INED
Institut national d'études démographiques

Quand certaines generations ont une mortalite differente de celle que l'on pourrait attendre
Author(s): John Wilmoth, Jacques Vallin, Graziella Caselli
Source: *Population (French Edition)*, 44e Année, No. 2 (Mar. - Apr., 1989), pp. 335-376
Published by: Institut National d'Études Démographiques
Stable URL: http://www.jstor.org/stable/1533586
Accessed: 03/04/2009 17:08

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=ined.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.

Institut National d'Études Démographiques is collaborating with JSTOR to digitize, preserve and extend access to *Population (French Edition)*.

http://www.jstor.org

QUAND CERTAINES GÉNÉRATIONS ONT UNE MORTALITÉ DIFFÉRENTE DE CELLE QUE L'ON POURRAIT ATTENDRE *

Les Anciens expliquaient la mort par trois raisons au moins : la vieillesse, les vicissitudes du moment et les conséquences d'une vie rangée ou dissolue. Les Modernes ne voient pas les choses très différemment, mais ils les mesurent plus précisément : la vieillesse par l'âge, les effets du moment par l'année, et l'accumulation des expériences par l'appartenance à une cohorte. Il ne reste plus qu'à séparer ces trois causes de leur action sur le quotient de mortalité à un âge donné, dans une génération donnée à une époque donnée. C'est plus simple à dire qu'à faire, car les trois effets ne sont pas indépendants les uns des autres : connaître deux d'entre eux fixe le troisième.

John WILMOTH**, Jacques VALLIN*** *et* Graziella CASELLI**** *ont donc mis au point une procédure statistique qui permet d'estimer en bloc les trois effets, mettant ainsi en évidence le rôle du troisième, la génération, bien moins connu et bien plus surprenant que les deux premiers, âge et période.*

Introduction

La table de mortalité, généralement calculée pour caractériser les conditions du moment, mesure avant tout la variation des risques de décès avec l'âge. Passée la période délicate de la naissance et des très jeunes âges, ceux-ci augmentent d'une manière à peu près exponentielle, reflétant en quelque sorte la fragilisation croissante de l'organisme humain en raison de son vieillissement biologique.

---

* Une première version de ce texte a été présentée, en français, au colloque *Modélisation démographique*, organisé à Vérone en février 1987 par l'Association d'Économie Appliquée et, en anglais, à la *Conférence annuelle* de la Population Association of America tenue à la Nouvelle Orléans en avril 1987. Les différentes critiques émises à ces deux occasions nous ont été très utiles pour la mise au point du présent article et nous en remercions vivement les auteurs.

** Population Studies Center, Université de Michigan, Etats-Unis.

*** INED.

**** Dipartimento di Scienze Demografiche, Université « La Sapienza », Rome, Italie.

Cependant, les conditions de ce vieillissement biologique s'inscrivent dans un contexte sanitaire particulier qui évolue avec le temps. Deux siècles durant, dans les sociétés européennes, ce contexte n'a cessé de s'améliorer et la mortalité de reculer. C'est ainsi que la courbe des quotients de mortalité par âge, tout en conservant, grosso modo, la même structure, s'est profondément abaissée.

De ce fait, aucune table de mortalité du moment ne correspond à l'histoire vraie des générations. Celle-ci combine le facteur âge (vieillissement biologique) et le facteur temps (progrès sanitaire) à tel point que, pour certaines générations, la mortalité peut avoir diminué avec l'âge sur de longues tranches de vie. Mais cette histoire vraie des générations à chaque âge n'est pas elle-même le simple résultat croisé de ces deux premières composantes (âge, période); elle dépend aussi d'une troisième composante, propre à chaque génération. La vulnérabilité d'un individu à un âge donné ne tient pas seulement au potentiel vital théorique correspondant à cet âge combiné aux conditions sanitaires du moment mais aussi à la détérioration (ou à l'amélioration) de ce potentiel résultant de son vécu antérieur. Cette troisième composante, généralement appelée effet de génération, a souvent excité la curiosité des démographes sans que ceux-ci parviennent jamais à la satisfaire complètement.

Mettre en évidence cette troisième composante n'est en effet pas chose facile [15]. Il faudrait, pour bien faire, pouvoir séparer chaque génération en deux groupes (par exemple à la naissance) : un premier groupe (expérimental) qui, durant une certaine période, serait soumis aux risques extérieurs et un second groupe (témoin) qui, durant la même période, serait tenu à l'écart de ces risques et ne subirait donc que le vieillissement biologique théorique. A l'issue de cette période, les deux groupes seraient soumis aux conditions sanitaires du moment et leurs mortalités seraient alors comparées. Une telle comparaison ne sera sans doute jamais possible, même dans le cadre d'enquêtes suivies. Au mieux, dans le contexte statistique actuel, peut-on espérer comparer les différentes générations à travers une matrice de quotients de mortalité par âge et année de naissance ou d'observations. Cette situation elle-même n'est pas encore très fréquente. Les données nécessaires à la confection d'une telle matrice ne sont couramment publiées que par quelques pays et depuis les années cinquante. On dispose cependant de quelques collections de données, publiées ou reconstituées, de plus grande ampleur historique, notamment pour l'Italie [6,24], la France [28,30], le Japon [39], la Belgique [37], les Pays-Bas [9] et l'Autriche [39]. C'est sur la base de ces ensembles qu'ont été développés des modèles visant à distinguer la composante génération des composantes âge et période.

Sur cette base on devrait en réalité se contenter d'une bonne description de la structure de la matrice dans ses trois perspectives (ligne, colonne et diagonale) et tirer profit des aspérités de cette structure pour émettre des hypothèses sur les mécanismes qui lient les trois composantes de la mortalité (âge, période et génération). On ne peut notamment alors raisonner sur la composante de génération qu'au travers des différences significatives que l'on arrive à mettre en évidence pour des générations voisines. On ne pourra ainsi prendre en

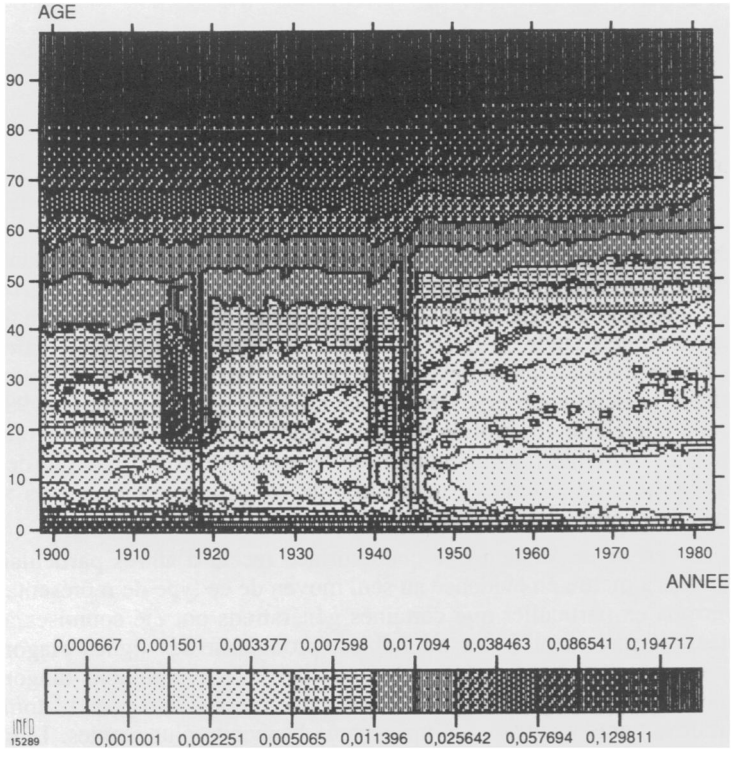

Figure 1. – Représentation en courbes de niveau des quotients de mortalité q_{ij} entre 0 et 99 ans, de 1899 à 1981. France, sexe masculin

compte qu'une fraction très marginale du poids total du vécu de chaque génération puisque l'on ne mesurera que l'effet différentiel des antécédents les plus singuliers.

La plupart des auteurs de modèles laissent planer un certain flou sur la distinction entre la structure à trois dimensions de la matrice et les trois dimensions de la mortalité. L'incertitude est souvent renforcée par la complexité du modèle ou des techniques d'ajustement utilisées. Qui plus est, la recherche de causalité va parfois au delà des limites imposées par la pauvreté inhérente aux informations contenues dans les données de base.

Nous utiliserons ici un modèle mis au point à l'occasion d'un travail antérieur [39,40], sensiblement différent de ses aînés. Ce modèle, relativement simple et flexible, prend résolument le parti d'une description simple mais systématique de la matrice des quotients de mortalité. Il n'en permet probablement que mieux de confirmer certaines hypothèses ou d'en formuler de nouvelles sur l'histoire particulière de certaines générations singulières. C'est ce que nous allons voir ici en explorant les données françaises portant sur la période 1899-1981.

I. Des générations singulières

A) Les données Une représentation simple de la matrice des données peut être obtenue au moyen d'un graphique en courbes de niveau [10,35]. Des travaux antérieurs ont déjà montré l'intérêt d'une telle approche pour la mise en évidence de certaines particularités affectant une surface tridimensionnelle de mortalité [7,8]. La représentation en courbes de niveau d'une telle surface est, bien que beaucoup plus approximative, de la même nature que l'étude de la structure des données de la matrice entreprise ici. La figure 1 donne une représentation en courbes de niveau des quotients de mortalité masculins observés en France de 1899 à 1981, entre 0 et 99 ans. Elle donne une claire description des principales caractéristiques générales de la surface de mortalité : augmentation de la mortalité avec l'âge (au-delà de 10 ans), baisse avec le temps ; mais aussi de ses principaux accidents : sur-mortalités de guerre.

Nous verrons cependant que cette surface recèle d'autres particularités, plus difficiles à mettre en évidence au seul moyen de ce type de représentation. Nous verrons en particulier que certaines générations ont été soumises à des mortalités anormalement fortes ou faibles, produisant des lignes diagonales de crête ou de creux sur la surface de mortalité. Ces ondulations diagonales sont cependant trop peu prononcées pour apparaître sur la figure 1, dominée par les mouvements verticaux ou horizontaux beaucoup plus amples. L'intérêt théorique que revêtent ces ondulations diagonales nous a conduit à utiliser un modèle paramétrique capable de nous donner une description plus précise des données observées.

B) Rappel de la méthode Le modèle utilisé donne une première description synthétique de l'évolution de la mortalité aux divers âges et de la structure par âge de la mortalité à différentes périodes. Après une transformation assez classique des quotients de mortalité, cette description s'appuie sur la structure rectangulaire de la matrice : chaque valeur est alors caractérisée par son appartenance à une ligne (âge) et à une colonne (période). Cependant, ce type de description de la matrice des données peut dans bien des cas être amélioré par la prise en considération systématique de la structure des résidus observés pour chaque génération.

Nous avons opté pour un modèle essentiellement additif et utilisé une transformation des quotients de mortalité q_{ij} de la forme :

$$f_{ij} = \log\left(\frac{q_{ij}}{1 - q_{ij}/2}\right)$$

où i et j représentent respectivement les lignes (âge) et les colonnes (période) de la matrice.

L'estimation actuarielle classique de la force de mortalité μ_{ij} étant elle-même de la forme :

$$\mu_{ij} = \frac{q_{ij}}{1 - q_{ij}/2}$$

cette transformation peut aussi être interprétée comme une transformation de la force de mortalité : $f_{ij} = \log(\mu_{ij})$[1].

La famille de modèles qui nous intéresse est de la forme :

$$f_{ij} = \alpha_i + \beta_j + \sum_{m=1}^{\rho} \phi_m \gamma_{im} \delta_{jm} + \theta_k + \epsilon_{ij}$$

où $k = j - i$. Les deux premiers termes $(\alpha_i + \beta_j)$ constituent la partie additive du modèle, qui rend compte des effets α de ligne et β de colonne. Les termes suivants, multiplicatifs, schématisent la structure rectangulaire de la part non additive du modèle. Leur nombre, ρ, dépend de leurs apports successifs à l'ajustement. Les effets "θ" de diagonale schématisent la part de non additivité propre à chaque diagonale de la matrice, le reliquat "ϵ" représentant les erreurs aléatoires. Le modèle est ajusté par itérations successives. On en trouvera en annexe une description plus complète.

Pour simplifier, nous appellerons "modèle AM" tout modèle de ce type contenant un terme additif et un nombre indéterminé de termes multiplicatifs (mais sans terme diagonal). Nous préciserons éventuellement le nombre de termes multiplicatifs du modèle en notant tout simplement "modèle AM(2)", "modèle AM(3)", ... ou "modèle AM(ρ)" selon qu'il y a $2, 3, \ldots$ ou ρ termes multiplicatifs. De la même manière nous appellerons "modèle AMD" (ou "modèle AM(ρ)D" pour une valeur spécifique ρ) tout modèle incluant un terme diagonal.

Par construction, ce type de modèle privilégie la structure rectangulaire des données et ne traite leur structure diagonale que comme un élément secondaire. Autrement dit, nous nous efforçons de décrire la structure rectangulaire de la matrice en utilisant de simples termes additifs et multiplicatifs indicés en fonction des lignes et des colonnes du tableau, et c'est seulement pour tenter d'améliorer l'ajustement que nous introduisons *in fine* un terme diagonal. Cette amélioration peut être interprétée, selon les cas, de deux manières différentes. Dans les cas où le modèle AM est ajusté avec un petit nombre de termes multiplicatifs (habituellement 2 ou 3, pour les matrices de quotients de mortalité examinées), le terme diagonal tend à synthétiser les excès ou déficits de mortalité propres à certaines générations et, de ce fait, la description est améliorée en ce sens qu'elle est plus *complète*. Dans les cas où,

---

[1] L'alternative, qui consisterait à choisir un modèle essentiellement multiplicatif et à utiliser tels quels les quotients q_{ij} de mortalité, doit être écartée car la variance du terme aléatoire d'un tel modèle serait loin d'être constante à travers la matrice. Comme on l'a montré ailleurs [3], la variance de μ_{ij} est à peu près égale à μ_{ij}/N_{ij} où N_{ij} est la population exposée au risque en ij. Ce rapport varie énormément avec l'âge (très petit aux jeunes âges, il devient très grand aux âges élevés). Au contraire, la variance de $\log(\mu_{ij})$ $(\simeq 1/\mu_{ij} N_{ij})$ est beaucoup plus stable.

au contraire, le modèle AM est ajusté avec un grand nombre de termes multiplicatifs (mettons, plus de trois), l'inclusion d'un terme diagonal tend à réduire le nombre de termes multiplicatifs nécessaires à une complète description et l'amélioration tient alors au fait que l'on aboutit à un modèle *plus simple.*

Le plus souvent, on remarque qu'à eux seuls les deux termes additifs du modèle expliquent une part écrasante de la variance totale de la matrice des données. Néanmoins, il semble que l'on soit généralement en présence d'une certaine dose de non additivité. En introduisant à la fois un ou plusieurs termes multiplicatifs et un terme diagonal, nous avons choisi de distinguer entre deux types de non additivité : les termes multiplicatifs illustrent la lente transformation, avec le temps, de la structure de la mortalité par âge, tandis que le terme diagonal met en évidence les générations qui ont subi une mortalité sensiblement plus forte ou plus faible que normale. Il serait évidemment possible de remplacer le terme diagonal par des termes multiplicatifs supplémentaires, mais une telle stratégie présente plusieurs inconvénients (cf. [39] pour une discussion plus complète). D'une part elle serait plus dispendieuse, puisque, généralement, plusieurs termes multiplicatifs seraient nécessaires pour remplacer le terme diagonal. D'autre part, et c'est plus important, nous ne serions pas en mesure de distinguer entre deux types de structures non additives relevant d'enchaînements causaux différents. En effet, la modification progressive de la forme de la courbe de mortalité par âge que décrivent les facteurs multiplicatifs peut tout aussi bien refléter des changements dans la mortalité par génération que des changements dans la mortalité du moment. A l'inverse, la mortalité anormalement forte ou faible par rapport à celle des générations encadrantes détectée par le terme diagonal peut être interprétée de façon moins ambiguë comme le seul résultat de différences plus ou moins nettes dans le vécu des générations successives.

Nous distinguons, dans ce qui suit, entre deux procédures d'ajustement, appelées, la première, *exploratoire* et la seconde, *complète*, utilisées à différentes étapes de l'analyse. L'ajustement complet suit l'intégralité de la procédure itérative décrite en annexe alors que l'ajustement exploratoire s'en tient automatiquement au résultat d'une seule itération. L'ajustement exploratoire revient ainsi à effectuer un premier ajustement sur la base d'un modèle AM et à dériver la composante diagonale θ_k des résidus de ce modèle. Si une telle approche a le mérite d'être simple à mettre en œuvre sur le plan du calcul, elle a aussi le défaut de reposer sur une hypothèse, implicite mais abusive, d'indépendance entre les différents termes du modèle. En pratique, on se rendra compte que la principale différence entre les ajustements exploratoires et complets est que le premier tend à sous-estimer l'amplitude (mais non la direction) de l'effet diagonal puisqu'il commence par faire absorber dans le cadre du modèle AM une partie de la structure diagonale par les termes multiplicatifs. Néanmoins, nous avons pu utilement recourir à ce premier type d'ajustement, à certains stades provisoires de notre analyse, à la fois pour alléger les calculs nécessaires et pour simplifier la présentation initiale de la méthode.

Notre modèle se distingue des modèles antérieurs d'analyse "âge-période-génération" (APG) [4,14,15,20] de deux points de vue importants. D'une

part, il offre une grande souplesse dans le choix du modèle le plus approprié au sein d'une famille de modèles. D'autre part, et c'est encore plus important, il participe d'une philosophie différente, en s'appuyant sur une approche purement descriptive pour tenter de contribuer à une meilleure compréhension théorique. Notre méthode est essentiellement narrative, historiographique : nous nous efforçons d'abord de fournir une description aussi détaillée que possible de l'évolution de la mortalité et seulement ensuite d'en rechercher une interprétation.

Il est vrai que ce modèle comporte, tout comme les modèles APG classiques, plusieurs contraintes d'identification (voir annexe). Il ne faut cependant pas laisser germer de malentendu à propos de réserves à faire à ce sujet. On ne doit en aucun cas présumer *a priori* que ces contraintes reflètent des relations sous-jacentes entre les facteurs explicatifs liés à l'âge, la période ou la génération. Elles ont été choisies au contraire pour mettre en valeur la structure additive prédominante de la matrice des données et rendre compte de l'existance de deux types différents de structures non additives. Nous avons à dessein évité les expressions "effets d'âge", "effets de période" ou "effets de génération", en nous référant de préférence aux paramètres du modèle et en utilisant dans un premier temps un jargon spécifique (du genre "effets de ligne et de colonne", "non additivité rectangulaire", "termes multiplicatifs", "effet diagonal", etc.). Quand il s'agira de donner plus de substance à ces concepts, on dira par exemple que les α représentent la structure moyenne de la mortalité par âge sur toute la période étudiée ou que les θ représentent la sur- (ou sous-) mortalité des générations. Comme nous l'avons dit en introduction, les modèles théoriques qui combinent les variables âge, période et génération sont confrontés au problème incontournable de l'inadéquation des données (absence de groupes témoins). C'est cette difficulté majeure qui est à l'origine des problèmes d'interprétation auxquels se heurtent les modèles APG traditionnels et qui a conduit à utiliser des contraintes d'identification supposant une connaissance préalable des relations causales liées à l'âge, la période ou la génération. Il nous semble au contraire qu'une approche descriptive permet d'éviter les ambiguïtés inhérentes à l'utilisation de ce type de contraintes arbitraires et devrait conduire à une interprétation plus claire des paramètres.

C) Premiers résultats Une première application du modèle exploratoire [40] a porté sur la mortalité masculine française de la période 1899-1981 [30]. Les données utilisées ici sont les quotients de mortalité par génération et par année d'âge de 0 à 89 ans : q_{ij} est donc la probabilité de décès entre les âges exacts i et $i+1$ pour la génération $j-i$ atteignant l'âge exact i durant l'année de calendrier j. Alors que, dans ce cas, l'âge i correspond précisément à l'intervalle d'âge entre deux anniversaires $(i, i+1)$, la période notée j chevauche au contraire deux années consécutives de calendrier (j et $j+1$).

1. La structure additive

La figure 2 présente les effets de ligne additifs qui caractérisent le profil moyen de mortalité par âge sur l'ensemble de la période 1899-1981, tandis que la figure 3 illustre les effets de colonne additifs qui reflètent l'évolution moyenne de la mortalité des divers âges de 1899 à 1981.

La figure 3 fait clairement apparaître les fortes perturbations dues aux deux guerres mondiales[2]. Autant ces perturbations peuvent être intéressantes par les différences qu'elles introduisent dans le passé des générations étudiées, autant elles peuvent rendre délicate la mise en évidence des effets diagonaux dans les périodes où elles sévissent. C'est pourquoi il nous a paru utile de restreindre une grande partie des analyses qui suivent à une période exempte de graves perturbations, en l'occurrence l'après-deuxième-guerre mondiale (1946-1981). A la figure 2, la ligne discontinue donne le profil moyen de mortalité par âge pour cette période restreinte. Ce profil diffère évidemment du précédent (ligne continue) en raison de la baisse générale de la mortalité mais aussi du fait que cette baisse a été particulièrement accusée aux jeunes âges et que les âges

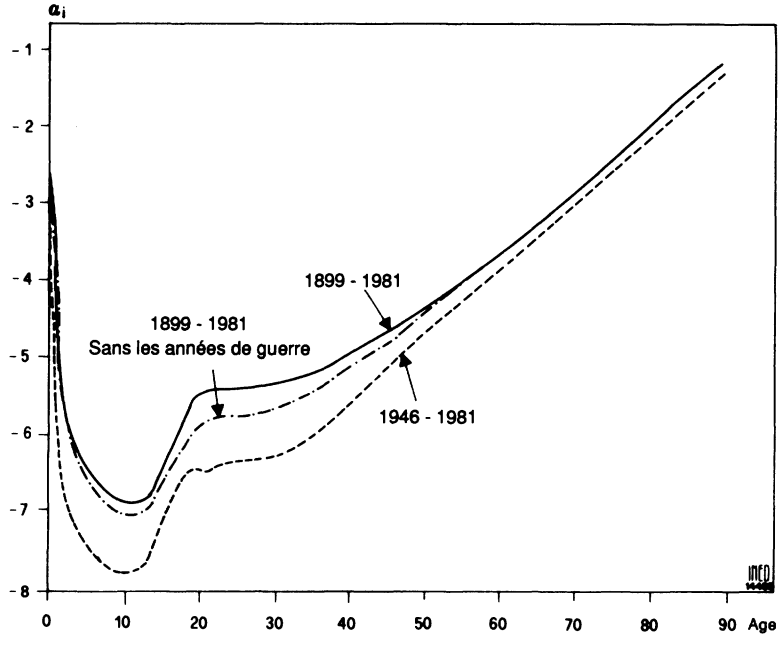

Figure 2. – Valeurs des effets de ligne additifs α_i obtenus pour le sexe masculin dans différents cas de figure

[2] Ici, 1913-1919 et 1939-1945, du fait que l'on utilise des quotients de mortalité par génération.

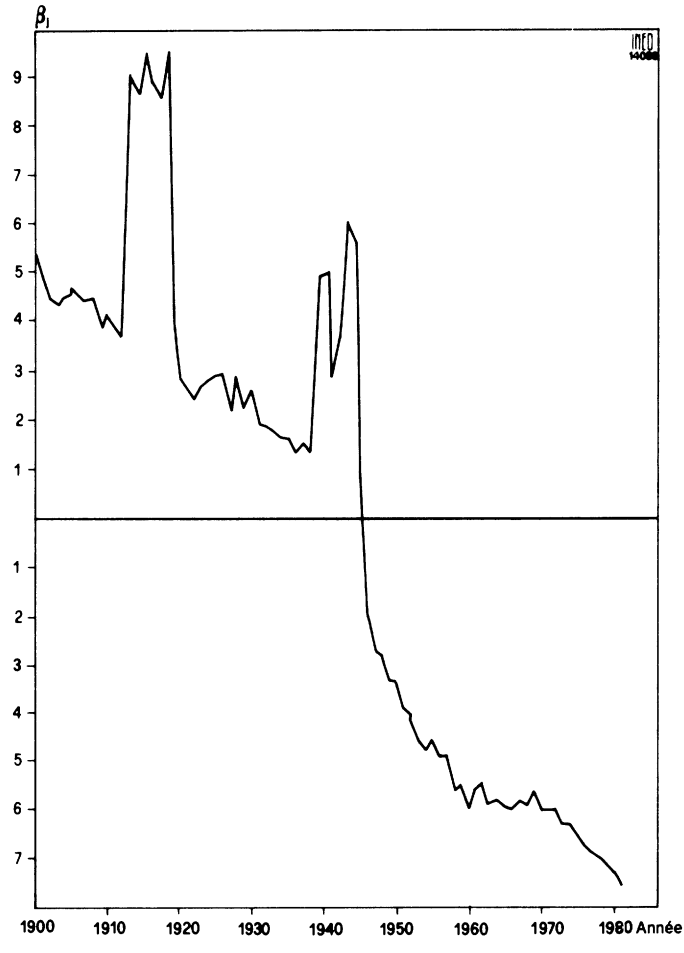

Figure 3. – Valeurs des effets de colonne additifs β_j obtenus pour le sexe masculin (analyse de la matrice complète)

adultes sont exempts, dans cette période, de toute surmortalité de guerre (sur le même graphique la ligne pointillée représente le profil moyen de mortalité par âge sur toute la période 1899-1981, mais après exclusion des années de guerre).

2. *Les termes multiplicatifs*

Une partie des déviations observées par rapport à ce schéma additif est prise en compte par les termes multiplicatifs du modèle. S'agissant de l'ensemble de la période 1899-1981, il nous a paru nécessaire de retenir trois termes multiplicatifs ($\rho = 3$) pour décrire la structure rectangulaire de la ma-

trice, alors que deux suffiront ($\rho = 2$) lorsque l'on se limitera à la période 1946-1981 (les critères de choix sont précisés en annexe). Il est difficile d'illustrer le rôle de ces termes multiplicatifs. La figure 4 donne une représentation en courbes de niveaux [35,40], en fonction des âges et des années de calendrier, des effets combinés de ces trois termes :

$$\hat{\phi}_1\hat{\gamma}_{i1}\hat{\delta}_{j1} + \hat{\phi}_2\hat{\gamma}_{i2}\hat{\delta}_{j2} + \hat{\phi}_3\hat{\gamma}_{i3}\hat{\delta}_{j3}$$

Les deux trames les plus foncées représentent les valeurs positives de ces effets multiplicatifs combinés et les deux trames les plus claires en représentent les valeurs négatives. Une valeur positive pour un point donné indique que la prise en compte des termes multiplicatifs majore la mortalité attendue, pour cet âge et cette période, du simple schéma additif. Une valeur négative indique au contraire que cette prise en compte minore la mortalité prédite par le modèle additif.

La figure 4 peut ainsi se lire soit en "ligne" soit en "colonne". Pour un âge donné, dans le contexte général de baisse de la mortalité, le fait de passer d'une zone foncée à une zone claire indique qu'à cet âge la diminution de la mortalité a été plus rapide que la moyenne ; au contraire, le passage d'une zone claire à une zone foncée indique une baisse plus lente, voire l'absence de baisse. De ce point de vue, la figure 4 est remarquable par sa division en quatre quartiers opposant en alternance les moins de 40 ans et les plus de 40 ans, avant et après la seconde guerre mondiale. Au-dessous de 40 ans la baisse de la mortalité a été plus rapide que la moyenne (et c'est particulièrement vrai au-dessous de 10 ans), tandis que le contraire s'est produit au-dessus de 40 ans.

De même, pour une année donnée, à mesure que l'on avance en âge, le passage d'une zone foncée à une zone claire indique que la mortalité augmente moins vite avec l'âge que dans le profil moyen; et inversement, si on passe d'une zone claire à une zone foncée. On retrouve évidemment, de ce point de vue aussi, la division de la figure 4 en quatre quartiers, illustrant en fait la transformation progressive du profil de mortalité par âge.

Au-delà de cette modification d'ensemble par rapport au schéma additif, la lecture de la figure 4 permet aussi de mettre en relief certaines particularités de l'ajustement par le modèle AM(3). Notons surtout en ligne l'évolution propre aux âges ayant payé le plus lourd tribut à la guerre. A ces âges la surmortalité de guerre produit une brusque augmentation des termes multiplicatifs tandis que, dans les périodes qui suivent, ceux-ci apparaissent en négatif, du simple fait que le cumul des résidus du modèle additif sur une même ligne (ou colonne) est nécessairement égal à zéro. Notons aussi le cas particulier des jeunes de 18-20 ans, qui, frappés dans les années récentes par une recrudescence de mortalité par accidents de la route, font parallèlement l'objet d'une augmentation des termes multiplicatifs dans les années soixante-dix.

De même, on remarque, en colonne, le cas particulier de la première guerre mondiale durant laquelle la surmortalité a été si forte entre 20 et 40 ans que les âges supérieurs à 50 ans sont affectés de termes multiplicatifs fortement

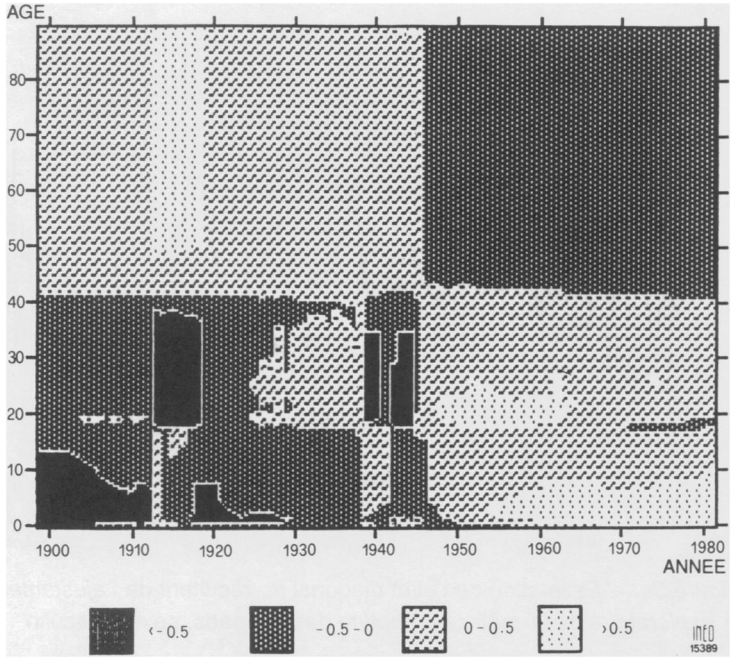

Figure 4. – Représentation en courbes de niveaux de la valeur cumulée des trois termes multiplicatifs $\sum_{m=1}^{3} \hat{\phi}_m \hat{\gamma}_{im} \hat{\delta}_{jm}$ utilisés pour l'ajustement rectangulaire de la matrice complète des données, sexe masculin

négatifs. On voit ainsi que les termes multiplicatifs offrent une possibilité de correction non négligeable du modèle additif, allant jusqu'à prendre en compte une grande part de certaines perturbations très conjoncturelles.

3. La structure diagonale résiduelle

Nous pouvons donc être à peu près assurés que la part non aléatoire des résidus laissés par cet ajustement rectangulaire de la matrice des données tient à sa structure diagonale, dont l'effet est estimé pour chaque génération par les termes θ_k. La figure 5a illustre l'effet diagonal ainsi estimé $\hat{\theta}_k$ dans le cas où l'on applique le modèle à l'ensemble de la période 1899-1981. Elle porte sur les générations 1810 à 1981 qui traversent le groupe d'âges 0-89 ans entre 1899 et 1981. Les estimations originales (en ligne discontinue) ont fait l'objet d'une procédure de lissage [36] (en ligne continue).

Bien entendu, les résultats sont peu significatifs aux extrémités du graphique pour les générations qui ne sont représentées que pour quelques années d'âges. Il est prudent d'ignorer les générations antérieures à 1830 ou postérieures à 1960, pour lesquelles les résultats apparaissent d'ailleurs clairement soumis à de fortes fluctuations aléatoires. De la génération 1830 à la

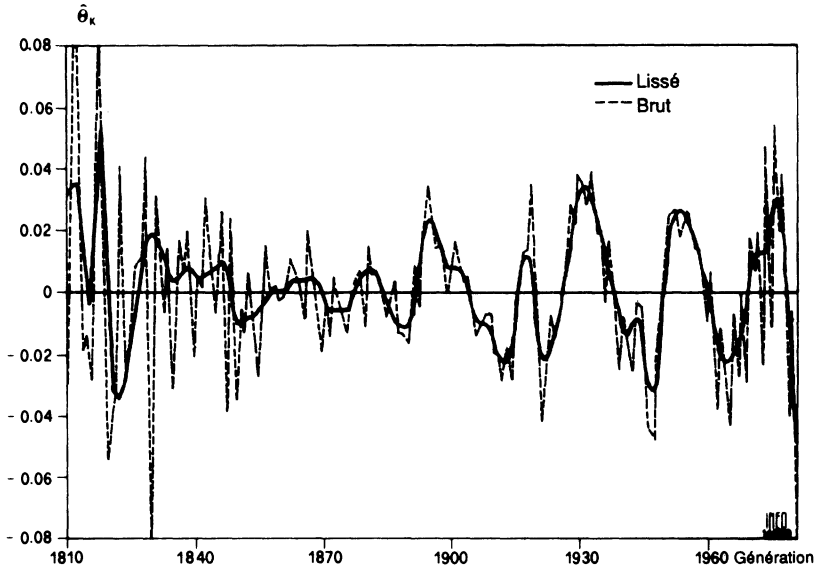

Figure 5a. – Estimation de l'effet diagonal $\hat{\theta}_k$ résultant de l'*ajustement exploratoire* de la *matrice complète* des données, sexe masculin

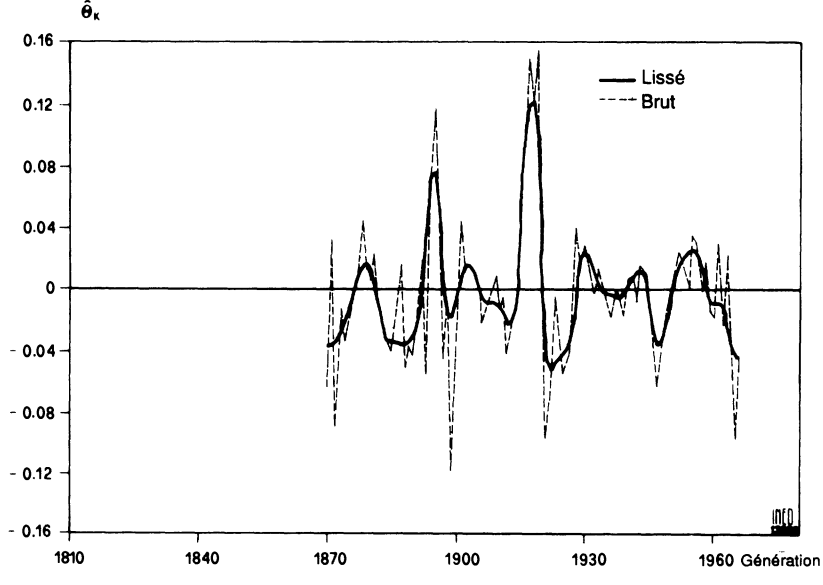

Figure 5b. – Estimation de l'effet diagonal $\hat{\theta}_k(15, 29)$ résultant de l'*ajustement exploratoire* de la *matrice complète* des données restreinte au *groupe d'âges 15-29 ans*, sexe masculin

génération 1890, aucun effet diagonal majeur n'est révélé par l'analyse. En revanche, à partir des générations nées à la fin du siècle passé, les valeurs estimées $\hat{\theta}_k$ de θ_k passent par des oscillations significatives[3].

Un premier ensemble de termes positifs concerne les générations 1894 à 1904 suivies par une série de valeurs négatives de 1905 à 1926 (entrecoupée cependant d'une poussée positive en 1917, 1918 et 1919). Une seconde oscillation commence ensuite avec des termes positifs de 1927 à 1938 puis négatifs de 1939 à 1949. Enfin une troisième oscillation semble s'amorcer avec une série de termes positifs observés pour les générations nées dans les années cinquante.

On sait que la valeur de $\hat{\theta}_k$ mesure approximativement le pourcentage moyen de surmortalité (ou de sous-mortalité) qui différencie chaque génération du niveau de mortalité de référence estimé par le modèle AM approprié (cf. annexe). On peut donc dire que les générations 1894 à 1904, 1927 à 1938 ou 1950 à 1959 se caractérisent par une plus grande vulnérabilité que celle que l'on attendrait de la simple combinaison de la structure par âge et de l'évolution tendancielle de la mortalité. Pour les deux premiers groupes de générations on retrouve sans aucun doute ici les conséquences à long terme des deux guerres mondiales déjà soulignées par de nombreux auteurs [1,3,4,7,8,16,25,28]. Le cas des générations nées dans les années cinquante pose davantage question.

Le modèle reposant par définition sur une somme nulle des effets ainsi mesurés[4], les générations affectées de termes négatifs n'apparaissent probablement bénéficier d'une sous-mortalité que par contraste avec la surmortalité des générations les plus vulnérables. La mortalité de référence estimée par le modèle AM est nécessairement le résultat de l'expérience moyenne des générations désavantagées et des générations avantagées et, de ce fait, toute évocation d'avantage ou de désavantage doit s'entendre ici en termes relatifs.

Pour des raisons que nous discuterons plus loin, nous avons choisi de concentrer notre attention sur les générations défavorisées, considérant que l'avantage des autres générations n'est que relatif au désavantage des premières. Une application de la même méthode à la mortalité toutes causes de plusieurs autres pays réalisée par ailleurs semble pouvoir donner quelque crédit à l'hypothèse inverse, laissant apparaître par exemple que certaines générations avantagées le sont grâce à la sélection opérée sur les individus ayant survécu à la guerre. Cependant, dans le cadre des données que nous étudions ici, la première hypothèse semble *a priori* plus plausible. Elle est également con-

---

[3] Nous avons montré par ailleurs [39] qu'il était fort improbable que de telles oscillations ne soient dues qu'à des variations aléatoires. Nous ne développerons donc pas davantage ici ce thème de la significativité de ces oscillations.

[4] En réalité c'est la somme pondérée qui est nulle :

$$\sum_k w_k \theta_k = 0,$$

où w_k est le nombre de quotients par âge disponibles pour la génération k.

forme aux conclusions d'études antérieures, démographiques [16] ou médicales [13].

La relation suggérée par la figure 5a entre les effets diagonaux et le vécu de certaines générations durant les guerres reste cependant assez floue. L'effet dû à la première guerre mondiale semble de moindre ampleur que celui qui résulte de la seconde alors que l'on avance généralement le contraire et le cas particulier des générations 1917, 1918 et 1919 semble bien exiger une explication propre.

Il se peut cependant, comme nous l'avons déjà suggéré, que le fait de soumettre à l'analyse l'ensemble de la période 1899 à 1981, incluant donc les années de guerre elles-mêmes, perturbe d'une certaine façon ces résultats. Bien que les termes multiplicatifs de l'ajustement rectangulaire prennent en compte une bonne part de ces fluctuations conjoncturelles, il peut en subsister un certain reliquat dont les effets rendent plus confuse l'estimation de l'effet diagonal. La guerre a en effet pour conséquence de transformer brutalement le profil par âge assez régulier en un profil très perturbé au niveau des classes d'âges mobilisées (en 1914 par exemple le quotient de mortalité passe de 5 pour mille à 17 ans à 82 pour mille à 18 ans). C'est sans doute au moins en partie en raison d'une prise en compte insuffisante de telles perturbations par l'ajustement rectangulaire que l'on remarque à la figure 5a des valeurs maximum de $\hat{\theta}_k$ pour les générations 1895 et 1919.

D'une manière plus générale, il serait plus logique, pour étudier les effets à long terme de certains événements sur le devenir des générations, de restrein-

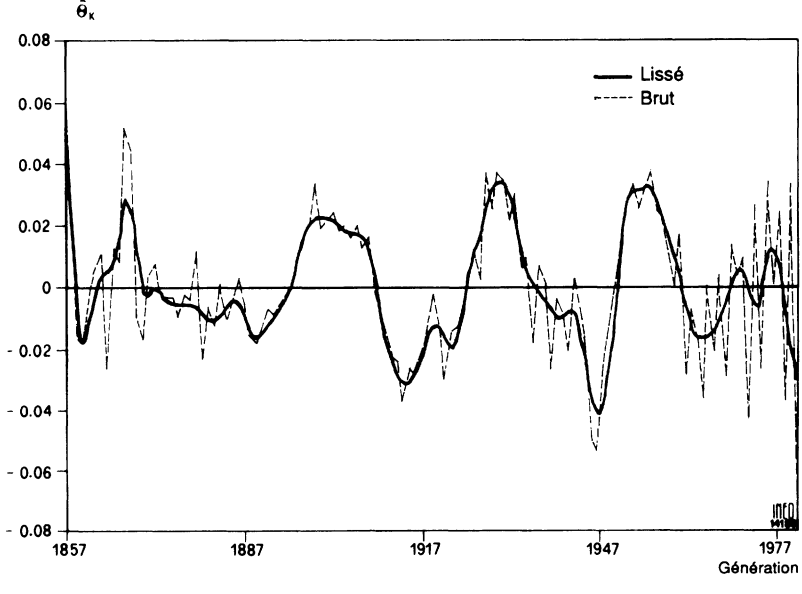

Figure 6. – Estimation de l'effet diagonal $\hat{\theta}_k$ résultant de l'*ajustement exploratoire* de la *matrice d'après-guerre (1946-1981)*, sexe masculin

dre l'analyse aux âges supérieurs à celui auquel surviennent ces événements. La situation est assez complexe dans la mesure où les divers événements qui semblent modifier la mortalité ultérieure de certaines générations interviennent à des âges très différents. Un compromis assez simple, permettant à la fois d'éliminer le bruit de fond des deux guerres mondiales et de parer aux inconvénients de l'analyse globale, consiste à ne retenir ici que la période postérieure à la seconde guerre mondiale. Cette solution présente en outre l'avantage de fonder l'analyse sur la partie des données dont la qualité est la plus fiable.

La figure 6 présente les résultats obtenus à partir de la matrice de données 1946-1981. Les trois oscillations apparaissent en effet plus clairement. Les effets à long terme de la première guerre mondiale prennent ici plus d'ampleur mais mettent en jeu un groupe de générations un peu différent : 1896 à 1909 au lieu de 1894 à 1904. Notons aussi que les générations 1917, 1918 et 1919 n'apparaissent plus ici en positif. Cependant la série négative concernant les générations 1910 à 1927 continue d'être interrompue par trois générations faisant contraste par des valeurs nettement moins négatives. Précisons toutefois qu'il ne s'agit plus tout à fait des mêmes générations : 1918, 1919 et 1920 au lieu de 1917, 1918 et 1919.

D) Hypothèses d'explication Cette lecture des figures 5 et 6 appelle un certain nombre d'explications qui peuvent différer d'un groupe de générations à l'autre. Il nous faut essayer de comprendre ce qui peut avoir déterminé la plus grande vulnérabilité des générations 1896-1909, 1927-1938 et 1950-1959 ainsi que le cas particulier des générations 1918, 1919 et 1920.

Dans le cas des deux premiers groupes de générations, le lien avec les deux guerres mondiales semble évident. Plusieurs mécanismes peuvent cependant entrer en jeu selon qu'il s'agit des générations qui ont fait la guerre, de celles qui ont vécu leur adolescence pendant les années difficiles ou enfin de celles qui y sont nées.

Le cas particulier des générations 1918, 1919 et 1920 peut faire référence à deux types d'explications : la participation à la seconde guerre mondiale mais aussi, puisque le maximum se situe en 1919, certains effets de la grippe espagnole de 1918.

Le problème posé par les générations nées dans les années cinquante est plus délicat. Elles n'ont été marquées par aucune crise d'envergure. On ne peut évidemment évoquer à ce propos ni la guerre d'Indochine (1950-1954) ni la guerre d'Algérie (1954-1962) qui n'ont fait du côté français qu'un nombre de victimes négligeable par comparaison aux deux guerres mondiales. Ces générations ont cependant vécu un moment très spécifique de l'histoire sanitaire française. C'est en effet dans les années cinquante que s'est brusquement généralisé l'accouchement en milieu hospitalier. Ce changement s'est accompagné d'un développement des interventions obstétricales qui a certainement

permis de sauver un grand nombre d'enfants qui autrement seraient morts à la naissance. Mais il se peut que d'une part les premières générations d'enfants ainsi sauvées soient restées particulièrement fragiles et que d'autre part ces interventions aient aussi provoqué un certain nombre de traumatismes obstétricaux. On remarque en effet que la mortalité par lésion obstétricale a augmenté dans cette période [33]. Dans un second temps au contraire, avec le perfectionnement des techniques d'intervention, l'hospitalisation des naissances a pu produire pleinement ses effets bénéfiques.

Ces hypothèses d'explication, et notamment la dernière, restent évidemment à vérifier. Il ne nous est pas possible d'en faire des affirmations mais nous allons essayer de voir, dans une deuxième partie, si elles résistent à une analyse plus complète prenant en compte les différences entre sexe et la répartition des effets diagonaux entre différents groupes d'âges. Nous avons pour ce faire subdivisé systématiquement la matrice 1946-1981 en quatre grands groupes d'âges (0-14 ans, 15-29 ans, 30-64 ans et 65-89 ans) et analysé les effets diagonaux à l'intérieur de chacune de ces subdivisions. Dans une troisième partie, nous examinerons enfin si la prise en compte de différentes causes de décès permet d'affiner ces explications.

II. Le résultat d'histoires différentes

Nous venons de présenter quelques résultats concernant certains avantages ou désavantages dont bénéficient ou pâtissent certaines générations masculines françaises en matière de mortalité, analysés soit à partir de l'ensemble de l'information disponible (1899-1981) soit sur une plage restreinte à l'après-guerre (1946-1981). On pouvait, par commodité, à ce stade de l'étude, se contenter de l'ajustement exploratoire. Il convient cependant maintenant d'employer, au prix de calculs plus longs et complexes, la procédure complète d'ajustement qui seule permet de minimiser la part de l'effet diagonal absorbée par les termes multiplicatifs et donc de juger de l'amplitude de cet effet. L'utilisation de la procédure complète est d'autant plus nécessaire que nous aurons à comparer les structures diagonales de plusieurs matrices (en l'occurrence, par sexe et par cause de décès), dont la modélisation pourrait faire appel à des nombres différents de termes multiplicatifs (conformément aux critères énoncés en annexe). Nous nous concentrerons cependant pour ce faire sur la seule portion de la matrice postérieure à la guerre (1946-1981), pour laquelle les risques d'interprétation sont moins grands.

A) Les effets à long terme de la guerre

1) Les générations ayant le plus participé aux deux guerres mondiales

Un grand nombre de générations ont participé à la première guerre mondiale : toutes les générations nées entre 1869 et 1899 ont été mobilisées à plus de 60% de leur effectif de 1914. Toutes celles nées entre 1876 et 1898

l'ont été à plus de 80% [28]. Cependant, parmi elles, certaines ont été plus massivement impliquées dans les combats que d'autres. Si l'on en juge par le nombre de morts sur le champ de bataille, c'est la génération 1894 qui l'a été le plus, perdant ainsi 27% de son effectif, suivie de près par les générations 1892 à 1895, toutes amputées de plus du quart. Les autres ont été, de ce point de vue, nettement moins touchées: il est vraisemblable que ces mêmes générations sont aussi celles qui ont compté, parmi les survivants, le plus grand nombre de blessés ou de gazés. Mais plus encore que les blessures de guerre, ce sont sans doute les conditions de vie sur le front et notamment les privations alimentaires qui ont le plus marqué les rescapés. De ce point de vue, ce sont probablement les plus jeunes soldats qui ont le plus souffert, dans la mesure où leurs besoins physiologiques étaient les plus grands. Il faut donc s'attendre à trouver le maximum d'effets à long terme dans les générations 1892 à 1899.

A la figure 5a (analyse de la matrice complète), on note effectivement un maximum d'effet diagonal pour les générations 1894-1895. De même, les générations qui ont le plus contribué à la seconde guerre mondiale sont les générations 1915 à 1919 et tout particulièrement les générations 1917-1918 dont plus de 60% ont été faits prisonniers [28]. Les plus jeunes étant là encore vraisemblablement les plus affectés, on peut s'attendre à ce que les effets à long terme marquent surtout les générations 1917, 1918 et 1919. C'est bien ce que nous avons déjà noté à la lecture de la figure 5a.

Cependant, comme nous l'avons déjà dit, la figure 5a est perturbée par l'insuffisance de l'ajustement rectangulaire de la matrice au niveau des deux guerres mondiales et comporte un risque de confusion entre les effets immédiats et les effets à long terme des chocs considérés. On peut s'en rendre compte en analysant la structure diagonale de chacune des quatre subdivisions de la matrice en fonction de l'âge. Pour le groupe d'âge 15-29 ans (figure 5b), on trouve un effet diagonal très élevé (de l'ordre de 7 à 15%) pour les générations 1894 à 1896 et pour les générations 1915 à 1919 qui ne peut guère s'expliquer que par les pertes de guerre elles-mêmes. Au contraire, cet effet diagonal est beaucoup plus modeste à 30-64 ans (de l'ordre de 1 à 3% pour les générations 1894 à 1896 et négatif pour les générations 1915 à 1919) et pratiquement nul à 65-89 ans.

De fait, si l'on se rapporte aux figures 6 (ajustement exploratoire) et 7a (ajustement complet), restreintes aux années 1946 à 1981 et donc aux seuls effets postérieurs à la seconde guerre mondiale, la génération la plus marquée par la première guerre mondiale est née en 1899 et celle qui a été la plus touchée par la seconde guerre mondiale (parmi celles qui y ont participé) est née en 1919 (encore que, après lissage, cela ne se voit plus guère à la figure 7a). Autrement dit, dans l'un et l'autre cas, il s'agit de la plus jeune des générations mobilisées et non pas de celles qui ont le plus participé aux combats.

S'agissant des effets à long terme des opérations militaires, on est tenté de penser que le sexe féminin n'est pas concerné et il peut être intéressant de comparer ici la situation des hommes à celle des femmes. On peut s'étonner a priori qu'à la figure 7b apparaisse pour celles-ci un mouvement analogue à celui de la figure 7a pour les hommes. Notons cependant que, pour la

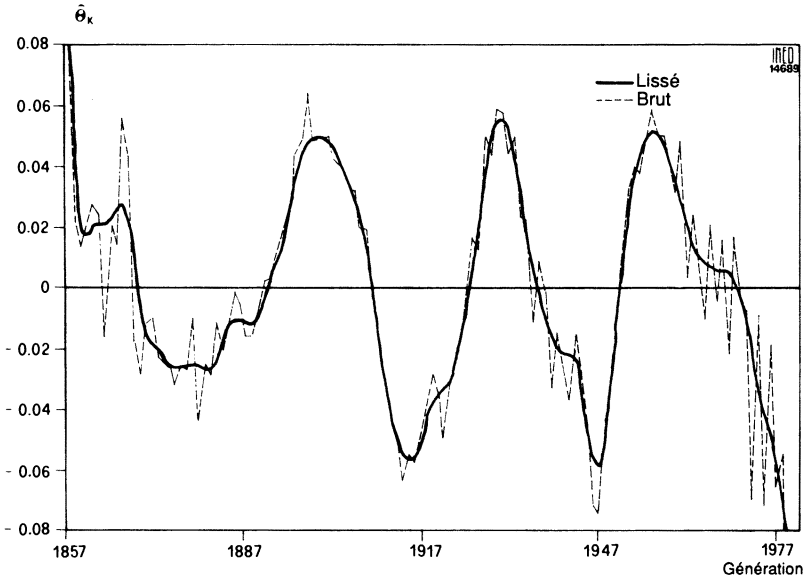

Figure 7a. – Estimation de l'effet diagonal $\hat{\theta}_k$ résultant de l'*ajustement complet* de la *matrice d'après-guerre (1946-1981)*, sexe masculin

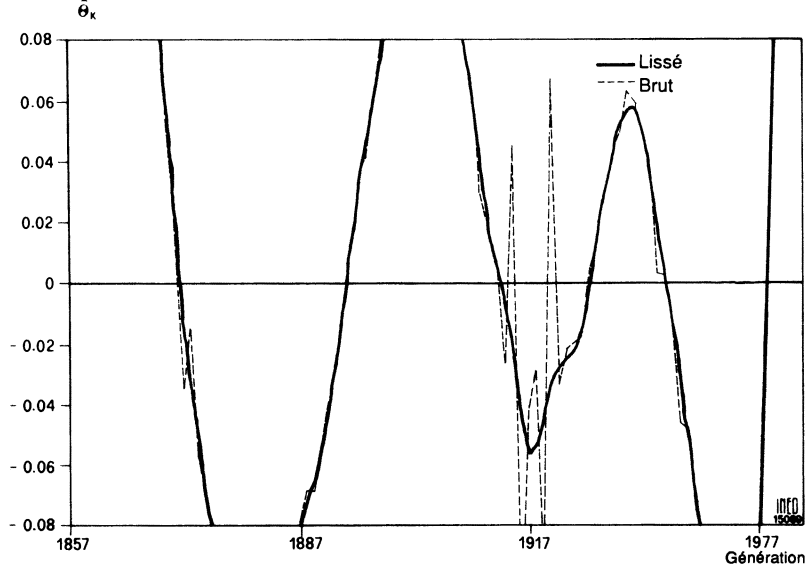

Figure 7b. – Estimation de l'effet diagonal $\hat{\theta}_k$ résultant de l'*ajustement complet* de la *matrice d'après-guerre (1946-1981)*, sexe féminin

première guerre mondiale, ce mouvement est d'ampleur un peu plus réduite et ne concerne pas exactement de la même façon les mêmes générations. Les générations féminines les plus touchées, qui s'étalent de 1891 à 1897, sont sensiblement plus âgées que les générations masculines 1896 à 1899. La relative vulnérabilité de ces générations féminines ne peut s'expliquer que par la détérioration des conditions générales de vie pendant la guerre, qui, bien qu'elle ait touché tout le monde a probablement plus durement affecté les femmes enceintes ou les jeunes mères, générations qui avaient pour la plupart de 20 à 25 ans entre 1914 et 1918.

Le cas de la seconde guerre mondiale est un peu plus complexe car on constate pour les femmes des valeurs relatives pour les générations 1918, 1919 et 1920 que l'on ne peut expliquer sans évoquer les effets de la grippe espagnole, comme nous le verrons plus bas.

2) Les générations ayant vécu leur adolescence pendant la guerre

Shiro Horiuchi [16] a plus particulièrement insisté sur la fragilisation des générations ayant vécu leur adolescence pendant l'une ou l'autre des deux guerres mondiales. On retrouve ici très clairement ce phénomène chez les hommes : tant pour les générations 1900 à 1908 que pour les générations 1928 à 1933 on observe des valeurs nettement positives de $\hat{\theta}_k$ (figure 7a) et cela est vrai quel que soit le groupe d'âges (au-dessus de 15 ans) auquel on se limite. Shiro Horiuchi n'avait pu cependant mettre ce phénomène en évidence que pour le sexe masculin et cela restait étonnant dans la mesure où l'on imaginait mal que la détérioration des conditions de vie des adolescents durant la guerre ait pu épargner les jeunes filles. La figure 7b montre ici que le phénomène de fragilisation de ces générations atteint aussi le sexe féminin. Il reste à expliquer pourquoi il est nettement moins grave pour les femmes que pour les hommes. Shiro Horiuchi a repris l'hypothèse, déjà proposée par M. Okubo [25], selon laquelle certaines déficiences alimentaires durant l'adolescence peuvent entraîner plus particulièrement chez les jeunes garçons un épaississement des artères créant par la suite un terrain favorable aux maladies de dégénérescence cardio-vasculaire.

Cette explication nous semble beaucoup trop spécifique. Il vaut probablement mieux rattacher cette moindre fragilisation des générations féminines à un phénomène plus général. Les femmes ont bénéficié depuis la seconde guerre mondiale d'une amélioration de leur situation sanitaire beaucoup plus grande et rapide que les hommes. Ceux-ci ont au contraire davantage porté le poids de l'industrialisation [29,31]. De ce fait, les contrastes entre générations masculines peuvent avoir été renforcés alors que ceux qui différenciaient les générations féminines s'atténuaient.

3) Les générations nées pendant la guerre

Compte tenu de la fragilité spécifique des nouveau-nés, on peut se demander si les enfants de la guerre ne subissent pas aux âges adultes les

conséquences lointaines de leur naissance dans des conditions particulièrement difficiles. C'est en effet ce qui a pu être constaté dans le cas des enfants italiens des deux sexes nés pendant la première guerre mondiale [3,4,7,8], mais une comparaison des données italiennes aux données françaises a également montré que ce phénomène n'était pas général et tenait vraisemblablement à une détérioration particulièrement accusée du système sanitaire et alimentaire italien pendant la première guerre [7]. On ne voit en tout cas ici aucune fragilisation particulière des générations françaises nées pendant la guerre, à l'exception des générations 1917 et 1918 pour lesquelles nous avons déjà avancé d'autres explications.

B) Les effets à long terme de la grippe espagnole

L'hiver 1918-1919 a été marqué par la grippe la plus sévère de l'histoire épidémiologique moderne, la grippe espagnole, suivie d'une résurgence de moindre ampleur en 1920. Nous avons déjà évoqué l'idée que la fragilisation des générations 1918, 1919 et 1920, plus nette pour le sexe féminin que pour le sexe masculin, pouvait en être, au moins en partie, une conséquence. Cela peut paraître paradoxal quand on sait que les individus survivant à une atteinte virale ont pour principale caractéristique d'être définitivement immunisés. On sait cependant aussi que, dans le cas de la grippe, cette immunisation, limitée à une forme provisoire d'un virus extrêmement changeant, est de peu d'intérêt. En revanche, il nous semble que la grippe espagnole peut avoir fragilisé certaines générations de deux manières différentes. D'une part, les nouveau-nés, particulièrement vulnérables, peuvent avoir développé des complications, notamment respiratoires, laissant subsister des séquelles. Mais d'autre part, comme toutes les maladies virales, la grippe espagnole a dû être source de nombre de malformations congénitales chez les femmes enceintes en début de gestation. La grippe espagnole ayant sévi pour l'essentiel à la fin de 1918 et plus modérément au début de 1919, les générations qui peuvent le plus en conserver ainsi la marque à long terme sont celles nées de 1918 à 1920, soit qu'elles aient moins d'un an au moment de l'épidémie soit qu'elles aient été conçues dans les trois mois précédents. De fait, ce sont ces générations qui, tant pour les femmes que pour les hommes, se distinguent fortement des générations encadrantes à la figure 7. Qui plus est, si l'on traite séparément les groupes d'âges au sein de la matrice complète (1899-1981), on se rend compte que la fragilisation de ces générations, et notamment de la génération 1919, était déjà particulièrement sensible entre 0 et 15 ans.

La génération masculine 1919 est certes aussi, parmi les générations combattantes de la seconde guerre mondiale, celle qui nous a semblé pâtir le plus des effets à long terme de la guerre. Elle cumule donc ces effets avec ceux de la grippe. La figure 7a dissimule pourtant en grande partie après lissage sa fragilisation alors que la figure 7b fait nettement ressortir celle des femmes nées la même année. On ne peut que souligner ici le fait que les deux graphiques ne sont pas, de ce point de vue, comparables. Chez les hommes les effets propres aux générations 1918, 1919 et 1920 interviennent dans la partie basse

de l'oscillation induite par les effets à long terme des deux guerres, beaucoup plus importants, et paraissent ainsi minorés. Au contraire, chez les femmes la grippe espagnole constituant le principal facteur de différenciation des générations, ils se dessinent beaucoup plus nettement.

C) Les générations nées dans les années cinquante

Nous avons émis l'hypothèse que la vulnérabilité des générations nées dans les années cinquante observée à la figure 7 pouvait être due à la brusque généralisation de l'accouchement en milieu hospitalier et au développement des interventions obstétricales. Cette hypothèse restera évidemment fragile tant que l'on ne disposera pour ces générations et les générations suivantes que d'une information très partielle (la génération 1959 n'avait que 22 ans en 1981).

On ne doit pas écarter totalement, là non plus, la possibilité de la rémanence d'un bruit de fond. En effet les générations 1950 ont été particulièrement affectées par la montée de la mortalité par accidents de la route observée chez les jeunes jusqu'au milieu des années soixante-dix. La figure 4 a cependant montré que les termes multiplicatifs ont permis de prendre en compte une bonne part de ce phénomène et il y a ici moins de risque que pour les perturbations dues à la guerre puisque la montée des accidents a été plus progressive et qu'elle n'a pas été suivie d'une baisse brutale.

De plus, quand on analyse les effets diagonaux séparément dans chaque groupe d'âge la vulnérabilité des générations cinquante apparaît dès le premier groupe d'âge (0-14 ans) et elle n'est pas particulièrement plus forte dans le groupe 15-29 ans, le plus affecté par la mortalité accidentelle.

Le phénomène est à peu près aussi net chez les femmes que chez les hommes. La figure 7b laisse apparaître des valeurs de $\hat{\theta}_k$ fortement positives pour les générations cinquante qui s'opposent clairement à celles des générations encadrantes même si celles-ci ne sont pas franchement négatives.

La généralisation de l'accouchement hospitalier qui, à nos yeux, est le principal responsable de ce phénomène n'est peut-être pas le seul en cause. C'est par exemple vers la fin des années cinquante que sont apparus les accidents congénitaux liés à l'usage de la thalidomine. Il se peut que l'usage de ce médicament, interdit en France depuis 1961, ait lui-même légèrement renforcé le handicap relatif de ces générations.

III. L'analyse par cause de décès explique-t-elle davantage ?

La fragilisation relative des générations ainsi touchées par certains événements traumatisants a sans doute, à long terme, des conséquences assez générales. On peut cependant se demander si elle n'entraîne pas une sensibilisation plus grande à certaines causes de décès qu'à d'autres, notamment en raison des processus étiologiques qui les déterminent. Nous reprenons

ici, pour ce faire, un reclassement étio-anatomique effectué à l'INED [33,34], en ne retenant que sept grandes catégories étiologiques et trois subdivisions anatomiques particulières :

1. processus infectieux ou parasitaire,
2. malnutrition, intoxication, allergie, etc.,
   a. système nerveux,
   b. foie et voies biliaires,
3. accidents et homicides,
4. tumeurs,
   a. appareil respiratoire,
5. processus héréditaire ou congénital,
6. dégénérescence,
7. suicide.

Les données utilisées proviennent d'un travail de reconstitution de séries cohérentes de décès par cause, de 1925 à 1978, entrepris à l'INED [33]. A ce niveau cependant, les quotients de mortalité utilisés n'ont pas la même précision que ceux qui composent la matrice toutes causes. D'une part la statistique de décès par cause à partir de laquelle ils ont été calculés est fondée sur un regroupement quinquennal d'âges. Les quotients par cause et par année d'âge ont été estimés en appliquant aux quotients toutes causes par année d'âge une interpolation des proportions de décès par cause et par groupe quinquennal d'âges [33]. Il s'agit en outre de quotients du moment (représentant exactement chaque année de calendrier mais chevauchant deux générations consécutives) et non plus de quotients par génération. Dans ce cas, la génération indicée k est en fait une combinaison des générations k et $k-1$. Cette dernière particularité ne présente dans la plupart des cas aucun inconvénient. Elle trouble cependant gravement les résultats pour les générations nées durant les années où la guerre a brutalement modifié le nombre mensuel de naissances. Elle introduit notamment une forte surestimation de l'effet diagonal pour les générations indicées 1915 et 1920 et une sous-estimation pour celles indicées 1916 et 1919. On peut donc dire par avance que cette analyse des causes de décès ne nous apportera aucune explication supplémentaire sur la fragilisation des générations 1918, 1919 et 1920.

A) Une singularité assez générale

Les générations pour lesquelles nous venons de mettre en évidence une relative fragilisation se révèlent en fait être singulières pour la plupart des causes de décès. Cela est particulièrement vrai pour le sexe masculin qui fait aussi l'objet des effets diagonaux les plus importants.

Le tableau 1 présente, pour chacun des groupes de générations masculines pour lesquels nous avons remarqué une fragilisation relative, à l'exception du groupe 1918-1920, la valeur moyenne de $\hat{\theta}_k$, autrement dit le pourcentage approximatif de surmortalité estimée par rapport à la mortalité attendue de la structure rectangulaire de la matrice. A l'exception des accidents et du suicide, pour les générations 1895-1899, chacune des grandes catégories étiologiques utilisées présente des valeurs nettement positives pour les trois

premiers groupes de générations. La situation des générations nées dans les années cinquante est un peu moins évidente. On note cependant pour ces générations une valeur fortement positive pour l'infection et, plus légèrement, pour la dégénérescence. Il n'est pas étonnant de ne pas trouver le même relief pour les tumeurs s'agissant des générations dont la plus ancienne n'atteint pas 28 ans à la fin de la période étudiée. Il est a priori plus surprenant de constater la même absence de relief pour le processus héréditaire ou congénital alors que nous avons émis l'hypothèse d'une fragilisation liée à des malformations congénitales ou à des accidents obstétricaux. Cependant il n'est pas du tout évident qu'au moment de la déclaration du décès, la relation avec ces antécédents soit faite d'une manière suffisamment explicite pour apparaître dans la statistique des causes de décès. En revanche s'agissant de générations encore jeunes il est curieux d'observer une valeur positive pour la dégénérescence. Il est vraisemblable qu'une partie des séquelles d'atteintes congénitales aient été assimilées à des processus de dégénérescence.

Ces exceptions confirmant en quelque sorte la règle, ce qui frappe le plus à la lecture de tableau 1 est donc la similitude des reliefs diagonaux observés chez les hommes pour chacune des grandes catégories étiologiques. Ce résultat n'était pas évident à l'avance mais il peut s'expliquer. On comprend bien que des générations affaiblies pour une raison ou pour une autre résistent plus difficilement à des processus comme l'infection, le cancer ou la dégénérescence,

TABLEAU 1. — EFFET DIAGONAL MOYEN ($100\hat{\theta}_k$) RELEVÉ POUR QUATRE GROUPES DE GÉNÉRATIONS FRAGILISÉES : SEXE MASCULIN

(Les valeurs ci-dessous représentent approximativement le pourcentage de surmortalité estimé par rapport à la mortalité attendue, cf. annexe).

Catégorie étio-anatomique	Groupe de générations singulières(1)				ρ
	1895-1899	1900-1908	1928-1933	1951-1959	
Toutes causes	4,2	3,7	5,1	4,0	2
1. Infection	1,1	3,7	0,3	13,1	2
2. Malnutrition, intoxication, etc.(2)	6,3	10,1	5,1	-	1
a) système nerveux(2)	-0,1	2,7	15,3	-	2
b) foie et voies biliaires(2)	3,0	-1,5	13,4	-	3
3. Accidents	-2,8	1,3	8,0	-1,4	2
4. Tumeurs	4,4	3,6	5,7	-4,2	2
a) appareil respiratoire	12,5	8,0	11,1	-	1
5. Processus héréditaire ou cong.(3)	-	-	-	-2,1	2
6. Dégénérescence	3,3	3,8	3,2	2,8	2
7. Suicide	-3,9	3,7	3,1	-	1

(1) A l'exclusion des générations 1918-1920, pour lesquelles les données par cause sont inadéquates.
(2) L'analyse est restreinte au groupe 25-89 ans, pour éviter qu'elle ne soit perturbée par les données trop erratiques observées avant 25 ans.
(3) L'analyse est restreinte ici au groupe 0-14 ans.

qui ont évidemment moins de prise sur un organisme en parfait état de fonctionnement et de défense. Il en va vraisemblablement de même de l'ensemble malnutrition, intoxication, etc. mais cela ne suffit pas à expliquer que, dans ce cas, l'effet diagonal soit beaucoup plus fort que dans celui de la mortalité toutes causes. Non seulement ces générations, parce qu'affaiblies, résistent moins aux processus de malnutrition ou d'intoxication, mais il faut sans doute admettre que, soit par leur comportement soit en raison de conditions de vies moins bonnes, elles sont en outre plus exposées que les autres à ces processus. De même, dans le cas du suicide et des accidents, et surtout pour les générations 1928-33 qui comptent parmi celles qui ont subi de plein fouet la montée des accidents de la route durant les années cinquante et soixante, on peut se demander si cette surmortalité tient seulement à de moindres chances de survie après un choc, à risque égal de violence, ou s'il tient aussi en partie à une plus grande exposition au risque. Les individus ayant subi dans leur passé de graves traumatismes ou vivant en permanence avec un handicap peuvent notamment avoir une plus grande propension au suicide ou une moindre aptitude à prévenir les accidents.

S'agissant des phénomènes de comportement, on ne peut s'empêcher d'évoquer aussi le tabagisme. Au sein des tumeurs, nous avons isolé les tumeurs de l'appareil respiratoire directement liées à la consommation de tabac. L'effet diagonal est sensiblement plus prononcé pour ces tumeurs que pour les autres, et ce notamment pour les générations ayant participé aux deux guerres mondiales. C'est, certes, parmi les poilus de la Grande Guerre que la consommation de tabac s'est pour la première fois généralisée en France mais il se peut aussi, d'une manière plus générale, que, comme pour le suicide, la propension au tabagisme soit plus grande chez les individus acceptant mal telle ou telle forme de handicap.

B) Mais quelques cas particuliers tout de même

Cette relative uniformité des principales causes de décès dans l'analyse des effets diagonaux ne va cependant pas sans nuance. D'une part, au sein de l'ensemble malnutrition, intoxication, etc., l'alcoolisme joue un rôle particulier et très différent d'un groupe de générations à l'autre. D'autre part, ce que nous constatons à propos du sexe masculin ne s'étend pas au sexe féminin. Enfin un nouveau groupe de générations singulières se dessine.

1) Malnutrition et alcoolisme

La figure 8a représente, pour les hommes, les valeurs $\hat{\theta}_k$ pour l'ensemble de la catégorie étiologique malnutrition, intoxication, etc. résultant d'une analyse restreinte au groupe 25-89 ans. La surmortalité à long terme des générations les plus affectées par les deux guerres mondiales apparaît très clairement. On remarque cependant que pour la première guerre mondiale l'éventail de générations ainsi touchées est plus large que celui auquel nous avons fait référence jusqu'à présent : les valeurs de $\hat{\theta}_k$ sont ici systématiquement positives de la

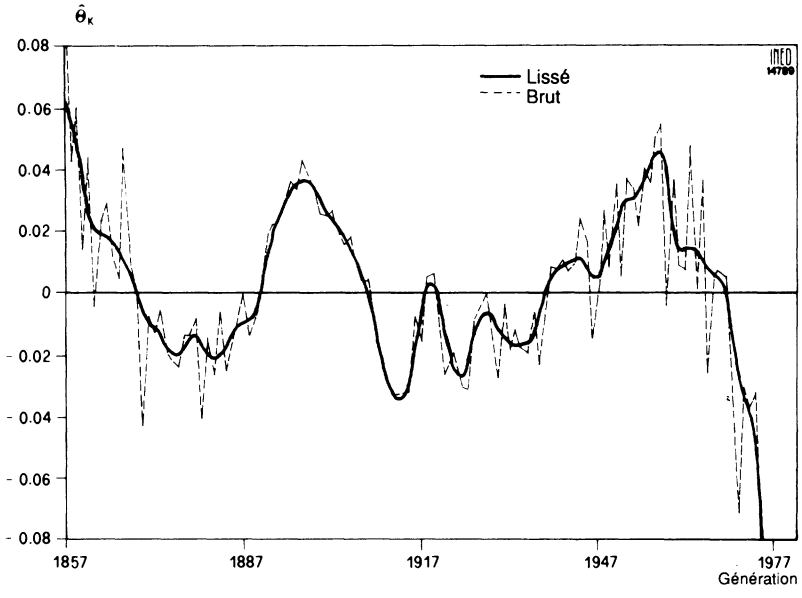

Figure 8a. – Estimation de l'effet diagonal $\hat{\theta}_k$ pour la catégorie étiologique *malnutrition, intoxication, etc.*, à partir des données d'après-guerre (1946-1978), sexe masculin

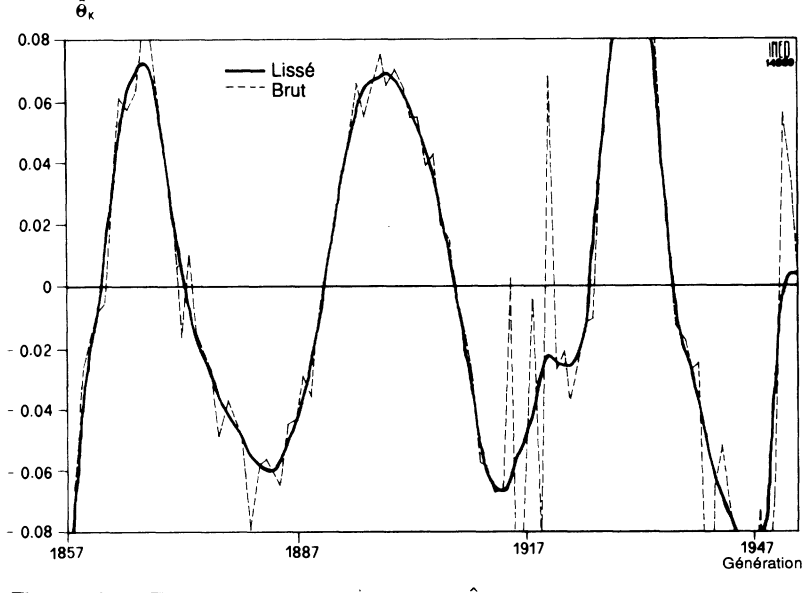

Figure 8b. – Estimation de l'effet diagonal $\hat{\theta}_k$ pour la catégorie étiologique *malnutrition, intoxication, etc.*, à partir des données d'après-guerre (1946-1978), sexe féminin

génération 1894 à la génération 1914. De plus un groupe de générations (1867-1868) que nous n'avons encore jamais mentionné apparaît ici lui-même affecté d'une surmortalité par malnutrition. Même si ces générations se rapprochent d'une zone peu significative du graphique, le contraste avec les générations encadrantes est tellement grand qu'il est peu probable qu'il puisse s'expliquer par le simple hasard.

On sait que l'ensemble étiologique malnutrition, intoxication, etc. est assez hétérogène[5]. Il recouvre en réalité pour l'essentiel deux types de causes spécifiques de mortalité : la misère physiologique par dénutrition et l'ensemble des affections liées à l'abus d'alcool. La part de chacune de ces deux pathologies a très fortement évolué dans le temps. Autrefois, c'était la misère physiologique qui dominait. Aujourd'hui, c'est l'alcoolisme.

Cela permet de comprendre pourquoi au tableau 1 l'effet diagonal est très fort pour chacun des trois premiers groupes de générations quand on considère l'ensemble de la catégorie étiologique et qu'il n'apparaît que progressivement pour prendre son plein effet dans le troisième groupe de générations lorsqu'on isole les atteintes du système nerveux (essentiellement la psychose alcoolique) et du foie (cirrhose alcoolique). La mortalité liée à l'alcool a très fortement augmenté en France dans les quinze années qui ont suivi la seconde guerre mondiale et, bien qu'elle ait atteint son maximum dans les années soixante, elle reste depuis très élevée [31,33]. Les générations affaiblies par la première guerre mondiale n'ont connu qu'assez tardivement cette montée de l'alcoolisme alors que les générations 1928-1933 en subissent pleinement les effets.

Concernant l'alcool lui-même on note une importante différence entre les atteintes du système nerveux et celles du foie. Dans les deux cas (figures 9a et 9b) l'effet à long terme de la seconde guerre est bien visible. En revanche, celui de la première est absent (ou légèrement négatif) pour la cirrhose du foie, alors qu'il apparaît assez nettement pour la psychose alcoolique. Cependant, dans ce dernier cas il s'agit surtout des générations 1905-1910, et donc d'un groupe nettement plus jeune que celui auquel nous faisons généralement référence. Le développement de la cirrhose du foie exige une longue imprégnation alcoolique alors que la psychose alcoolique est plus souvent la conséquence d'effets plus rapides de l'alcool (*delirium tremens*). On comprend ainsi que l'effet diagonal observé pour la cirrhose ne concerne que les générations affaiblies par la seconde guerre alors que celui que l'on observe pour la psychose alcoolique atteint aussi les plus jeunes des générations affaiblies par la première, les générations les plus anciennes restant en dehors de ce phénomène typique de l'après-deuxième-guerre.

L'effet diagonal observé pour la catégorie étiologique malnutrition, etc. dans les générations 1894-1904 tient fort peu à l'alcoolisme et doit donc être lié à la misère physiologique. Il en va de même de l'effet observé pour les générations 1867-1868. En revanche, pour les générations 1928-1933, les valeurs moyennes de $\hat{\theta}_k$ pour les atteintes alcooliques sont nettement plus fortes que

[5] Le principe de cette catégorie est de regrouper les affections se caractérisant par une réponse de l'organisme au contact (interne ou externe) avec des substances étrangères.

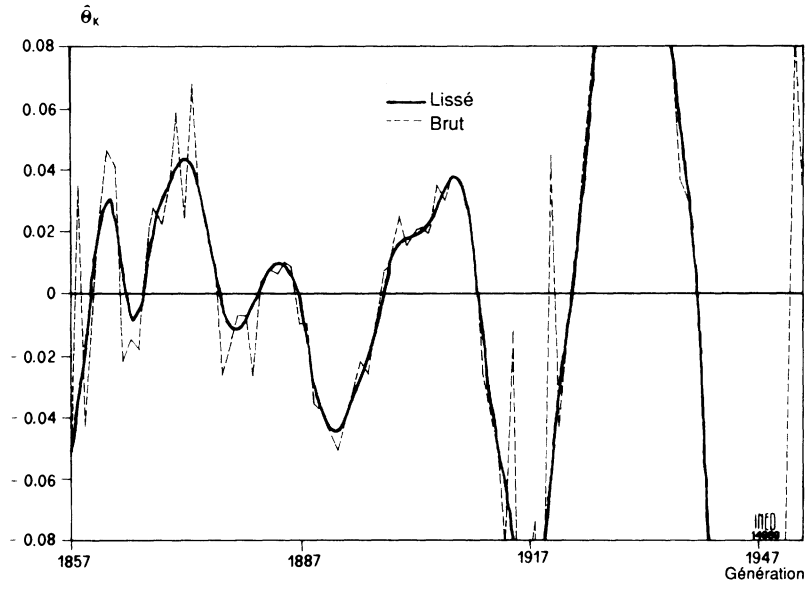

Figure 9a. – Estimation de l'effet diagonal $\hat{\theta}_k$ pour les atteintes du *système nerveux* par *malnutrition, intoxication, etc.*, à partir des données d'après-guerre (1946-1978), sexe masculin

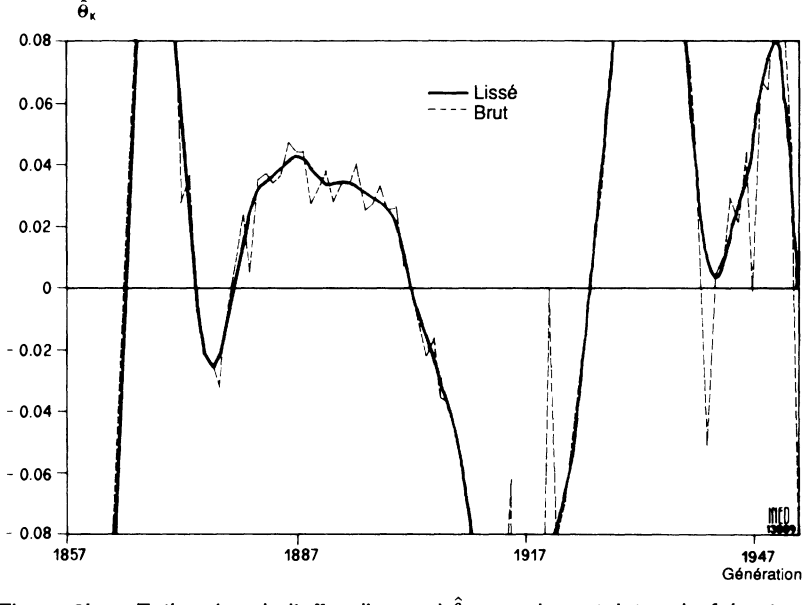

Figure 9b. – Estimation de l'effet diagonal $\hat{\theta}_k$ pour les atteintes du *foie et voies biliaires* par *malnutrition, intoxication, etc.*, à partir des données d'après-guerre (1946-1978), sexe masculin

pour l'ensemble de la catégorie étiologique (tableau 1), ce qui souligne le rôle spécifique de l'alcool dans la surmortalité de ces générations.

2) Le cas des femmes

Moins apparente au niveau de la mortalité totale, la disparité entre les générations féminines est aussi variable d'une cause à l'autre. On est notamment frappé à la lecture du tableau 2 par le fait que même pour les groupes de générations dont la fragilisation n'est pas très marquée on observe une surmortalité assez importante pour certaines causes. Les deux exemples les plus notoires concernent d'une part la catégorie étiologique malnutrition, intoxication, etc. pour laquelle on observe une valeur nettement positive de $\hat{\theta}_k$ dans les trois groupes de générations affectées par les guerres et d'autre part les générations nées dans les années cinquante pour lesquelles la surmortalité déjà significative au total apparaît encore plus clairement pour l'infection et la dégénérescence.

La figure 8b qui illustre, dans le cas des femmes, les valeurs de $\hat{\theta}_k$ pour la catégorie étiologique malnutrition, intoxication, etc. offre en effet un profil très semblable à celui de la figure 8a pour les hommes. Bien que se situant à un niveau beaucoup plus bas pour les femmes que pour les hommes, la mortalité par alcoolisme a suivi exactement la même évolution pour les deux sexes et tout ce que nous avons dit pour les hommes à propos des rôles respectifs de la misère physiologique et de l'alcoolisme s'applique vraisemblablement aussi aux femmes. On retrouve de plus le cas particulier déjà évoqué à propos des hommes des générations 1867-1868. Reste à comprendre pourquoi chez

TABLEAU 2 - EFFET DIAGONAL MOYEN RELEVÉ POUR QUATRE GROUPES DE GÉNÉRATIONS FRAGILISÉES : SEXE FÉMININ

(Les valeurs ci-dessous représentent approximativement le pourcentage de surmortalité estimé par rapport à la mortalité attendue, cf. annexe).

Catégorie étiologique	Groupe de générations singulières(1)				ρ
	1891-1899	1900-1908	1928-1933	1951-1959	
Toutes causes	3,6	1,6	-1,3	3,0	2
1. Infection	-10,2	-5,3	2,8	7,1	2
2. Malnutrition, intoxication, etc.(2)	6,5	4,0	9,1	-	2
3. Accidents	-3,7	-4,9	2,8	3,9	2
4. Tumeurs	-2,0	-2,1	-0,9	-3,2	2
5. Processus héréditaire ou cong.(3)	-	-	-	-0,4	2
6. Dégénérescence	7,0	4,6	-4,6	7,0	2
7. Suicide	-1,8	1,1	-1,0	-	1

(1) A l'exclusion des générations 1918-1920, pour lesquelles les données par cause sont inadéquates.
(2) L'analyse est restreinte au groupe 25-89 ans, pour éviter qu'elle ne soit perturbée par les données trop erratiques observées avant 25 ans.
(3) L'analyse est restreinte ici au groupe 0-14 ans.

les femmes ce relief diagonal apparaît plus nettement pour cette catégorie étiologique que pour les autres. Cela reste assez mystérieux.

S'agissant des générations féminines nées dans les années cinquante la prise en considération des causes de décès est particulièrement intéressante dans la mesure où elle confirme l'existence de certaines surmortalités spécifiques là où on pouvait hésiter à affirmer celle d'une surmortalité globale. Qui plus est, on retrouve parmi les causes faisant l'objet de cette surmortalité les trois principales causes déjà soulignées pour le sexe masculin : infection, accidents et dégénérescence.

3) Le cas des générations 1867-1868

La figure 8 a mis en relief, tant pour les femmes que pour les hommes, un effet diagonal non négligeable portant sur les générations 1867-1868. Une lecture attentive de la figure 6 permettait d'emblée de constater, quel que soit le sexe, une valeur élevée de $\hat{\theta}_k$ pour les générations 1867 et 1868. Il ne nous avait cependant pas paru prudent de nous y apesantir dès le départ compte tenu de leur emplacement dans une zone extrême de la matrice de données. Une fois encore l'analyse des causes de décès permet de vérifier l'existence de situations dont la singularité n'était pas absolument évidente au niveau de la mortalité totale. On retrouve en effet cette particularité des générations nées autour des années 1867-1868 avec l'infection et, quoique moins nettement, avec la dégénérescence. Peut-on imaginer que l'épidémie de choléra de 1867-1868 ait pu avoir un effet à long terme sur les générations qui l'ont subie pendant leur première enfance? Ce n'est qu'une hypothèse à vérifier.

En guise de conclusion

Sans nullement prétendre clore définitivement le débat, on peut tenter d'avancer quelques conclusions sur l'apport de cette nouvelle méthode de description des données, sur la validité des reliefs ainsi mis en évidence par le modèle utilisé et sur l'hypothèse fondamentale retenue pour leur interprétation.

L'exploration systématique de la structure de la matrice des quotients de mortalité nous a permis non seulement de retrouver les caractéristiques particulières déjà mises en évidence par de nombreux auteurs mais aussi de faire apparaître de nouvelles aspérités affectant d'autres générations singulières. Ainsi avons-nous pu souligner la surmortalité relative des générations nées ou conçues pendant la grippe espagnole de 1918-1919 ou de celles nées dans les années cinquante et même évoquer le cas particulier des générations nées aux alentours de 1867-1868. Il nous semble donc que la méthode utilisée ici présente au moins l'avantage de décrire plus complètement les données disponibles. Par ailleurs l'extension de l'analyse aux causes de décès a permis de constater qu'au-delà des effets majeurs affectant toutes les causes et apparaissant donc clairement au niveau de la mortalité totale il existe aussi des phénomènes plus spécifiques qui, bien que limitant leurs effets à certaines causes, n'en sont pas moins significatifs.

On peut cependant se demander si cette description reflète convenablement la réalité. De ce point de vue, on doit s'interroger tant sur les données elles-mêmes que sur la méthode. S'agissant des données on peut être assez confiant dans la mesure où la même méthode appliquée à d'autres pays ayant connu des histoires semblables fait ressortir les mêmes phénomènes [39]. Ajoutons à cela que les données françaises nous paraissent particulièrement bien adaptées à ce genre d'analyse dans la mesure où elles reposent sur un bon enregistrement des décès par année d'âge et année de naissance et sur une série abondante de recensements à des dates irrégulières[6]. S'agissant de la méthode, nous avons vu qu'elle confirme généralement les résultats déjà mis en évidence par d'autres auteurs avec d'autres méthodes à partir des mêmes données. Mais, qui plus est, une application préalable [39] portant soit sur un réseau original de tables types de mortalité soit sur ce même réseau modifié pour y introduire des aspérités (régulières ou aléatoires), a permis de lever les doutes : aucun effet diagonal n'est apparu ni dans le réseau original ni dans le réseau perturbé aléatoirement alors que les effets diagonaux simulés sont au contraire bien ressortis.

Il reste que nous avons délibérément préféré l'hypothèse fondamentale d'une fragilisation de générations soumises à certains types de traumatismes et se singularisant dans notre analyse par des valeurs positives de $\hat{\theta}_k$ à celle d'une sélection d'autres générations ayant subi d'autres types de traumatismes et se distinguant ici par des valeurs négatives de $\hat{\theta}_k$. Nous avons ainsi, par exemple, préféré considérer que les générations ayant vécu leur adolescence durant la guerre 1914-1918 ont été fragilisées plutôt que de retenir l'idée d'une sélection par leur surmortalité infantile des générations nées pendant la guerre. Comme nous l'avons cependant noté, l'hypothèse d'un possible effet de sélection ne peut être totalement écartée lorsque l'on compare entre plusieurs pays belligérants les générations qui ont participé aux deux guerres mondiales. Il ne nous semble cependant pas que dans le cas présent cette hypothèse puisse fournir une explication aux différences observées et il serait nécessaire d'approfondir l'analyse comparative pour affirmer qu'elle est réellement à la clé de certaines différences dans d'autres pays.

Insistons enfin sur le fait que les effets diagonaux que nous avons mis ici en rapport avec l'histoire propre de certaines générations singulières ne représentent en réalité que la fraction émergée de l'iceberg que constituent les effets de génération au sens complet du terme. Cependant cette approche de l'hétérogénéité entre générations nous a permis de mieux comprendre l'importance de cette troisième composante et une partie des relations qu'elle entretient avec la structure et l'évolution de la mortalité.

John WILMOTH, Jacques VALLIN et Graziella CASELLI.

---

[6] Durant la période de 37 ans que nous avons plus particulièrement explorée, six recensements ont eu lieu : 1946, 1954, 1962, 1975 et 1982. Reposant sur des intervalles intercensitaires aussi irréguliers, les reconstitutions de population par âge au premier janvier de chaque année risquent fort peu d'être perturbées même aux âges élevés par une quelconque attraction pour les âges ronds.

ANNEXE MÉTHODOLOGIQUE

Cette annexe fournit une brève présentation de la méthode employée dans le texte principal. Une description plus complète de cette dernière a été donnée ailleurs [39] et le lecteur peut aussi se référer à d'autres publications traitant du problème que pose l'analyse de matrices non additives [11,12,14,17,22,27].

Comme nous le disons dans l'article, la famille de modèles qui nous intéresse ici est de la forme :

$$f_{ij} = \alpha_i + \beta_j + \sum_{m=1}^{\rho} \phi_m \gamma_{im} \delta_{jm} + \theta_k + \epsilon_{ij} \qquad (1)$$

où $i = 1, 2, ..., I$ représente les années d'âge, $j = 1, 2, ..., J$ les années de calendrier et $k = j - i$ les générations. L'application de cette famille de modèles à l'analyse d'une matrice pose trois questions : le choix du modèle, la procédure d'ajustement (qui appellera également une discussion spécifique sur la décomposition en valeurs singulières) et l'interprétation de l'effet diagonal.

1. Le choix du modèle

Il nous faut choisir, au sein de la famille définie ci-dessus, le modèle adéquat en déterminant le nombre approprié, m, de termes mutiplicatifs. On peut y parvenir en ne considérant au départ qu'une famille de modèles AM, faisant donc abstraction du terme diagonal θ_k.

Nous déterminons le nombre de termes multiplicatifs nécessaires en essayant successivement plusieurs modèles AM, en jugeant, pour chaque valeur de ρ, de la qualité de l'ajustement obtenu et de la structure du résidu qui en résulte. Il y a ainsi deux méthodes distinctes de détermination du modèle AM adéquat : l'une considère la part de variance expliquée par les modèles successifs, l'autre s'attache à la taille moyenne des résidus en ligne et en colonne obtenue dans chaque cas.

On percevra mieux ces deux critères de choix à travers un exemple. Considérons la matrice de mortalité masculine française pour les années 1946-1981, de 0 à 89 ans. Le tableau 3 donne la part de variance dont rendent compte successivement les modèles AM(0) à AM(3). Il est clair que cette matrice est très largement additive, puisque le terme additif explique à lui seul plus de 99% de la variance totale. Cependant, le fait que les deux premiers termes multiplicatifs expliquent chacun plus de 60% de la variance restante indique l'existence d'une structure non additive distincte caractérisant certaines colonnes et certaines lignes, dont rendent assez bien compte ces deux termes multiplicatifs. La chute brutale du pouvoir explicatif des termes multiplicatifs à partir du troisième montre au contraire qu'au-delà du deuxième terme multiplicatif, le résidu ne recèle plus guère de structure rectangulaire et il nous semble judicieux de retenir dans ce cas un modèle AM(2).

L'expérience a montré que, pour que le résidu comporte une structure non additive significative, la prise en compte d'un terme multiplicatif supplémentaire doit expliquer au moins 50% de la variance restante. C'est cependant une règle d'usage qu'il ne faut pas prendre au pied de la lettre mais qu'il convient d'appliquer en tenant compte de la structure des résidus résultant des différents modèles, notre second critère de choix.

TABLEAU 3. — PART DE LA VARIANCE EXPLIQUÉE PAR DIFFÉRENTS MODÈLES DANS LE CAS DE LA MORTALITÉ MASCULINE FRANÇAISE DE 0 À 89 ANS ET DE 1946 À 1981

Dernier terme du modèle	Part (%) de la variance résiduelle	Part (%) de la variance totale	Part cumulée(%) de la variance totale
Terme additif	99,496	99,496	99,496
1er terme multiplicatif	65,444	0,330	99,826
2e terme multiplicatif	62,528	0,109	99,935
3e terme multiplicatif	23,026	0,015	99,950

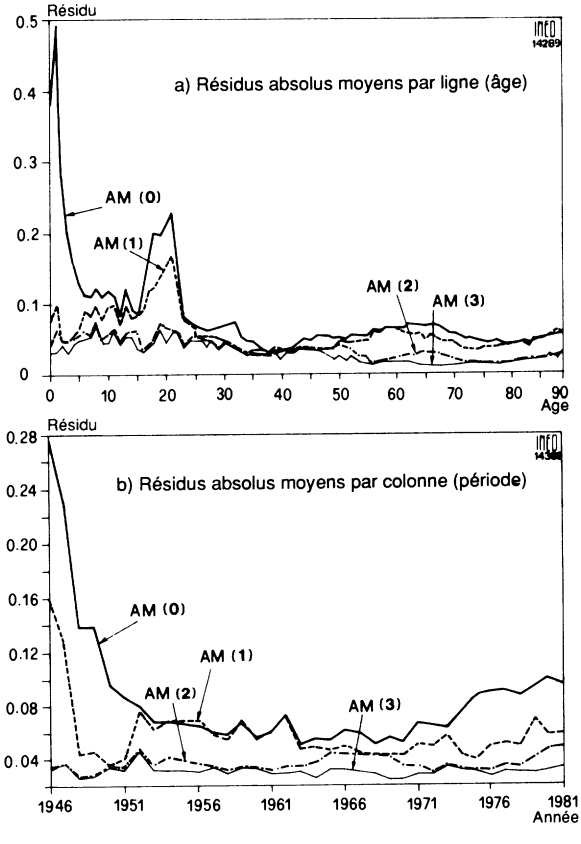

Figure 10. – Résidus résultant de l'application de différents modèles (AM(0) à AM(3)) à la mortalité masculine française 1946-1981, de 0 à 89 ans

La figure 10 montre les modifications successives du résidu absolu moyen par ligne (figure 10a) et par colonne (figure 10b) obtenues à mesure que l'on ajoute au modèle le premier puis le deuxième puis le troisième terme multiplicatif. Avec le seul terme additif, le modèle rend assez bien compte de la structure par âge sauf avant 22 ou 23 ans. Le premier terme multiplicatif semble rendre compte de la non additivité observée aux plus jeunes âges tandis que le second rend compte de celle qui marque les âges situés autour de 20 ans. S'agissant des années de calendrier, la non additivité caractérise surtout les premières années (1946-1950) et les dernières (1979-1981). Elle est presque entièrement absorbée par les deux premiers termes multiplicatifs.

Nous devons aussi juger de ce deuxième critère en examinant la distribution des résidus absolus moyens par ligne et par colonne.

La figure 11 donne une représentation en forme d'arbre[7] de la distribution des

a) Modèle additif simple

Nombre (n=90)		Résidus (unité= 0,01)
5	03	13367
25	04	00012345566677788899
45	05	00112334444455566889
45	06	03444445666677899999
26	07	25
24	08	446
21	09	5
	10	
20	11	001348
14	12	33
12	13	1
	14	
11	15	18
9	16	0
	17	
	18	
8	19	69
6	20	19
	21	
	22	
4	23	1
	HI	0,269 0,340 0,494

b) Modèle AM (1)

Nombre (n=90)		Résidus (unité= 0,01)
25	03	00001344455555667777788889
(23)	04	00112334444556777788999
42	05	01245556678888
28	06	0134444
21	07	2389
17	08	0144
13	09	1157799
7	10	0
6	11	7
5	12	33
3	13	7
	14	
2	15	4
1	16	7

c) Modèle AM (2)

Nombre (n=90)		Résidus (unité= 0,01)
13	01	4445666677889
31	02	001223344445566779
(28)	03	00000012233444555667778888889
31	04	0011257778889
18	05	00455788
10	06	02233445
2	07	25

15189

Figure 11. – Représentation en arbre de la distribution des résidus en ligne pour trois versions du modèle

[7] La distribution en forme d'arbre [27] est semblable au traditionnel histogramme mais avec l'avantage de contenir plus d'informations et de pouvoir être facilement construite avec une simple machine à écrire. La partie principale en est en effet constituée par un histogramme renversé sur le côté et construit de la façon suivante : après avoir opéré le changement d'unité adéquat, chaque observation est repérable, en face de l'une des valeurs entières portées sur la colonne immédiatement à gauche de la ligne verticale par sa première décimale mentionnée à droite de la ligne verticale. Les unités de "branche" et de "feuille" peuvent être utilisées pour retrouver les valeurs originales des données : à la figure 11a, la branche 15 a deux feuilles qui correspondent respectivement aux valeurs 0,151 et 0,158. Les valeurs marquées "HI" ou "LO", placées sous cette partie principale du graphique, sont celles des observations extrêmes (dans leur échelle d'origine). La première colonne de gauche donne le nombre d'observations repérées à chaque ligne, cumulées jusqu'à la moitié en partant du haut puis cumulées en partant du bas : elle sert à déterminer la médiane ou les quantiles de la distribution.

résidus absolus moyens par ligne à trois étapes successives de l'ajustement [27]. Avec le terme additif seul (figure 11a), la majorité des résidus absolus moyens se concentre entre les valeurs 0,03 et 0,07. Néanmoins quelques-uns dépassent 0,10 et les trois valeurs extrêmes signalées sous la distribution dépassent 0,25. Cette très longue queue de distribution est fortement réduite lorsqu'on ajoute un premier terme multiplicatif (figure 11b) mais il faut l'intervention du second terme multiplicatif pour qu'elle soit complètement résorbée (figure 11c).

Ainsi, dans cet exemple, tant le premier que le second critère conduisent à retenir deux termes multiplicatifs pour rendre compte de la structure rectangulaire de la matrice. On a aussi remarqué [39] qu'il pouvait être utile d'examiner les ensembles de paramètres γ_{im} et δ_{jm} pour déterminer le type de structure absorbée par chacun des termes multiplicatifs successifs mis en œuvre. Il est alors possible de vérifier que les troisième et quatrième termes multiplicatifs décrivent en fait une part de la structure diagonale, qu'il est préférable de réserver ce rôle au terme diagonal θ_k, que ces troisième et quatrième termes multiplicatifs peuvent être éliminés lorsqu'on utilise un modèle AMD.

Il est en outre assez rassurant de constater que la procédure itérative que nous employons pour ajuster le modèle est assez robuste par rapport à la valeur choisie pour ρ, dès lors que ρ est assez grand. Dans le cas qui nous intéresse ici, l'algorithme produit des estimations semblables de θ_k pour les valeurs 2, 3 ou 4 de ρ, en diminuant graduellement l'amplitude des termes multiplicatifs inutiles (en l'occurrence, les troisième et quatrième) et en redonnant au fur et à mesure à la structure diagonale le relief qui lui manquait à la première itération. En cas de doute sur le choix de la valeur de ρ, il vaut donc mieux s'aventurer un peu vers des valeurs trop grandes que de s'en tenir à des valeurs trop petites.

2. *La méthode d'ajustement*

Pour discuter du choix du modèle, nous avons écarté la question de savoir comment obtenir, à chaque étape de la procédure, l'estimation des paramètres du modèle AM pour une valeur donnée de ρ. Les ajustements de type AM utilisés ici sont fondés sur la minimisation de la somme des carrés des erreurs et sont donc relativement simples à effectuer. On trouvera ailleurs une discussion des avantages comparés de cette méthode des moindres carrés et de méthodes alternatives [11,39].

Le modèle AM peut s'écrire :

$$f_{ij} = \alpha_i + \beta_j + \sum_{m=1}^{\rho} \phi_m \gamma_{im} \delta_{jm} + \epsilon_{ij} \qquad (2)$$

et suppose les contraintes d'identification suivantes[8] :

[8] Le choix de certaines de ces contraintes est évidemment un peu arbitraire bien qu'il ne puisse d'aucune manière changer l'interprétation des résultats comme c'est parfois le cas dans d'autres modèles APG (âge-période-génération) plus classiques. Par exemple, on peut décider une somme nulle pour les α plutôt que pour les β. Cela changerait l'interprétation de chacun de ces deux paramètres mais non celle de la description de la structure additive de la matrice que donne leur combinaison. On peut dire la même chose des γ et des δ. En particulier, les contraintes d'orthogonalité sont plutôt une commodité qu'une nécessité : une matrice de rang ρ peut être décomposée d'un nombre infini de manières et la décomposition orthogonale (impliquée par la décomposition en valeurs simples) est loin d'être la meilleure solution.

$$\sum_j \beta_j = 0, \tag{3}$$

$$\sum_i \gamma_{im} = \sum_j \delta_{jm} = 0, \tag{4}$$

$$\sum_i \gamma_{im}^2 = \sum_j \delta_{jm}^2 = 1, \tag{5}$$

$$\sum_i \gamma_{im}\gamma_{im'} = \sum_j \delta_{jm}\delta_{jm'} = 0, \tag{6}$$

$$\phi_1 > \phi_2 > \ldots > \phi_\rho, \tag{7}$$

où $1 \leq m \leq \rho$ et $m \neq m'$.

Compte tenu de ces contraintes, il est simple de vérifier que les solutions par les moindres carrés sont données pour α_i et β_j, par :

$$\hat{\alpha}_i = \frac{1}{J}\sum_j f_{ij} \tag{8}$$

et

$$\hat{\beta}_j = \frac{1}{I}\sum_i (f_{ij} - \hat{\alpha}_i).$$

Dès lors, en définissant

$$r_{ij} = f_{ij} - \hat{\alpha}_i - \hat{\beta}_j, \tag{9}$$

les termes multiplicatifs sont déduits des ρ premiers termes de la décomposition en valeurs singulières (DVS, voir ci-dessous) de la matrice des résidus (r_{ij}). Cette étape peut s'écrire :

$$(\hat{\phi}_m, \hat{\gamma}_{im}, \hat{\delta}_{jm}) = DVS_\rho(r_{ij}). \tag{10}$$

Plus précisément, les $\hat{\phi}_m$ sont les ρ valeurs singulières, tandis que $\hat{\gamma}_{im}$ et $\hat{\delta}_{jm}$ sont respectivement les éléments des vecteurs singuliers de gauche et de droite correspondants. L'ajustement d'un modèle AMD par les moindres carrés est plus compliqué et requiert une démarche itérative. De plus, nous devons imposer à l'effet diagonal la contrainte suivante[9] :

$$\sum_k w_k\theta_k = \sum_k k w_k\theta_k = 0, \tag{11}$$

où w_k est égal au nombre d'observations figurant sur la k^e diagonale.

[9] Les équations qui suivent indiquent que le niveau et la pente des $\hat{\theta}_k$ sont nuls. Plus précisément, la droite d'ajustement linéaire par les moindres carrés des $\hat{\theta}_k$ coincide avec l'axe horizontal.

Les équations normales peuvent s'écrire :

$$\hat{\alpha}_i = \frac{1}{J}\sum_j (f_{ij} - \hat{\theta}_k) \tag{12}$$

$$\hat{\beta}_j = \frac{1}{I}\sum_i (f_{ij} - \hat{\alpha}_i - \hat{\theta}_k) \tag{13}$$

$$(\hat{\phi}_m, \hat{\gamma}_{im}, \hat{\delta}_{jm}) = DVS_\rho(f_{ij} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\theta}_k) \tag{14}$$

$$\hat{\theta}_k = \frac{1}{w_k}\sum_{(k)} (f_{ij} - \hat{\alpha}_i - \hat{\beta}_j - \sum_{m=1}^{\rho} \hat{\phi}_m \hat{\gamma}_{im} \hat{\delta}_{jm}), \tag{15}$$

où l'on utilise $\sum_{(k)}$ pour représenter la somme de tous les éléments de la k^e diagonale. Les solutions sont obtenues par itération comme suit :

$$\alpha_i^{(n)} \leftarrow \frac{1}{J}\sum_j (f_{ij} - \theta_k^{(n-1)}) \tag{16}$$

$$\beta_j^{(n)} \leftarrow \frac{1}{I}\sum_i (f_{ij} - \alpha_i^{(n)} - \theta_k^{(n-1)}) \tag{17}$$

$$(\phi_m^{(n)}, \gamma_{im}^{(n)}, \delta_{jm}^{(n)}) \leftarrow DVS_\rho(f_{ij} - \alpha_i^{(n)} - \beta_j^{(n)} - \theta_k^{(n-1)}) \tag{18}$$

$$\theta_k^{(n)} \leftarrow \frac{1}{w_k}\sum_{(k)} (f_{ij} - \alpha_i^{(n)} - \beta_j^{(n)} - \sum_{m=1}^{\rho} \phi_m^{(n)} \gamma_{im}^{(n)} \delta_{jm}^{(n)}) \tag{19}$$

où l'on définit $\theta_k^{(0)} \equiv 0$ quel que soit k, afin de démarrer le calcul à $n = 1$. Nous arrêtons l'itération quand les différences entre deux estimations successives de θ_k deviennent suffisamment petites. Nous calculons la moyenne pondérée des différences absolues observées entre n et $n-1$:

$$\frac{\sum_k w_k |\theta_k^{(n)} - \theta_k^{(n-1)}|}{\sum_k w_k} = \frac{1}{IJ}\sum_k w_k |\theta_k^{(n)} - \theta_k^{(n-1)}|, \tag{20}$$

et nous mettons fin à l'algorithme quand cette moyenne devient inférieure à un niveau prédéterminé (0,0005 dans le cas des analyses présentées ici). Il est également utile de fixer un nombre maximum d'itérations. Nous avons choisi 25.

Il faut noter que les contraintes imposées par les équations (3), (4), (5), (6) et (10) ne suffisent pas à identifier un jeu unique d'estimations des paramètres. Il y a généralement une infinité de solutions possibles permettant de satisfaire aux équations normales (11), (12), (13) et (14) [39]. A l'intérieur de ce lot de solutions possibles (dont toutes satisfont aux équations normales et donc minimisent la somme des carrés des erreurs), les estimations données par l'algorithme ci-dessus ont la propriété supplémentaire de minimiser la somme

$$\sum_k w_k \theta_k^2. \tag{21}$$

Nous devons cependant préciser que si ce résultat a pu être observé empiriquement avec une grande constance [39], nous n'avons pas encore pu démontrer mathématiquement que ledit algorithme converge toujours vers la solution qui minimise l'expression (20).

Plus importante que la rigueur mathématique, l'interprétation d'une telle propriété doit retenir notre attention. En termes simples : alors qu'il y a une infinité de décompositions de type AMD équivalentes possibles pour une même matrice de données initiale (équivalentes en ce sens que toutes ont le même ρ et que toutes minimisent la somme des carrés des erreurs), l'algorithme que nous employons aboutit à celle qui donne les poids les plus faibles au terme diagonal. Cela veut dire que nous avons choisi de décrire la matrice aussi complètement que possible à travers deux dimensions (lignes et colonnes) et que nous n'avons introduit de troisième dimension (diagonales) que pour les structures qui ne pourraient que très difficilement être expliquées par les deux premières dimensions. Comme nous le faisons remarquer dans l'article, le terme diagonal a été en partie introduit pour réduire le nombre de termes multiplicatifs nécessaires à une description complète des données. Une fois cette réduction obtenue, cependant, nous n'en avons pas moins choisi de privilégier le rôle des termes additifs et multiplicatifs afin de n'isoler à travers le terme diagonal que des structures reflétant sans ambiguïté l'influence de facteurs propres aux générations.

3. La décomposition en valeurs singulières

Nous avons fait référence au point précédent à la procédure bien connue de **décomposition en valeurs singulières (DVS)** de matrices réelles [12]. Résumons ici quelques-unes de ses propriétés élémentaires avant de mettre en avant deux procédés de calcul. Soit $\mathbf{Y}$ une matrice réelle quelconque de rang r. $\mathbf{Y}$ peut s'entendre comme la somme de r matrices de rang 1 d'une infinité de façons. Une des manières d'écrire une telle matrice est fournie par la DVS, sous la forme :

$$\mathbf{Y} = \sum_{a=1}^{r} \lambda_a \mathbf{p}_a \mathbf{q}_a', \tag{22}$$

où les nombres $\lambda_1, \lambda_2, ..., \lambda_r$ sont les *valeurs singulières* de $\mathbf{Y}$, tandis que les vecteurs $\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_r$ et les vecteurs $\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_r$ sont les *vecteurs singuliers* gauches et droits. Les vecteurs singuliers forment deux ensembles orthonormaux. Les contraintes (5) et (6) sont ainsi satisfaites, puisque les éléments des vecteurs $\mathbf{p}_a$ et $\mathbf{q}_a$ deviennent respectivement les γ_{im} et δ_{jm} de notre modèle. De la même façon, les λ_a deviennent ϕ_m et $\mathbf{Y}$ représente aussi bien la matrice (r_{ij}) que la matrice $(r_{ij} - \hat{\theta}_k)$.

Il est possible d'obtenir une DVS de deux manières différentes. Une première méthode conduit aux valeurs propres et aux vecteurs propres soit de $\mathbf{YY}'$ soit de $\mathbf{Y}'\mathbf{Y}$; les valeurs singulières λ_a sont données par les racines carrées des valeurs propres positives de l'une ou l'autre matrice, tandis que les vecteurs singuliers droit et gauche $\mathbf{p}_a$ et $\mathbf{q}_a$ coïncident respectivement avec les vecteurs propres des matrices $\mathbf{YY}'$ et $\mathbf{Y}'\mathbf{Y}$.

Puisque λ_a, $\mathbf{p}_a$ et $\mathbf{q}_a$ suivent la relation :

$$\lambda_a = \mathbf{p}_a' \mathbf{Y} \mathbf{q}_a, \tag{23}$$

il suffit de trouver un seul jeu de vecteurs singuliers et de calculer le second en applicant la formule (22).

Un second procédé de DVS s'avère très utile quand on n'a besoin que des tout premiers termes de la décomposition. On peut en effet calculer $\mathbf{p}_1$ et $\mathbf{q}_1$ par itérations en employant les équations suivantes :

$$\mathbf{p}_1^{(n)} \leftarrow \mathbf{Y}\mathbf{q}_1^{(n-1)}, \tag{24}$$

$$\mathbf{p}_1^{(n)} \leftarrow \mathbf{p}_1^{(n)} / \| \mathbf{p}_1^{(n)} \|, \tag{25}$$

$$\mathbf{q}_1^{(n)} \leftarrow \mathbf{p}_1^{\prime(n)}\mathbf{Y}, \tag{26}$$

$$\mathbf{q}_1^{(n)} \leftarrow \mathbf{q}_1^{(n)} / \| \mathbf{q}_1^{(n)} \|, \tag{27}$$

où, pour un $s \times 1$ vecteur $\mathbf{x}$, la valeur $\|\mathbf{x}\|$ indique la valeur normée de $\mathbf{x}$ définie comme suit :

$$\| \mathbf{x} \| = \sqrt{x_1^2 + x_2^2 + \ldots + x_s^2} \tag{28}$$

On choisit arbitrairement, pour le vecteur $\mathbf{q}_1^{(0)}$, un point de départ $\mathbf{q}_1^{(0)} \equiv 0$. Quand on atteint la convergence, λ_1 est donné par l'équation (22). Ayant trouvé λ_1, $\mathbf{p}_1$ et $\mathbf{q}_1$ de cette façon, on forme $\mathbf{Y} - \lambda_1 \mathbf{p}_1 \mathbf{q}_1'$ et on répète l'opération pour trouver λ_2, $\mathbf{p}_2$ et $\mathbf{q}_2$ et ainsi de suite.

Il faut ici souligner les ressemblances et les différences entre la DVS et l'analyse en composantes principales. Dans le premier cas, la matrice $\mathbf{Y}$ est décomposée selon la formule (21), dans le second, la décomposition est effectuée sur la matrice des covariances $\mathbf{YY}'$[10]. Les relations entre les valeurs et les vecteurs propres de $\mathbf{YY}'$ (ou de $\mathbf{Y}'\mathbf{Y}$) et les paramètres de la DVS ont été décrits ci-dessus et l'équivalence mathématique des deux techniques est évidente. Cependant [12], la DVS a sur l'analyse en composantes principales deux principaux avantages. En premier lieu, la DVS (seconde approche) requiert moins de calculs. Deuxièmement, la DVS nous livre davantage d'information dans la mesure où les vecteurs $\mathbf{p}_a$ sont associés à des vecteurs $\mathbf{q}_a$ (propres à la DVS), avec lesquels ils forment des paires de vecteurs qui donnent une description conjointe de la matrice $\mathbf{Y}$ à la fois plus accessible et plus complète.

4. L'interprétation des effets diagonaux

Nous définissons deux estimations, ou valeurs attendues, de f_{ij} comme suit :

$$\hat{f}_{ij} = \hat{\alpha}_i + \hat{\beta}_j + \sum_{m=1}^{\rho} \hat{\phi}_m \hat{\gamma}_{im} \hat{\delta}_{jm} \tag{29}$$

et

$$\dot{f}_{ij} = \hat{f}_{ij} + \hat{\theta}_k = \hat{\alpha}_i + \hat{\beta}_j + \sum_{m=1}^{\rho} \hat{\phi}_m \hat{\gamma}_{im} \hat{\delta}_{jm} + \hat{\theta}_k. \tag{30}$$

[10] La matrice $\mathbf{YY}'$ n'est pas nécessairement une matrice de covariance pour n'importe quel $\mathbf{Y}$ mais elle en est une dans le cas présent puisque les lignes (et les colonnes) de $\mathbf{Y}$ ont été centrées par soustraction de leur moyenne dans le calcul du terme additif.

La dernière, $\dot{f}_{ij}$, est la meilleure prédiction de f_{ij}, fondée sur l'ajustement complet; d'un autre côté, $\hat{f}_{ij}$ représente le niveau de mortalité à l'âge i pour la période j que l'on pourrait attendre en l'absence de perturbations longitudinales. L'ampleur de ces perturbations est donnée par $\hat{\theta}_k$, qui représente l'avantage ou le désavantage relatif de la k^e génération. Sa valeur correspond à l'ampleur moyenne des différences séparant les valeurs observées f_{ij} et les estimations des valeurs attendues $\hat{f}_{ij}$ en l'absence de perturbations longitudinales, puisque l'équation (14) peut aussi s'écrire :

$$\hat{\theta}_k = \frac{1}{w_k} \sum_{(k)} (f_{ij} - \hat{f}_{ij}). \tag{31}$$

Comme la transformation de f_{ij} est une approximation du logarithme de la force de mortalité μ_{ij}, nous pouvons aussi définir :

$$\hat{\mu}_{ij} = \exp(\hat{f}_{ij}) \quad \text{et} \quad \dot{\mu}_{ij} = \exp(\dot{f}_{ij}), \tag{32}$$

comme estimateurs de μ correspondant respectivement à $\hat{f}$ et à $\dot{f}$. De là, on tire de la dernière équation du paragraphe précédent :

$$\log(\dot{\mu}_{ij}) = \log(\hat{\mu}_{ij}) + \hat{\theta}_k \ , \tag{33}$$

ce qui implique que :

$$\dot{\mu}_{ij} = \hat{\mu}_{ij} \exp(\hat{\theta}_k) \ , \tag{34}$$

et donc

$$\dot{\mu}_{ij} \simeq \hat{\mu}_{ij}(1 + \hat{\theta}_k) \ ,$$

de telle sorte que $100\hat{\theta}_k$ est une approximation du pourcentage moyen de déviation observée par rapport à l'estimation $\hat{f}_{ij}$.

RÉFÉRENCES

[1] Boleslowski (Lech). – Roznice w umieralnosci miedzy generacjami jaka skutek wojen swiatowych. *Studia Demograficzne*, vol. 4, n° 82, 1985, pp. 51–71.

[2] Brillinger (D.). – The natural variability of vital rates associated statistics. *Biometrics*, vol. 42, n° 4, 1986, pp. 693–734.

[3] Caselli (Graziella). – Les effets de génération dans les différences de mortalité et leur évolution. *In : Mesure et analyse de la mortalité : nouvelles approches.* Éd. par Vallin (Jacques), D'Souza (Stan), et Palloni (Alberto) p. 251-272. – Paris, INED, PUF, 1988. (Travaux et documents, Cahier 119).

[4] Caselli (Graziella) et Capocaccia (Riccardo). – The impact of early mortality on adult mortality : an age, period and cohort analysis. *Population Studies*, vol. 43, n° 1, 1989. – (à paraître).

[5] Caselli (Graziella) et Egidi (Viviana). – *Nouvelles tendances de la mortalité en Europe.* – Strasbourg, Conseil de l'Europe, 1981. – (Étude démographique n° 5).

[6] Caselli (Graziella) et Greco (Bruno). – Aggiornamento delle tavole di mortalità per generazione di Natale e Bernassola, periodo 1965-1979. – Rome, Dipartimento di scienze demografiche, 1983. – (Non publié, disponible sur bande magnétique).

[7] Caselli (Graziella), Vallin (Jacques), Vaupel (James) et Yashin (Anatoli). – Age-specific mortality trends in France and in Italy since 1900 : period and cohort effects. *European Journal of Population*, n° 3, 1987, pp. 33–60.

[8] Caselli (Graziella), Vaupel (James) et Yashin (Anatoli). – Mortality in Italy : contours of a century of evolution. *Genus*, vol. XLI, n° 1-2, 1985, pp. 39–55.

[9] Centraal bureau voor de statistiek. – *Sterftetafels voor Nederland.* – s'-Gravenhage, CBS, (publication quinquennale).

[10] Delaporte (Pierre). – *Évolution de la mortalité en Europe depuis l'origine des statistiques de l'état civil : tables de mortalité de génération.* – Paris, Imprimerie Nationale, 1941. – (Statistique Générale de la France, Études démographiques n° 2).

[11] Emerson (J. D.) et Wong (G. Y.). – Resistant nonadditive fits for two-way tables. In : *Exploring data tables, trends and shapes.* Éd. par Hoaglin (D.), Mosteller (F.), et Tukey (J.). – New York, John Wiley and sons, 1985.

[12] Good (I.). – Some applications of the singular value decomposition of a matrix. *Technometrics*, vol. 4, n° 11, 1969, pp. 823–831.

[13] Hearst (N.), Newman (T.) et Hulley (S.). – Delayed effects of the military draft on mortality. *New England Journal of Medicine*, vol. 10, n° 314, 1986, pp. 620–624.

[14] Hobcraft (John) et Gilks (Walter). – Age, period and cohort analysis. *In : Methodologies for the collection and analysis of mortality.* Éd. par Vallin (Jacques), Pollard (John) et Heligman (Larry), pp. 245–264. – Liège, Ordina Éditions, 1984.

[15] Hobcraft (John), Menken (Jane) et Preston (Samuel). – Age, period and cohort effects in demography : a review. *Population Index*, vol. 48, n° 1, 1982, pp. 4–43.

[16] Horiuchi (Shiro). – The long term impact of war on mortality : old-age mortality of First World War survivors in the Federal Republic of Germany. *Bulletin démographique des Nations Unies*, n° 15, 1983, pp. 80–92.

[17] Mandel (J.). – A new analysis of variance model for nonadditive data. *Technometrics*, n° 13, 1971, pp. 1–18.

[18] Manton (Kenneth), Stallard (Eric) et Vaupel (James). – Alternative estimates of the heterogeneity of mortality risks among the aged. *Mathematical Demography*, 1988. – (A paraître).

[19] Manton (Kenneth), Stallard (Eric) et Vaupel (James). – Methods for comparing the mortality experience of heterogeneous populations. *Demography*, n° 18, 1981.

[20] Mason (W.) et Fienberg (S.) (édité par). – *Cohort analysis in social research : beyond the identification problem.* – New York, Academic Press, 1985.

[21] McMillen (M. M.) et Nam (Charles). – *Mortality crossovers by causes of death and race in the U.S. in the 1970's.* – Liège, UIESP, 1985. – (Communication présentée au *Congrès international de la population, Florence, juin 1985).*

[22] McNeil (D. R.) et Tukey (J. W.). – Higher-order diagnosis of two-way tables, illustrated on two sets of demographic empirical distributions. *Biometrics*, vol. 31, 1975, pp. 487–510.

[23] Mosteller (F.) et Tukey (J.). – *Data analysis and regression.* – Reading, MA, Addison-Wesley, 1977.

[24] Natale (Marcelo) et Bernassola (Amedeo). – *La mortalità per causa nelle regioni italiane, tavole per contemporanei e per generazione 1790-1964.* – Rome, Università di Roma, Istituto di demografia, 1973. – (N° 25).

[25] Okubo (Masakazu). – *Increase in mortality of middle-aged males in Japan.* – Tokyo, Nupri, 1981. – (Nupri research paper, Series N° 3).

[26] Ryder (Norman). – Components of temporal variations in American fertility. In : *Demographic patterns in developed societies.* Éd. par Hiorns (R.), pp. 15–54. – Londres, Taylor and Francis, 1980.

[27] Tukey (J. W.). – *Exploratory data analysis.* – Reading, MA, Addison-Wesley, 1977.

[28] Vallin (Jacques). – *La mortalité par génération en France depuis 1899.* – Paris, INED, PUF, 1973. – (Travaux et documents, Cahier n° 63).

[29] Vallin (Jacques). – Sex patterns of mortality : a comparative study of model life tables and actual situations with special reference to the cases of Algeria and France. In : *Sex differentials in mortality : trends, determinants and consequences.* Éd. par Lopez (Alan) et Ruzicka (Lado), pp. 443–476. – Canberra, Australian National University, 1983.

[30] Vallin (Jacques). – *Tables de mortalité du moment et par génération 1899-1981 mise à jour provisoire des tables annexes du Cahier 63.* – Paris, INED, 1984. – (3 volumes).

[31] Vallin (Jacques). – Tendances récentes de la mortalité française. *Population*, n° 1, 1983, pp. 77–105.

[32] Vallin (Jacques) et Lopez (Alan) (édité par). – *La lutte contre la mort : influence des politiques sociales et des politiques de santé sur l'évolution de la mortalité.* – Paris, INED, PUF, 1985. – (Travaux et documents, Cahier 108).

[33] Vallin (Jacques) et Meslé (France). – *Les causes de décès en France de 1925 à 1978.* – Paris, INED, PUF, 1988. – (Travaux et document, Cahier n° 115).

[34] Vallin (Jacques) et Nizard (Alfred). – Les causes de décès en France : pour une typologie simple et homogène, application à la période 1968-1974. *Population*, n° 3, 1978, pp. 547–608.

[35] Vaupel (James), Gambill (Bradley), Yashin (Anatoli) et Bernstein (A.). – *Contour maps of demographic surfaces.* – Laxenburg, Autriche, IIASA, 1985. – (Working paper).

[36] Velleman (P.) et Hoaglin (D.). – *Applications, basics and computing of exploratory data analysis.* – Belmont, CA, Wadsworth, 1981.

[37] Veys (Dion). – *Cohort survival in Belgium in the past 150 years : principles and methods.* – New York and London, Plenum press, 1978.

[38] Wilmoth (John R.). – *Age-period-cohort analysis of mortality data : an explanatory approach.* – Princeton, Princeton University, 1986. – (Communication présentée à la Conférence annuelle de la Population Association of America, San Francisco, avril 1986).

[39] Wilmoth (John R.). – *On the statistical analysis of large arrays of demographic rates.* – Princeton, Princeton University, 1988. – (Thèse de PhD).

[40] Wilmoth (John R.) et Caselli (Graziella). – *A simple model for the statistical analysis of large arrays of mortality data : rectangular vs. diagonal structure.* – Laxenburg, Autriche, IIASA, 1987. – (Working paper, n° WP-87-58).

WILMOTH John, VALLIN Jacques et CASELLI Graziella. — **Quand certaines générations ont une mortalité différente de celle que l'on pourrait attendre.**

L'analyse de la matrice française de quotients de mortalité par année d'âge et année de calendrier de la période 1899-1981 à partir d'un modèle descriptif simple permet d'isoler une composante diagonale qui sans prétendre rendre compte de la totalité de l'influence de l'histoire propre de chaque génération, permet cependant de souligner la singularité de certaines générations dont l'histoire diffère sensiblement de celle des générations voisines.

Se trouvent ainsi clairement confirmés, les effets à longs termes des deux guerres mondiales déjà mis en évidence par divers auteurs mais d'autres apparaissent également : effets à long termes de la grippe espagnole de 1918, effets négatifs de la généralisation de l'accouchement hospitalier dans les années cinquante.

Poursuivie au niveau de huit grandes catégories étiologiques, l'analyse montre que toutes les causes participent à la surmortalité des générations singulières, mais certains processus jouent un rôle particulier (malnutrition et alcoolisme, infection, dégénérescence, notamment).

WILMOTH John, VALLIN Jacques and CASELLI Graziella. — **Differences between observed and expected generation mortality.**

This analysis of the matrix of mortality rates in France by age and calendar year between 1899 and 1981 is based on a simple descriptive model that makes it possible to isolate diagonals in the matrix, which show the characteristics of specific cohorts whose mortality history differs from that of neighbouring cohorts, without, however, pretending that all the significant factors have been taken into account.

The long term effects of the two World Wars, previously demonstrated by various authors, are clearly confirmed, as are those of the epidemic of Spanish influenza in 1918, as well as the reductions that followed the increases in the proportions of institutional confinements during the 1950's.

After establishing eight major aetiological categories, the analysis shows that all the causes listed contributed to the high mortality rates found in specific generations, but that some factors played a specific part (notably, malnutrition and alcoholism, infection and ageing).

WILMOTH John, VALLIN Jacques y CASELLI Graziella. — **Cuando ciertas generaciones tienen una mortalidad diferente a la que se podría suponer.**

El análisis de la matriz francesa de probabilidades de muerte por año de edad y año calendario del periodo 1899-1981 a partir de un modelo descriptivo simple, permite aislar una componente diagonal que sin pretensión de explicar la totalidad de la influencia de la historia propia de cada generación, permite sin embargo hacer resaltar la singularidad de ciertas generaciones cuya historia difiere notoriamente de la de generaciones continguas.

Se hallan así claramente confirmados los efectos a largo plazo de las dos guerras mundiales ya puestas en evidencia por otros autores, pero aparecen igualmente otras : efectos a largo plazo de la influenza española de 1918, efectos negativos de la generalización del parto en hospital durante los años cincuenta.

Realizado a nivel de ocho grandes categorías etiológicas, el análisis muestra que todas las causas participan en la sobremortalidad de las generaciones identificadas, pero ciertos procesos juegan un papel especial (sobre todo, malnutrición, alcoholismo, infecciones, degeneración).