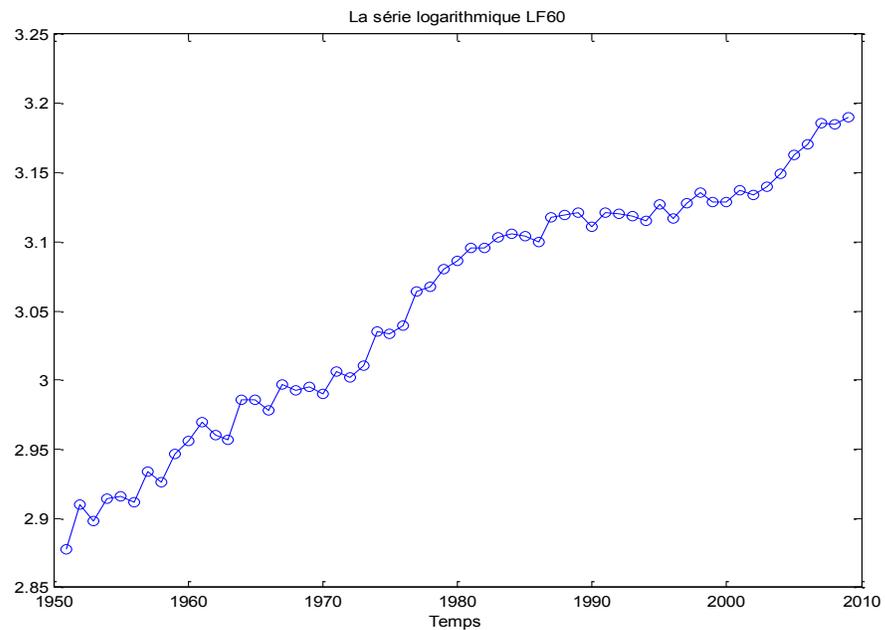
Figure 3 : Analyse de la fonction d'autocorrélation de l'indice de longévité féminine à 60 ans aux Pays-Bas



La série présente une non stationarité témoignée sur le premier graphe par une tendance exponentielle à la hausse et confirmée par une lente décroissance de la fonction d'autocorrélation.

Nous retenons la série logarithmique pour minimiser l'effet temps observé ci-dessus.

Figure 4 : Série logarithmique de la longévité féminine à 60 ans aux Pays-Bas



Cette dernière semble toujours non stationnaire. Un test Dickey Fuller, tenant compte de la tendance haussière observée, nous indique la présence d'une racine unitaire.

Nos séries sont donc stationnarisées en retenant les différences premières et ne comptent désormais que 58 observations.

4.2.1. Modèle sous-jacent aux données de longévité des « Femmes » aux Pays Bas :

Comme nous l'avons déjà expliqué, les caractéristiques statistiques des données de longévité nous incitent à écarter la modélisation indice par indice au profit d'une approche plus récapitulative. Nous exploitons des corrélations dépassant la barre de 0,8 entre les différentes variables analysées pour définir le modèle sous-jacent aux données du *LifeMetrics Index*.

Rappelons qu'une ACP classique ne tient compte que des corrélations générales inter-indices et que nous avons recours à la M-SSA pour permettre à notre modèle d'exploiter également les dépendances entre nos variables de longévité à différents retards.

Nous commençons d'ailleurs par présenter les résultats de l'ACP classique pour justifier le recours à la M-SSA.

Tableau 2 : ACP classique : Variances expliquées & variances cumulées

	% de variance expliquée	% cumulé de la variance expliquée
CP(1)	93,87	93,87
CP(2)	3,36	97,23
CP(3)	2,52	99,75
CP(4)	0,15	99,90
CP(5)	0,06	99,96

Les deux premières composantes principales expliquent plus que 99% de la variance des données de longévité parmi lesquels 94 % sont expliqués par la première composante seulement. Remarquons que ce résultat rejoint la principale conclusion

du modèle Lee & Carter où la seule composante temporelle retenue de la SVD résume 93 % de l'information analysée.

Nous aurions pu arrêter l'ACP à ce niveau, retenir la première composante et donc proposer un modèle unifactoriel pour résumer la surface de longévité féminine aux Pays Bas. Cependant, les caractéristiques statistiques de nos données, tel qu'expliqué plus tôt, nous encouragent à soupçonner que le potentiel informatif de notre base de données n'est ni convenablement exploité ni bien reflété par cette conclusion.

Nous réarrangeons nos indices de longévité selon la technique M-SSA. Nous appliquons l'ACP par la suite.

Nous procédons à la détermination du retard optimal pour la mise en place de la M-SSA : Il s'agit du même retard retenu si, éventuellement, nous aurions modélisé nos 9 variables selon un VAR (un modèle autorégressif vectoriel). Plusieurs critères de sélection peuvent être utilisés : Akaike, Hannan- Quinn, Schwarz, ... Nous retenons le critère de Schwarz, le retard optimal est celui qui en minimise la fonction. Nous calculons les valeurs associées aux retards respectifs de 1 à 5 ; le retard optimal est celui d'ordre 2.

Donc après recours à la M-SSA, chaque tranche d'âge (chaque indice de longévité à 58 observations) se trouve représentée par trois variables. Soient trois séries chronologiques correspondant respectivement à la série des 56 premières observations, la série regroupant les observations de 2 à 57 et la série des observations de 3 à 58. L'ACP que nous effectuons pour ressortir les dimensions globales de la longévité s'applique ainsi à une matrice de données de taille (27*56).

Tableau 3 : ACP après M-SSA : Variances expliquées & variances cumulées

	% de variance expliquée	% cumulé de la variance expliquée
CP(1)	45,97	45,97
CP(2)	39,05	85,02
CP(3)	11,00	96,03
CP(4)	1,66	97,67
CP(5)	1,31	98,99

Figure 5 : Nombre optimal de CP à retenir – critère de décision : « Scree-test »

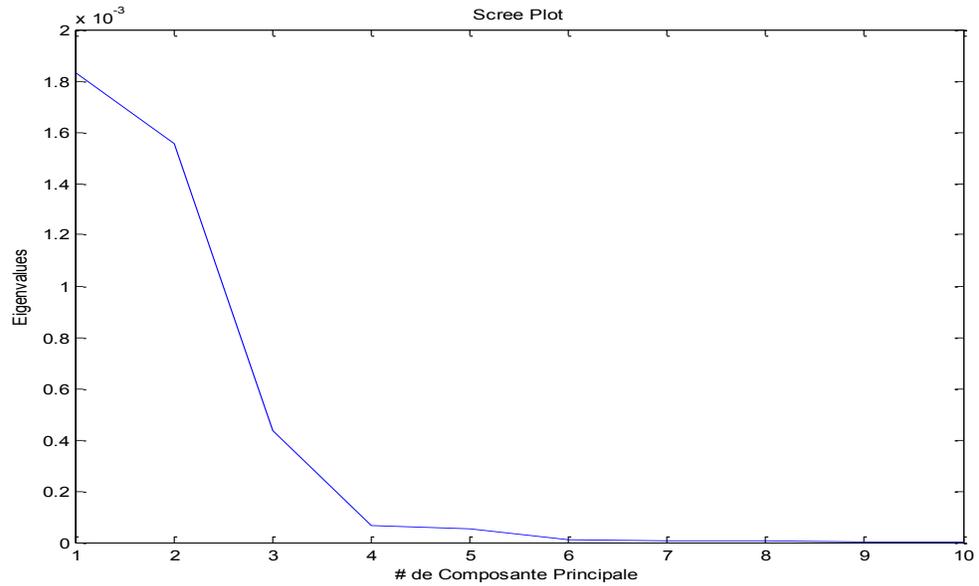
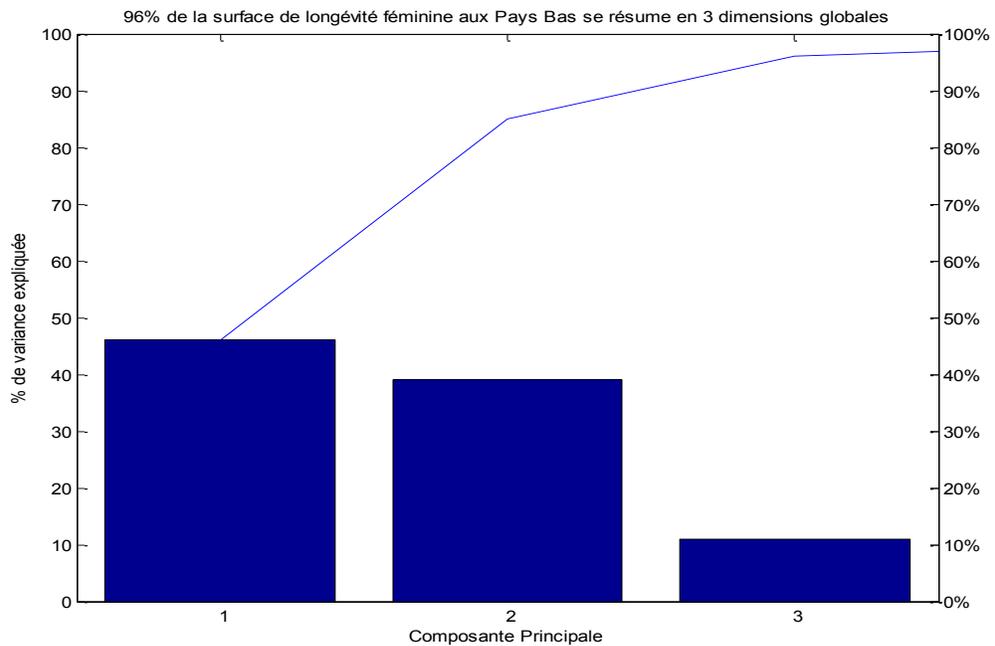


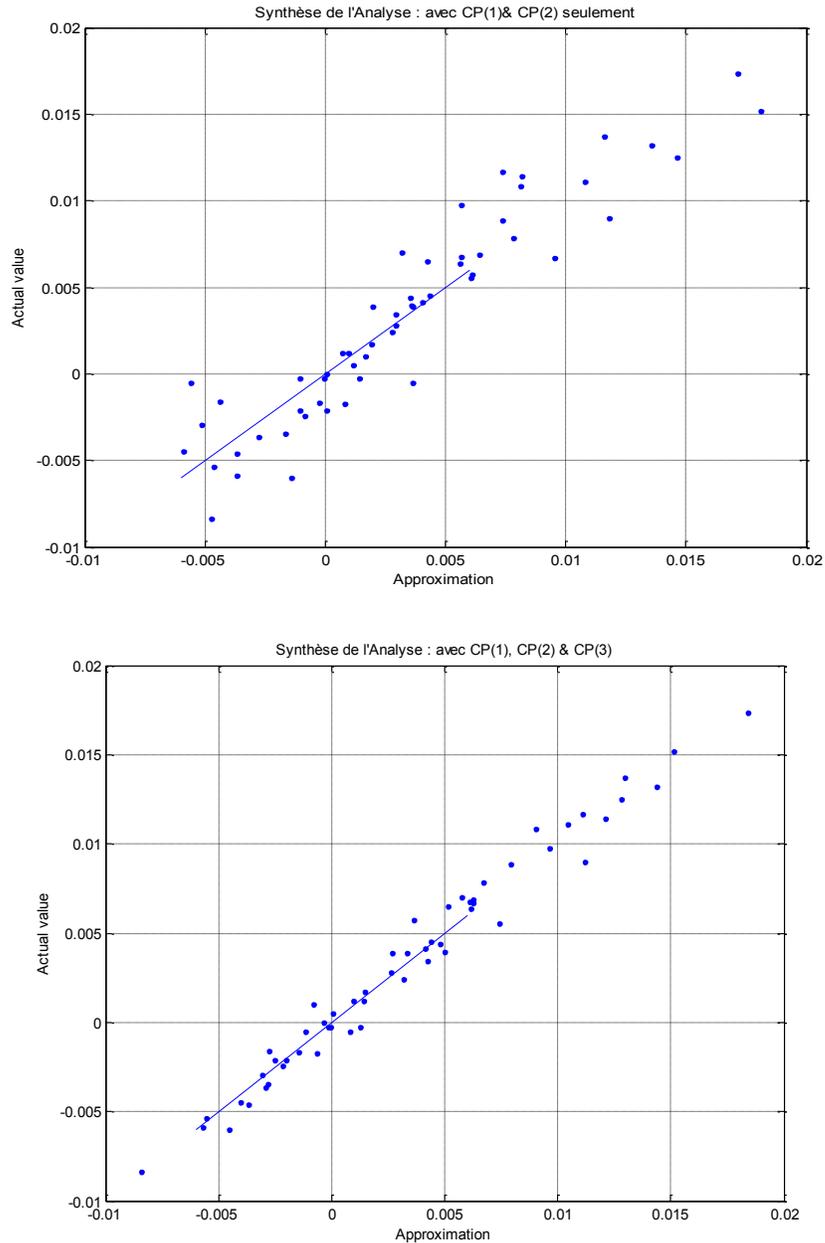
Figure 6 : Nombre optimal de CP à retenir – critère de décision : « % cumulé de la variance expliquée »

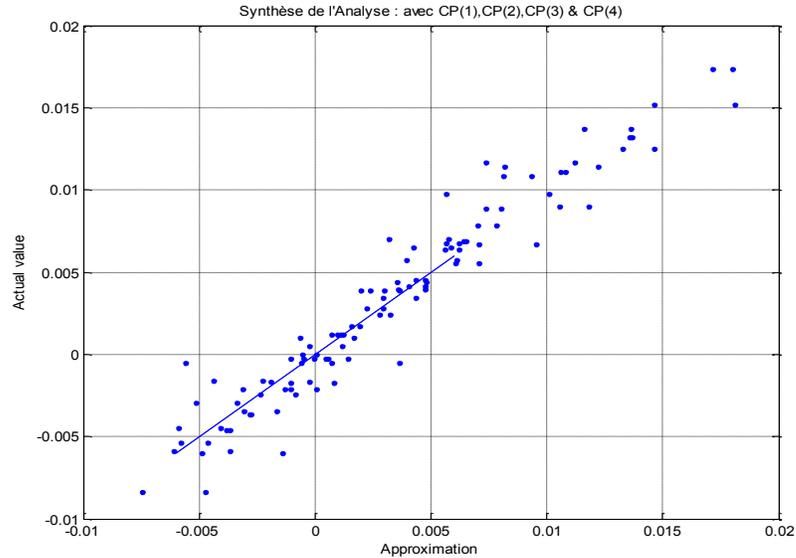


L'ACP après M-SSA résume la surface de longévité du *LifeMetrics Index* en trois composantes principales. La CP(1) restitue 46 % de la variance totale des données analysées contre 39 % et 11 % pour la CP(2) et la CP(3) respectivement donc un total de 96 % de la variance totale.

Une analyse graphique ainsi qu'une analyse de parcimonie nous permettrait de confirmer la conclusion ci-dessus, soit la validité d'un modèle sous-jacent à trois facteurs.

Figure 7 : Évaluation graphique de l'ACP après M-SSA





Retenir seulement les deux premiers axes résultant de l'ACP ne restitue qu'une partie limitée de l'information initialement analysée. En revanche, aller au-delà des trois premières composantes auxquelles les règles de décisions nous indiquent de se limiter ne contribue que marginalement à l'amélioration de la reconstitution du nuage des données.

Nous utilisons le critère d'information d'Akaike pour effectuer une analyse de parcimonie et confirmer la conclusion d'un modèle à trois facteurs.

$$AIC(k) = -N * \text{Log}(N) + N * \text{Log}(RSS) + 2 * k$$

Tableau 4 : Nombre optimal de CP à retenir – critère de décision : « Akaike »

CP retenues	k = 1	k = 2	k* = 3	k = 4	K = 5	K = 6
AIC	72.4405	68.6496	59.3122	59.9872	59.9550	59.0762

Le modèle retenant les 3 premières CP détient effectivement le meilleur rapport « rendement /coût de modélisation ».

Donc en notant x ($x = 40, 45, \dots, 80$) un indicateur de la tranche d'âge, nous pouvons écrire :

$$\text{Log}(e_x) = \text{Log}(E_x) + b_x^1 CP(1) + b_x^2 CP(2) + b_x^3 CP(3) \quad \text{pour tout } x=40, 45, \dots, 80$$

Où :

- E_a est égal à la moyenne des espérances de vie, à l'âge x pour les « T » années de l'indice de longévité.
- $CP(1)$, $CP(2)$ et $CP(3)$ sont les mêmes pour tous les âges,
- b_x^1 , b_x^2 et b_x^3 varient d'une tranche d'âge à l'autre et sont identifiés grâce aux *outputs* de l'ACP. Rappelons que chaque indice de longévité est représenté par trois variables dans la matrice analysée en composantes principales. Nous retenons les coefficients associés à la version la plus récente de chaque indice, soit l'échantillon contenant les observations de 3 à 58.

Tableau 5 : Paramètres de la relation établie entre la longévité et ses composantes principales

	CP(1)	CP(2)	CP(3)
LF (40)	0,0672	-0,1061	0,0406
LF (45)	0,0757	-0,1191	0,0450
LF (50)	0,0869	-0,1385	0,0485
LF (55)	0,1003	-0,1652	0,0583
LF (60)	0,1190	-0,1962	0,0616
LF (65)	0,1440	-0,2499	0,0752
LF (70)	0,1773	-0,3031	0,0912
LF (75)	0,2163	-0,4077	0,1089
LF (80)	0,2714	-0,5398	0,1491

La structure sous-jacente à la surface de longévité étant identifiée et résumée en trois dimensions globales, nous procédons maintenant à la définition du modèle de projection.

4.2.2. Modèle de projection :

Après avoir défini les dimensions globales (axes principaux) de la longévité féminine aux Pays Bas, nous procédons à l'estimation de leurs processus générateurs que nous extrapolons afin de prévoir l'évolution future de l'espérance de vie. Pour ce faire et tout comme Lee & Carter, nous avons recours à la méthode de Box & Jenkins. Cette dernière repose essentiellement sur une modélisation des données par un processus $ARMA(p, q)$ où la partie autorégressive notée AR est une combinaison linéaire finie en t des valeurs passées du processus et la partie moyenne mobile notée MA est une combinaison linéaire finie en t des valeurs passées d'un bruit blanc.

Chacune des composantes principales est donc identifiée par un modèle qui s'écrit comme suit :

$$X_t = \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$$

Où : les paramètres φ_i et θ_i sont constants et les termes d'erreurs ε_i sont indépendants du processus (ε_i sont *iid* selon une loi normale de moyenne nulle).

Nous modélisons respectivement 60%, 70% et 80% de l'échantillon des données et faisons appel à la statistique de U_THEIL pour déterminer la fenêtre d'identification optimale.

$$U = \sqrt{\frac{\sum (y_{t+h} - y_{t+h}^*)^2}{\sum (y_{t+h} - y_t)^2}}$$

Une plus faible valeur de U indique une meilleure performance des prévisions. Nous en calculons une moyenne pour chaque pourcentage retenu de l'échantillon des données. Les résultats écartent le choix 60 % mais ne permettent pas de trancher entre les deux autres intervalles. Cependant, la fenêtre d'identification à 80 % assure une meilleure performance des prévisions à long terme ce qui nous encourage à la retenir comme fenêtre optimale.

L'analyse retient donc 45 observations pour chaque variable respectivement. Les résultats se présentent comme suit :

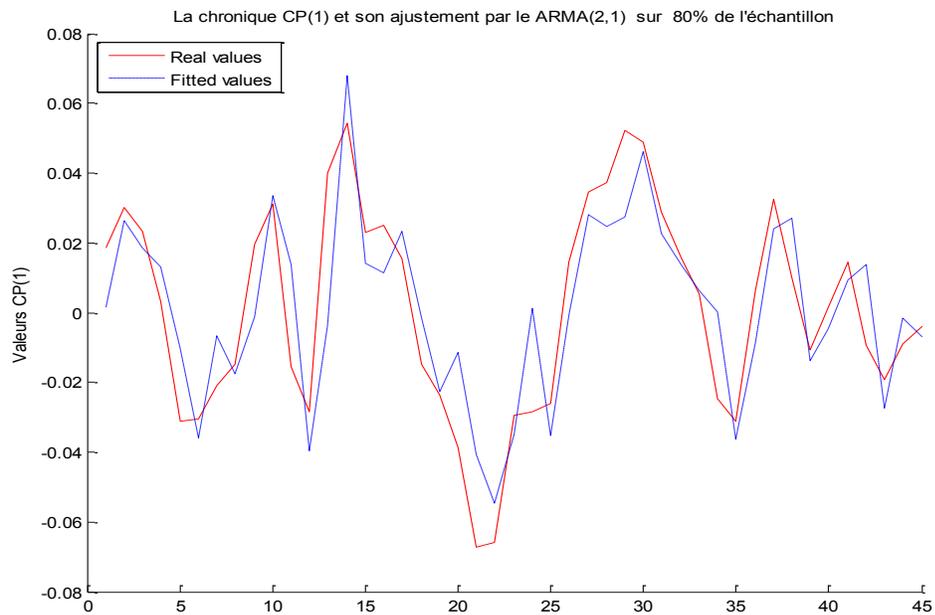
A- Analyse de la première composante principale :

Tableau 6 : Estimation du processus sous-jacent à la CP(1)

Mean: ARMAX(2,1,0); Variance: GARCH(0,0)			
Parameter	Value	Standard Error	T Statistic
C	7.883e-005	0.0044376	0.0178
AR(1)	-0.81699	0.25845	-3.1611
AR(2)	-0.46128	0.2106	-2.1903
MA(1)	-0.28094	0.26628	-1.0551
K	0.0010876	0.00019322	5.6287

Seule la partie autorégressive du ARMA semble être significative.

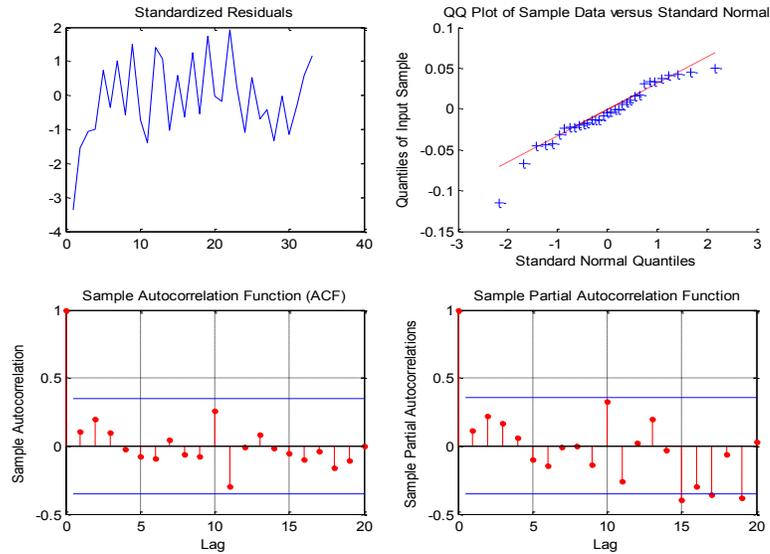
Figure 8 : Qualité d'ajustement de la CP(1) par un ARMA(2,1)



Le graphique montre qu'il n'y a pas une grande différence entre la chronique CP(1) et son ajustement par un ARMA(2,1).

Pour la validation du modèle, nous en analysons les résidus.

Figure 9 : Validation du modèle spécifié pour la CP(1) : Analyse des résidus du ARMA(2,1)



L'analyse de la fonction d'auto-corrélation (ACF) et des résidus standardisés nous permet de confirmer la stationnarité des innovations du modèle retenu.

Nous calculons également la statistique de Ljung-Box qui présente une P-value égale à 0.6274 et donc supérieure à 5% nous permettant de conclure que les résidus forment un bruit blanc. D'où la validation du modèle spécifié pour la CP(1).

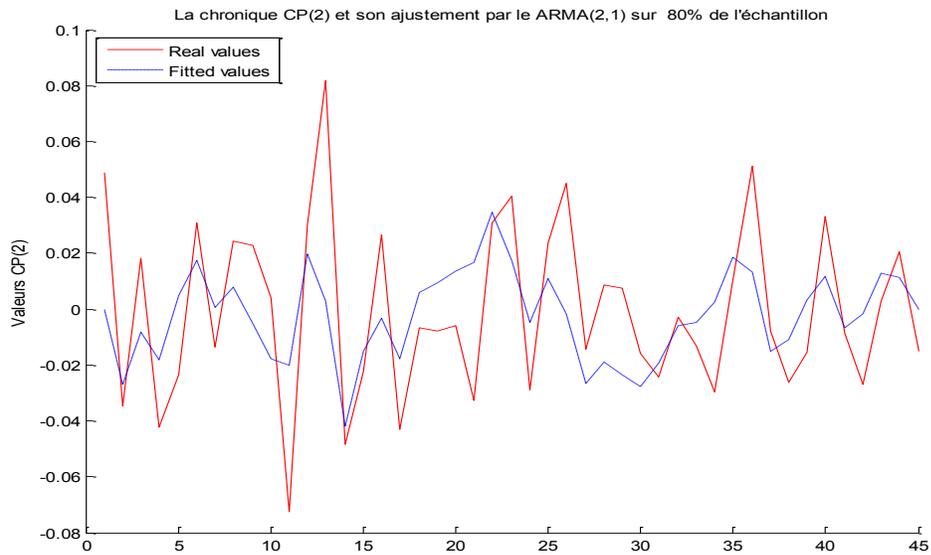
B- Analyse de la deuxième composante principale :

Tableau 7 : Estimation du processus sous-jacent à la CP(2)

Mean: ARMAX(2,1,0); Variance: GARCH(0,0)			
Parameter	Value	Standard Error	T Statistic
C	-0.0003212	0.0038184	-0.0841
AR(1)	-0.40698	0.19513	-2.0857
AR(2)	-0.65015	0.13187	-4.9301
MA(1)	-0.24963	0.24054	-1.0378
K	0.00094255	0.00021297	4.4256

Encore une fois et au seuil de 5%, seule la partie autorégressive du modèle spécifié semble être significative.

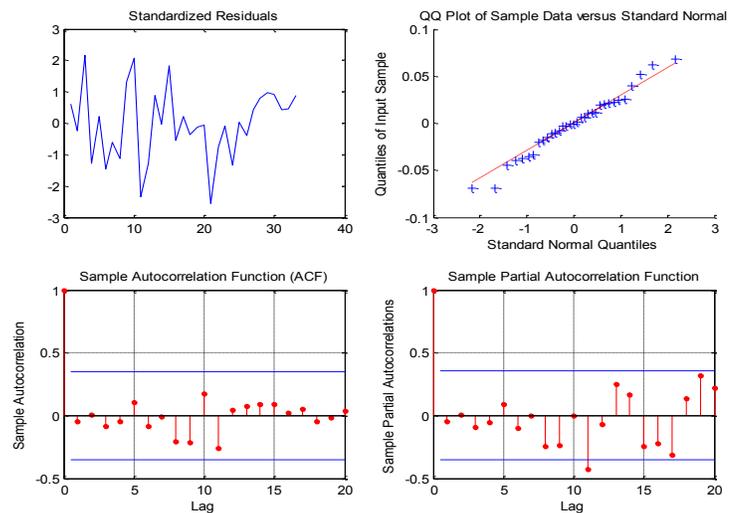
Figure 10 : Qualité d'ajustement de la CP(2) par un ARMA(2,1)



L'ajustement n'est pas aussi bon que celui de la CP(1), il est cependant acceptable.

Nous analysons les résidus du ARMA(2,1) pour en confirmer la validation comme modèle spécifié pour la CP(2).

Figure 11 : Validation du modèle spécifié pour la CP(2) : Analyse des résidus du ARMA(2,1)



L'analyse de la fonction d'auto-corrélation (ACF) et des résidus standardisés nous permet de confirmer la stationnarité des innovations du modèle.

Nous calculons également la statistique de Ljung-Box dont une P-value égale à 0.7952 (et donc supérieure à 5 %) nous permet d'affirmer que la série des résidus est effectivement un bruit blanc. D'où la validation du modèle.

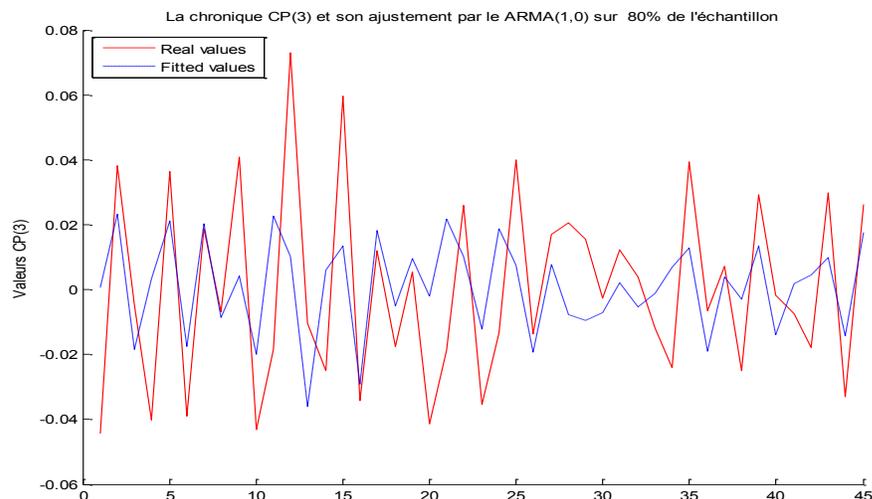
C- Analyse de la troisième composante principale :

Tableau 8 : Estimation du processus sous-jacent à la CP(3)

Mean: ARMAX(1,0,0); Variance: GARCH(0,0)			
Parameter	Value	Standard Error	T Statistic
C	8.7339e-005	0.0035996	0.0243
AR(1)	0.51519	0.17184	2.9981
K	0.00054599	0.00012786	4.2703

Les tstats indiquent que l'AR (1) identifié pour la CP (3) est effectivement significatif.

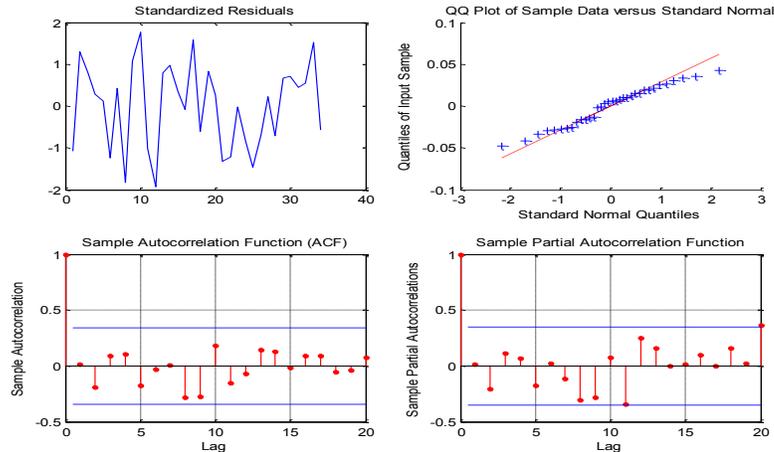
Figure 12 : Qualité d'ajustement de la CP(3) par un ARMA(1,0)



L'ajustement réussi à capter l'évolution générale de la CP(3).

Encore une fois, nous avons recours à l'étude des résidus en vue d'une éventuelle approbation du modèle.

Figure 13 : Validation du modèle spécifié pour la CP(3) : Analyse des résidus du ARMA(1,0)



Les graphes de la fonction d'auto-corrélation (ACF) et celui des résidus standardisés nous permettent d'affirmer la stationnarité des résidus. En outre, la statistique de Ljung-Box présente une P-value égale à 0.6337, et donc supérieure à 5%, nous permettant de conclure que les résidus forment un bruit blanc. D'où la validation du modèle retenu pour la CP(3).

Ayant défini la structure sous-jacente à la longévité féminine aux pays bas et ayant réussi à estimer les processus générateurs de ses dimensions globales (i.e., les processus générateurs des composantes principales), nous procédons à la prévision de l'espérance de vie résiduelle par tranche d'âge.

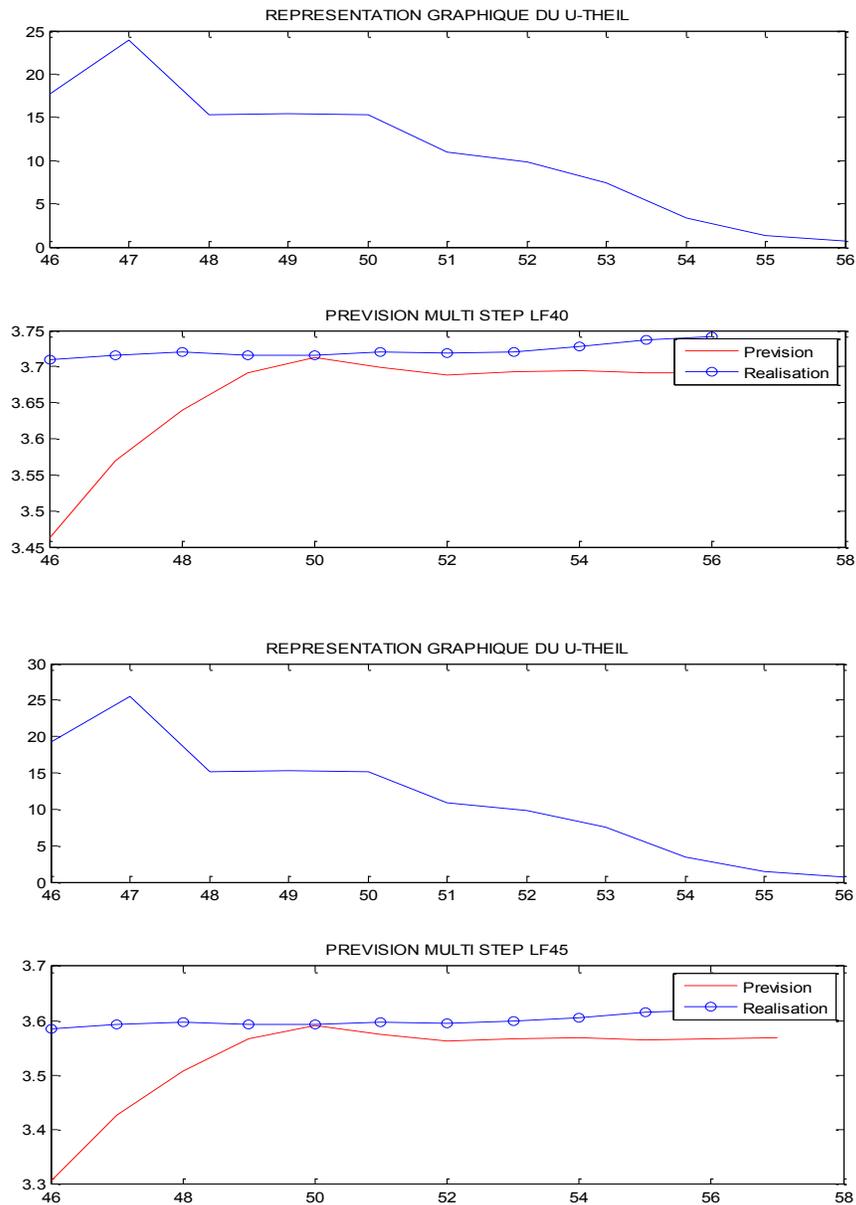
4.2.3. Prévision de l'espérance de vie résiduelle par tranche d'âge :

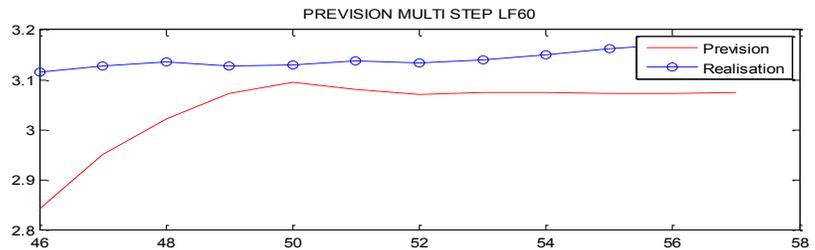
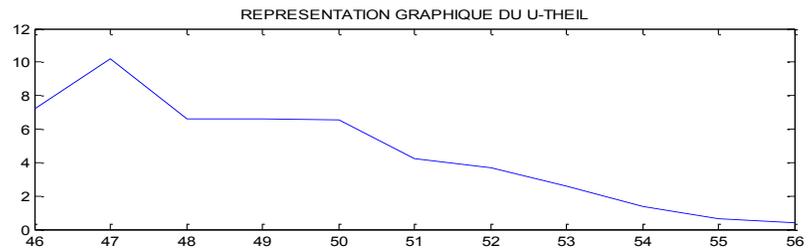
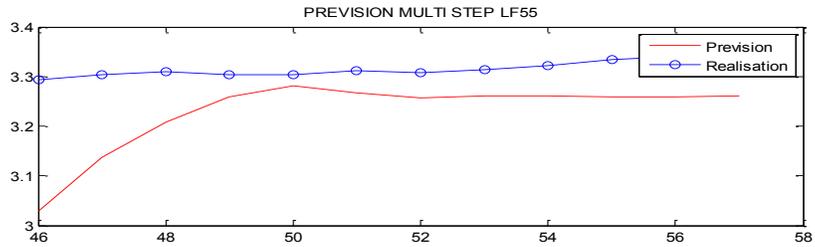
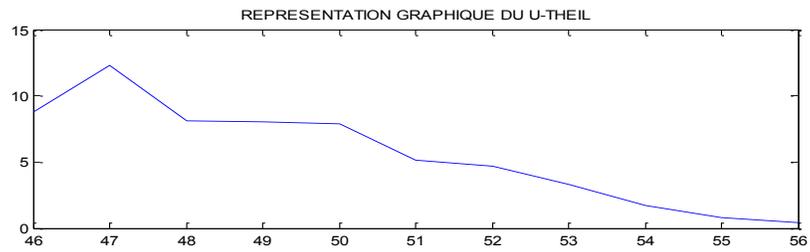
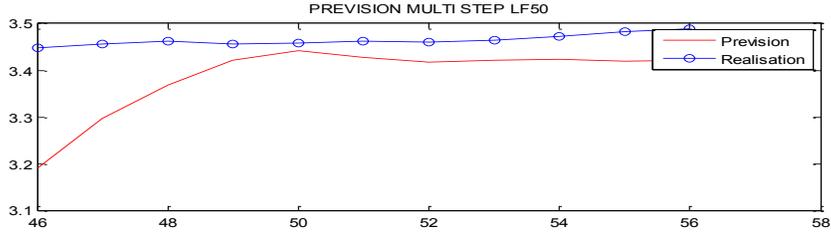
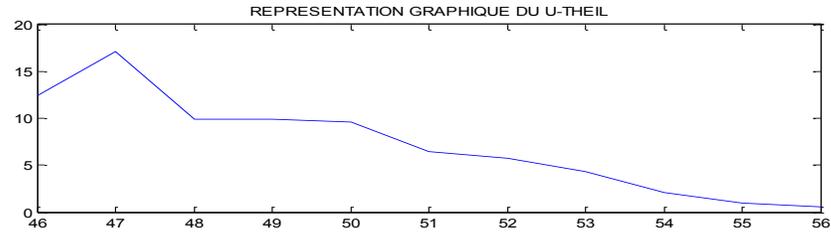
Nous joignons les résultats d'estimation des modèles ARMA aux résultats de l'ACP et procédons à la prévision de l'évolution de l'espérance de vie résiduelle par tranche d'âge.

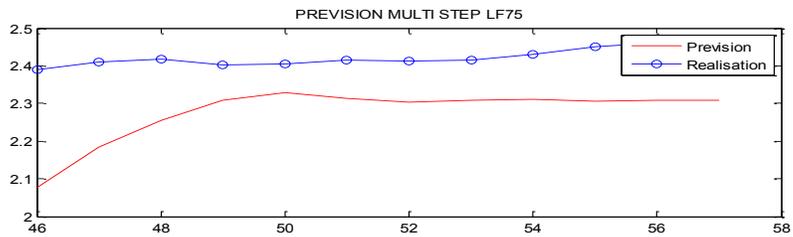
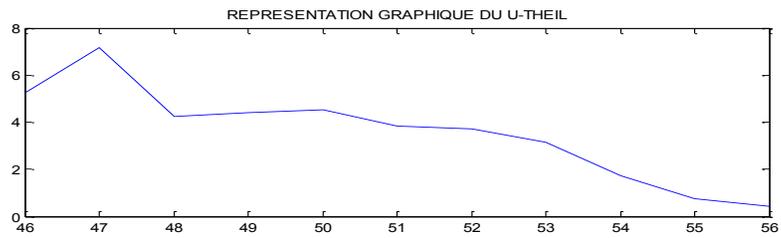
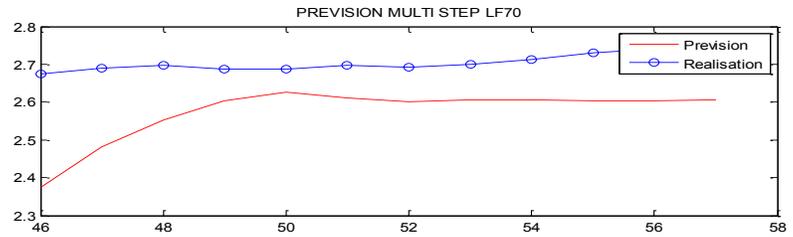
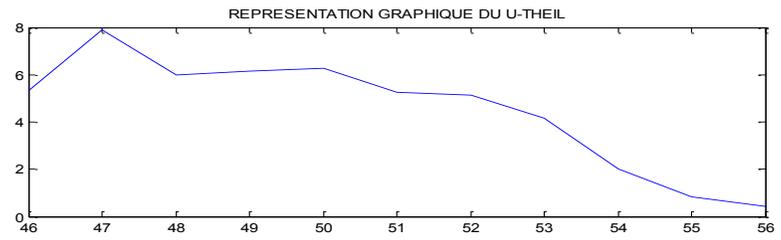
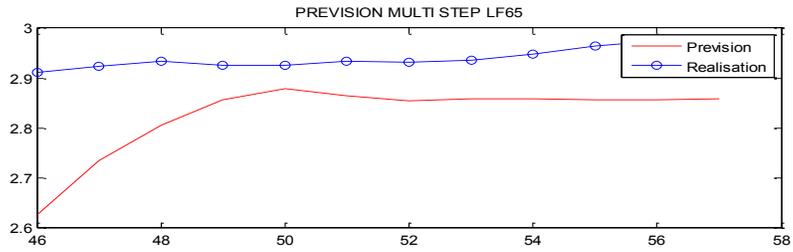
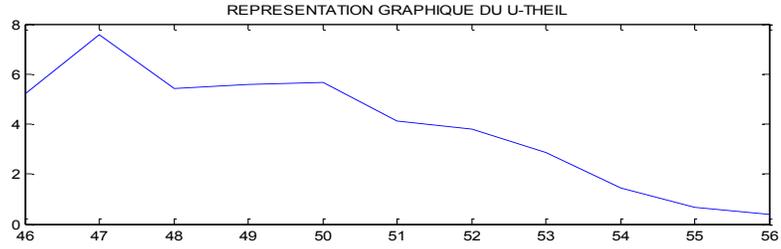
Rappelons que les deux étapes de notre modélisation nous permettent de relier les indices originaux de longévité aux composantes principales identifiées par la relation suivante : $Log(e_x) = Log(E_x) + b_x^1 CP(1) + b_x^2 CP(2) + b_x^3 CP(3)$

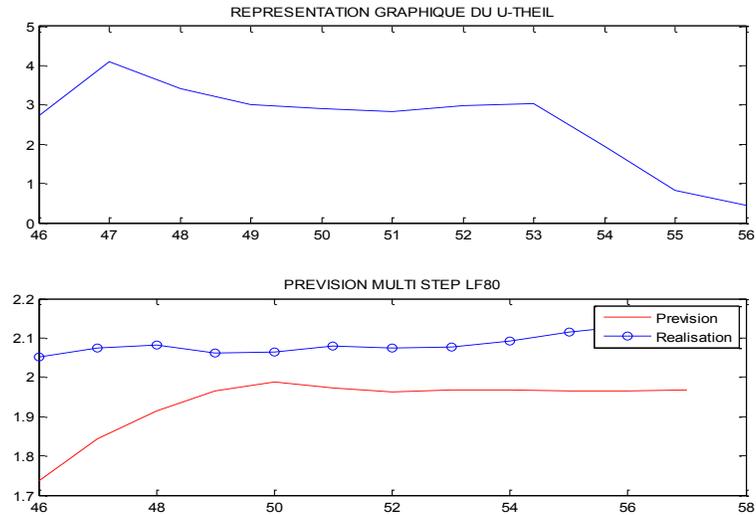
Nous fournissons deux outputs par tranche d'âge : la représentation graphique des coefficients U_THEIL et le graphe des prévisions versus les réalisations.

Figure 14 : Prévisions versus réalisations & qualité du modèle, par tranche d'âge de 40 à 80 ans









Globalement, notre approche performe mieux à long terme qu'à court terme. La courbe des coefficients U_THEIL qui démarre à une forte valeur, pour tous les âges, se dégrade à moyen terme et fini souvent par s'annuler ou presque. Notre modèle semble se stabiliser à long terme.

Connaissant les caractéristiques statistiques des indices de longévité, nous osons confirmer le « bon sens » de ce résultat. En effet, les séries d'espérance de vie peuvent avoir des évolutions divergentes à court terme mais finissent toujours par se rejoindre à long terme.

Autrement dit, les liens que notre approche réussit à résumer en dimensions globales (ou composantes principales) correspondraient à des relations implicites de long terme entre les indices de longévité ce qui expliquerait que notre modèle performe mieux à cet horizon-ci.

Remarquons également que nos prévisions de l'espérance de vie résiduelle convergent asymptotiquement vers la réalité sans jamais la dépasser. Ce résultat rejoint celui de l'approche Lee & Carter, ce qui nous rassure quant à la fiabilité de notre tentative de reproduction (quoi que légèrement modifiée) de ce modèle référence.

Au terme de ce chapitre, l'objectif d'une nouvelle approche de projection de la longévité est atteint. Le modèle factoriel extrapolatif que nous proposons constitue effectivement une alternative simple et contribuerait potentiellement à l'amélioration de la prévision de l'espérance de vie humaine.

Notre approche présente cependant le même inconvénient reproché à la plupart des modèles extrapolatifs. La méthode est purement objective et totalement déconnectée des avancées des approches non-extrapolatives. Nous essayons de remédier à cette faiblesse en proposant d'interpréter les dimensions globales de la longévité identifiées statistiquement plus tôt. L'idée est de procéder à un rapprochement entre nos composantes principales et les variables clés des approches causales et explicatives. Notre modèle factoriel extrapolatif pourrait ainsi être adapté et potentiellement modifié en un modèle structurel construit autour des principaux facteurs à risques et tiendrait compte des principales causes d'amélioration de l'espérance de vie humaine. Le modèle de projection remplacerait les dimensions globales de la longévité, identifiées grâce à l'ACP de son historique, par des « substituts explicatifs », identifiés grâce au rapprochement proposé.

Les résultats de cette stratégie d'interprétation sont présentés dans la section suivante.

Chapitre 5 : Identification des dimensions globales de la longévité :

Nous nous sommes limités, jusqu'à ce stade de l'étude, à une attitude plutôt objective lors de l'analyse de nos données de longévité. Si les variables originales du *LifeMetrics Index* ont une interprétation « native », les dimensions globales de la longévité que nous identifions à travers les composantes principales n'ont à priori qu'une définition mathématique. Dans ce mémoire, nous avons certes cherché à tirer profit des avantages associés à l'aspect objectif de l'approche extrapolative mais, tel qu'expliqué plus tôt, nous tenons à contourner les inconvénients qui lui sont associés.

Pour ce faire, nous faisons appel aux méthodes explicatives ainsi qu'aux méthodes causales de projection de la mortalité. L'idée est de faire un rapprochement entre les principales variables des approches subjectives et les axes principaux de la longévité identifiés par notre méthode. Une analyse des corrélations entre les variables de part et d'autre permettrait de réaliser le rapprochement en question. L'idée est de permettre à notre modèle de prendre en compte l'avis d'expert sur l'évolution présumée de la longévité et d'être sensible aux impacts potentiels des avancées médicales et des changements des styles de vie. Remarquons que le recours aux approches explicatives permet d'identifier des facteurs à risque, des variables tels que des facteurs économiques ou environnementaux susceptibles d'affecter l'évolution de la longévité humaine. Les approches causales, quant à elles, détectent les facteurs à l'origine des décès et permettent de modéliser la longévité d'un point de vue biomédical.

Nous entamons cette tentative d'identification des composantes principales en s'assurant que les orientations générales des approches à confronter sont effectivement comparables. Les résultats des comparaisons préliminaires des approches respectives se trouvent en effet favorables à la stratégie d'identification proposée. A priori, notre décision de limiter l'analyse aux âges les plus élevés est soutenue. Les approches explicative et causale confirment qu'aux Pays-Bas comme dans la plupart des pays développés, l'augmentation actuellement en cours de l'espérance de vie est principalement due à la diminution de la mortalité chez les personnes âgées et très âgées.

En outre, nous apprenons que l'évolution commune de la longévité en Europe a fait place depuis les années 60 à des divergences donnant naissance à trois groupes de pays avec des évolutions distinctes de l'espérance de vie humaine. Le premier est «à forte convergence» incluant la Suisse et la France et caractérisé par une augmentation incessante de la longévité. Le deuxième est le groupe «à faible convergence» incluant le Royaume-Uni et la Belgique et caractérisé par un ralentissement durant les années 60. Le troisième groupe, dit «divergent», compte le Danemark, la Norvège et les Pays-Bas dans ses rangs et se démarque par des périodes de stagnation, voire même des diminutions passagères de l'espérance de vie. L'impact de ces évolutions disparates continue à se ressentir aujourd'hui et se traduit par des écarts d'amélioration la longévité frôlant les douze ans chez les hommes et sept ans chez les femmes lorsqu'on compare certains pays de l'Est de l'Europe à la France ou à la Suisse. Les études consultées confirment la persistance de ces divergences et des causes qui les ont produites, du moins pour un horizon de court à moyen terme.

Nous faisons donc l'hypothèse que la compréhension des facteurs, qui auraient causé/amorcé la divergence aux Pays-Bas, serait une piste efficace pour identifier les composantes principales de son histoire de longévité. Une consultation de la littérature disponible nous apprend que pour le cas de l'Europe, les raisons des divergences en question sont encore mal connues. Elles se rapporteraient probablement à des phénomènes/crises de l'environnement social, économique, environnemental, ... éventuellement similaires dans les pays du même groupe mais peu d'études ont cherché à le prouver. Cependant, nous apprenons deux faits intéressants en ce qui concerne les Pays-Bas. Le premier c'est qu'une bonne part de l'irrégularité est attribuable à l'évolution du phénomène de tabagisme. Le deuxième est que l'évolution statistique de la longévité aux Pays-Bas est assez similaire à celle aux États-Unis. A l'échelle mondiale, les deux pays appartiennent au même groupe, le groupe dit «divergent» tel que décrit ci-haut. A l'encontre des Pays-Bas, la divergence de l'évolution de l'espérance de vie humaine est largement débattue aux États-Unis. Le sujet est couvert par une littérature abondante.

Une des études les plus importantes est celle réalisée par le Conseil National de Recherche «*The National Research Council*» sous mandat de l'Institut National sur le vieillissement «*The National Institute on Aging*». Cette dernière, tout comme la littérature «européenne», met en évidence le fait que les Pays-Bas et le Danemark se distinguent comme les deux SEULS autres pays (à revenu élevé et considérés comme d'avant-garde) qui ont récemment des résultats insuffisants par rapport à l'amélioration de l'espérance de vie. On y examine un certain nombre de facteurs de risque dont l'évolution pourrait expliquer l'histoire de la longévité aux États-Unis et aux pays à tendances comparables, notamment les Pays-Bas et le Danemark. L'étude a permis «d'établir la force de l'évidence» pour un certain nombre d'explications fréquemment utilisées telles que les inefficacités relatives aux systèmes de santé et les comportements malsains de la population notamment concernant le tabagisme, la suralimentation ou encore l'insuffisance d'exercice physique.

On affirme donc que le tabagisme est responsable d'une bonne partie de l'évolution divergente de la longévité aux USA, aux Pays-Bas ainsi qu'au Danemark. Selon «*The National Research Council*», ce phénomène constitue un candidat naturel sur la liste des facteurs de risque parce qu'il représente une cause majeure de mortalité. Un indicateur efficace permettant de mesurer/retracer l'évolution du «tabagisme» (autre que le nombre des fumeurs dans le pays s'il est répertorié) serait donc le nombre de décès liés à ce phénomène : Il s'agit des décès ayant pour cause le cancer du poumon⁶. Nous pensons que le facteur «tabagisme» aurait une forte corrélation avec la première composante principale de notre modèle.

«La performance du système de santé» est un autre facteur déterminant de l'histoire de la longévité aux Pays-Bas. Le meilleur indicateur que nous pouvons retenir pour mesurer la performance en question est «le pourcentage annuel du PIB imputables

⁶ Le cancer du poumon est la première cause dans la catégorie des décès associés au tabagisme. Plusieurs autres maladies figurent sur la liste en question cependant : les divers autres types de cancer, de maladies coronariennes, de maladies pulmonaires obstructives chroniques (MPOC) ainsi que les maladies cérébro-vasculaires. Pour des raisons de simplification, nous nous limitons aux décès causés par le cancer du poumon. Source : "Centers for Disease Control and Prevention, 2004".

aux dépenses de santé». Ce facteur pourrait avoir une liaison significative avec la deuxième composante principale de notre modèle sinon la troisième.

L'évolution des maladies cardiaques et l'évolution du nombre des diabétiques, comme indicateurs potentiels de l'évolution des conditions de santé, sont deux autres facteurs de risque cités par l'étude américaine. Nous analysons les corrélations respectives qu'ils pourraient avoir avec la deuxième et troisième composante principale.

L'obésité et la surcharge pondérale aux Pays-Bas, tout comme aux États-Unis ainsi que dans la grande majorité des pays développés, est un déterminant potentiel de l'évolution de l'espérance de vie. Deux indicateurs pourraient être retenus pour mesurer l'effet de cette variable : le pourcentage d'obésité et son évolution dans le temps sinon les dépenses annuelles de santé imputables au surpoids (les chiffres annuels en valeur ou en pourcentage)⁷. Ce facteur pourrait avoir une liaison significative avec la deuxième composante principale de notre modèle sinon la troisième.

Pour concrétiser le rapprochement entre les variables réelles suggérées par les études consultées et les composantes principales détectées par notre modèle, nous devons disposer des données et des séries chronologiques appropriées. Pour ce faire, nous avons recours à des bases de données telles que celle de l'organisation mondiale de la santé « *The WHO Database* », « *Eurostat* » et « *Statistics Netherlands* ». Remarquons que lorsqu'à cause de l'indisponibilité des données, il nous est impossible de retracer un substitut explicatif tel que suggéré ci-haut, nous nous contentons d'une approximation. Par exemple, au lieu d'avoir la série des « décès ayant pour cause le cancer de poumon » (comme indicateur associé au phénomène du tabagisme), nous retenons les seules informations disponibles : « les décès ayant pour cause le cancer (tout type) » et « les décès ayant pour cause la pneumonie ».

⁷ Rapport OMS, Obésité : Prévention et prise en charge de l'épidémie mondiale

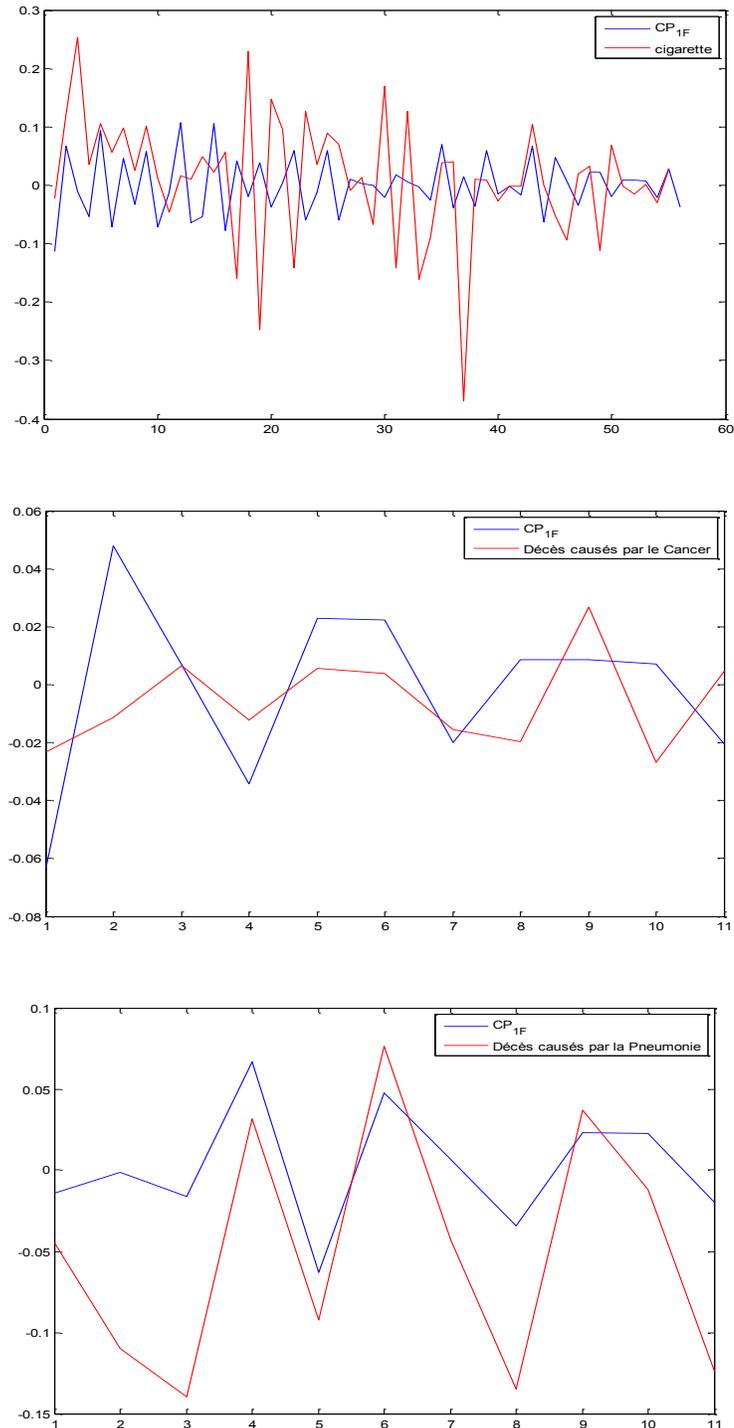
Au final, les variables que nous retenons pour le rapprochement avec les composantes principales de notre modèle sont les suivantes :

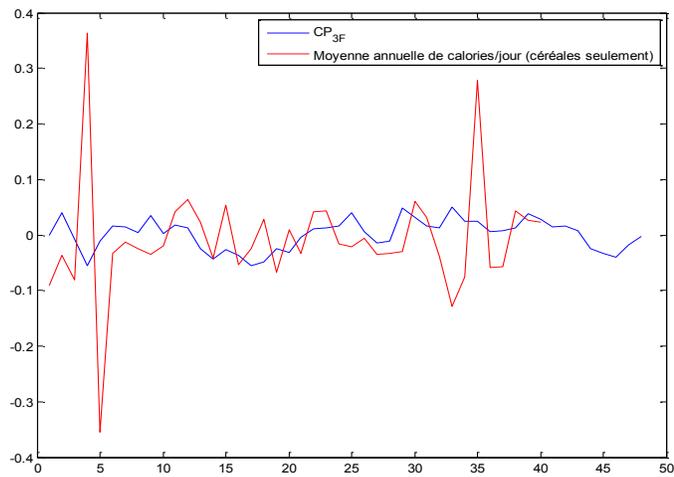
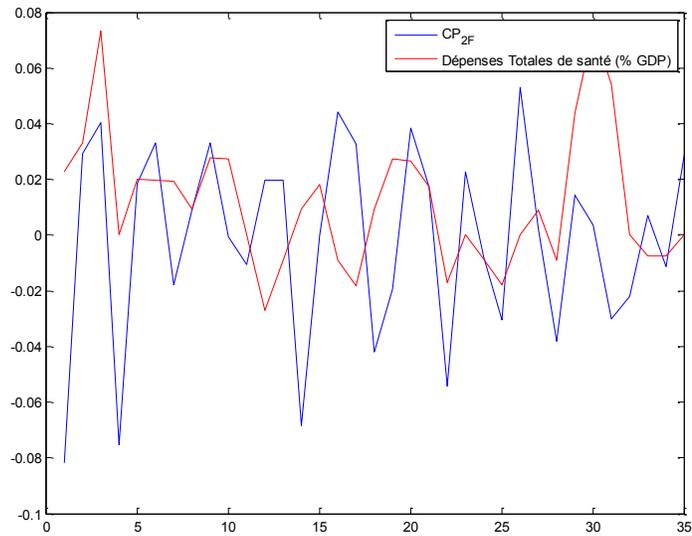
- Consommation moyenne du tabac au Pays-Bas :
 - Nombre de cigarettes (pièces) / année : 1946 – 2009, *Statistics Netherlands*,
 - Quantité de tabac (kg) / année : 1946 – 2009, *Statistics Netherlands*.
- Nombre de décès pour cause de :
 - Cancer : 1999 – 2010, *WHO database*,
 - Pneumonie : 1999 – 2010, *WHO database*.
- Dépenses de santé :
 - Dépenses totales de santé / habitant : 1995 – 2011, *WHO database*,
 - Dépenses gouvernementales de santé / habitant: 1995 – 2011, *WHO database*,
 - Dépenses totales de santé (% GDP) : 1972 – 2012, *Statistics Netherlands*.
- Conditions générales de santé :
 - Pourcentage de la population atteint de diabète : 2001 – 2011, *Statistics Netherlands*,
 - Nombre de décès pour cause d'Infarctus Du Myocarde (IDM) : 1994 – 2010, *WHO database*.
- Obésité :
 - IMC (≥ 25.0) par sexe : 2001 – 2011, *OECD Library*,
 - Consommation de céréales (calories / jour) : 1961 – 2012, *WHO database*.

Nous analysons les corrélations de chacune des composantes principales avec ses « substituts explicatifs » potentiels tel que suggéré ci-dessus. Nous calculons donc des coefficients de corrélations entre chaque CP et la/les variables (s) explicative(s) qui lui est (sont) associée(s).

Remarquons que la représentation graphique des couples de variables « CP vs variable explicative retenue » sur un même plan, nous confirme souvent l'hypothèse de la dépendance/corrélation. En voici quelques exemples :

Figure 15 : Corrélations entre les CP et les « substituts explicatifs potentiels »





Les résultats de l'analyse de corrélation « par mesure de coefficient de corrélation » se présentent comme suit :

Tableau 9 : « Substituts explicatifs potentiels » de la CP(1)

<i>CP_1_Femmes</i>	Coefficient de corrélation	P-value (significativité)
Nombre de cigarettes (pièces) / fumeur / année	-0.1839	0.1683
Quantité de tabac (Kg) / fumeur / année	-0.0723	0.7397
Nombre de décès causés par le cancer (Femmes) / année	-0.3216	0.3243
Nombre de décès causés par la pneumonie (Femmes) / année	-0.8205	0.0024

Tableau 10 : « Substituts explicatifs potentiels » de la CP(2)

<i>CP_2_Femmes</i>	Coefficient de corrélation	P-value (significativité)
Dépenses totales de santé / habitant	-0.3157	0.3423
Dépenses gouvernementales de santé / habitant	-0.3098	0.1862
Dépenses totales de santé (% GDP)	-0.2971	0.0850
% de la population atteint du diabète	-0.4310	0.1743
Décès – IDM	-0.2127	0.6219
IMC (>=25.0)	0.3145	0.3872
Moyenne annuelle de calories / jour (céréales seulement)	-0.1952	0.3309

Tableau 11 : « Substituts explicatifs potentiels » de la CP(3)

<i>CP_3_Femmes</i>	Coefficient de corrélation	P-value (significativité)
Dépenses totales de santé / habitant	0.1998	0.3598
Dépenses gouvernementales de santé / habitant	0.1052	0.6832
Dépenses totales de santé (% GDP)	0.0284	0.8312
% de la population atteint du diabète	-0.7332	0.1127
Décès – IDM	-0.4357	0.1709
IMC (>=25.0)	-0.2734	0.3143
Moyenne annuelle de calories / jour (céréales seulement)	0.1917	0.1923

Trois aspects sont à analyser pour chacun des coefficients calculés ci-haut :

- Son signe (positif / négatif) et sa correspondance à nos attentes
- L'intensité de la liaison qu'il reflète
- Sa significativité

En général, le signe est le seul aspect dont les résultats semblent satisfaisants. En effet, en reliant les résultats de l'ACP (l'output *factor loadings* notamment) aux interprétations suggérées par les méthodes explicatives et causales, nous avons établi des attentes concernant les signes des corrélations éventuelles entre toute CP et son « substitut » potentiel.

Par exemple, la CP(1) est liée « positivement » à l'espérance vie : les coefficients reliant cette composante aux différentes tranches d'âge sont tous positifs⁸ indiquant que, pour tout âge x , $Log(e_x)$ est proportionnel au premier axe identifié. En revanche, l'interprétation potentielle de la CP(1) que nous retenons suppose que cette dernière est fortement reliée à des variables mesurant l'évolution du phénomène de tabagisme. Nous nous attendons donc à une liaison de signe négatif entre la première dimension de longévité et chacune des variables-substituts qui lui est associée.

La plupart de nos attentes en termes de « signe de coefficient de corrélation » sont validées.

En revanche, nos attentes en termes de « valeurs » et de « significativité » sont loin d'être satisfaites : l'analyse se traduit par des faibles coefficients de corrélation, souvent peu significatifs.

Nous sommes conscients que ces résultats pourraient provenir d'une mauvaise qualité d'interprétation : les méthodes explicatives et causales de la longévité retenues pour l'identification de nos CP ne sont pas les mieux réputées pour ce qui de l'efficacité. Le problème pourrait aussi résulter d'une mauvaise concordance entre l'interprétation théorique adoptée et la variable explicative qui lui est associée (faute de disponibilité des données). Par exemple, l'interprétation retenue pour la CP(1) stipule que cette dernière devrait représenter l'évolution du tabagisme aux Pays-Bas. Autrement dit, la variable qui aurait le plus de chance de correspondre à la CP(1) serait « le nombre des fumeurs / année aux Pays-Bas » (données désagrégées par sexe, idéalement). En revanche, faute de disponibilité de cette information, nous retenons des variables telles que « le nombre de cigarettes / fumeur / année » ou « la quantité de tabac / fumeur / année ». Ces séries seraient plus adaptées à l'analyse de l'évolution des habitudes de consommation des fumeurs aux Pays Bas plutôt qu'à l'analyse de l'évolution de l'ampleur du phénomène-tabagisme, d'où les faibles corrélations.

⁸ Voir tableau 5 ci-haut

Il nous était difficile de détourner cet obstacle de non disponibilité des données. Nous n'étions pas souvent en mesure de retenir la variable qui correspond parfaitement à l'interprétation théorique de chacune des CP. Cependant, nous reconnaissons que cette discordance n'est probablement pas la vraie raison derrière les résultats surprenants de notre analyse des corrélations (nous nous attendions à des coefficients de corrélation beaucoup plus élevés et plus significatifs suite à la visualisation des couples « CP – variable explicative associée » sur un même plan, respectivement).

En effet, notre objectif est d'étudier l'intensité de la liaison qui existerait éventuellement entre chaque CP et le « substitut-explicatif » qui lui est associé. Mesurer un coefficient de corrélation pour évaluer cette liaison suppose que l'on essaye de juger l'existence d'une relation linéaire entre les deux variables : le coefficient de corrélation indique uniquement une dépendance linéaire.

La discordance entre les résultats de l'analyse de corrélation et les représentations graphiques pourrait être due au fait que les liaisons en question correspondent plutôt à des corrélations exponentielles ou des relations sous forme de puissance. Le coefficient de corrélation n'est pas adapté pour les évaluer, le cas échéant.

Remarquons enfin qu'un bilan comparatif intégrant le signe, l'intensité et la significativité des coefficients respectifs privilégierait les « dépenses de santé aux Pays-Bas » comme « substitut-explicatif » pour la CP (2). La CP (3) serait plutôt associée aux « conditions générales de santé aux Pays-Bas ».

Chapitre 6 : Conclusion :

Nous avons consacré ce travail à l'étude du risque de longévité dans l'objectif ultime d'en améliorer la couverture. A cette fin, la première partie de ce mémoire consistait en un état du phénomène et des principales tentatives et techniques de *hedging* qui lui sont associées. Nous apprenons que l'ampleur de ce risque et l'urgence d'un marché transparent et liquide de longévité exigent un perfectionnement continu de sa modélisation : L'augmentation rapide et inattendue de l'espérance de vie et le dépassement des limites d'âges précédemment observées ont effectivement permis à la prévision de la longévité de gagner en importance au cours des dernières décennies.

Nous avons consulté les différents types des méthodes d'ajustement et de projection existantes. Nous avons observé brièvement qu'il existait un large choix de modèles qu'il convient de sélectionner soigneusement en fonction notamment de l'objectif des prévisions et de la disponibilité des données. Nous avons eu l'idée de proposer une méthode qui combine les avantages des différentes approches. Nous avons enfin abouti à un modèle factoriel extrapolatif représentatif des dimensions globales de la longévité : les dimensions mathématiques dans un premier temps et les dimensions réelles ensuite.

L'application de notre méthodologie aux données des Pays-Bas se conclue par une surface de longévité résumée en trois axes générateurs chez les femmes. Le modèle performe mieux avec les prévisions de long terme confirmant l'hypothèse stipulant que les relations identifiées grâce aux modèles factoriels de longévité sont souvent des relations de long terme.

Notre méthodologie est une reproduction, légèrement modifiée, de celle du modèle standard du Lee & carter notamment pour ce qui est de l'analyse économétriques des dimensions de longévité. Cependant, la structure du modèle sous-jacent spécifié par notre approche ressemble davantage à celle du modèle de Renshaw & Haberman (2003). Rappelons que ces auteurs remettent en question leurs propre choix d'une

modélisation distincte et univariée pour chacune des composantes temporelles de leur modèle et recommandent plutôt une analyse multivariée des dimensions globales de longévité. Nous adoptons cette même recommandation et proposons d'en faire une extension pour ce mémoire. La performance des prévisions générées par une analyse multivariée des composantes principales y serait comparée à la performance de notre modèle.

Comme la plupart des modèles standards, incluant ceux évoqués dans ce mémoire, notre approche s'applique à l'ensemble des données et ne retient que la distinction par sexe et âge. Des facteurs supplémentaires de la longévité pourraient cependant être pris en considération, l'effet de la génération et la distinction entre la population des assurés et autres. D'ailleurs, appliquer la méthode au contexte moins général de la population assurée et des portefeuilles d'assurance vie pourrait constituer un exercice intéressant à réaliser. Non seulement, il est raisonnable de penser que le modèle pourrait mieux performer dans le cadre exclusif de la population assurée (étant donné la plus forte homogénéité des données). Mais les prévisions des taux de survie qui pourraient être générées dans ce contexte seraient définitivement mieux adaptées au développement des outils de couverture du risque de longévité.

Au final, nous rappelons que l'approche de modélisation de longévité proposée par ce mémoire constitue une base fondamentale pour une modélisation stochastique des engagements d'assurance, seule technique capable de fournir une « *fair value* » du passif; cette problématique pourrait également faire l'objet d'un complément utile à ce travail.

Bibliographie

Articles :

- ❖ Blake D., Cairns A-J., Dowd K. & MacMinn R., 2006, *Longevity Bonds : Financial Engineering, Valuation and hedging*, The journal of Risk and Insurance n°73.
- ❖ Bongaarts J., 2004, *Long-Range Trends in Adult Mortality : Models and Projection Methods*, Population Council WP 192.
- ❖ Cairns A-J., Blake D., Dowd K., 2006. *A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty : Theory and Calibration*, Journal of Risk and Insurance.
- ❖ CMIB : Continuous Mortality Investigation Bureau, 2005, The Faculty of Actuaries and the Institute of Actuaries, *Working paper No.15*.
- ❖ Hassami H. & Mahmoudvand R., 2013, *Multivariate Singular Spectrum Analysis: A General View And New Vector Forecasting Approach*, International Journal of Energy and Statistics.
- ❖ Lee R-D. & Carter L., 1992, *Modeling and Forecasting U.S Mortality*, Journal of the American Statistical Association n°87.
- ❖ Lee R-D. & Nault F., 1993, *Modeling and Forecasting Provincial Mortality in Canada*, Papier présenté lors du “World Congress of the IUSSP, Montreal, Canada”.
- ❖ Lee R. & Miller T, 2001, *Evaluating the performance of the Lee-Carter method for forecasting mortality*, Demography Review n°38.
- ❖ Planchet F., Lelieur V., 200, *Utilisation des méthodes de Lee-Carter et Log-Poisson pour l'ajustement de tables de mortalité dans le cas de petits échantillons*, Institut des Actuaire, Bulletin Français d'Actuariat Bulletin n°14.
- ❖ Renshaw A-E. & Haberman S., 2003, *Lee-Carter mortality forecasting with age-specific enhancement*, disponible en ligne: <http://www.sciencedirect.com/science/article/pii/S0167668703001380>
- ❖ White K-M., 2002, *Longevity Advances in High-Income Countries*, Population and Development Review n°28

Livres :

- ❖ Girosi, F. & G. King, 2007, *Demographic Forecasting*, Draft book manuscript, Disponible en ligne : <http://gking.harvard.edu/files/smooth.pdf>
- ❖ Jeroen, V.K, 2010, *Econométrie des Séries Temporelles, notes du cours 6-837-07, Département d'Économie Appliquée*, Ecoles des Hautes Etudes Commerciales, Montréal.

Sites Internet :

- ❖ ActuarialCNAM, Site des auditeurs et anciens auditeurs en actuariat du CNAM :
<http://www.actuariacnam.net/spip.php?article215>
- ❖ AXA : Le Global Forum for Longevity
<http://longevity.axa.com/fr/index.php>
- ❖ Cours d'introduction à l'Analyse Factorielle en général, UQAM :
<http://www.er.uqam.ca/nobel/k22761/geo3052/facto.html>
- ❖ Cours d'introduction à l'Analyse en Composantes Principales, UQTR :
<http://www.uqtr.ca/cours/srp-6020/acp/acp.pdf>
- ❖ LifeMetrics library, JP Morgan :
<http://www.jpmorgan.com/pages/jpmorgan/investbk/solutions/lifemetrics/library>
- ❖ LLMA, Life and Longevity Markets Association : <http://www.llma.org/about.html>
- ❖ Principes de la M-SSA :
<http://www.environnement.ens.fr/IMG/file/AndreasGrothPDF/Oldenburg2010.pdf>
<http://www.worldscientific.com/doi/pdf/10.1142/S2335680413500051>
- ❖ Principes et Interprétations de l'ACP :
<http://geo.polymtl.ca/~marcotte/g1q3402/chapitre3.pdf>
- ❖ Interprétation des CP : Bases de données de l'OMS
http://www.who.int/healthinfo/mortality_data/en/
- ❖ Interprétation des CP : Centers for Disease Control and Prevention
<http://www.cdc.gov/DiseasesConditions/>
- ❖ Interprétation des CP : Eurostat
http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database
- ❖ Interprétation des CP : Rapport OMS (Obésité : Prévention et prise en charge de l'épidémie mondiale) :
<http://books.google.ca/books?id=8FYYUXJx150C&pg=PA95&lpg=PA95&dq=ob%C3%A9sité%20aux+pays-bas&source=bl&ots=uFFZC5ns2T&sig=UcqxXNTV2wYmXG-86E3bTep97Mg&hl=fr&sa=X&ei=uhpSUZDADKS6yAGE7ICwDg&ved=0CC4Q6AEwBQ#v=onepage&q=ob%C3%A9sité%20aux%20pays-bas&f=false>
- ❖ Interprétation des CP : Statistiques Pays-Bas
<http://www.cbs.nl/en-GB/menu/home/default.htm>
<http://statline.cbs.nl/StatWeb/publication/?VW=T&DM=SLNL&PA=37154&D1=49-54&D2=47-110&HD=120125-1608&HDR=T&STB=G1>