

CHRONIQUES DÉMOGRAPHIQUES DES NAISSANCES : LONGUE MÉMOIRE OU CHANGEMENT DE RÉGIME ?

Ahdi Noomen AJMI,¹ Mohamed BOUTAHAR²

RÉSUMÉ – *De rares études démographiques récentes utilisent des modèles à longues mémoires pour l'analyse des chroniques démographiques. Cela dit les examens de telles chroniques mettent en évidence des ruptures importantes analysées de longue date par les démographes. Comme en économie une question peut être posée : a-t-on affaire à une dynamique linéaire à longue mémoire ou à de simples changements de régime pour ces chroniques.*

MOTS CLÉS – Changement de régimes, Démographie, Longue mémoire, Processus FARIMA.

SUMMARY – Demographic time series: long memory or regime switching?

A few works have been recently advocated the use of long memory concept in modelling demography time series. However, the study of such series shows evidence for regime switching which was pointed out by demographers many years ago. As in economic framework, one ask the following question, what is the genuine process describing such series: long memory or regime switching?

KEYWORDS – ARFIMA process, Demography, Long memory, Structural changes.

1. INTRODUCTION

L'étude des modèles à longue mémoire et leurs applications constituent actuellement un des domaines populaires de la recherche statistique. Depuis une bonne quinzaine d'années la littérature est abondante et riche en résultats. Le lecteur peut se référer aux ouvrages généraux de Beran [1994], Deniau *et al.* [2000], Doukhan *et al.* [2003] et Robinson [2003]. Les domaines d'applications utilisant les modèles à longue mémoire n'ont cessé de se développer ces dernières années, touchant des domaines variés. Pour illustrer cette affirmation citons quelques travaux couvrant un large spectre : Hurst [1951], Hosking [1984] en hydrologie, Mandelbrot [1962], Lo [1991], Willinger *et al.* [1999] en finance, Hassler et Wolters [1995] en macroéconomie, Graf [1983], Beran et Terrin [1996] en géophysique, Taqqu *et al.* [1997] en télécommunication, ainsi que deux domaines d'applications plus récents : la psychologie, Wagenmaker *et al.* [2004] Torre *et al.* [2006] et la démographie, Gil-Alana [2003] et Mishra et Urbain [2005].

¹Laboratoire Business & Economic Statistics MODeling (BESTMOD), Institut Supérieur de Gestion, 41 rue de la liberté, Cité Bouchoucha, Le Bardo 2000, Tunis, TUNISIE, noomen.ajmi@isg.rnu.tn.

²Département de mathématiques, Université de la Méditerranée, Faculté des Sciences de Luminy, case 901, 163 avenue de Luminy 13288 Marseille Cedex 9, boutahar@univmed.fr.

Dans ce dernier article Mishra et Urbain étudient l'existence de relations entre les chroniques de population et la croissance économique de 151 pays. La lecture de ce travail original nous a incité à examiner des modèles de longue mémoire pour des chroniques démographiques de naissance concernant quelques pays. Nous avons évidemment commencé par étudier la chronique des naissances française (cf. Figure 1). On remarque d'abord la présence de ruptures nettes dans cette dynamique de naissances. Une question que peut se poser le statisticien est la suivante: le phénomène de persistance détecté statistiquement sur de telles chroniques n'est-il pas un artéfact dû à la présence de ces changements de régime ? La question a déjà été soulevée et étudiée en économie, par exemple chez Diebold et Inoue [2001] qui ont montré que le changement de régime induit la longue mémoire, Boutahar et Jouini [2007] qui confirment la présence des deux phénomènes dans le processus de l'inflation américaine, Dufrénot et al. (2008) considèrent la longue mémoire et le changement de régime lisse de type STAR (Smooth Transition AutoRegressive) pour modéliser la dynamique de taux de change.



FIGURE 1. La chronique des naissances françaises

Notre étude va porter sur la comparaison de performance des modèles à longue mémoire et à changement de régime pour dix chroniques démographiques de naissances³. Il ne s'agit pas d'une étude fine sur le plan démographique conduisant à des interprétations en terme de fécondité ou de naissances, des changements de régime ou de la longue mémoire. Plus modestement, du point de vue du statisticien, nous tentons de mettre en évidence les difficultés d'un choix de modèle. Les dix chroniques n'ont pas été choisies arbitrairement, mais simplement parce que leur historique était suffisant pour pouvoir parler de longue mémoire. Un examen rapide des graphiques de ces chroniques (cf. Figure 2) met en évidence pour certaines d'entre elles, comme pour la chronique française, d'importantes ruptures dues à la guerre de 1939-45 et au choc pétrolier du début des années 1970 mais aussi des différences importantes dans leurs dynamiques. Comme nous le verrons au cours de l'étude empirique certaines chroniques sont stationnaires d'autres pas. C'est sur les

³Sources : "National Center for Health Statistics" et Institut National des Études Démographiques (INED), www.ined.fr.

chroniques stationnaires de 4 pays (États-Unis, France, Norvège et Suisse) que nous effectuerons les comparaisons de modèles.

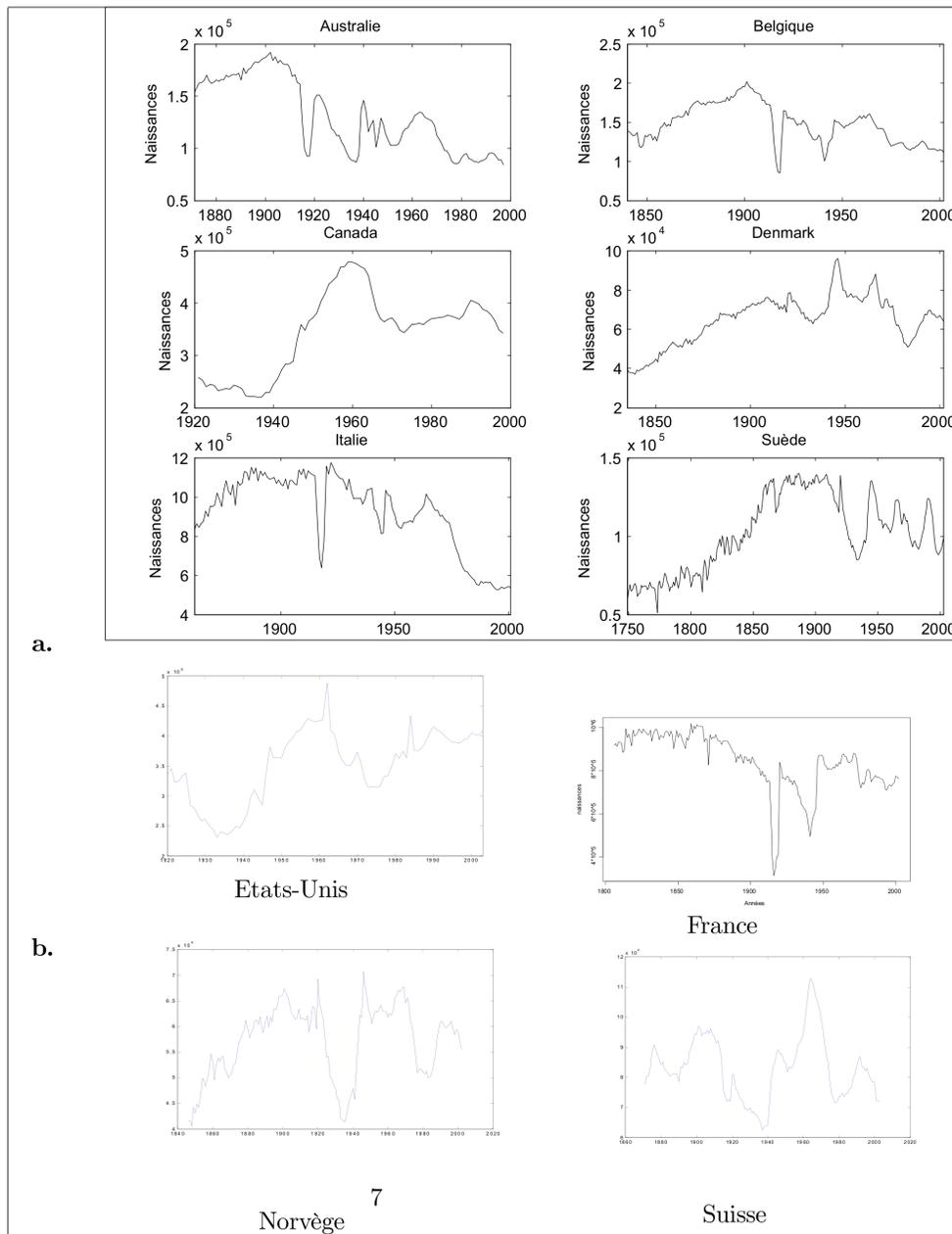


FIGURE 2. Chroniques des naissances. a) non stationnaires, b) stationnaires

Les études théoriques et méthodologiques aussi bien des changements de régime que de la persistance prise en compte par des modèles dits à longue mémoire se sont développées dans la communauté statistique. Les chroniques issues de modèles à longue mémoire et à changement de régime et leurs premières analyses peuvent présenter des ressemblances troublantes. À titre d'illustration, nous avons simulé un modèle à longue mémoire $FARIMA(0, 0.4, 0)$ défini dans la Section 2 et un modèle à seuil de type (1). La Figure 3 (B) et (D) montre une forte persistance

dans la fonction d'autocorrélation de la chronique simulée à partir d'un modèle à changement de régime qui laisse penser à l'existence de la longue mémoire. De plus, la chronique $FARIMA(0, 0.4, 0)$ simulée (cf. Figure 3 (A)) montre l'existence d'une tendance locale qui peut laisser croire à un modèle avec changement de régime.

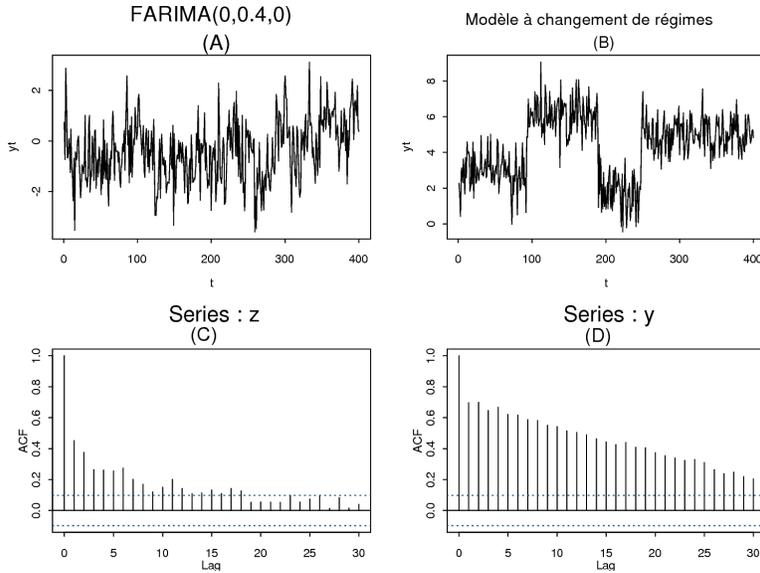


FIGURE 3. Longue mémoire ou changement de régime ?

L'étude des chroniques à longue mémoire a pour origine l'article de Mandelbrot et Van-Ness [1968] sur le bruit fractionnaire, elle a connu ensuite une grande popularité grâce aux modèles $FARIMA$ introduits au début des années 1980.

La littérature sur les changements de régime est abondante (cf. par exemple [Chow, 1960 ; Yao, 1988 ; Bai et Perron, 1998]). Beaucoup de travaux se concentrent sur l'instabilité de la moyenne (moment d'ordre 1). En effet, un modèle largement considéré est celui de régression linéaire avec m dates de rupture ($m + 1$ régimes) :

$$Y_t = \mu_j + u_t, \quad \text{si } T_{j-1} < t \leq T_j, \quad (1)$$

pour $j = 1, \dots, m + 1$, $T_0 = 0$ et $T_{m+1} = T$, Y_t est la variable dépendante et ε_t est le terme d'erreur supposé stationnaire.

Le nombre m de ruptures ainsi que les dates ($T_j, 1 \leq j \leq m$) peuvent être prédéterminés, si on admet qu'une variable exogène joue un rôle dominant dans les changements de la moyenne du processus Y_t . Par exemple l'effet des deux guerres mondiales (1914-1918 et 1939-1945) sur les chroniques de naissances des pays européens ou encore l'effet de la crise économique de 1929 et le choc pétrolier de 1973 sur les variables macroéconomiques (inflation, PIB...) américaines. Dans le cas général m et ($T_i, 1 \leq i \leq m$) sont supposées inconnues.

Davidson et Sibbertsen [2005] considèrent le modèle (1). En supposant que les durées $\tau_j = T_j - T_{j-1}$ sont des variables aléatoires i.i.d dont la probabilité des queues se comporte comme une puissance (i.e. $P(\tau_j > c) \simeq c^{-\alpha} L(c)$, avec L une fonction

qui varie lentement), ils montrent que la fonction d'autocorrélation de Y_t décroît de façon hyperbolique, ce qui est caractéristique de la présence de longue mémoire.

Selon ce résultat, si on veut distinguer entre longue mémoire stationnaire et changement de régime apparaissant à des dates arbitraires, la tâche peut être très difficile. L'objet de ce travail constitue une illustration de cette difficulté.

Cet article est organisé de la manière suivante : la Section 2 est consacrée à une introduction aux processus *FARIMA* avec des exemples simples, la Section 3 présente quelques méthodes d'estimation du paramètre de longue mémoire et les procédures de test utilisées au cours de l'analyse empirique. L'illustration du conflit longue mémoire et changement de régime sur les chroniques des naissances est présentée dans la Section 4.

2. PROCESSUS À LONGUE MÉMOIRE

Deux grandes classes de processus sont derrière la popularité des modèles à mémoires longues : le bruit gaussien fractionnaire et le modèle *FARIMA*. Dans cette section, on va s'intéresser à la présentation de la deuxième classe de modèle, i.e., les modèles *FARIMA* et nous renvoyons au livre de Beran [1994] pour une étude des processus de la première classe.

2.1. LES MODÈLES FARIMA

Une manière de présenter les processus à longue mémoire consiste à les définir à partir de propriétés de la fonction d'autocovariance. On considère un processus à temps discret $\mathbf{Y}=(Y_t)_{t \in \mathbb{Z}}$ tel que :

$$\begin{aligned} E(Y_t) &= \mu \\ \gamma(j) &= \text{cov}(Y_t, Y_{t+j}) < \infty \quad \forall j \in \mathbb{Z}, \forall t \in \mathbb{Z} \end{aligned} \quad (2)$$

c'est-à-dire un processus stationnaire au second ordre.

On définit la fonction d'autocorrélation ρ d'un tel processus par:

$$\rho(j) = \frac{\gamma(j)}{\gamma(0)} \quad \forall j \in \mathbb{Z}$$

La fonction d'autocovariance peut-être considérée comme la transformée de Fourier d'une mesure positive finie sur $[-\pi, +\pi[$, soit :

$$\gamma(j) = \int_{-\pi}^{+\pi} e^{ij\lambda} \mu(d\lambda) \quad \forall j \in \mathbb{Z} \quad (3)$$

appelée mesure spectrale du processus \mathbf{Y} .

Pour les détails sur ce paragraphe, on peut se reporter à une littérature abondante sur le sujet, par exemple à Priestley [1981] et Gouriéroux et Monfort [1990].

Si la mesure μ possède une densité, celle-ci est définie par :

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{+\infty} \gamma(j)e^{-ij\lambda}. \quad (4)$$

La mémoire d'un processus peut-être caractérisée par des propriétés de sa fonction d'autocovariance ou de sa densité spectrale.

Un processus est dit à longue mémoire si :

$$\begin{aligned} \sum_{j=-\infty}^{+\infty} |\gamma(j)| &= +\infty, \\ f(0) &= +\infty. \end{aligned} \quad (5)$$

L'autocovariance entre des observations éloignées du processus, bien que petite, n'est pas négligeable.

Par opposition un processus à courte mémoire sera caractérisé par la propriété suivante :

$$\begin{aligned} \sum_{j=-\infty}^{+\infty} |\gamma(j)| &< +\infty, \\ f(0) &< +\infty. \end{aligned} \quad (6)$$

Afin de nous placer dans le contexte théorique de nos modélisations ultérieures nous allons rapidement rappeler la définition d'une classe populaire de processus à mémoire courte qui nous permettra ensuite de présenter une famille simple de processus à longue mémoire. Il s'agit des processus autorégressifs à moyenne mobile ARMA(p,q) introduits par Box-Jenkins [1970] et définis par :

$$Y_t + a_1 Y_{t-1} + \dots + a_p Y_{t-p} = \varepsilon_t + c_1 \varepsilon_{t-1} + \dots + c_q \varepsilon_{t-q} \quad (7)$$

où p et q sont des entiers non négatifs appelés respectivement ordre de la partie autorégressive (AR) et la partie moyenne mobile (MA), ε_t est un bruit blanc caractérisé par :

$$\begin{aligned} E(\varepsilon_t) &= 0, \quad \forall t \in \mathbb{Z}, \\ \text{cov}(\varepsilon_t, \varepsilon_{t+k}) &= 0 \text{ si } k \neq 0, \\ \text{var}(\varepsilon_t) &= \sigma^2. \end{aligned}$$

On peut écrire (4) sous la forme :

$$A(B)Y_t = C(B)\varepsilon_t \quad (8)$$

où :

$$A(B) = 1 + a_1 B + \dots + a_p B^p, C(B) = 1 + c_1 B + \dots + c_q B^q,$$

avec la convention suivante : $B^j Y_t = Y_{t-j}$. Pour qu'un tel processus soit stationnaire au second ordre les racines du polynôme $A(z)$ doivent être toutes de module supérieur à 1.

À partir de cette représentation $ARMA(p, q)$, on peut distinguer deux cas particuliers de modèles : les modèles autorégressifs $AR(p)$ et les modèles moyennes mobiles $MA(q)$.

Un modèle $AR(p)$ est défini par :

$$Y_t + a_1 Y_{t-1} + \dots + a_p Y_{t-p} = \varepsilon_t$$

ou encore :

$$A(B)Y_t = \varepsilon_t.$$

De même, un modèle $MA(q)$ s'écrit comme suit :

$$Y_t = \varepsilon_t + c_1 \varepsilon_{t-1} + \dots + c_q \varepsilon_{t-q}$$

ou encore

$$Y_t = C(B)\varepsilon_t.$$

Un processus $ARMA(p, q)$ possède une densité spectrale définie par :

$$f(\lambda) = \frac{\sigma^2 |C(e^{-i\lambda})|^2}{2\pi |A(e^{-i\lambda})|^2} \quad (9)$$

Comme l'ont proposé initialement Hosking [1981] et Granger-Joyeux [1980], auxquels nous renvoyons pour plus de détails, nous présentons maintenant une classe simple de processus à longue mémoire. Ces auteurs se placent dans le cadre des processus autorégressifs à moyenne mobile intégré ($ARIMA$) introduits par Box-Jenkins [1971]. Ils définissent une classe de processus $ARIMA$ fractionnaires ($FARIMA$) définis par :

$$(1 - B)^d A(B)Y_t = C(B)\varepsilon_t, \quad d \in \left] -\frac{1}{2}, \frac{1}{2} \right[\subset \mathbb{R} \quad (10)$$

où d , paramètre de longue mémoire, est aussi appelé degré d'intégration du processus. La densité spectrale d'un tel processus est donnée par :

$$f(\lambda) = \frac{\sigma^2 |C(e^{-i\lambda})|^2}{2\pi |A(e^{-i\lambda})|^2} |1 - e^{-i\lambda}|^{-2d}$$

avec un comportement original au voisinage de $\lambda = 0$ pour les longues mémoires.

- Si $0 < d < \frac{1}{2}$ $\lim_{\lambda \rightarrow 0} f(\lambda) = +\infty$ (longue mémoire).
- Si $-\frac{1}{2} < d < 0$ $\lim_{\lambda \rightarrow 0} f(\lambda) = 0$ (mémoire intermédiaire).

Un tel processus est alors stationnaire au second ordre, sa fonction d'autocovariance est telle que :

$$\gamma(j) \sim Cj^{d-1} \quad \text{avec } C \neq 0 \quad \text{si } j \rightarrow \infty \quad (11)$$

$$\gamma(j) \rightarrow 0 \quad \text{avec une vitesse hyperbolique.} \quad (12)$$

Alors que pour un processus ARMA(p,q) :

$$|\gamma(j)| < Cr^j \quad C > 0 \text{ et } 0 < r < 1 \quad (13)$$

$$\gamma(j) \rightarrow 0 \quad \text{avec une vitesse exponentielle.} \quad (14)$$

Dans (10) on note que :

- Si $d = 0$ donc Y_t est un processus ARMA, on a évidemment $\sum_{j=-\infty}^{+\infty} |\gamma(j)| < +\infty$.
- Si $0 < d < 1/2$ alors $\sum_{j=-\infty}^{+\infty} |\gamma(j)| = +\infty$ (longue mémoire).
- Si $-1/2 < d < 0$ alors $\sum_{j=-\infty}^{+\infty} |\gamma(j)| < +\infty$ (mémoire intermédiaire ou antipersistante).
- Si $d \in \mathbb{N} - \{0\}$ alors Y_t est un processus autorégressif à moyenne mobile intégré *ARIMA* (non stationnaire).

Dans la classe des processus ARMA(p,q) on peut même trouver des processus tels que $f(0) = 0$ comme par exemple pour le processus défini par : $Y_t = \varepsilon_t - \varepsilon_{t-1}$.

2.2. EXEMPLES

On va présenter et simuler quelques exemples simples de chroniques satisfaisant les modèles précédents. Les graphiques des chroniques, des densités spectrales et des fonctions d'autocorrélation sont présentés en Annexe.

EXEMPLE 1: *ARMA à basses fréquences.*

Soit le modèle à mémoire courte à basses fréquences suivant :

$$(1 - 0.5B) Y_t = (1 - 0.4B) \varepsilon_t$$

dans ce cas la densité spectrale est donnée par :

$$f(\lambda) = \frac{1}{2\pi} \frac{|1 - 0.4e^{-i\lambda}|^2}{|1 - 0.5e^{-i\lambda}|^2} \quad -\pi \leq \lambda \leq \pi.$$

EXEMPLE 2 : *ARMA à hautes fréquences.*

Soit le modèle à mémoire courte à hautes fréquences suivant :

$$(1 + 0.5B) Y_t = (1 - 0.4B) \varepsilon_t$$

dans ce cas la densité spectrale est donnée par :

$$f(\lambda) = \frac{1}{2\pi} \frac{|1 - 0.4e^{-i\lambda}|^2}{|1 + 0.5e^{-i\lambda}|^2} \quad -\pi \leq \lambda \leq \pi.$$

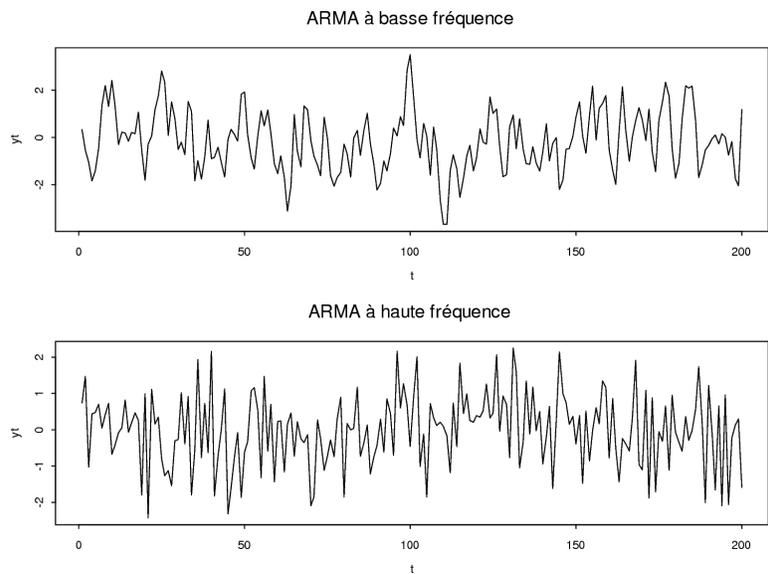


FIGURE 4. Les chroniques simulées des modèles ARMA

EXEMPLE 3 : $FARIMA(1, 0.4, 1)$.

Soit le modèle à longue mémoire suivant :

$$(1 - B)^{0.4} (1 + 0.5B) Y_t = (1 - 0.4B) \varepsilon_t$$

dont la densité spectrale est :

$$f(\lambda) = \frac{1}{2\pi} \frac{|1 - 0.4e^{-i\lambda}|^2}{|1 + 0.5e^{-i\lambda}|^2} |1 - e^{-i\lambda}|^{-0.8} \quad -\pi \leq \lambda \leq \pi.$$

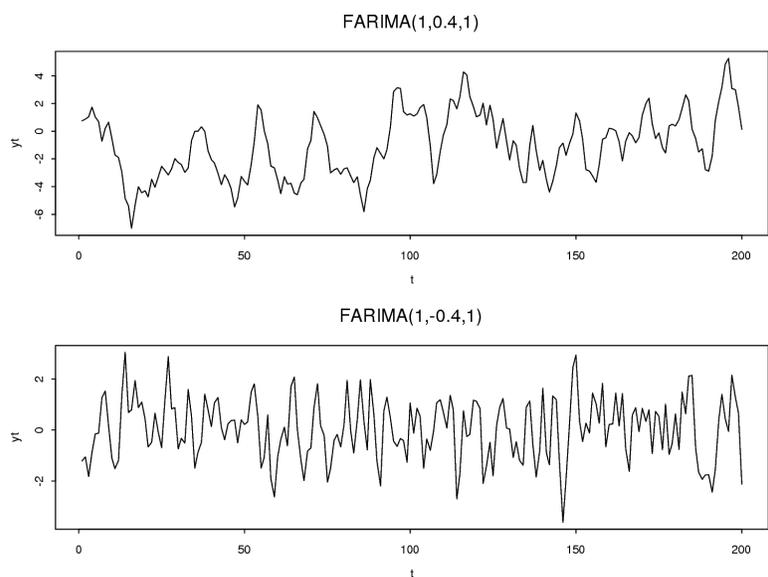


FIGURE 5. Les chroniques simulées des modèles FARIMA

EXEMPLE 4 : $FARIMA(1, -0.4, 1)$.

Soit le modèle à mémoire intermédiaire suivant :

$$(1 - B)^{-0.4} (1 + 0.5B) Y_t = (1 - 0.4B) \varepsilon_t$$

dont la densité spectrale est :

$$f(\lambda) = \frac{1}{2\pi} \frac{|1 - 0.4e^{-i\lambda}|^2}{|1 + 0.5e^{-i\lambda}|^2} |1 - e^{-i\lambda}|^{0.8} \quad -\pi \leq \lambda \leq \pi.$$

EXEMPLE 5 : $ARIMA(1, 1, 1)$.

Soit le modèle non stationnaire suivant :

$$(1 - B)(1 + 0.5B) Y_t = (1 - 0.4B) \varepsilon_t$$

ce processus ne possède pas de densité spectrale.

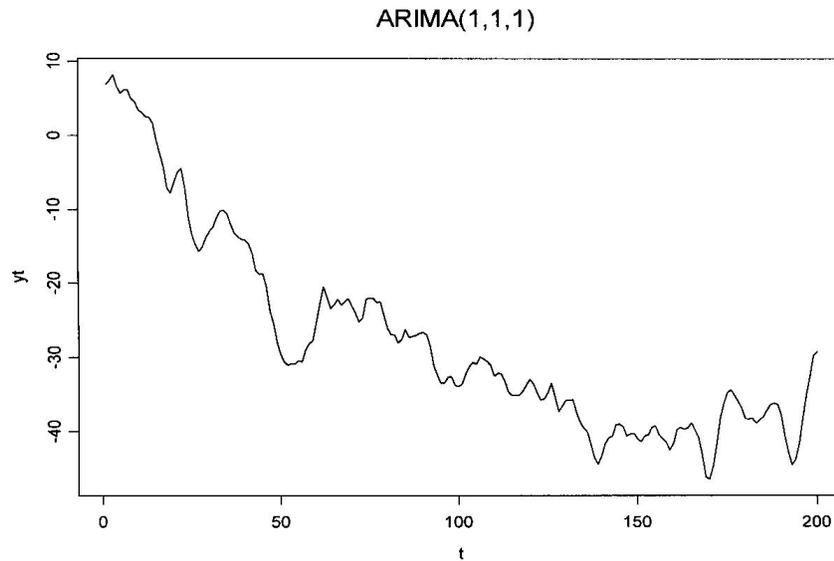


FIGURE 6. Les chroniques simulées des modèles ARIMA

COMMENTAIRES DES GRAPHIQUES :

a) Les chroniques :

D'après les Figures 4 et 5, nous constatons qu'à l'exception de l'exemple 5 (cas du modèle $ARIMA(1,1,1)$) toutes les chroniques fluctuent autour de zéro. Ajoutant à ceci que les modèles ARMA à hautes fréquences et le modèle FARIMA $(1,-0.4,1)$ à mémoire intermédiaire fluctuent plus rapidement que le modèle ARMA à basse fréquence et le modèle à longue mémoire FARIMA $(1,0.4,1)$. D'après la Figure 6, le modèle ARIMA présente une tendance ou un non retour à la moyenne théorique zéro.

b) *Les fonctions d'autocorrélations :*

La Figure 7 indique que les fonctions d'auto-corrélations des modèles ARMA à basse ou à hautes fréquences et le modèle ARIMA antipersistant décroissent d'une manière exponentielle, alors que celle du modèle FARIMA à longue mémoire persiste et décroît lentement.

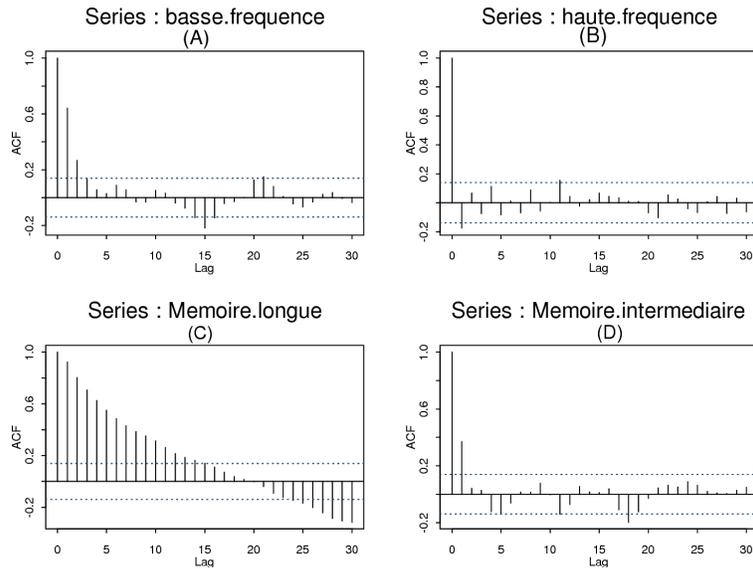


FIGURE 7. Les fonctions d'autocorrélations

c) *Les densités spectrales :*

La Figure 8 indique que la densité spectrale du modèle FARIMA (1,0.4,1) explose. Cependant, les densités des autres modèles sont bornées. La densité spectrale du modèle FARIMA antipersistant s'annule en zéro.

La mise en oeuvre de ces modèles pour des applications nécessite d'abord de procéder à des estimations. Il s'agit d'estimer le paramètre de longue mémoire d , les entiers p et q donnant l'ordre des parties autorégressive et moyenne mobile du modèle et enfin les paramètres $(a_j)_{1 \leq j \leq p}$ et $(c_j)_{1 \leq j \leq q}$. Les méthodes d'estimation de p , q de même que celles des paramètres $(a_j)_{1 \leq j \leq p}$ et $(c_j)_{1 \leq j \leq q}$ se trouvent dans la plupart des livres consacrés à l'étude des chroniques. Il n'en est pas de même de l'estimation du paramètre d à laquelle nous consacrons un paragraphe.

3. ESTIMATION ET TESTS DES PROCESSUS À LONGUE MÉMOIRE

Dans ce qui suit, nous allons présenter deux méthodes d'estimation : une méthode à deux étapes et une méthode à une seule étape. Pour un exposé plus complet sur les méthodes d'estimation, nous renvoyons à Boutahar *et al.* [2006].

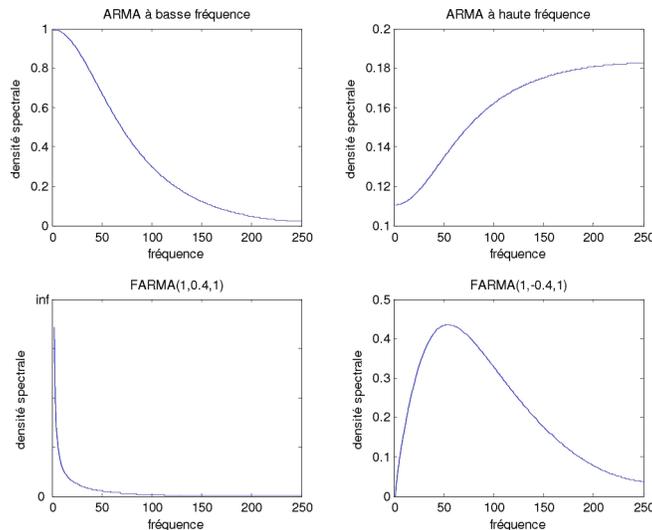


FIGURE 8. Les densités spectrales

3.1. MÉTHODES À DEUX ÉTAPES

Cette méthode est utilisée pour estimer un $FARIMA(p, d, q)$ en se basant sur un échantillon (Y_1, \dots, Y_T) . Elle consiste, dans sa première phase, à estimer le degré d'intégration d et appliquer à la chronique le filtre $(1 - B)^d$. Dans la deuxième phase, on modélise la chronique résiduelle par un $ARMA(p, q)$. Le paramètre d peut être estimé en utilisant les méthodes semi-paramétriques. Ces méthodes sont devenues populaires au début des années quatre-vingt par les travaux de Geweke et Porter-Hudak [1983], (GPH). Elles se basent sur l'expression de la densité spectrale du processus $FARIMA(p, d, q)$ lorsque les fréquences tendent vers zéro. Plusieurs travaux se sont succédés pour améliorer⁴ le biais de l'estimateur GPH dû au choix du paramètre de troncature m_T . On se limite ici à la présentation de cette méthode. Pour d'autres méthodes d'estimation semi-paramétriques nous renvoyons à Boutahar *et al.* [2006].

Pour présenter cette méthode, remarquons que l'on peut écrire la densité spectrale du processus $FARIMA(p, d, q)$ sous la forme :

$$f(\lambda) = \{4 \sin^2(\lambda/2)\}^{-d} f_u(\lambda)$$

où $f_u(\lambda) = \frac{\sigma^2}{2\pi} \frac{|C(e^{-i\lambda})|^2}{|A(e^{-i\lambda})|^2}$ est la densité spectrale du processus ARMA $u_t = (1 - B)^d Y_t$. On peut écrire l'égalité suivante :

$$\ln \{f(\lambda)\} = \ln \{f_u(0)\} - d \ln \{4 \sin^2(\lambda/2)\} + \ln \{f_u(\lambda)/f_u(0)\}.$$

Soit :

$$I(\lambda) = \frac{1}{2\pi T} \left| \sum_{t=1}^T Y_t e^{-\lambda t} \right|^2$$

⁴On peut citer Robinson [1995(b)].

le périodogramme de la chronique, un estimateur naïf de la densité spectrale (cf. [Priestley, 1981]) du processus. Ce périodogramme est évalué en les fréquences de Fourier $\lambda_j = \frac{2\pi j}{T}, 1 \leq j \leq m_T$, où m_T est le paramètre de troncature.

On a alors :

$$\ln \{I(\lambda_j)\} = \ln \{f_u(0)\} - d \ln \{4 \sin^2(\lambda_j/2)\} + \ln \{f_u(\lambda_j)/f_u(0)\} + \ln \{I(\lambda_j)/f(\lambda_j)\}.$$

L'estimateur GPH exige les deux conditions suivantes :

- i) Pour les fréquences suffisamment basses $\ln \{f_u(\lambda_j)/f_u(0)\}$ est négligeable (cette condition est trivialement vérifiée pour les ARMA)
- ii) Les variables aléatoires $\ln \{I(\lambda_j)/f(\lambda_j)\}, 1 \leq j \leq m_T$, sont asymptotiquement indépendantes et identiquement distribuées (i.i.d.)

Sous ses deux conditions on obtient le modèle de régression suivant :

$$\ln \{I(\lambda_j)\} = \alpha - d \ln \{4 \sin^2(\lambda_j/2)\} + e_j,$$

où $(e_j)_{1 \leq j \leq m_T}$ est une suite i.i.d. de moyenne $-c$, et de variance $\pi^2/6, c = 0.57721\dots$ est la constante d'Euler. Posons :

$$R_j = -\ln \{4 \sin^2(\lambda_j/2)\}, \tag{15}$$

alors l'estimateur GPH est donné par :

$$\hat{d}_{GPH} = \frac{\sum_{j=1}^{m_T} (R_j - \bar{R}) \ln \{I(\lambda_j)\}}{\sum_{j=1}^{m_T} (R_j - \bar{R})^2}, \quad \bar{R} = \frac{1}{m_T} \sum_{j=1}^{m_T} R_j. \tag{16}$$

Geweke et Porter-Hudak [1983] ont montré la convergence en probabilité et la normalité asymptotique de cet estimateur pour $-0.5 < d < 0$ et en supposant que le paramètre de troncature $m_T = T^\delta, 0 < \delta < 1$.

Robinson [1995(b)] montre la convergence en probabilité ainsi que la normalité asymptotique de cet estimateur pour $-0.5 < d < 0.5$ et sous des conditions de régularités assez fortes.

3.2. MÉTHODES EN UNE SEULE ÉTAPE

Ces méthodes consistent à estimer simultanément les paramètres $d, (a_j)_{1 \leq j \leq p}$ et $(c_j)_{1 \leq j \leq q}$ du modèle. On présente ici la méthode du maximum de vraisemblance considérée parmi les méthodes les plus efficaces et les plus répandues permettant d'estimer conjointement ces paramètres, cf. [Boutahar *et al.* [2006] pour d'autres méthodes.

L'estimateur de maximum de vraisemblance exact $\hat{\theta}$ d'un processus *FARIMA* (p, d, q) est donné par :

$$\hat{\theta} = \arg \max_{\theta \in \Theta} L_T(Y, \theta),$$

avec $L_T(Y, \theta)$ est la fonction log vraisemblance définie par :

$$L_T(Y, \theta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma(\theta)| - \frac{1}{2} Y^t \Sigma^{-1}(\theta) Y,$$

où $\Sigma(\theta)$ est la matrice d'autocovariance de $Y = {}^t(Y_1, \dots, Y_T)$ et $\theta = (\sigma_\varepsilon^2, d, a_1, \dots, a_p, c_1, \dots, c_q)$ est le vecteur des paramètres. Autrement dit, $\hat{\theta}$ est la solution de l'équation suivante:

$$\begin{aligned} \frac{\partial}{\partial \theta_j} L_T(Y, \theta) &= -\frac{1}{2} \frac{\partial}{\partial \theta_j} \ln |\Sigma(\theta)| - \frac{1}{2} Y^t \left[\frac{\partial}{\partial \theta_j} \Sigma^{-1}(\theta) \right] Y \\ &= 0, \end{aligned}$$

avec $j = 1, \dots, p + q + 2$.

Dahlhaus [1989] a établi la consistance, la normalité asymptotique ainsi que l'efficacité de cet estimateur.

3.3. TESTS

3.3.1. Test de longue mémoire

Pour tester la présence de longue mémoire on considère⁵ le test semi-paramétrique, basé sur l'estimateur *GPH*.

Pour un *FARIMA*(p, d, q), le test à effectuer est $H_0 : d = 0$ contre $H_1 : d \neq 0$.

D'après Geweke et Porter-Hudak [1983], la statistique :

$$\frac{1}{\pi} \left(6 \sum_{j=1}^{m_T} (R_j - \bar{R})^2 \right)^{1/2} \hat{d}_{GPH}$$

est distribuée sous H_0 suivant une loi $N(0, 1)$, où les R_j sont données par (15).

3.3.2. Tests d'autocorrélations des résidus

TEST DE BOX-PIERCE : Ce test permet de tester l'hypothèse nulle qui suppose que les résidus sont des bruits blancs, i.e. toutes leurs autocorrélations sont nulles, contre l'hypothèse alternative de présence de corrélations. Box et Pierce [1970] proposent la statistique Q suivante :

$$Q(l) = T \sum_{j=1}^l \hat{\rho}(j)^2$$

où $\hat{\rho}(j)$ est l'estimateur empirique de la fonction d'autocorrélation $\rho(j)$.

Sous l'hypothèse nulle, $Q(l)$ est distribuée asymptotiquement selon une $\chi^2(l)$, loi de Khi-deux à l degrés de liberté.

TEST DE LUJUNG-BOX : Ljung-Box [1978] ont proposé une modification de la statistique Q :

$$MQ(l) = T(T+2) \sum_{j=1}^l \frac{\hat{\rho}(j)^2}{T-j}$$

Sous l'hypothèse nulle, $MQ(l)$ est distribuée asymptotiquement selon une $\chi^2(l)$.

⁵Pour d'autres tests de présence de longue mémoire nous renvoyons le lecteur à Boutahar [2006].

3.3.3. Test de changement de régime : test de Chow

C'est un test classique qui permet de savoir si les coefficients sont identiques au cours de deux ou plusieurs sous-périodes. C'est à Chow [1960] que l'on doit un article novateur et très influant à l'époque de sa parution dans lequel il présente une statistique en F habituelle pour tester l'instabilité des paramètres. Le problème consiste à diviser l'échantillon d'observations en deux parties. Pour chacun de ces deux échantillons on ajuste un modèle de régression de vecteurs de paramètres β_1 et β_2 de même dimension K . Donc, il s'agit de tester :

$$\begin{cases} H_0 : \beta_1 = \beta_2 \\ \text{contre} \\ H_1 : \beta_1 \neq \beta_2 \end{cases} .$$

La statistique de Chow est donnée par :

$$F = \frac{(SSR - SSR_1 - SSR_2) / K}{(SSR_1 + SSR_2) / (T - 2K)} \quad (17)$$

où SSR est la somme des carrés des résidus sous H_0 , SSR_1 est la somme des carrés des résidus à partir de la régression linéaire du premier régime, SSR_2 est la somme des carrés des résidus à partir de la régression linéaire du deuxième régime et T est le nombre total des observations.

4. APPLICATION À DES CHRONIQUES DÉMOGRAPHIQUES

4.1. MODÉLISATION LONGUE MÉMOIRE

On considère les chroniques de naissances de dix pays (France, Belgique, Italie, Australie, Danemark, Suède, Norvège, Suisse, Canada, États-Unis) en fréquences annuelles sur des périodes différentes (cf. Tableau 1). Ce panier de pays a été choisi selon la longueur de leurs chroniques afin de tester convenablement la présence de longue mémoire.

Pays	Périodes des données
Australie	1871-1997
Belgique	1840-2002
Canada	1921-2003
Danemark	1835-2002
États-Unis	1920-2003
France	1806-2002
Italie	1862-2001
Norvège	1846-2002
Suède	1749-2003
Suisse	1871-2003

Table 1. Présentation des chroniques des naissances pour les différents pays.

Dans ce qui suit, on donne les résultats de l'analyse sur les différentes chroniques, en commençant par la stationnarité, puis on teste la présence de longue mémoire,

ensuite on présente les modèles identifiés et enfin on expose les résultats des tests d'autocorrélations sur les résidus.

4.1.1. Étude de la stationnarité

En se basant sur le test de Dickey-Fuller Augmenté introduit par Dickey et Fuller [1981], nous testons l'hypothèse d'existence de racine unitaire, soit $H_0 : d = 1$ contre $H_1 : d \neq 1$, l'hypothèse alternative de non existence. Les résultats du test pour les différents pays sont résumés dans le Tableau 2.

Pays	Stationnarité
Australie	Rejetée
Belgique	Rejetée
Canada	Rejetée
Danemark	Rejetée
Italie	Rejetée
Suède	Rejetée
États-Unis	Acceptée à 10%
France	Acceptée à 10%
Norvège	Acceptée à 5%
Suisse	Acceptée à 10%

Résultats du test de Dickey-Fuller augmenté(*ADF*).

Table 2. Test de stationnarité des chroniques.

Conformément aux résultats obtenus, on peut rejeter l'hypothèse nulle de racine unitaire pour la France, la Norvège, la Suisse et les États-Unis. Ce sont les quatre chroniques qui feront l'objet de comparaison de modèles.

4.1.2. Test de la longue mémoire

Au regard des fonctions d'autocorrélation pour les différents pays, on constate une décroissance lente pour la France, la Norvège, la Suisse et les États-Unis. Tout ceci peut constituer un indicateur de l'existence de longue mémoire. Nous allons vérifier cette hypothèse de longue mémoire à l'aide du test *GPH* précédent. L'examen du Tableau 3 suggère la présence de longue mémoire pour chacun des quatre pays. Nous proposons et estimons un modèle FARIMA(p,d,q) pour les chroniques des naissances de ces quatre pays.

Pays	Test <i>GPH</i>
France	3.0194
Suisse	3.4327
Norvège	2.9382
États-Unis	4.1156

Valeurs de la statistique *GPH* pour les pays où nous avons détecté la présence de la longue mémoire. La statistique *GPH* est significative à 1% pour les quatre pays.

Table 3. Test de mémoire longue.

4.1.3. Tests d'autocorrélation des résidus

Les résultats des tests sont présentés dans le Tableau 5.

Conformément aux résultats obtenus, quatre pays (France, Norvège, Suisse et États-Unis) sur les dix présentent de la persistance dans leurs chroniques des naissances. De plus, cette modélisation est validée par l'absence d'autocorrélation des résidus en appliquant les tests de Ljung-Box et Box-Pierce.

4.1.4. Modèles retenus

La modélisation de la chronique des naissances pour les différents pays a permis de retenir les modèles présentés dans le Tableau 4.

Pays	Ljung-Box	Box-Pierce
France	0.9269	0.9373
Suisse	0.2447	0.2902
Suède	0.7203	0.695
États-Unis	0.6907	0.7615

Les p-values avec lesquelles nous acceptons l'hypothèse nulle d'absence d'autocorrélations des résidus jusqu'à l'ordre 10.

Table 4. Analyse des résidus.

Les modèles $FARIMA(p, d, q)$ retenus sont comme suit :

- Pour la France :

$$(1 - B)^{0.401} Y_t = (1 - 0.55B - 0.48B^2 - 0.49B^3 - 0.27B^4) \varepsilon_t$$

- Pour les États Unis :

$$(1 - B)^{0.493} Y_t = (1 - 1.11B - 0.48B^2 - 0.43B^3 - 0.25B^4 - 0.48B^5 - 0.15B^6) \varepsilon_t$$

- Pour la Norvège :

$$\begin{aligned} & (1 - B)^{0.32} (1 + 1.37B - 0.47B^2) Y_t \\ &= (1 + 1.16B - 0.34B^2 - 0.16B^3 - 0.02B^4 + 0.18B^5 - 0.73B^6) \varepsilon_t \end{aligned}$$

- Pour la Suisse :

$$(1 - B)^{0.47} Y_t = (1 - 0.79B - 0.82B^2 - 0.63B^3 - 0.49B^4) \varepsilon_t$$

4.2. MODÉLISATION AVEC CHANGEMENT DE RÉGIME

La modélisation avec changement de régime est présentée de manière détaillée pour la chronique française: méthodologie, application des tests et modélisation des résidus. Pour les chroniques des trois autres pays, seuls les modèles retenus et les comparaisons sont présentés.

Pays	Modèles retenus
France	FARIMA(0,0.401,4)
Norvège	FARIMA(2,0.32,6)
États-Unis	FARIMA(0,0.493,6)
Suisse	FARIMA(0,0.47,4)

Table 5. Modèles retenus pour les différentes chroniques à longue mémoire.

4.3. CAS DE LA FRANCE

L'utilisation des modèles à changement de régime résulte de l'importance des points de rupture dans la chronique visualisée par la Figure 9. En effet, d'après le graphique nous remarquons quatre périodes où il y a une chute notable des naissances. La première correspond à la période 1867-1872 qui a été marquée par la guerre contre la Prusse (1870), plusieurs épidémies comme la variole, la rougeole et la dysenterie⁶.

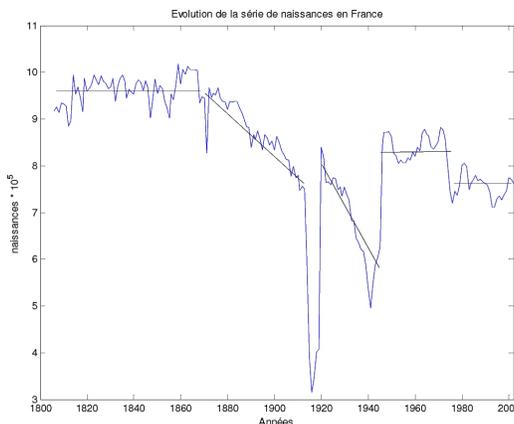


FIGURE 9. Changement de régime dans la chronique des naissances en France

La deuxième chute se situe aux alentours de la première guerre mondiale (1914-1918) qui a provoqué un effondrement temporaire de la fécondité en France. La troisième période (1938-1945) est relative à la deuxième guerre mondiale dont l'effet sur les naissances en France a été moins important que celui de la première. Enfin, une chute de la fécondité a été repérée pendant la période (1970-1975) où l'âge des femmes lors de la naissance du premier enfant a augmenté significativement⁷.

Pour chaque période on va déterminer la date de rupture la plus significative en appliquant le test de Chow pour chaque année et en choisissant comme date de rupture celle qui correspond à la plus petite p-value.

La modélisation avec changement de régime pour la chronique des naissances française est donnée par la formule (1) avec $m = 4$, $\mu_t = a + bt$, soit :

$$y_t = a_i + b_i t + u_t \quad \text{si } T_{i-1} < t \leq T_i, \quad (18)$$

⁶Pour plus de détails sur l'évolution des naissances en France pendant le XIX^e siècle, cf. Dupâquier [1988].

⁷Pour plus de détails sur l'évolution des naissances en France pendant le XX^e siècle, cf. Daguet [2002].

avec $T_0 = 1806$; $T_1 = 1870$; $T_2 = 1918$; $T_3 = 1945$; $T_4 = 1973$ et $T_{m+1} = 2002$.

H_0	H_1	Statistique F	p-value	Résultats
] 1806 ; 2002]]1806 ; 1870] et] 1870 ; 2002]	11.683	0	rupture en 1970
] 1870 ; 2002]]1870 ; 1918] et] 1918 ; 2002]	41.125	0	rupture en 1918
] 1918 ; 2002]]1918 ; 1945] et] 1945 ; 2002]	66.423	0	rupture en 1945
] 1945 ; 2002]]1945 ; 1973] et] 1973 ; 2002]	23.019	0	rupture en 1973

Une p-value < 0.01 indique qu'on rejette l'hypothèse H_0 , donc la présence de rupture.

Table 6. Résultats du test de Chow

L'autocorrélation des résidus (Figure 10) obtenus à partir de l'estimation du modèle (18) permet de dire qu'il y a une dépendance de court terme. Pour cela, on modélise les résidus $\hat{u}_t, 1 \leq t \leq T$, pour les cinq régimes par un $ARMA(p, q)$. La sélection de p et q est faite en se basant sur le critère d'information AIC d'Akaike [1970]. En effet, l'examen des AIC des différents modèles nous a permis de retenir un $ARMA(2, 2)$ comme modèle pour les résidus $\hat{u}_t, 1 \leq t \leq T$. Cette modélisation est validée par l'absence d'autocorrélation des résidus retenue par les tests de Ljung-Box [1978] et Box-Pierce [1970] (cf. Tableau 7 et Figure 11).

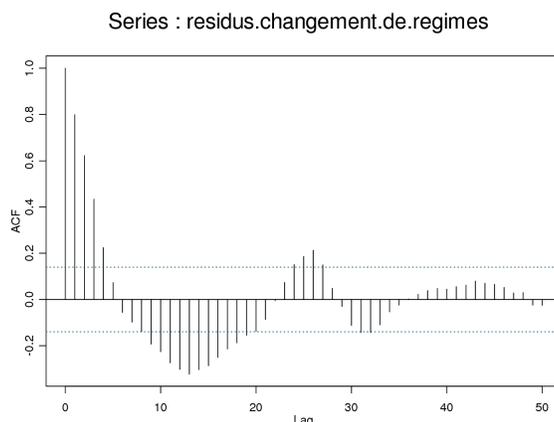


FIGURE 10. La fonction d'autocorrélation des résidus

Le modèle retenu pour la modélisation à changement de régime se présente comme suit⁸ :

$$\left\{ \begin{array}{l} y_t = a_i + b_it + u_t \text{ si } T_{i-1} < t \leq T_i; \\ T_0 = 1806, T_1 = 1870, T_2 = 1918, T_3 = 1945, T_4 = 1973, T_{m+1} = 2002; \\ a_1 = 12.66, a_2 = 38.12, a_3 = 34.09, a_4 = 11.75, a_5 = 15.69; \\ b_1 = 0.0006, b_2 = -0.013, b_3 = -0.0107, b_4 = 0.001, b_5 = -0.0011; \\ (1 + 1.37B - 0.58B^2) u_t = (1 + 0.75B - 0.33B^2) \varepsilon_t. \end{array} \right.$$

⁸Le modèle est présenté pour la chronique des naissances en France que nous réduisons à une échelle de 10^5 .

l	Ljung-Box	Box-Pierce
1	0.928	0.928
2	0.680	0.685
3	0.521	0.531
4	0.678	0.688
5	0.712	0.724
6	0.454	0.477
7	0.487	0.513
8	0.474	0.504
9	0.485	0.518
10	0.569	0.602
11	0.559	0.597
12	0.578	0.618

Les p-values calculées permettent d'accepter l'hypothèse nulle d'absence d'autocorrélation jusqu'à l'ordre l .

Table 7. Résultats des tests d'autocorrélations sur les résidus du modèle à changement de régime.

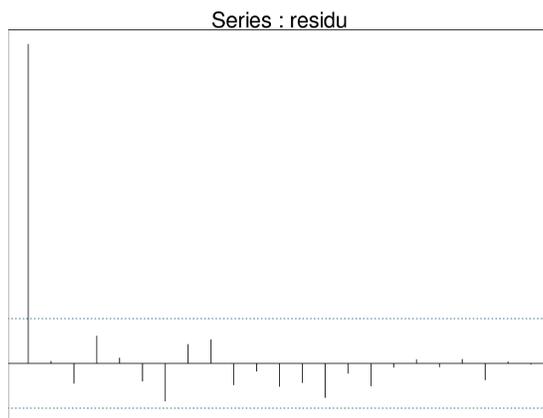


FIGURE 11. La fonction d'autocorrélation des résidus du modèle ARMA(2,2)

4.3.1. Comparaison

L'étude de la chronique des naissances pour la France par un modèle à longue mémoire FARIMA repose essentiellement sur la décroissance hyperbolique de la fonction d'autocorrélation et le test de longue mémoire de GPH. De plus, les sauts observables dans cette chronique laissent entendre une possible modélisation avec changement de régime. En se basant sur le test de Chow, cinq régimes ont été détectés en concordance avec des faits historiques déjà cités dans le document. Toutefois, une comparaison de ces deux modélisations semble être intéressante pour voir quel modèle sera plus convenable pour la chronique des naissances pour la France. Pour cela nous calculons la somme des carrés des résidus pour chacun des modèles, soit SSR_{LM} pour les longues mémoires et SSR_{CR} pour les changement de régime. La comparaison des performances est donnée par le rapport de ces deux quantités soit :

$$\frac{SSR_{LM}}{SSR_{CR}} = 1.14$$

Donc statistiquement la modélisation avec changement de régime semble mieux décrire la chronique des naissances en France.

4.4. CAS DES ÉTATS-UNIS

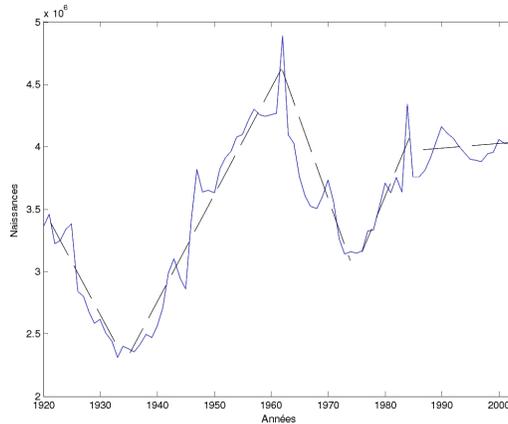


FIGURE 12. Changement de régime dans la chronique des naissances aux États-Unis

Le modèle retenu pour la modélisation à changement de régimes se présente comme suit:

$$\left\{ \begin{array}{l} y_t = a_i + b_i t + u_t \text{ si } T_{i-1} < t \leq T_i; \\ T_0 = 1920, T_1 = 1934, T_2 = 1962, T_3 = 1975, T_4 = 1984, T_{m+1} = 2003; \\ a_1 = 1713.66, a_2 = -1694.52, a_3 = 1474.12, a_4 = -2139.21, a_5 = -139.21; \\ b_1 = -0.8744, b_2 = 0.8876, b_3 = -0.7307, b_4 = 1.0986, b_5 = 0.0897; \\ (1 + 0.414B - 0.207B^2) u_t = \varepsilon_t. \end{array} \right.$$

4.4.1. Comparaison

La comparaison des performances est donnée par le rapport de ces deux quantités soit :

$$\frac{SSR_{LM}}{SSR_{CR}} = 2.486$$

Donc, statistiquement la modélisation avec changement de régime semble mieux décrire la chronique des naissances aux États-Unis.

4.5. CAS DE LA NORVÈGE

Le modèle retenu pour la modélisation à changement de régime se présente comme suit:

$$\left\{ \begin{array}{l} y_t = a_i + b_i t + u_t \text{ si } T_{i-1} < t \leq T_i; \\ T_0 = 1846; T_1 = 1894; T_2 = 1935; T_3 = 1973, T_{m+1} = 2002; \\ a_1 = -7.0459, a_2 = 10.6855, a_3 = -8.5631, a_4 = -5.1834; \\ b_1 = 0.0041, b_2 = -0.0053, b_3 = 0.0047, b_4 = 0.0029; \\ (1 + 0.752B + 0.0602B^2) u_t = \varepsilon_t. \end{array} \right.$$

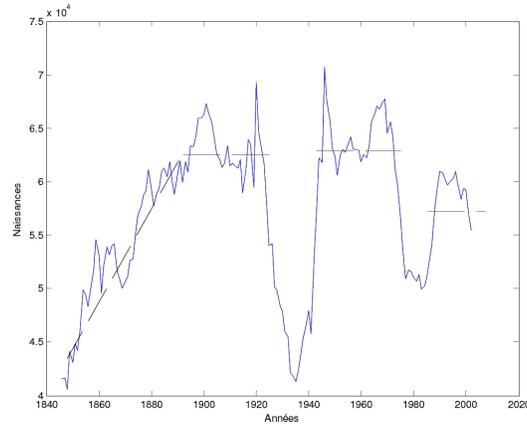


FIGURE 13. Changement de régime dans la chronique des naissances pour la Norvège

4.5.1. Comparaison

La comparaison des performances est donnée par le rapport de ces deux quantités soit :

$$\frac{SSR_{LM}}{SSR_{CR}} = 0.68$$

Donc statistiquement la modélisation à longue mémoire semble mieux décrire la chronique des naissances pour la Norvège.

4.6. CAS DE LA SUISSE

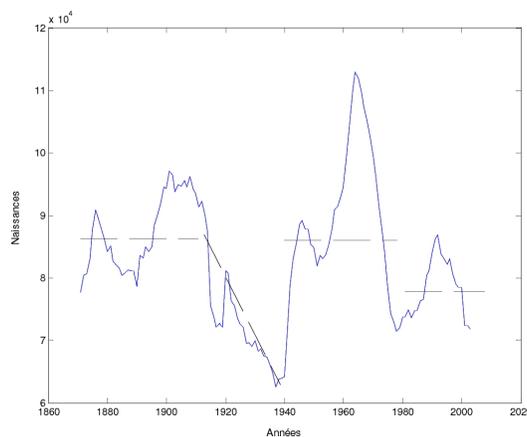


FIGURE 14. Changement de régime dans la chronique des naissances pour la Suisse

Le modèle retenu pour la modélisation à changement de régime se présente comme suit:

$$\left\{ \begin{array}{l} y_t = a_i + b_i t + u_t \text{ si } T_{i-1} < t \leq T_i; \\ T_0 = 1871; T_1 = 1913; T_2 = 1940; T_3 = 1979, T_{m+1} = 2003; \\ a_1 = -5.5304, a_2 = 12.7947, a_3 = -1.5051, a_4 = -0.9454; \\ b_1 = 0.0034, b_2 = -0.0063, b_3 = 0.0012, b_4 = 0.0009; \\ (1 + 1.851B - 0.902B^2) u_t = (1 + 1.013B - 0.180B^2) \varepsilon_t. \end{array} \right.$$

4.6.1. Comparaison

La comparaison des performances est donnée par le rapport de ces deux quantités soit :

$$\frac{SSR_{LM}}{SSR_{CR}} = 0.56$$

Donc statistiquement la modélisation à longue mémoire semble mieux décrire la chronique des naissances pour la Suisse.

5. CONCLUSION

Nous avons présenté une classe populaire de modèles à longue mémoire (les *FARIMA*) permettant de modéliser les persistances observées pour quelques pays comme les États-Unis, la France, la Norvège et la Suisse.

La modélisation longue mémoire a un avantage au niveau de la simplicité du modèle estimé, (estimation des paramètres $d, (a_j)_{1 \leq j \leq p}$ et $(c_j)_{1 \leq j \leq q}$) et a montré sa supériorité pour les chroniques de la Norvège et de la Suisse. Cependant, la comparaison entre les modèles à longue mémoire et avec changement de régime a montré la supériorité de ce dernier pour la modélisation de la chronique des naissances de la France et surtout des États-Unis. La représentation avec changement de régime est souvent lourde à réaliser, (estimation du nombre de régimes m , des dates de rupture T_i , et les paramètres des sous-régimes a_i, b_i) ce qui nécessite l'utilisation de plusieurs procédures, néanmoins le modèle avec changement de régime offre une interprétation plus claire des ruptures liées notamment à des faits historiques. En conclusion les deux concepts longue mémoire et changement de régime semblent être deux outils concurrents dans le sens où ils capturent la dynamique d'une chronique avec des performances qui peuvent être semblables.

Remerciements. Les auteurs remercient le Comité Mixte pour la Coopération Universitaire franco-tunisienne (CMCU) pour l'aide accordée à Ahdi Noomen Ajmi, ainsi que les deux rapporteurs anonymes grâce auxquels le travail a pu être amélioré.

BIBLIOGRAPHIE

- AKAIKE H., "Fitting autoregression for prediction", *Annals of the Institute of Statistical Mathematics* 21, 1970, p. 243-47.
- BAI J., PERRON P., "Estimating and testing linear models with multiple structural changes", *Econometrica* 66, 1998, p. 47-78.

- BERAN J., "Statistical methods for data with long-range dependence", *Statistical Science* 7 (4), 1992, p.404-427.
- BERAN J., *Statistics for long-memory processes*, Chapman & Hall, 1994.
- BERAN J., TERRIN N., "Testing for a change of the long-memory parameter", *Biometrika* 83 (3), 1996, p. 627-638.
- BOUTAHAR M., *Comparaison des tests non-paramétriques et semi-paramétriques dans la détection de la longue mémoire*, preprint, 2006.
- BOUTAHAR M., MARIMOUTOU V., NOUIRA L., "Estimation methods of the long memory parameter: Monte Carlo analysis and application", *Journal of Applied Statistics* 34 (3), 2006, p. 261-301.
- BOUHATAR M., JOUINI J., "Structural change and long memory approaches: methodology and numerical illustration" [soumis à *Economic modelling*, 2007].
- BOX G.E.P., JENKINS G.M., *Time series analysis forecasting and control*, San Francisco, Holden Day, 1976.
- BOX G.E.P., PIERCE D.A., "Distribution of residual autocorrelations in autoregressive-integrated moving-average time series models", *J. Amer. Stat. Assoc* 65, 1970, p. 1509-1526.
- CHOW G., "Tests of equality between sets and regression coefficients in two linear regressions", *Econometrica* 28, 1960, p. 591-605.
- DAGUET F., *Un siècle de fécondité française*, Collection INSEE résultats 8, 2002.
- DAHLHAUS R., "Efficient parameter estimation for self-similar processes", *The Annals of Statistics* 17, 1989, p. 1749-1766.
- DAVIDSON J., SIBERTSEN P., "Generating schemes for long memory processes: regimes, aggregation and linearity", *Journal of Econometrics* 128 (2), 2005, p. 253-282.
- DENIAU C., DOUKHAN P. OPPENHEIM G. RENAULT E., "Théorèmes limites et longue mémoire en statistique", *Statistical Inference for Stochastic Processes* 3, Special Issue, 2000.
- DICKEY D.A., FULLER W.A., "Likelihood ratio statistics for autoregressive time series with a unit root", *Econometrica* 49, 1981, p. 1057-1072.
- DIEBOLD F. X., INOUE A., "Long memory and regime switching", *Journal of Econometrics* 105, 2001, p. 131-159.
- DOUKHAN P., OPPENHEIM G., TAQQU M. S., *Theory and applications of long-range dependence*, Basel (Suisse), Birkhäuser, 2003.
- DUFRENOT G., LARDIC S., MATHIEU L., MIGNON V, PEGUIN A., "Explaining the european exchange rates deviations: long memory or nonlinear adjustment?", *Journal of International Financial Markets, Institutions and Money* 18 (2), 2008 [forthcoming].
- DUPÂQUIER J., *Histoire de la population française*, tome 3 : De 1789 à 1914, Paris, Presses Universitaires de France, 1988.
- GEWEKE J. PORTER-HUDAK S., "The estimation and application of long-memory time series models", *Journal of Time Series Analysis* 4, 1983, p. 221-238.
- GIL-ALANA L.A., "A fractional integration analysis of the population in some OECD countries", *Journal of Applied Statistics* 30 (10), 2003, p. 1-13.
- GRAF W.L., "The arroyo problem: paleohydrology and paleohydraulics in the short term", Gregory K. G. (ed.), *Background to paleohydrology*, New York, Wiley, 1983, p. 279-302.
- GRANGER C. W. J., JOYEUX R., "An introduction to long-memory time series models and fractional differencing", *Journal of Time Series Analysis* 1, 1980, p. 15-29.

- GOURIEROUX C., MONFORT A., *Séries Temporelles et modèles dynamiques*, Paris, Economica, 1990.
- HASSLER U., WOLTERS J., "Long memory in inflation rates: international evidence", *Journal of Business and Economic Statistics* 13, 1995, p. 37-45.
- HOSKING J. R. M., "Fractional differencing", *Biometrika* 68, 1981, p. 165-176.
- HOSKING J. R. M., "Modeling persistence in hydrological time series using fractional differencing", *Water Resources Research* 20, 1984, p. 1898-1908.
- HURST H. E., "Long-term storage capacity of reservoirs", *Transactions of the American Society of Civil Engineers* 116, 1951, p. 770-799.
- LO A., MACKINLAY C., "The size and power of the variance ratio test in finite samples: a Monte Carlo investigation", *Journal of Econometrics* 40, 1989, p. 203-238.
- LO A., "Long-term memory in stock market prices", *Econometrica* 59, 1991, p. 1279-1313.
- LJUNG G. M., BOX G. E. P., "On a measure of lack of fit in time series models", *Biometrika* 65, 1978, p. 553-564.
- MANDELBROT B. B., "Sur certains prix spéculatifs : faits empiriques et modèle basé sur les processus stables additifs de Paul Lévy", *Comptes Rendus Académie des Sciences* 254, 1962, p. 3968-3970.
- MANDELBROT B. B., VAN NESS J. W., "Fractional brownian motions, fractional noises and applications", *SIAM Review* 10, 1968, p. 422-437.
- MISHRA T. K., URBAIN J. P., "Stochastic demographic dynamics, economic growth and long-memory", 2006 [soumis à publication].
- PRIESTLEY M. B., *Spectral analysis of time series*, Academic Press, 1981.
- ROBINSON P. M., "Semiparametric analysis of long-memory time series", *The Annals of Statistics* 22, 1994, p. 515-539.
- ROBINSON P. M., "Log-periodogram regression of time series with long range dependence", *The Annals of Statistics* 23, 1995(a), p. 1048-1072.
- ROBINSON P. M., "Gaussian semiparametric estimation of long range dependence", *The Annals of Statistics* 23, 1995(b), p. 1630-1661.
- ROBINSON P. M., *Time series with long memory*, Oxford, Oxford University Press, 2003.
- TAQQU M. S., WILLINGER W., SHERMAN R., WILSON D., "Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level", *IEEE/ACM Transactions on Networking* 5 (1), 1997, p. 71-96.
- TORRE K., DELIGNIERES D., LEMOINE L., "Detection of long-range dependence and estimation of fractal exponents through ARFIMA modelling", *British Journal of mathematical and Statistical Psychology* [in press].
- WAGENMAKERS E. J., FARREL S., RATCLIFF R., "Estimation and interpretation of 1/f anoise in human cognition", *Psychonomic Bulletin and Review* 11 (4), 2004, p. 579-615.
- WILLINGER W., TAQQU M. S., TEVEROVSKY V., "Stock market prices and long memory dependance", *Finance and stochastics* 3, 1999, p. 1-13.
- YAO Y. C., "Estimating the number of change-points via Schwarz' criterion", *Statistics and Probability Letters* 6, 1988, p. 181-189.