JOURNAL DE LA SOCIÉTÉ FRANÇAISE DE STATISTIQUE

SANDRINE LARDIC
VALÉRIE MIGNON

La mémoire longue en économie : une revue de la littérature

Journal de la société française de statistique, tome 140, nº 2 (1999), p. 5-48
<http://www.numdam.org/item?id=JSFS_1999__140_2_5_0>

LA MÉMOIRE LONGUE EN ÉCONOMIE : UNE REVUE DE LA LITTÉRATURE

Sandrine LARDIC *, Valérie MIGNON **

RÉSUMÉ

Avec la reconnaissance de la complexité de la dynamique de nombreuses séries économiques et financières, l'utilisation de nouvelles procédures économétriques est apparue nécessaire. En particulier, diverses études récentes tendent à montrer l'existence de phénomènes de non linéarité et de mémoire longue sur certains marchés financiers. Afin de prendre en compte ces deux caractéristiques, Granger et Joyeux (1980) et Hosking (1981) ont élaboré les processus ARFIMA. Nous proposons ici une étude théorique des phénomènes de mémoire longue au travers de ces processus. Après avoir rappelé la définition des processus ARFIMA, nous passons en revue les différentes techniques d'estimation. Les récents travaux relatifs aux extensions des processus ARFIMA sont également abordés avant de dresser un panorama de la littérature orientée vers les applications.

Mots clés — mémoire longue, processus ARFIMA, méthodes d'estimation, processus GARMA, processus FIGARCH.

ABSTRACT

In order to take into account the complexity of a great number of economic and financial time series, new econometric methods have been developed. Specially, various recent analyses show the presence of two types of phenomena on financial markets : non linearity and long-term memory In order to take into account these characteristics, Granger and Joyeux (1980) and Hosking (1981) have introduced ARFIMA processes The aim of this paper is to present a theoretical study of long-term memory phenomena by means of such processes. After having recalled the definition of ARFIMA processes, we review the various estimation procedures. Recent developments of ARFIMA processes are also presented. Finally, we survey empirical results.

Keywords — long-term memory, ARFIMA processes, estimation methods, GARMA processes, FIGARCH processes.

Codification au Journal of Economic Literature : C22.
Classification AMS : 90A20, 62M10.

* CCF, Direction de la Recherche et de l'Innovation, et MODEM.
** Université de Valenciennes et MODEM, Université Paris X – Nanterre, 200 avenue de la République, 92001 Nanterre Cedex, Tél. : 01 40.97.77 84

1. INTRODUCTION

L'hypothèse d'indépendance des séries temporelles est dans la plupart des cas uniquement une approximation de la véritable structure de corrélation des séries. D'importantes corrélations pour de faibles retards peuvent parfois être détectées et des processus à mémoire courte, comme par exemple les processus ARMA, peuvent suffire à modéliser la structure de dépendance des séries.

Il existe cependant de nombreux exemples de données où les corrélations prises seules sont faibles, mais de somme extrêmement élevée. Le périodogramme de ces séries présente un pic dans le spectre à la fréquence zéro. De manière équivalente, dans le domaine temporel, les autocorrélations de ces séries diminuent très lentement. Un tel phénomène peut être le reflet de plusieurs possibilités, en particulier : la non stationnarité et la stationnarité avec dépendance à long terme. Nous nous intéresserons ici uniquement à la seconde explication par le biais d'une étude sur les processus à mémoire longue[1].

L'étude de la mémoire longue s'est initialement développée dans le domaine de l'hydrologie avec les travaux fondateurs de Hurst (1951) sur les crues du Nil; travaux qui ont par la suite été approfondis par Mandelbrot (1965, 1972), puis, plus récemment par Kashyap et Eom (1988). La mémoire de long terme a également fait l'objet d'analyses de la densité du trafic dans les réseaux à grande vitesse, mais aussi en géophysique, astronomie et économie. Les analyses de données astronomiques par Newcomb (1886) et Jeffreys (1939), de la physique atmosphérique par Graf (1983) et de données chimiques par Student (1927) font ressortir la présence de fortes autocorrélations. Smith (1938) et Whittle (1956) mettent également en avant un tel phénomène sur diverses séries agricoles[2]. Dans le domaine économique, plus spécifiquement, on peut par exemple citer l'étude du comportement du produit réel par Diebold et Rudebusch (1989) et Sowell (1992), l'analyse du comportement du revenu disponible et de l'hypothèse de revenu permanent par Diebold et Rudebusch (1991), la question de la prévisibilité de la rentabilité des titres et de l'hypothèse de marchés efficients (Lo, 1991; Lardic et Mignon, 1996, 1997; Mignon, 1996, 1998), les variances «bornées» pour la structure à terme des taux d'intérêt et l'hypothèse d'anticipation de Hicks (Shea, 1991), les dynamiques des taux de change réels et l'hypothèse de parité des pouvoirs d'achat (Diebold, Husted et Rush, 1991; Cheung et Lai, 1993; Lardic, 1997), les comportements des salaires réels et l'hypothèse de substitution intertemporelle (Hassett, 1990) ou encore les dynamiques des taux de change nominaux (Cheung, 1993)[3].

---

1. Pour plus de détails sur les deux premières explications, on pourra se reporter à Lardic et Mignon (1997). Notons simplement que les séries temporelles fortement persistantes se comportent de façon très semblable aux séries non stationnaires, notamment par la présence de cycles et des changements de niveau de tous ordres de taille (sur les problèmes de non stationnarité, Cf. Lardic (1992, 1996, 1997))

2. On pourra consulter l'ouvrage de Peitgen et Saupe (1988) qui présente des applications des processus à mémoire longue dans les domaines les plus divers.

3. Signalons également, pour des références plus récentes, le numéro spécial de *Statistical Inference for Stochastic Processes* (à paraître) reprenant les actes de la XIX^{e} rencontre

Les processus à mémoire longue peuvent être définis de façon équivalente dans le domaine temporel et le domaine spectral :

- Dans le domaine temporel, les processus à mémoire longue sont caractérisés par une fonction d'autocorrélation décroissant hyperboliquement au fur et à mesure que le retard s'accroît, alors que celle des processus à mémoire courte décroît exponentiellement.
- Dans le domaine fréquentiel, les processus à mémoire longue sont caractérisés par une densité spectrale s'accroissant sans limite quand la fréquence tend vers zéro.

De nombreux processus répondent à une telle définition, comme le bruit gaussien fractionnaire en temps continu ou en temps discret[4]. Nous nous intéresserons ici uniquement à une catégorie particulière de processus à mémoire longue : les processus ARFIMA (autorégressifs, moyenne mobile, fractionnairement intégrés). Par rapport aux autres modèles à mémoire longue, les processus ARFIMA présentent l'avantage d'être relativement facilement applicables en économie dans la mesure où ils constituent une extension directe des processus ARIMA usuels. En outre, si comme le mouvement brownien fractionnaire et le bruit gaussien fractionnaire, ils visent à modéliser le comportement de long terme des séries temporelles, les processus ARFIMA se distinguent par la prise en compte conjointe de la dynamique de court terme.

Dans un premier temps, nous nous attacherons à la représentation de la mémoire longue au travers de la définition des processus ARFIMA avant de détailler de la manière la plus exhaustive possible les diverses procédures visant à estimer les paramètres d'intérêt de ces processus. Nous présenterons ensuite les récents développements des processus ARFIMA visant, d'une part à tenir compte du comportement saisonnier des séries temporelles et, d'autre part, à modéliser la structure de dépendance de long terme des séries de volatilité. Nous terminerons cet article par une présentation des apports empiriques des études sur la mémoire longue.

2. REPRÉSENTATION DE LA MÉMOIRE LONGUE : LES PROCESSUS ARFIMA

Traditionnellement, l'étude des séries temporelles s'est centrée autour d'une alternative : la présence d'une racine unitaire, indiquant une non stationnarité de la série d'une part, et l'absence d'une telle racine unitaire impliquant que la série est stationnaire d'autre part. Du point de vue des modélisations de type Box et Jenkins, ces deux cas correspondent respectivement aux cas des processus ARIMA(p,d,q) et ARMA(p,q) standards. Ainsi, la présence

---

franco belge sur « Théorèmes limites et longue mémoire en statistique », 19 21 novembre 1998, Luminy.

4. Sur ces processus, on pourra consulter Samorodnitsky et Taqqu (1994), Mandelbrot et Wallis (1968, 1969a, 1969b, 1969c), Mandelbrot et Van Ness (1968) et Mandelbrot et Taqqu (1972). Une présentation sommaire est donnée dans Lardic et Mignon (1997) et Mignon (1998).

d'une racine unitaire ($d = 1$ dans les modèles ARIMA) fait référence au phénomène de mémoire infinie, puisque la non stationnarité en différence renvoie à des conséquences permanentes de toute perturbation, et l'absence de racine unitaire ($d = 0$) correspond au cas de mémoire courte, voire d'absence de mémoire. Ces modélisations classiques ne prennent pas en compte les cas intermédiaires, à savoir l'existence d'un paramètre de différenciation d fractionnaire. Or, la présence d'un tel coefficient non entier est particulièrement intéressante puisqu'elle caractérise typiquement les processus à mémoire longue. Ainsi, nous nous proposons ici d'étudier une généralisation des processus ARIMA standards où l'exposant de différenciation d n'est plus un entier mais un réel. Ces processus sont appelés ARFIMA et ont été introduits par Granger et Joyeux (1980) et Hosking (1981). Ils présentent l'intérêt de tenir compte à la fois du comportement de court terme de la série au travers des paramètres autorégressifs et moyenne mobile et du comportement de long terme par le biais du paramètre d'intégration fractionnaire.

Nous nous proposons de présenter ci-après les principales caractéristiques des processus ARFIMA en commençant par le processus le plus simple, ARFIMA$(0, d, 0)$, puis en généralisant ce processus pour donner lieu aux ARFIMA(p, d, q). Nous mettrons ensuite en avant la relation théorique existant entre l'exposant de Hurst et le paramètre d'intégration fractionnaire en ayant recours au bruit gaussien fractionnaire.

2.1. Définition des processus ARFIMA

Le processus ARFIMA le plus simple est le processus ARFIMA$(0, d, 0)$ ou bruit fractionnaire de paramètre d en temps discret. Il se définit comme suit :

$$(1 - L)^d X_t = u_t \tag{1}$$

où L est l'opérateur retard et u_t est un bruit blanc[5].

Le développement « binomial » est donné par :

$$\begin{aligned}\nabla^d = (1 - L)^d &= 1 - dL - \frac{d(1-d)}{2!} L^2 - \frac{d(1-d)(2-d)}{3!} L^3 - \dots \\ &= \sum_{j=0}^{\infty} \pi_j L^j\end{aligned} \tag{2}$$

$$\text{où } \pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} = \prod_{0<k\leqslant j} \left(\frac{k-1-d}{k}\right) \quad j = 0, 1, \dots \tag{3}$$

Γ étant la fonction eulérienne de seconde espèce.

Le processus ainsi défini est le processus ARFIMA$(0, d, 0)$. Les propriétés de ce modèle sont largement développées dans Hosking (1981), nous n'en reprendrons ici que les bases principales.

5. On rappelle qu'un bruit blanc est un processus (non nécessairement gaussien) non autocorrélé d'espérance nulle et de variance finie

Soit $\{X_t\}$ un processus ARFIMA$(0,d,0)$. Sous la condition $-\frac{1}{2} < d < \frac{1}{2}$,

- Le processus $\{X_t\}$ est stationnaire et inversible,
- La fonction d'autocovariance est donnée par :

$$\gamma_k = \frac{\Gamma(1-2d)\Gamma(k+d)}{\Gamma(d)\Gamma(1-d)\Gamma(k+1-d)} \tag{4}$$

$$\text{avec } \gamma_k \sim \frac{\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)}\, k^{2d-1} \text{ quand } k \longrightarrow \infty \tag{5}$$

- La fonction de densité spectrale s'écrit :

$$f(\omega,d) = \left(2\sin\frac{\omega}{2}\right)^{-2d} \quad \text{avec } 0 < \omega \leqslant \pi \tag{6}$$

$$\text{et } \lim_{\omega \to 0} f(\omega,d) = \omega^{-2d} \tag{7}$$

Les autocorrélations diminuent à un taux hyperbolique, donc nettement plus lentement que les autocorrélations des processus ARMA (qui décroissent à un taux géométrique). La densité spectrale n'est pas limitée à une valeur finie à la fréquence zéro. Ainsi, du fait de ces deux caractéristiques, les processus ARFIMA sont appelés processus à mémoire longue.

Le processus ARFIMA$(0,d,0)$ est un cas particulier des processus ARFIMA (p,d,q) qui peuvent être définis comme suit :

$$\Phi(L)X_t = \Theta(L)\varepsilon_t \tag{8}$$

où $\varepsilon_t = \nabla^{-d}u_t$

u_t est un bruit blanc de variance σ^2

$\Phi(L)$ et $\Theta(L)$ sont des polynômes retard de degrés p et q respectivement.

Le développement de la formule ci-dessus donne :

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = \varepsilon_t + \theta_1\varepsilon_{t-1} + \ldots + \theta_q\varepsilon_{t-q} \tag{9}$$

$$\text{avec } \varepsilon_t = u_t + du_{t-1} + \frac{d(d+1)}{2}\, u_{t-2} + \frac{d(d+1)(d+2)}{6}\, u_{t-3} + \ldots \tag{10}$$

Bien sûr, comme dans le cas du processus ARFIMA$(0,d,0)$, les processus ARFIMA(p,d,q) sont des processus à mémoire longue stationnaires et inversibles lorsque $d \in \left]-\frac{1}{2},\frac{1}{2}\right[$ et $d \neq 0$.

L'existence d'une relation entre l'exposant de Hurst (défini ci-dessous) et le paramètre d'intégration fractionnaire permet de classifier les séries temporelles en fonction de leur structure de dépendance ainsi que nous le rappelons ci-après.

2.2. Structure de la mémoire : exposant de Hurst et processus ARFIMA

Lors de son analyse sur les crues du Nil, Hurst (1951) a introduit une statistique permettant de détecter la présence de phénomènes de mémoire longue. Cette statistique, appelée analyse R/S, est particulièrement intéressante dans la mesure où elle donne lieu à un coefficient, appelé exposant de Hurst, permettant de classifier les séries temporelles en fonction de la nature de leur mémoire.

2.2.1. Analyse R/S et exposant de Hurst

La statistique R/S se définit comme l'étendue (R) des sommes partielles des écarts d'une série temporelle à sa moyenne divisée par son écart type (s_T). Ainsi, soit une série temporelle X_t, $t = 1, \ldots, T$, de moyenne $\overline{X}_T$, la statistique R/S, notée ici Q_T, s'écrit :

$$Q_T = R/s_T = \frac{1}{\left[\frac{1}{T}\sum_{j=1}^{T}(X_j - \overline{X}_T)^2\right]^{1/2}} \left[\underset{1 \leqslant k \leqslant T}{\mathrm{Max}} \sum_{j=1}^{k}(X_j - \overline{X}_T) - \underset{1 \leqslant k \leqslant T}{\mathrm{Min}} \sum_{j=1}^{k}(X_j - \overline{X}_T)\right] \quad (11)$$

Cette statistique est asymptotiquement proportionnelle à T^H (voir Hurst, 1951), où la constante H, $0 < H < 1$, est appelée exposant de Hurst. L'exposant de Hurst est ainsi donné par[6] :

$$H \cong \frac{\log Q_T}{\log T} \quad (12)$$

L'exposant de Hurst est particulièrement intéressant dans la mesure où sa valeur permet de classifier les séries temporelles en fonction de leur structure de dépendance. Le paragraphe suivant propose d'établir une telle classification en mettant en avant la relation existant entre exposant de Hurst et paramètre d'intégration fractionnaire des processus ARFIMA.

6. Notons qu'il existe de nombreuses autres estimations de l'exposant de Hurst. On peut par exemple citer l'analyse R/S modifiée (Lo, 1991), la méthode de la variance agrégée, la méthode des valeurs absolues de la variance agrégée, la méthode des résidus ou encore la procédure de Higuchi (1988) basée sur l'utilisation de la dimension fractale On trouvera une présentation de ces diverses techniques dans Taqqu, Teverovsky et Willinger (1995). Une application aux séries de rentabilités boursières figure dans Mignon (1998)

2.2.2. Relation entre exposant de Hurst et paramètre d'intégration fractionnaire

Ainsi que nous l'avons précédemment souligné, les premiers processus à mémoire longue ont été développés en temps continu. En particulier, Mandelbrot et Van Ness (1968) ont introduit une généralisation du mouvement brownien ordinaire, appelé mouvement brownien fractionnaire, afin de tenir compte de la dépendance de long terme entre les observations. Ainsi, le mouvement brownien fractionnaire d'exposant H, noté $B_H(t,\cdot)$, est donné par :

$$B_H(t,\cdot) = \frac{1}{\Gamma\left(H+\frac{1}{2}\right)} \left\{ \int_0^t (t-s)^{H-\frac{1}{2}} dB(s,\cdot) \right\} \tag{13}$$

où H est l'exposant de Hurst, $0 < H < 1$, et $B(s,\cdot)$ est le mouvement brownien ordinaire de variance unitaire.

Les incréments X_t du mouvement brownien fractionnaire :

$$X_t = B_H(t,\cdot) - B_H(t-1,\cdot) \tag{14}$$

forment un processus gaussien stationnaire appelé bruit gaussien fractionnaire en temps discret (Mandelbrot et Wallis, 1969c). La fonction d'autocovariance de ce processus est donnée par :

$$\gamma_k = \frac{1}{2}\left[|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right] \tag{15}$$

avec $\gamma_k \sim H(2H-1)k^{2H-2}$ quand $k \longrightarrow \infty$ (16)

En reprenant les formulations des fonctions d'autocovariance du bruit gaussien fractionnaire et des processus ARFIMA, on peut établir une relation entre le paramètre d des processus ARFIMA et l'exposant de Hurst H. En effet, bien qu'il s'agisse d'un processus à temps discret et d'un processus à temps continu, comparons la limite asymptotique de ces deux fonctions (relations (5) et (16)) ; on constate qu'elles ont la même puissance de déclin hyperbolique. Ainsi, en liant les deux exposants :

$$2H - 2 = 2d - 1 \tag{17}$$

nous obtenons la relation :

$$d = H - \frac{1}{2} \tag{18}$$

Dès lors, il est possible d'effectuer une classification des séries temporelles en fonction des valeurs du paramètre d :

- Si $0 < d < \frac{1}{2}$ le processus ARFIMA est un processus stationnaire à mémoire longue. Les autocorrélations sont positives et diminuent hyperboliquement vers 0 lorsque le retard augmente. La densité spectrale est concentrée autour des faibles fréquences (cycles lents), elle tend vers l'infini lorsque la fréquence tend vers zéro. On est face à un processus persistant.

- Si $d = 0$, le processus ARFIMA se réduit au processus ARMA standard et ne présente aucune structure de dépendance à long terme.
- Si $-\frac{1}{2} < d < 0$ le processus est anti-persistant. Les autocorrélations alternent de signe et la densité spectrale est dominée par des composantes de haute fréquence (la densité spectrale tend vers zéro lorsque la fréquence tend vers zéro).

Ces définitions étant posées, la section suivante propose de dresser un aperçu des différentes méthodes d'estimation des processus ARFIMA.

3. LES MÉTHODES D'ESTIMATION DES PROCESSUS ARFIMA

Les méthodes d'estimation des processus ARFIMA peuvent être regroupées en deux catégories selon que l'on estime ou non simultanément tous les paramètres de la représentation ARFIMA(p, d, q). On distingue ainsi les méthodes en deux étapes, qui sont les plus anciennes, et les méthodes en une étape[7].

3.1. Les méthodes en deux étapes

Les méthodes en deux étapes peuvent être réunies en trois techniques principales dont la caractéristique commune est d'estimer dans une première étape le paramètre d'intégration fractionnaire d puis, dans une seconde étape, d'estimer par les méthodes usuelles de séries temporelles les paramètres autorégressif et moyenne mobile de la représentation ARMA(p, q) de la série transformée. Les trois techniques que nous présentons se distinguent uniquement par la méthode de calcul du paramètre de différenciation fractionnaire d.

3.1.1. Méthode basée sur l'exposant de Hurst

La méthode la plus simple consiste à estimer d par l'exposant de Hurst. Nous avons en effet précédemment rappelé le lien entre l'exposant de Hurst et le paramètre d'intégration fractionnaire d : $\widehat{d} = \widehat{H} - \frac{1}{2}$. Cette technique, malgré sa facilité de mise en oeuvre, souffre selon nous de deux inconvénients majeurs. D'une part, ne disposant pas des écart-types associés à l'estimation de l'exposant de Hurst, ni de la loi de probabilité de ce dernier, nous ne pouvons effectuer de test de significativité sur la valeur de d. D'autre part, il est évident que s'il existe un biais dans le calcul de l'exposant de Hurst,

7. Outre les références bibliographiques données par la suite, le lecteur intéressé par les aspects statistiques relatifs aux techniques d'estimation pourra également se reporter aux travaux de Akonom (1988), Azaıs et Lang (1993) et Hardouin (1993).

celui-ci se répercutera alors inévitablement sur l'estimation du paramètre d'intégration fractionnaire. Or, précisément, on peut noter à la suite des résultats de simulations d'Hosking (1984), que l'exposant de Hurst estimé par l'analyse R/S est biaisé à la hausse dès lors que H est inférieur à 0,7 et à la baisse lorsque H est supérieur à 0,7.

3.1.2. Les méthodes spectrales

Il est possible de construire des estimateurs de d basés sur le comportement spectral aux basses fréquences de la série temporelle. Selon Olshen (1967), la densité spectrale a de meilleures propriétés échantillonnales que la fonction d'autocorrélation, dans le cas où les coordonnées spectrales aux fréquences $2\pi j/T$, $j = 1, \dots, T/2$ sont asymptotiquement non corrélées.

Deux procédures alternatives d'estimation exploitent la forme particulière de la densité spectrale des séries temporelles à intégration fractionnaire : Janacek (1982) estime d par intégration numérique du log périodogramme, Geweke et Porter-Hudak (1983) s'appuient sur le comportement de la densité spectrale autour de zéro [8]. En utilisant les fréquences basses, une régression univariée du log-périodogramme sur le log de la fréquence est estimée.

3.1.2.1. La méthode de Geweke et Porter-Hudak (1983)

La méthode de Geweke et Porter-Hudak (1983) ou « méthode de régression », est basée sur la forme de la densité spectrale :

$$f(\lambda) = |1 - e^{-i\lambda}|^{-2d} f_u(\lambda) \tag{19}$$

où $f_u(\lambda) = \dfrac{\sigma^2}{2\pi} \dfrac{|\Theta(e^{-i\lambda})|^2}{|\Phi(e^{-i\lambda})|^2}$ est la densité spectrale du processus ARMA(p, q) : $u_t = \nabla^d X_t$.

Si l'on prend le logarithme de l'expression ci-dessus, on obtient :

$$\ln f(\lambda) = \ln f_u(0) - d \ln|1 - e^{-i\lambda}|^2 + \ln \left[\frac{f_u(\lambda)}{f_u(0)}\right] \tag{20}$$

En remplaçant λ par la fréquence de Fourier $\lambda_j = \dfrac{2\pi_j}{T} \in (0, \pi)$ et en ajoutant $\ln I(\lambda_j)$ – le périodogramme de la série X – de part et d'autre de l'égalité, on obtient :

$$\ln I(\lambda_j) = \ln f_u(0) - d \ln|1 - e^{-i\lambda_j}|^2 + \ln \left[\frac{I(\lambda_j)}{f(\lambda_j)}\right] + \ln \left[\frac{f_u(\lambda_j)}{f_u(0)}\right] \tag{21}$$

où le périodogramme est défini par :

$$I(\lambda) = \frac{1}{2\pi T} \left| \sum_{j=1}^{T} X_j e^{ij\lambda} \right|^2 \tag{22}$$

8. On pourra également se reporter à Robinson (1994, 1995b) pour des approfondissements sur la procédure de Geweke et Porter Hudak.

Si λ_j est proche de zéro, le dernier terme est négligeable comparé aux autres termes à droite de l'égalité et l'on réécrit alors l'équation sous la forme d'une équation de régression linéaire simple :

$$Y_j = a + bZ_j + \xi_j \tag{23}$$

où $j = 1, 2, \ldots, m$ où m correspond aux ordonnées du périodogramme[9].

$Y_j = \ln I(\lambda_j)$

$a = \ln f_u(0)$

$b = -d$

$Z_j = \ln|1 - e^{-i\lambda_j}|^2$

$\xi_j = \ln \dfrac{I(\lambda_j)}{f(\lambda_j)}$

On estime alors simplement b et d par les moindres carrés ; l'estimateur étant fourni par[10] :

$$\widehat{d} = -\left[\sum_{j=1}^{m}(Z_j - \overline{Z})^2\right]^{-1}\left[\sum_{j=1}^{m}(Z_j - \overline{Z})(Y_j - \overline{Y})\right] \tag{24}$$

Geweke et Porter-Hudak (1983) montrent alors que, quand $-1/2 < d < 1/2$, la loi de l'estimateur $\widehat{d}$ de d tend vers une loi normale quand $T \longrightarrow \infty$:

$$\widehat{d} \sim N\left(d, \pi^2\left[6\sum_{j=1}^{m}(Z_j - \overline{Z})^2\right]^{-1}\right)$$

3.1.2.2. La méthode de Janacek (1982)

Janacek propose une technique de détection et d'estimation de d basée sur le logarithme du spectre puissance. Cette technique présente deux avantages : en premier lieu, le spectre n'est pas dépendant d'un modèle ajusté et en second lieu c'est une technique complémentaire aux méthodes du domaine temporel.

Soit une série temporelle X_t qui différenciée d fois conduit à une série stationnaire Z_t ayant un spectre rationnel $f_Z(\lambda)$. Le spectre de la série X_t s'écrit alors sous forme logarithmique :

$$\ln f_X(\lambda) = -d \ln[2(1 - \cos\lambda)] + \ln f_Z(\lambda) \tag{25}$$

Si l'on introduit de plus la fonction de poids $W(\lambda)$ définie par :

$$W(\lambda) = -\frac{1}{2}\ln[2(1 - \cos\lambda)] \tag{26}$$

9. On retient en général un nombre d'ordonnées égal à la racine carrée du nombre d'observations.

10. L'estimateur des moindres carrés ordinaires de b est naturellement donné par le rapport de la covariance empirique entre les Y_j et Z_j à la variance empirique des Z_j.

alors :

$$\ln f_X(\lambda) = 2dW(\lambda) + \ln f_Z(\lambda) \tag{27}$$

Par ailleurs, puisque

$$W(\lambda) = \sum_{K=1}^{\infty} \frac{1}{K} \cos K\lambda \tag{28}$$

on peut écrire, en utilisant les résultats des séries de Fourier :

$$\frac{1}{\pi}\int_0^{\pi} W(\lambda)\ln f_X(\lambda)d\lambda = d\sum_{K=1}^{\infty}\frac{1}{K^2} + \sum_{K=1}^{\infty}\frac{a_K}{2K} \tag{29}$$

et $f_Z(\lambda) = a_0 + a_1 \cos\lambda + a_2 \cos 2\lambda + a_3 \cos 3\lambda + \ldots$ (30)

Comme on peut le constater, le terme de droite de cette dernière équation peut être estimé sans spécifier un modèle particulier.

Si comme Geweke et Porter-Hudak (1983), Janacek part de la définition du log-spectre, il dévie de la procédure de Geweke et Porter-Hudak quant au calcul du paramètre d'intégration fractionnaire d : Janacek (1982) analyse les coefficients de Fourier de $f_Y(\lambda)$, c'est-à-dire les coefficients a_0, a_1, a_2, ... où :

$$a_K = \frac{1}{2\pi}\int_0^{\pi} \ln\left[2\pi f_X(\lambda)\right]\cos K\lambda \, d\lambda \tag{31}$$

a_0 est relié à l'erreur quadratique moyenne de prévision minimale σ^2 comme l'ont montré Kolmogorov (1940) ou Hannan (1970). Par ailleurs, Kolmogorov a également montré que l'erreur quadratique moyenne de prévision à un pas en avant σ_K^2 est donnée par :

$$\sigma_K^2 = \sigma^2(1 + b_1^2 + b_2^2 + \ldots + b_K^2)$$

où • σ^2 est l'erreur quadratique moyenne de prévision minimale :

$$\ln\sigma^2 = \frac{1}{\pi}\int_0^{\pi} \ln(2\pi)f(\lambda)d\lambda$$

• les coefficients b_K sont liés aux coefficients a_K par la relation :

$$\exp\left[\frac{1}{2}(a_1 z + a_2 z^2 + \ldots)\right] = 1 + b_1 z + b_2 z^2 + \ldots$$

En outre, on a les relations suivantes :

$$\begin{cases} b_1 = c_1 \\ b_2 = c_2 + \dfrac{c_1^2}{2!} \\ b_3 = c_3 + c_1 c_2 + \dfrac{c_1^3}{3!} \\ \ldots \end{cases} \tag{32}$$

avec $c_K = \dfrac{a_K}{2}$ (33)

Ainsi, estimer les coefficients a_K revient à estimer la série des coefficients c_K. Dès lors, après avoir estimé ces paramètres, on pourra en déduire un estimateur de d à partir de la relation (29).

Janacek (1982) montre que les coefficients estimés $\widehat{c}_K$ sont donnés par :

$$\widehat{c}_K = \frac{1}{\tau} \sum_{p=1}^{\tau-1} \ln I(p) \cos K\lambda_p + \frac{1}{2\tau} \left[\ln I(0) - \delta_T \ln I(\pi) \right] \tag{34}$$

où

$I(\lambda)$ est le périodogramme de la série X_t,

τ est la partie entière du rapport $\left[(T-1)/2\right]$, T étant le nombre d'observations,

δ_T est tel que $\begin{cases} \delta_T = 1 & \text{si } T \text{ est pair} \\ \delta_T = 0 & \text{sinon} \end{cases}$

La série des coefficients $\widehat{c}_K$ est asymptotiquement normale et :

$$\widehat{c}_K = \frac{d + c_K}{K} + e_K \tag{35}$$

où e_K est une suite de variables aléatoires indépendantes, et $d = 0$ si la série est stationnaire.

Dès lors, en reprenant l'équation (29) et compte tenu de la normalité asymptotique des coefficients $\widehat{c}_K$, on en déduit que le membre de gauche de la relation (29) :

$$S = \frac{1}{\pi} \int_0^{\pi} W(\lambda) \ln f_X(\lambda) d\lambda \tag{36}$$

suit une loi normale de moyenne : $d \sum_{K=1}^{\infty} \frac{1}{K^2} + \sum_{K=1}^{\infty} \frac{c_K}{K}$.

On en déduit alors l'estimateur $\widehat{d}_M$ de d :

$$\widehat{d}_M = \frac{S - \sum_{K=1}^{M} \frac{\widehat{c}_K}{K}}{\sum_{K=1}^{M} \frac{1}{K^2}} \tag{37}$$

où M est un paramètre de troncature ayant pour objet d'approximer la somme infinie.

Cet estimateur est asymptotiquement normal :

$$\widehat{d}_M \sim N \left(\frac{d + \sum_{K=1}^{M} \frac{c_K}{K}}{\sum_{K=1}^{M} \frac{1}{K^2}}, \frac{\Psi'(1)}{T \sum_{K=1}^{M} \frac{1}{K^2}} \right) \tag{38}$$

où $\Psi(x) = \frac{d}{dx} \ln \Gamma(x)$ et $\Psi^{(k)}(x) = \frac{d^k}{dx^k} \Psi(x)$.

Janacek (1982) montre enfin que si l'on choisit un paramètre de troncature M suffisamment grand de sorte que les c_K deviennent négligeables, alors $\widehat{d}_M$ est un estimateur sans biais de d.

3.1.3. Extension des méthodes spectrales au cas multivarié

La troisième technique en deux étapes est celle de Shea (1991). Cette procédure consiste à étendre le modèle de régression utilisé par Geweke et Porter-Hudak (1983) au cas multivarié et à estimer le paramètre d'intégration fractionnaire, non plus par les moindres carrés ordinaires, mais par la technique du maximum de vraisemblance.

Soit un vecteur de dimension N tel que :

$$\vec{X}_t = \overline{\Psi}(L) \times \vec{\varepsilon}_t \tag{39}$$

où $\overline{\Psi}(L)$ est une matrice de polynômes retards en L

$\vec{\varepsilon}_t$ est un vecteur d'erreurs aléatoires.

$\vec{X}_t$ est généré par un système à mémoire longue si chaque polynôme dans $\overline{\Psi}(L)$ contient, en commun, le filtre d'intégration $(1-L)^{-d}$ avec $d \neq 0$. Dans ses applications, Shea (1991) considère que le vecteur $\vec{X}_t$ est composé d'un taux d'intérêt à court terme et d'un taux à long terme.

Dans la procédure de Geweke et Porter-Hudak (1983), nous avons vu que le log-périodogramme, après simplification, s'écrivait :

$$\ln I(\lambda) = \ln\left[\frac{\sigma^2}{2\pi}\, f_u(0)\right] - d\,\ln\left[4\,\sin^2\left(\frac{\lambda}{2}\right)\right] + \ln\left[\frac{I(\lambda)}{f(\lambda)}\right] \tag{40}$$

On obtient ainsi une régression qui permet d'estimer le paramètre d'intégration fractionnaire d à partir des moindres carrés ordinaires.

Shea (1991) propose dans un premier temps d'étendre cette procédure au cas multivarié. En fait, l'idée sous-jacente à la technique de Geweke et Porter-Hudak est que, pour une série stationnaire, les ordonnées du périodogramme $I(\lambda)$ sont indépendamment et identiquement distribuées. Ceci permet alors de traiter le terme $v = \ln\left[\frac{I(\lambda)}{f(\lambda)}\right]$ comme le résidu de l'équation ci-dessus. Shea (1991) affirme également que les ordonnées des périodogrammes croisés et les ordonnées des périodogrammes entre des variables stationnaires ne sont pas corrélées (*Cf.* Shea, 1991, p.297).

Soit $X_t^{(k)}$ la $k$ième composante du vecteur $\vec{X}_t$ et soit $f_u^{(k)}(\lambda)$ la densité spectrale de la composante stationnaire de $X_t^{(k)}$. En regroupant les logarithmes des densités spectrales pour tous les $X_t^{(k)}$ en une équation similaire au modèle de Geweke et Porter-Hudak, on obtient :

$$\begin{aligned}\ln I^{(k)}(\lambda) = {} & \sum_{j=1}^{N} \ln\left[\frac{\sigma^2}{2\pi}\, f_u^{(j)}(\lambda)\right] \gamma_k \\ & - d\,\ln\left[4\,\sin^2\left(\frac{\lambda}{2}\right)\right] + \sum_{j=1}^{N} \ln\left[\frac{f_u^{(j)}(\lambda)}{f_u^{(j)}(0)}\right] \gamma_k + v\end{aligned} \tag{41}$$

où $I^{(k)}(\lambda)$ représente le périodogramme de la $k$ième composante et $\gamma_k = \begin{cases} 1 & \text{si } j = k \\ 0 & \text{sinon} \end{cases}$

Cette équation constitue donc une variante du modèle de Geweke et Porter-Hudak (1983), dans laquelle une variable dichotomique est intégrée afin de tenir compte des différences entre les composantes spectrales de haute fréquence des N éléments de $\vec{X}_t$ (même si elles ont toutes un filtre commun aux fréquences basses, à savoir $(1-L)^d$).

Tout comme dans le modèle de Geweke et Porter-Hudak (1983), le terme $\sum_{j=1}^{N} \ln \left[\frac{f_u^{(j)}(\lambda)}{f_u^{(j)}(0)} \right] \gamma_k$ est négligeable, puisqu'il tend vers zéro aux basses fréquences.

On obtient finalement l'équation suivante :

$$\ln I^{(k)}(\lambda) = \sum_{j=1}^{N} \ln \left[\frac{\sigma^2}{2\pi} f_u^{(j)}(\lambda) \right] \gamma_k - d \ln \left[4 \sin^2 \left(\frac{\lambda}{2} \right) \right] + v \qquad (42)$$

Cette équation peut alors être considérée comme une régression dans laquelle d est estimé par les moindres carrés ordinaires. Mais, il est également possible d'estimer d par le maximum de vraisemblance. Dans ce cas, Shea (1991) suggère de maximiser la log-vraisemblance conditionnellement à v. L'estimateur du maximum de vraisemblance de d (conditionnellement à la vraisemblance de v) est obtenu en résolvant l'équation[11] :

$$\frac{\sum_{j=1}^{N} \sum_{i=1}^{g(T)} 4 \sin^2 \left(\frac{\lambda_i}{2} \right) \exp \left[\ln I^{(j)}(\lambda_i) + \widehat{d} \times 4 \sin^2 \left(\frac{\lambda_i}{2} \right) \right]}{\sum_{j=1}^{N} \sum_{i=1}^{g(T)} \exp [\ln I^{(j)}(\lambda_i) + \widehat{d} \times 4 \sin^2 \left(\frac{\lambda_i}{2} \right)]} = \frac{\sum_{i=1}^{g(T)} 4 \sin^2 \left(\frac{\lambda_i}{2} \right)}{g(T)} \qquad (43)$$

ce qui peut se faire facilement par de simples méthodes numériques. Comme dans la procédure de Geweke et Porter-Hudak (1983), Shea (1991) pose : $g(T) = T^\alpha$ avec $0 < \alpha < 1$.

L'intérêt de ces techniques en deux étapes est qu'elles permettent une estimation du paramètre d sans avoir à estimer les paramètres des polynômes retards autorégressif et moyenne mobile. Elles souffrent néanmoins d'un inconvénient majeur dans la mesure où les coordonnées du périodogramme, utilisées dans le calcul de $\widehat{d}$ sont influencées par la dynamique de court terme de la série et donnent en conséquence une mesure erronée du long terme. En effet, la régression utilisée dans la méthode de Geweke et Porter-Hudak (1983) entre le logarithme de la fréquence et le logarithme de la densité spectrale est valable pour toutes les fréquences si la série est une série purement à intégration fractionnaire, c'est-à-dire en l'absence de termes autorégressif et moyenne

11. Pour plus de détails concernant les calculs, voir Shea (1991), pp 297 298.

mobile. Dès lors que des termes autorégressif et/ou moyenne mobile apparaissent, la régression de Geweke et Porter-Hudak n'est plus correcte. Geweke et Porter-Hudak (1983) ont cependant montré que dans un tel cas la relation de régression reste valide approximativement pour des fréquences au voisinage de zéro. En pratique, le problème inhérent à la méthode en deux étapes vient de la présence d'un comportement de court terme de la série que le paramètre de différenciation fractionnaire d n'est pas à même de prendre en compte. Intuitivement, l'absence de paramètres prenant explicitement en considération les dynamiques de court terme dans la première étape risque d'induire une estimation inconsistante du paramètre de différenciation fractionnaire. C'est pourquoi nous nous proposons maintenant de présenter les méthodes d'estimation en une étape.

3.2. Les méthodes en une étape

Parmi les méthodes d'estimation en une étape, on peut distinguer celles basées sur une approximation de la méthode du maximum de vraisemblance et celles fondées sur le maximum de vraisemblance exact.

Une maximisation directe de la log-vraisemblance a été utilisée par McLeod et Hipel (1978) pour estimer les paramètres d'un bruit gaussien fractionnaire et peut également être employée pour des processus ARFIMA(p, d, q). Hosking (1984) souligne cependant que cette méthode est très coûteuse du point de vue du temps de calcul. Dans ces circonstances, il semble dans un premier temps naturel de chercher à remplacer la vraisemblance par une approximation qui pourra être plus rapidement évaluée et de chercher à maximiser cette vraisemblance approximée.

Concernant les méthodes du maximum de vraisemblance approché, nous distinguerons la procédure de Li et Mc Leod (1986), où l'approximation provient du fait qu'ils tronquent la somme infinie définissant le développement binomial et les procédures de Brockwell et Davis (1991) ou Fox et Taqqu (1986) utilisant une approximation de la fonction de log-vraisemblance fournie par la fonction de Whittle.

La méthode du maximum de vraisemblance exacte repose quant à elle sur la procédure élaborée par Sowell (1992).

3.2.1. La méthode du maximum de vraisemblance approché par troncature

La méthode suggérée par Li et Mc Leod (1986) est une procédure basée sur le maximum de vraisemblance approché. Les auteurs proposent en effet de tronquer la somme infinie qui définit $(1-L)^d$ et d'utiliser ensuite les procédures standards d'estimation des séries temporelles.

Ainsi, dans le modèle ARFIMA(p, d, q) suivant :

$$\Phi(L)(1-L)^d(X_t-\mu)=\Theta(L)\varepsilon_t \tag{44}$$

où μ est la moyenne de X_t, on pose :

$$w_t=(1-L)^d(X_t-\mu) \tag{45}$$

On obtient alors le processus suivant :

$$\Phi(L)w_t = \Theta(L)\varepsilon_t \tag{46}$$

$\{w_t\}$ suit ainsi un processus ARMA(p, q) de moyenne nulle. Dès lors, au lieu de maximiser la vraisemblance d'un modèle ARFIMA(p, d, q) par rapport aux paramètres autorégressif, moyenne mobile, variance du bruit blanc, μ et d, Li et Mc Leod (1986) suggèrent de maximiser la vraisemblance d'un modèle ARMA(p, q) par rapport au vecteur contenant les paramètres autorégressif, moyenne mobile et variance du bruit blanc.

Il s'agit d'une méthode de maximum de vraisemblance approché car le calcul de w_t à partir de X_t nécessite d'effectuer une approximation de la somme infinie définissant $(1 - L)^d$:

$$(1 - L)^d(X_t - \mu) = \nabla^d_\infty(X_t - \mu) = \sum_{j=0}^{\infty} \pi_j(X_{t-j} - \mu) \tag{47}$$

$$\text{avec } \pi_j = (-1)^j \frac{\Gamma(1 + d)}{\Gamma(1 + j)\Gamma(1 + d - j)} \tag{48}$$

Ce calcul nécessite les valeurs non observées X_0, X_{-1}, X_{-2}, ... Dès lors, Li et Mc Leod (1986) suggèrent l'approximation suivante :

$$\nabla^d_M(X_t - \mu) = \sum_{j=0}^{t+M-1} \pi_j(X_{t-j} - \mu) \tag{49}$$

Le choix du paramètre de troncature est crucial, puisque Hosking (1984) montre au moyen de simulations que l'estimation des paramètres dépend de la valeur de M. Il propose alors de choisir M tel que la série puisse être modélisée par un processus AR(M) prenant en compte la majeure partie de l'autocorrélation. Selon Li et Mc Leod (1986), pour une valeur de M suffisamment importante et pour une valeur de d inférieure à $1/2$, la différence entre le vrai modèle (sans tronquer la somme infinie) et le modèle approché (avec troncature) est négligeable.

Cependant, le fait de tronquer la somme infinie et d'utiliser par la suite les méthodes standards des séries temporelles, revient en fait à une procédure en deux étapes. La méthode de Li et Mc Leod souffre en conséquence des critiques adressées à ces procédures.

3.2.2. *La méthode du maximum de vraisemblance approché par la fonction de Whittle*

Fox et Taqqu (1986) proposent d'utiliser une autre approximation de la fonction de vraisemblance gaussienne. L'idée sous-jacente est que la technique du maximum de vraisemblance exact est certainement la plus efficace, mais c'est également la plus difficile à mettre en œuvre. C'est pourquoi Fox et Taqqu ont suggéré de se ramener à une approximation. Leurs estimateurs sont convergents et asymptotiquement normaux.

La vraisemblance gaussienne de la série X pour le processus général :

$$\Phi(L)\nabla^d(X_t - \mu) = \Theta(L)u_t \tag{50}$$

où Φ et Θ sont des polynômes retards d'ordre p et q respectivement et $-1/2 < d < 1/2$, peut être exprimée comme :

$$L(\beta,\sigma^2) = (2\pi\sigma^2)^{-T/2}(r_0,\dots,r_{T-1})^{-1/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^{T}\frac{(X_j-\widehat{X}_j)^2}{r_{j-1}}\right\} \tag{51}$$

où $\beta = (d,\phi_1,\dots,\phi_p,\theta_1,\dots,\theta_q)'$,

$\widehat{X}_j$, $j = 1,\dots,T$, correspond à la série des prévisions avec un pas unitaire,

$r_{j-1} = \sigma^{-2}E(X_j - \widehat{X}_j)^2 \quad j = 1,\dots,T.$

Les estimateurs du maximum de vraisemblance $\widehat{\beta}$ et $\widehat{\sigma}^2$ sont obtenus en maximisant la vraisemblance $L(\beta,\sigma^2)$ par rapport à β et à σ^2.

Brockwell et Davis (1991) indiquent que cette maximisation aboutit à :

$$\widehat{\sigma}^2 = T^{-1}S(\widehat{\beta}) \quad \text{où} \quad S(\widehat{\beta}) = \sum_{j=1}^{T}\frac{(X_j-\widehat{X}_j)^2}{r_{j-1}} \tag{52}$$

et $\widehat{\beta}$ est la valeur de β qui minimise :

$$l(\beta) = \ln\left(\frac{S(\beta)}{T}\right) + T^{-1}\sum_{j=1}^{T}\ln(r_{j-1}) \tag{53}$$

Pour u_t gaussien, Yajima (1985) a montré que, dans le cas où $p = q = 0$ et $d > 0$, on a (voir également Robinson, 1995a) :

$$\widehat{\beta} \sim N(\beta, T^{-1}W^{-1}(\beta)) \tag{54}$$

où $\underset{(p+q+1,p+q+1)}{W(\beta)}$ est une matrice dont l'élément (j,k) est :

$W_{jk}(\beta) = \dfrac{1}{4\pi}\displaystyle\int_{-\pi}^{\pi}\frac{\partial\ln g(\lambda;\beta)}{\partial\beta_j}\,\frac{\partial\ln g(\lambda;\beta)}{\partial\beta_k}\,\mathrm{d}\lambda$ et $\sigma^2 g(\cdot;\beta)/2\pi$ est la densité spectrale du processus.

Le comportement asymptotique de $\widehat{\beta}$ pour $d < 0$ est cependant inconnu.

Le calcul de la fonction $l(\beta)$ est long, surtout pour des valeurs élevées de T. Aussi, Brockwell et Davis (1991) suggèrent d'utiliser l'approximation suivante de $l(\beta)$:

$l_a(\beta) = \ln \frac{1}{T} \sum_j \left(\frac{I_T(\omega_j)}{g(\omega_j;\beta)} \right)$ la somme étant sur toutes les fréquences de Fourier ω_j non nulles.

Hannan (1973) et Fox et Taqqu (1986) ont montré que l'estimateur $\widetilde{\beta}$ qui minimise $l_a(\beta)$ est convergent et si $d > 0$, il a la même distribution limite que $\widehat{\beta}$. La variance du bruit blanc est estimée par :

$$\widetilde{\sigma}^2 = \frac{1}{T} \sum_j \left(\frac{I_T(\omega_j)}{g(\omega_j;\widetilde{\beta})} \right) \tag{55}$$

Brockwell et Davis (1991) remarquent cependant que l'approximation $l_a(\beta)$ ne prend pas en compte le terme $T^{-1} \sum_{j=1}^{T} \ln(r_{j-1})$. Bien que ce terme tende vers zéro lorsque T tend vers l'infini, cet oubli peut avoir des conséquences non négligeables sur la minimisation de $l(\beta)$. Les auteurs suggèrent alors d'approximer ce terme par $T^{-1} \sum_{j=1}^{T} \ln g(\omega_j;\beta)$.

On obtient en conséquence la seconde approximation de $l(\beta)$, soit :

$$l_b(\beta) = l_a(\beta) + T^{-1} \sum_j \ln g(\omega_j;\beta) \tag{56}$$

Brockwell et Davis (1991) proposent ainsi une technique visant à estimer les paramètres β et σ^2 du modèle fractionnairement intégré en maximisant l'approximation de Whittle, L_w, de la fonction de vraisemblance. Une telle opération revient à minimiser :

$$-2\ln(L_w) = T\ln(2\pi) + 2T\ln\sigma + \sigma^{-2} \sum_j \left(\frac{I_T(\omega_j)}{g(\omega_j;\beta)} \right) + \sum_j \ln g(\omega_j;\beta) \tag{57}$$

où I_T correspond au périodogramme,

$\sigma^2 \cdot g$ est la densité spectrale du modèle et

la somme est sur toutes les fréquences de Fourier non nulles.

3.2.3. La méthode du maximum de vraisemblance exact

Sowell (1992a, b) dérive la fonction de vraisemblance non conditionnelle exacte pour une série temporelle gaussienne stationnaire à intégration fractionnaire. Soit x_t une telle série et X_T un échantillon de T observations tel que $X_T = [x_1 x_2 \ldots x_T]'$. Alors X_T suit une loi normale de moyenne nulle et de matrice de variance covariance Σ et a pour fonction de densité :

$$f(X_T, \Sigma) = (2\pi)^{-T/2}|\Sigma|^{-1/2}\exp\left(-\frac{1}{2}\ X_T'\Sigma^{-1}X_T\right) \tag{58}$$

Du fait de la stationnarité de la série, la matrice de covariance a une forme de Toeplitz :

$\Sigma = \left[\gamma(i-j)\right]$ avec $i, j = 1, 2, ..., T$.

L'estimation des paramètres du modèle par le maximum de vraisemblance nécessite l'évaluation de la fonction de vraisemblance pour un ensemble donné de paramètres. Il faut donc écrire la matrice de covariance, ou la fonction d'autocovariance, en fonction des paramètres du modèle. Cette paramétrisation de la fonction d'autocovariance est déduite de la densité spectrale de x_t et se calcule par :

$$\gamma(s) = \frac{1}{2\pi}\int_0^{2\pi} f_z(\lambda)e^{i\lambda s}d\lambda \tag{59}$$

Ainsi, le calcul de la fonction d'autocovariance nécessite le calcul de la densité spectrale de x_t. Celle-ci est calculée en deux étapes : on évalue tout d'abord la densité spectrale de $u_t = (1-L)^d x_t$, puis, on calcule la densité spectrale de x_t. Alors u_t est généré par le processus ARMA(p, q) suivant : $\Phi(L)u_t = \Theta(L)\varepsilon_t$ où ε_t est le résidu associé au modèle ARFIMA(p, d, q).

La représentation de Wold de u_t est naturellement donnée par :

$$u_t = \Theta(L)\Phi^{-1}(L)\varepsilon_t \tag{60}$$

et l'on peut écrire $\Phi(x)$ comme :

$$\Phi(x) = \prod_{j=1}^{p}(1 - \rho_j x) \tag{61}$$

où $|\rho_n| < 1$ pour $n = 1, 2, \ldots, p$.

La densité spectrale de u_t s'écrit alors :

$$f_u(\lambda) = \frac{|\Theta(\omega)|^2}{|\Phi(\omega)|^2}\sigma^2 = |\Theta(\omega)|^2\sigma^2\prod_{j=1}^{p}(1 - \rho_j\omega)^{-1}(1 - \rho_j\omega^{-1})^{-1} \tag{62}$$

avec $\omega = e^{i\lambda}$

ce qui peut encore s'écrire :

$$f_u(\lambda) = \sigma^2 \sum_{l=-q}^{q} \Psi(l)\omega^l \sum_{j=1}^{p} \omega^p \xi_j \left[\frac{\rho_j^{2p}}{1-\rho_j\omega} - \frac{1}{1-\rho_j^{-1}\omega}\right] \tag{63}$$

où

$$\xi_j = \left[\rho_j \prod_{i=1}^{p}(1-\rho_i\rho_j) \prod_{m\neq j}(\rho_j - \rho_m)\right]^{-1} \tag{64}$$

$$\Psi(l) = \sum_{s=\max[0,\ l]}^{\min[q,q-l]} \theta_s\theta_{s-l} \tag{65}$$

L'équation définissant $f_u(\lambda)$ constitue ainsi une représentation de la densité spectrale d'un processus ARMA(p,q) dans lequel les racines du polynôme autorégressif sont simples.

Partant de la densité spectrale de u_t, on peut alors écrire la densité spectrale de x_t :

$$f_x(\lambda) = (1-\omega)^{-d}(1-\omega^{-1})^{-d} f_u(\lambda) \tag{66}$$

soit :

$$f_x(\lambda) = \sigma^2 \sum_{l=-q}^{q} \sum_{j=1}^{p} \Psi(l)\xi_j \left[\frac{\rho_j^{2p}}{1-\rho_j\omega} - \frac{1}{1-\rho_j^{-1}\omega}\right] (1-\omega)^{-d}(1-\omega^{-1})^{-d}\omega^{p+1} \tag{67}$$

Cette équation représente la forme générale de la densité spectrale d'une série temporelle stationnaire générée par un processus ARFIMA(p,d,q) dans lequel les racines du polynôme autorégressif sont simples.

Le calcul de la fonction d'autocovariance $\gamma(s)$ s'effectue alors en substituant la valeur de $f_x(\lambda)$ dans l'équation définissant $\gamma(s)$. On obtient ainsi (Cf. Sowell, 1992, p.173) :

$$\gamma(s) = \sigma^2 \sum_{l=-q}^{q} \sum_{j=1}^{p} \Psi(l)\xi_j C(d, p+1-s, \rho_j) \tag{68}$$

avec, en posant $h = p + l - s$:

$$C(d,h,\rho) = \frac{1}{2\pi}\int_0^{2\pi} \left[\frac{\rho^{2p}}{(1-\rho e^{-i\lambda})} - \frac{1}{(1-\rho^{-1}e^{-i\lambda})}\right] \times (1-e^{-i\lambda})^{-d}(1-e^{i\lambda})^{-d}e^{-i\lambda h}d \tag{69}$$

La forme utilisée pour les calculs s'écrit :

$$C(d,h,\rho) = \frac{\Gamma(1-2d)\Gamma(d+h)}{\Gamma(1-d+h)\Gamma(1-d)\Gamma(d)} \times \left[\rho^{2p}F(d+h,1;1-d+h;\rho) + F(d-h,1;1-d-h;\rho) - 1\right]$$

où $F(a, b; c; x)$ est la fonction hypergéométrique définie par :

$$F(a, b; c; x) = \sum_n \frac{\Gamma(a+n)\Gamma(b+n)\Gamma(c)}{\Gamma(a)\Gamma(b)\Gamma(c+n)\Gamma(n+1)} x^n \tag{70}$$

Ainsi, après avoir calculé la matrice de covariance au moyen des développements précédents, il est possible d'évaluer la log vraisemblance par rapport à tous les paramètres du modèle et donc d'estimer le processus ARFIMA(p, d, q).

3.3. Une brève comparaison des méthodes d'estimation

L'intérêt de la méthode du maximum de vraisemblance est qu'elle utilise toute l'information à court comme à long termes concernant le comportement des séries puisque sont estimés simultanément les paramètres autorégressif, moyenne mobile et le paramètre de différenciation fractionnaire d. Divers auteurs se sont alors attachés à comparer les diverses techniques d'estimation des processus ARFIMA par le biais de simulations de Monte Carlo (voir notamment Hosking (1984), Li et Mc Leod (1986)). Les travaux récents cherchent plus précisément à évaluer l'apport de la technique du maximum de vraisemblance relativement à la procédure de Geweke et Porter-Hudak (1983). Sowell (1992b) a réalisé une simulation dans laquelle il montre la supériorité en termes de précision de la méthode du maximum de vraisemblance exact. Plus précisément, il souligne que lorsque la vraie spécification du processus est connue, la méthode du maximum de vraisemblance donne des estimations plus précises que la procédure de Geweke et Porter-Hudak. Bien évidemment, en pratique la spécification n'est pas connue avec certitude et doit être estimée. Dans une telle situation, Sowell écrit que : «*C'est toujours une question ouverte que de déterminer la performance relative du maximum de vraisemblance par rapport à la procédure de Geweke et Porter-Hudak lorsque la spécification doit être estimée*» (Sowell, 1992b, pp. 300-301).

Diebold et Rudebusch (1989) ont adopté à ce sujet une position relativement tranchée en défaveur des méthodes en une étape. En référence à la procédure de Geweke et Porter-Hudak, ils notent ainsi que, puisque l'estimateur du paramètre d'intégration fractionnaire est efficace, il s'ensuit que les estimateurs de la seconde étape des paramètres autorégressifs et moyenne mobile sont également efficaces. Les auteurs soulignent alors l'avantage de cette procédure d'estimation semi paramétrique dans la mesure où la distribution asymptotique de l'estimateur du paramètre d'intégration fractionnaire ne dépend pas des paramètres de nuisance $\phi^{-1}(L)\theta(L)$. Pour finir, Diebold et Rudebusch (1989) notent que si les procédures d'estimation alternatives, telles que l'estimation par le maximum de vraisemblance, ont les «bonnes» propriétés statistiques lorsque la spécification du modèle est connue, il n'en reste pas moins qu'elles peuvent être inefficaces si le modèle n'est pas correctement spécifié.

Quoi qu'il en soit, les simulations menées par les différents auteurs (voir notamment Cheung et Diebold (1994) et Sowell (1992)) tendent à montrer que même si la méthode du maximum de vraisemblance exact est la plus difficile à mettre en oeuvre, il n'en reste pas moins qu'elle constitue la méthode d'estimation la plus précise des processus ARFIMA[12].

4. LES EXTENSIONS DES PROCESSUS ARFIMA

On se propose ici de présenter les principales extensions des processus ARFIMA. Nous scinderons notre exposé en deux paragraphes. Le premier paragraphe sera relatif aux extensions des processus ARFIMA au cas saisonnier. Le paragraphe suivant aura pour objet de rendre compte des développements les plus récents en matière de modélisation de la volatilité des séries économiques par le biais des processus à mémoire longue.

4.1. Prise en compte des comportements saisonniers

Tout comme les processus ARIMA ont été étendus aux processus SARIMA (ARIMA saisonniers), il est assez naturel de chercher à étendre les processus ARFIMA afin de tenir compte d'un comportement périodique de la série. A cette fin, Hosking (1981, 1984) avait suggéré deux types d'extensions possibles :

- l'introduction d'un filtre $(1 - \phi L + L^2)^d$
- l'introduction d'un opérateur de différenciation saisonnière fractionnaire du type $(1 - L^s)^d$

La première idée a été suivie par Gray, Zhang et Woodward (1989) qui ont élaboré les processus GARMA. La seconde extension a été développée par Porter-Hudak (1990).

4.1.1. Les processus GARMA

Gray, Zhang et Woodward (1989) proposent une extension des processus ARFIMA afin de modéliser le comportement cyclique persistant au moyen des pôles du spectre qui ne se situent pas à la fréquence zéro. Ceci permet ainsi de prendre en compte le comportement saisonnier de la mémoire. Les auteurs ont alors repris l'idée initialement proposée par Hosking (1981) d'un processus de la forme :

$$\Phi(L)(1 - 2uL + L^2)^\lambda X_t = \Theta(L)\varepsilon_t \tag{71}$$

Ce modèle, appelé processus GARMA(p, u, λ, q), n'a cependant, jusqu'aux travaux de Gray, Zhang et Woodward (1989), jamais été étudié en raison de la complexité due à l'inversion du polynôme $(1 - 2uL + L^2)^\lambda$. Les processus du type (71) peuvent cependant être étudiés au moyen des polynômes de Gegenbauer[13] définis comme suit :

12. Pour une étude récente relative aux simulations de Monte Carlo concernant les méthodes d'estimation des processus ARFIMA, on pourra se reporter à Hauser (1998).

13 On pourra trouver une étude récente des propriétés et de l'estimation des processus de Gegenbauer dans Chung (1996).

Soit $\lambda \neq 0$ et $|Z| < 1$, alors pour $|u| \leqslant 1$ on définit le polynôme de Gegenbauer $C_n^\lambda(u)$ par :

$$(1 - 2uZ + Z^2)^{-\lambda} = \sum_{n=0}^{\infty} C_n^\lambda(u) Z^n \tag{72}$$

$$\text{soit } C_n^\lambda(u) = \sum_{k=0}^{[n/2]} (-1)^k \frac{\Gamma(\lambda + n - k)}{\Gamma(\lambda)} (2u)^{n-2k} \big[k!(n-2k)!\big]^{-1} \tag{73}$$

où $[n/2]$ est la partie entière de $n/2$.

Considérons un processus X_t défini par :

$$X_t = \sum_{k=0}^{\infty} C_k^\lambda(u) \varepsilon_{t-k} \tag{74}$$

où ε_t est un bruit blanc $(0, \sigma_\varepsilon^2)$ et λ est un nombre réel différent de zéro.

En utilisant le polynôme de Gegenbauer, on peut écrire :

$$X_t = \sum_{k=0}^{\infty} C_k^\lambda(u) L^k \varepsilon_t = (1 - 2uL + L^2)^{-\lambda} \varepsilon_t \tag{75}$$

Si X_t est inversible, on a finalement :

$$(1 - 2uL + L^2)^\lambda X_t = \varepsilon_t \tag{76}$$

Ce processus est appelé processus de Gegenbauer de paramètre λ et est un processus stationnaire à mémoire longue si $|u| < 1$ et $0 < \lambda < \frac{1}{2}$ ou si $|u| = 1$ et $0 < \lambda < \frac{1}{4}$.

Il est alors possible d'étendre les processus de Gegenbauer en incluant des termes autorégressifs et moyenne mobile. De tels processus sont appelés processus GARMA.

Ainsi, un processus GARMA(p, u, λ, q) s'écrit :

$$\Phi(L)(1 - 2uL + L^2)^\lambda X_t = \Theta(L) \varepsilon_t \tag{77}$$

Les conditions de stationnarité, d'inversibilité et de présence de mémoire longue sont identiques à celles d'un processus de Gegenbauer. On a en outre les propriétés suivantes :

- La fonction d'autocorrélation d'un processus GARMA(p, u, λ, q), lorsque le retard k tend vers l'infini, est telle que :
  - $\rho_k \sim k^{4\lambda - 1}$ pour $u = 1$ et $0 < \lambda < \frac{1}{4}$
  - $\rho_k \sim (-1)^k k^{4\lambda - 1}$ pour $u = -1$ et $0 < \lambda < \frac{1}{4}$

- $\rho_k \sim C \cdot k^{2\lambda-1}\cos(k\omega_0)$ pour $|u| < 1$ et $0 < \lambda < \frac{1}{2}$, ω_0 étant la fréquence de Gegenbauer et C une constante indépendante de k.

- La densité spectrale d'un processus GARMA(p, u, λ, q) répond aux conditions suivantes, $f(\omega)$ étant la densité spectrale d'un processus de Gegenbauer :
  - si $|u| < 1$, $\lim\limits_{\omega \to \omega_0} (\omega - \omega_0)^{2\lambda} f(\omega)$ existe et est finie.
  - si $|u| = 1$, $\lim\limits_{\omega \to 0} (\omega)^{4\lambda} f(\omega)$ existe et est finie.

On peut remarquer que, lorsque $|u| = 1$, le processus GARMA se réduit au processus ARFIMA(p, d, q) avec $\lambda = \frac{d}{2}$.

Les processus ARFIMA ont donc été étendus aux processus GARMA afin de tenir compte d'un comportement périodique de la mémoire longue. Une autre possibilité consiste à utiliser un modèle multiplicatif dérivé du modèle SARIMA (*Seasonal* ARIMA) traditionnel de Box et Jenkins. Nous proposons de nommer ces modèles SARFIMA (*Seasonal* ARFIMA).

4.1.2. Les processus SARFIMA

Porter-Hudak (1990) propose l'utilisation d'un processus ARFIMA incluant deux composantes de différenciation fractionnaire : une composante saisonnière et une composante non saisonnière. Ainsi, un processus SARFIMA peut être défini comme suit :

$$\Phi(L)(1 - L)^{d_1}(1 - L^s)^{d_2} X_t = \Theta(L)\varepsilon_t \tag{78}$$

où d_1 et d_2 appartiennent à l'intervalle $\left]-\frac{1}{2}, \frac{1}{2}\right[$

Ce modèle permet ainsi de rendre compte de la décroissance hyperbolique des autocorrélations grâce au paramètre d_1 et également de la décroissance hyperbolique des autocorrélations aux retards saisonniers au travers du paramètre d_2. En d'autres termes, les processus SARFIMA exhibent un comportement aux fréquences saisonnières similaire à celui des processus ARFIMA à la fréquence nulle.

4.2. Mémoire longue et volatilité : les processus FIGARCH[14]

Des études récentes se sont attachées au comportement de long terme de la variance conditionnelle des séries financières : Baillie, Bollerslev et Mikkelsen (1996), Bollerslev et Mikkelsen (1996), Baillie (1996) et Ding et Granger (1996). L'idée sous-jacente à ces études réside dans les travaux de Bollerslev et Engle (1986) qui ont mis en évidence la persistance des chocs sur la

14. Ce paragraphe reprend divers développements figurant dans Lardic et Mignon (1999b)

variance conditionnelle au moyen de processus GARCH[15] intégrés (processus IGARCH). Ces processus rendent compte des conséquences permanentes des chocs sur la volatilité. Afin d'étudier la possibilité de conséquences durables, et non forcément permanentes, des chocs sur la volatilité, Baillie, Bollerslev et Mikkelsen (1996) ont introduit une classe de modèles plus flexibles : les processus FIGARCH (GARCH fractionnairement intégrés).

Rappelons qu'un processus GARCH(p,q) peut s'écrire[16] :

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2 \tag{79}$$

où σ_t^2 est la variance conditionnelle, $\alpha_0 > 0$, $\alpha_i \geqslant 0$, $\beta_j \geqslant 0$, $i = 1, \ldots, q$ et $j = 1, \ldots, p$.

Un tel processus peut encore s'écrire, sous la forme d'une représentation ARCH infinie :

$$\sigma_t^2 = \alpha_0 \left[1 - \beta(1)\right]^{-1} + \alpha(L)\left[1 - \beta(L)\right]^{-1} \varepsilon_t^2 \tag{80}$$

En posant $\lambda(L) = \alpha(L)\left[1 - \beta(L)\right]^{-1}$, on a ainsi :

$$\sigma_t^2 = \alpha_0 \left[1 - \beta(1)\right]^{-1} + \lambda(L)\varepsilon_t^2 \tag{81}$$

Les processus GARCH(p,q) sont des processus à mémoire courte puisque l'effet d'un choc sur la variance conditionnelle décroît à un taux exponentiel.

Un processus GARCH(p,q) peut en outre s'interpréter comme un processus ARMA(m,p), avec $m = \max(p,q)$, sur le carré des innovations. Ainsi, en posant :

$$u_t = \varepsilon_t^2 - \sigma_t^2 \tag{82}$$

on déduit de l'équation (79) la formulation suivante :

$$\left[1 - \alpha(L) - \beta(L)\right]\varepsilon_t^2 = \alpha_0 + \left[1 - \beta(L)\right]u_t \tag{83}$$

Par conséquent, lorsque le polynôme retard $\left[1 - \alpha(L) - \beta(L)\right]$ contient une racine unitaire, le processus GARCH(p,q) devient un processus GARCH intégré, noté IGARCH(p,q) :

$$\Phi(L)(1 - L)\varepsilon_t^2 = \alpha_0 + \left[1 - \beta(L)\right]u_t \tag{84}$$

avec $\Phi(L) = \left[1 - \alpha(L) - \beta(L)\right](1 - L)^{-1}$ (85)

15. Les processus GARCH, développés par Bollerslev (1986), constituent une généralisation des processus ARCH consistant à introduire une partie moyenne mobile dans l'équation de la variance conditionnelle On rappelle par ailleurs que les processus de type ARCH sont basés sur une paramétrisation endogène de la variance conditionnelle et permettent ainsi de modéliser la variabilité de la volatilité des séries au cours du temps

16. On pourra se reporter à Bollerslev, Chou, Jayaraman et Kroner (1991) pour une revue de la littérature sur la classe des modèles ARCH.

Les processus FIGARCH constituent un cas intermédiaire entre les processus GARCH et les processus IGARCH et se déduisent de la relation (84) en remplaçant l'opérateur $(1-L)$ par l'opérateur $(1-L)^d$, où d est le paramètre d'intégration fractionnaire. Ainsi, un processus FIGARCH(p, d, q) s'écrit :

$$\Phi(L)(1-L)^d\varepsilon_t^2 = \alpha_0 + \left[1-\beta(L)\right]u_t \tag{86}$$

les racines des polynômes $\Phi(L)$ et $\left[1-\beta(L)\right]$ étant à l'extérieur du disque unité.

On constate alors que, si $d = 0$, le processus FIGARCH(p, d, q) se réduit à un processus GARCH(p, q) et, si $d = 1$, le processus FIGARCH devient un processus GARCH intégré. On retrouve ici clairement l'analogie entre les processus FIGARCH sur l'équation de la variance conditionnelle et les processus ARFIMA pour l'équation de la moyenne.

En remplaçant u_t par sa valeur en fonction de σ_t^2, on peut encore écrire l'équation (86) comme :

$$\left[1-\beta(L)\right]\sigma_t^2 = \alpha_0 + \left[1-\beta(L)-\Phi(L)(1-L)^d\right]\varepsilon_t^2 \tag{87}$$

L'équation de la variance est alors donnée par :

$$\sigma_t^2 = \alpha_0\left[1-\beta(1)\right]^{-1} + \left[1-\left[1-\beta(L)\right]^{-1}\Phi(L)(1-L)^d\right]\varepsilon_t^2 \tag{88}$$

soit finalement :

$$\sigma_t^2 = \alpha_0\left[1-\beta(1)\right]^{-1} + \lambda(L)\varepsilon_t^2 \tag{89}$$

avec

$$\begin{aligned}\lambda(L) &= \left[1-\left[1-\beta(L)\right]^{-1}\Phi(L)(1-L)^d\right] \\ &= \lambda_1 L + \lambda_2 L^2 + \ldots\end{aligned} \tag{90}$$

et $\lambda_k \geqslant 0$, $k = 1, 2, \ldots$

Baillie, Bollerslev et Mikkelsen (1996) notent alors que les conséquences d'un choc sur la variance conditionnelle du processus FIGARCH(p, d, q) décroissent à un taux hyperbolique lorsque $0 \leqslant d < 1$. Ainsi, par analogie avec les processus ARFIMA, la dynamique de long terme de la volatilité est prise en compte au travers du paramètre d'intégration fractionnaire d, la dynamique de court terme étant quant à elle modélisée au travers des paramètres GARCH traditionnels. Notons enfin que les processus FIGARCH ont également connu diverses extensions afin de tenir compte des phénomènes d'asymétrie. Bollerslev et Mikkelsen (1996) ont ainsi proposé la classe des processus FIGARCH exponentiels (FIEGARCH) afin de rendre compte de l'asymétrie des chocs sur la variance conditionnelle.

5. LES APPORTS EMPIRIQUES DES PROCESSUS À MÉMOIRE LONGUE

A la suite des nombreux travaux de Mandelbrot, les études sur la mémoire de long terme ont été développées dans le domaine économique : « *Entre chroniques économiques et Nilotes, la première ressemblance consiste dans le comportement cyclique non périodique qui les caractérise toutes deux* » [Mandelbrot (1973), p. 350]. Cette ressemblance peut être due au fait, couramment évoqué par Mandelbrot, que les phénomènes physiques exercent un effet direct sur l'économie.

On se propose ici de dresser un aperçu des principaux résultats obtenus dans la littérature économique empirique. On commencera par présenter les conclusions relatives aux séries financières avant de s'attacher aux applications dans le domaine macroéconomique.

5.1. Applications en finance

La recherche d'une structure de dépendance de long terme dans les séries financières est particulièrement intéressante dans la mesure où elle peut être clairement reliée à la théorie de l'efficience des marchés financiers (voir notamment Mignon, 1998). Selon cette théorie, un marché est informationnellement efficient si l'ensemble des informations pertinentes à l'évaluation des actifs financiers qui y sont cotés se trouve instantanément reflété dans les cours. Parmi les trois formes de l'efficience définies par Fama (1970, 1991), l'efficience au sens faible implique que les séries de rentabilités boursières sont caractérisées par une absence de mémoire, c'est à dire que l'on ne peut pas prévoir les rentabilités futures à partir des rentabilités passées. En d'autres termes, les cours boursiers suivent un processus de marche aléatoire et les séries de rentabilités boursières répondent à un processus de bruit blanc. La présence de mémoire longue dans les séries de rentabilités boursières constitue dès lors un défi important vis à vis de l'efficience des marchés financiers. En effet, la mémoire longue induit l'existence d'un écart durable entre le cours et sa valeur fondamentale[17], ce qui renvoie à la présence d'un lent délai d'ajustement des prix à l'information. Un tel délai paraît dès lors difficilement conciliable avec la théorie de l'efficience des marchés financiers. On commencera par dresser une revue de la littérature des résultats obtenus dans le domaine financier avant de mettre en exergue les conséquences de la mémoire longue du point de vue de la théorie financière.

5.1.1. Une revue de la littérature empirique

Outre les divers travaux de Mandelbrot suggérant d'utiliser la valeur de l'exposant de Hurst comme mesure du degré d'imperfection d'un marché, les études initiales portant sur les séries financières sont dues à Greene et Fielitz

17. La valeur fondamentale d'une action est définie comme la somme actualisée des anticipations rationnelles des dividendes futurs.

(1977) sur les rentabilités des actions et Booth, Kaen et Koveos (1982) sur les taux de change. Nous présenterons ci-après les principaux résultats obtenus en fonction de la nature des séries financières étudiées.

5.1.1.1. Les séries de rentabilités boursières

Les travaux pionniers de Greene et Fielitz (1977) portent sur 200 séries de rentabilités d'actions cotées au NYSE (*New York Stock Exchange*) sur la période s'étalant de décembre 1963 à novembre 1968. L'application de l'analyse R/S sur ces séries conduit les auteurs à conclure en termes d'une structure de dépendance de long terme et à remettre en question l'hypothèse d'efficience des marchés financiers : pour la plupart des séries, l'exposant de Hurst est en effet compris entre 0,501 et 0,660. Peters (1991 et 1994), en appliquant l'analyse R/S, conclut également à la présence de mémoire longue dans diverses séries de rentabilités boursières (par exemple, la valeur estimée de l'exposant de Hurst est de 0,78 pour l'indice boursier américain SP 500 sur la période janvier 1950 à juillet 1988) et suggère que de telles séries seraient prévisibles. L'étude plus récente de Berg et Lyhagen (1996), concernant les rentabilités boursières suédoises et basée sur l'estimation de processus ARFIMA par la méthode de Geweke et Porter-Hudak, fait ressortir la présence d'une mémoire de long terme sur les séries mensuelles et quotidiennes. Citons également les travaux de Sadique et Silvalpulle (1998) concernant les rentabilités des indices boursiers de sept pays : Australie, Japon, Corée, Malaisie, Singapour, Nouvelle Zélande et Etats Unis. L'application de la procédure de Geweke et Porter-Hudak et de divers tests de mémoire longue conduit les auteurs à retenir l'hypothèse de mémoire longue pour les rentabilités boursières coréenne et néo-zélandaise. Une telle conclusion est également obtenue pour le marché émergent grec par Barkoulas, Baum et Travlos (1996) estimant le paramètre d'intégration fractionnaire des processus ARFIMA par la méthode spectrale de Geweke et Porter-Hudak. Signalons enfin les travaux de Taqqu *et al.* (1995) et Willinger *et al.* (1999) utilisant diverses techniques pour estimer l'exposant de Hurst dans le domaine boursier.

Globalement, les résultats brièvement exposés ci-dessus basés sur l'application de l'analyse R/S conduisent à penser que les séries boursières sont caractérisées par une structure de dépendance de long terme. Cependant, l'analyse R/S traditionnelle ne constitue pas un test statistique et est sensible à la mémoire courte. Ces deux limites ont conduit Lo (1991) à introduire la statistique R/S modifiée qui, tout en restant sensible à la mémoire longue, est invariante sous une classe générale de processus à mémoire courte. Cette statistique diffère de la statistique R/S traditionnelle uniquement par le fait qu'elle tient compte, non plus uniquement de la somme des variances des termes individuels, mais inclut également les autocovariances. Lo (1991) applique alors l'analyse R/S modifiée sur les rentabilités des actions américaines et remet en question tous les résultats antérieurs portant sur les séries de rentabilités boursières puisqu'il montre qu'en aucun cas, on ne peut rejeter l'hypothèse nulle de mémoire courte. Ainsi, les études antérieures auraient conduit à des résultats biaisés en faveur d'une mémoire longue artificielle qui ne serait en réalité qu'une mémoire courte. Un tel résultat est confirmé par les travaux de Berg et

Lyhagen (1996) appliquant l'analyse R/S modifiée aux rentabilités de l'indice boursier suédois. Néanmoins, utilisant la technique R/S modifiée et diverses procédures d'estimation des processus ARFIMA, Lardic et Mignon (1996, 1997) mettent en avant la présence d'une structure de dépendance de long terme dans les rentabilités boursières quotidiennes des indices TOPIX, SBF 250, TSE 300 et BCI (voir tableau 1 en annexe) et confirment ainsi l'étude de Lo (1991) concernant l'absence de mémoire longue dans les rentabilités de l'indice boursier américain (SP 500).

5.1.1.2. Les séries de taux de change

Booth, Kaen et Koveos (1982) étudient les séries de taux de change sur deux sous périodes : période de changes fixes (1965-1971) et période de changes flottants (1973-1979). Les auteurs trouvent qu'en régime de changes fixes, les taux de change \$/DM (Dollar / Mark allemand), \$/£ (Dollar / Livre anglaise) et \$/FF (Dollar / Franc français) sont caractérisés par un phénomène d'anti-persistance ($H < 1/2$) alors qu'en régime de changes flottants, ils exhibent une mémoire de long terme positive : pour ces trois séries, les valeurs estimées de l'exposant de Hurst sont respectivement de 0,4821, 0,4825 et 0,4566 en changes fixes et de 0,5425, 0,6889 et 0,5273 en changes flottants. Le phénomène d'anti-persistance qui pourrait paraître a priori peu probable en économie a aussi été mis en évidence par Peters (1994) sur la volatilité. Cheung (1993) a également mené une étude sur diverses séries de taux de change. En appliquant la procédure de Geweke et Porter-Hudak, Cheung montre qu'un grand nombre de séries de taux de change sont caractérisées par une mémoire de long terme significative : \$/£ (H = 0,6334), \$/DM (H = 0,5464), \$/FS (Dollar / Franc suisse, H = 0,5477) \$/FF (H = 0,9707) et \$/JY (Dollar / Yen japonais, H = 0,550) sur la période 1974-1987. Ces valeurs témoignent ainsi de la présence d'une structure de dépendance de long terme dans les diverses séries de taux de change, structure particulièrement forte dans le cas du change \$/FF. La connaissance du phénomène de mémoire longue ne permet cependant pas d'effectuer de bonnes prévisions, puisque l'auteur montre que les modèles ARFIMA ne permettent pas de faire de meilleures prévisions hors échantillon que le simple modèle de marche aléatoire[18].

De nombreux autres travaux ont également été réalisés sur les taux de change dans le cadre de la vérification de l'hypothèse de parité des pouvoirs d'achat. Dans ce contexte, Diebold, Husted et Rusch (1991) ont estimé divers processus ARFIMA au moyen de la méthode de Fox et Taqqu (1986) sur les séries annuelles de taux de change de six pays débutant au milieu du XIXe siècle. Leurs résultats montrent que les chocs exogènes sur les taux de change ont des conséquences très durables, mais non permanentes, et valident ainsi l'hypothèse de parité des pouvoirs d'achat à long terme. Les travaux de

---

18. On rappelle que, lorsque l'on effectue des prévisions hors échantillon, la période prévisionnelle n'inclut pas la période d'estimation · les informations non connues ne sont donc pas utilisées pour la prévision Signalons en outre que la comparaison des prévisions entre plusieurs modèles est menée à l'aide de critères statistiques (racine de l'erreur quadratique moyenne, erreur absolue moyenne, écart type résiduel, etc) qu'il convient de minimiser

Cheung et Lai (1993) semblent également valider une telle hypothèse puisqu'ils mettent en avant l'existence d'une relation de cointégration fractionnaire (voir *infra*) entre les taux de change nominaux et les prix relatifs durant la période de changes fixes. Les travaux plus récents de Baum, Barkoulas et Caglayan (1999a, 1999b) semblent toutefois remettre en cause une telle hypothèse pour les séries de taux de change de 18 pays sur la période 1974-1995 dans la mesure où ils montrent que les taux de change ne sont pas caractérisés par la présence d'une mémoire longue, mais par la présence d'une mémoire infinie. Les auteurs valident ainsi l'hypothèse de martingale pour le processus d'évolution des taux de change et concluent en termes d'efficience des marchés des changes.

5.1.1.3. Les séries de taux d'intérêt

Les séries de taux d'intérêt ont également fait l'objet de travaux appliqués portant sur la recherche d'une structure de dépendance de long terme. Dans ce cadre, signalons l'étude de Shea (1991) qui applique la méthode de Geweke et Porter-Hudak pour estimer divers processus ARFIMA sur un ensemble de séries de taux d'intérêt. De même, Backus et Zin (1993) mettent en avant l'existence d'une structure de dépendance de long terme dans diverses séries de taux d'intérêt et étudient les implications de la présence d'une mémoire longue dans le contexte de la structure par terme des taux d'intérêt. Le même type d'étude est mené par Lardic et Mignon (1999b) pour les séries de prime de terme du Japon, du Royaume Uni, de la France et des Etats Unis. Utilisant la procédure de Geweke et Porter-Hudak ainsi que la méthode du maximum de vraisemblance exact, les auteurs montrent que ces séries sont caractérisées par la présence d'un paramètre d'intégration fractionnaire positif et significativement différent de zéro.

5.1.1.4. La modélisation de la volatilité des séries financières

Les travaux les plus récents en matière de recherche d'une structure de dépendance de long terme en finance portent principalement sur le processus d'évolution de la volatilité, c'est-à-dire de la variance conditionnelle des séries. L'intérêt vient ici du fait que la volatilité constitue une approximation du risque des séries financières : en conséquence, une modélisation adéquate de la volatilité peut apporter des éléments à la compréhension de l'évolution du risque inhérent aux actifs financiers. On peut notamment citer les travaux de Baillie, Bollerslev et Mikkelsen (1996) visant à modéliser le taux de change quotidien DM/$ au moyen de processus FIGARCH ou encore ceux de Bollerslev et Mikkelsen (1996) appliquant les processus FIGARCH exponentiels (FIEGARCH) aux séries de rentabilités boursières. Ces divers travaux mettent en avant la présence d'une mémoire de long terme dans le processus d'évolution de la volatilité des séries financières. Un tel résultat est par ailleurs confirmé sur diverses séries de taux de change par Harvey (1993), de rentabilités boursières par Ding et Granger (1996), Crato et de Lima (1994) et Lobato et Savin (1996) ou encore de taux d'intérêt par Lardic et Mignon (1999b).

5.1.2. *Mémoire longue et efficience des marchés financiers*

On se propose ici de présenter les principales implications des phénomènes de mémoire longue vis à vis de l'hypothèse d'efficience des marchés financiers (voir Mignon, 1998).

5.1.2.1. L'existence d'autocorrélations significatives et prévisions

En premier lieu, la présence d'une mémoire de long terme dans une série de rentabilités boursières indique l'existence d'autocorrélations significatives à long terme. Par conséquent, les cours boursiers ne suivent pas une marche aléatoire et les rentabilités ne répondent pas à un processus de bruit blanc. On sait cependant que le rejet de la marche aléatoire n'implique pas le rejet de l'hypothèse d'efficience des marchés[19]. On peut simplement dire que la propriété de mémoire longue suggère que les observations d'un processus à mémoire longue contiennent des informations utiles pour la détermination des valeurs futures de la série. Par conséquent, un « bon » processus à mémoire longue devrait conduire à de bonnes prévisions à long terme. Nous savons en effet que ce qui importe pour les prévisions à long terme, c'est la partie du spectre aux basses fréquences. Or, pour un processus de type ARMA la densité spectrale est finie et positive lorsque la fréquence tend vers zéro. Pour un processus ARIMA, elle est proportionnelle à l'inverse de la racine carrée de la fréquence. Par conséquent, ces deux types de processus ne conduisent pas à une modélisation adéquate aux basses fréquences. Les processus à mémoire longue ne donnent donc pas nécessairement de meilleures prévisions à court terme, mais ils devraient conduire à de meilleures prévisions à long terme où la modélisation aux basses fréquences est essentielle. On peut alors penser que le fait de pouvoir effectuer des prévisions correctes à long terme permet d'élaborer une stratégie rémunératrice sur les marchés financiers. Intuitivement, il serait donc possible de faire mieux qu'avec la marche aléatoire.

Lardic et Mignon (1999a) ont cherché à tester une telle supposition sur le marché des changes. Cette étude, portant sur les taux de change mensuels du dollar exprimé en six devises étrangères (mark, yen, dollar canadien, lire, livre et franc), a mis en avant la présence d'une mémoire de long terme pour trois de ces séries à l'aide de la technique du maximum de vraisemblance exact : \$/\$ canadien, \$/FF et \$/Lire (voir tableau 2 en annexe). Afin de savoir si la présence de mémoire longue permettait de battre la marche aléatoire, les auteurs ont réalisé des prévisions dynamiques[20] et comparé les capacités prédictives des processus ARFIMA, de la marche aléatoire et de

19. On rappelle qu'un test de l'hypothèse d'efficience est nécessairement un test de l'hypothèse jointe de la validité du modèle de formation des cours boursiers et de l'efficience. En effet, l'hypothèse d'efficience est toujours testée conjointement avec un certain modèle de formation des cours, qui peut être par exemple le modèle de marche aléatoire Dès lors, le rejet du modèle de marche aléatoire n'implique pas nécessairement le rejet de l'hypothèse d'efficience dans la mesure où le modèle de marche aléatoire peut être inadéquat pour décrire l'évolution des cours boursiers alors même que l'hypothèse d'efficience ne serait pas remise en cause

20. On rappelle que des prévisions dynamiques consistent à utiliser, pour la prévision, les valeurs prévues et non pas les vraies valeurs de la série. Par exemple, si on veut prévoir

trois modèles structurels couramment retenus dans les études sur les taux de change[21]. L'idée initiale consistait à reprendre l'étude de référence de Meese et Rogoff (1983) montrant que les modèles structurels et divers processus ARMA conduisaient systématiquement à de « moins bonnes » prévisions que la marche aléatoire. Leur étude, d'une part n'incluait pas les processus ARFIMA, et d'autre part portait sur les taux de change \$/DM, \$/yen et \$/£, c'est à dire sur des séries pour lesquelles Lardic et Mignon (1999a) avaient décelé uniquement une mémoire de court terme. Afin de déterminer si la supériorité de la marche aléatoire sur les autres modèles était due à l'absence de dynamique à long terme, Lardic et Mignon (1999a) ont mené des prévisions dynamiques roulantes sur les séries de taux de change caractérisées par une mémoire longue pour des horizons allant de un mois à douze mois. Les auteurs ont alors montré que pour la série \$/\$ canadien, la modélisation ARFIMA conduisait à de meilleures prévisions que la marche aléatoire, et ce quel que soit l'horizon de prévision. Toutefois, pour les deux autres séries, les prévisions ARFIMA ne supplantent celles de la marche aléatoire qu'à horizon court (un mois). Même si les résultats sont quelque peu différents selon les séries, cette étude suggère néanmoins que certains taux de change seraient prévisibles, même à long terme, ce qui jette un doute sur l'hypothèse d'efficience des marchés supports.

Une seconde implication de la mémoire de long terme concerne les phénomènes de cointégration et de retour à la moyenne (*mean reversion*).

5.1.2.2. Cointégration, retour à la moyenne et mémoire longue

La recherche d'une structure de dépendance de long terme dans les séries financières peut se révéler particulièrement pertinente vis à vis des phénomènes de cointégration et de retour à la moyenne.

a/ Cointégration et cointégration fractionnaire

La théorie de l'efficience des marchés financiers est basée sur une définition de la valeur fondamentale d'un actif fondée sur le modèle d'actualisation des dividendes futurs :

$$P_t = \sum_{j=0}^{\infty} \frac{1}{(1+r)^{j+1}} \, E[D_{t+j}|I_t] \tag{91}$$

où P_t est le cours en t, D_t les dividendes en t, r le taux d'intérêt (supposé ici constant) et $E[\cdot|I_t]$ est l'opérateur d'espérance conditionnelle à l'ensemble d'information I_t disponible en t.

---

en t une série pour la date $t+3$, on utilise les valeurs prévues pour les dates $t+1$ et $t+2$. Lorsque l'on réestime les paramètres du modèle de façon régulière dans le temps (par exemple à chaque fois qu'une nouvelle donnée est observée) et que l'on mène des prévisions dynamiques, les prévisions résultantes sont des prévisions dynamiques roulantes

21. Il s'agit du modèle monétaire à prix flexibles de Frenkel Bilson, du modèle monétaire à prix visqueux de Dornbusch Frankel et du modèle monétaire à prix rigides incorporant le compte courant de Hooper Morton

Dès lors, si ce modèle constitue une bonne représentation de l'évolution des cours boursiers – ce qui est généralement admis – alors la relation entre les cours et les dividendes doit être stationnaire : les cours et les dividendes doivent être cointégrés[22]. Si tel n'est pas le cas, il existe un écart durable entre le cours et sa valeur fondamentale. On peut même dire que cet écart dure infiniment. En effet, si le résidu de la relation de cointégration[23] n'est pas stationnaire, on peut supposer qu'il suit un processus de type ARIMA avec d entier (en général $d = 1$). En d'autres termes, l'absence de relation de cointégration indique que le cours ne revient jamais vers sa valeur fondamentale, ce qui remet bien évidemment fortement en cause l'hypothèse d'efficience des marchés.

Cependant, la restriction aux cas de résidu intégré d'ordre 0 (relation stationnaire) et de résidu intégré d'ordre 1 (relation non stationnaire) peut paraître arbitraire. Ainsi, de la même façon que les processus ARIMA peuvent être généralisés au travers des processus ARFIMA, ne peut-on pas étendre le concept de cointégration au cas fractionnaire ? L'idée serait alors que les cours et les dividendes soient fractionnairement cointégrés, de sorte que les résidus de la relation de cointégration suivent un processus ARFIMA. Cette notion de cointégration fractionnaire, évoquée par Granger (1986), nous paraît importante d'un point de vue économique puisqu'elle a pour conséquence l'existence d'une relation d'équilibre de long terme. En effet, dans ce cas, les erreurs tendent à retourner vers la moyenne, même si ce retour ne s'effectue qu'après un temps relativement long. En ce sens, les alternatives de résidus intégrés d'ordre 0 (cointégration) et intégrés d'ordre 1 (absence de cointégration) sont trop restrictives. Pour que le processus d'erreurs exhibe un comportement de retour à la moyenne, il n'est pas nécessaire qu'il soit intégré d'ordre 0 : les processus à mémoire longue, tels que les ARFIMA, affichent également un tel comportement. En outre, les processus ARFIMA incluent le cas $d = 1$ comme cas particulier : alors que l'effet d'un choc persiste à l'infini dans le cas d'un processus intégré d'ordre 1, il s'élimine, même lentement, pour un processus intégré d'ordre d, avec $-1/2 < d < 1$. En résumé, si les cours et les dividendes sont fractionnairement cointégrés, l'effet d'un choc sur le système s'annulera à long terme, et le modèle d'actualisation des dividendes sera vérifié. L'existence d'une relation de cointégration fractionnaire aurait alors pour conséquence la vérification, à long terme, de l'hypothèse d'efficience des marchés.

Le concept de cointégration fractionnaire a été récemment appliqué sur un système de séries de taux d'intérêt à long terme de cinq pays industrialisés (Etats-Unis, Canada, Allemagne, Royaume Uni et Japon) par Baum, Barkoulas et Oguz (1998). Cette étude fait ressortir le fait que ces taux d'intérêt forment un système fractionnairement cointégré. Plus précisément, alors même

22. On rappelle que les N composantes d'un vecteur X_t sont dites cointégrées à l'ordre (d, b) avec $0 < b \leqslant d$ si toutes les composantes de X_t sont intégrées d'ordre d et s'il existe un vecteur α ($\alpha \neq 0$) de taille $(N, 1)$ tel que $z_t = \alpha' X_t$ est intégré d'ordre $(d - b)$. Ce vecteur α est le vecteur de cointégration Rappelons également qu'une série est dite intégrée d'ordre d s'il est nécessaire de la différencier d fois pour la rendre stationnaire

23. La relation de cointégration est du type . $P_t = a + bD_t + \varepsilon_t$. Tester l'hypothèse d'absence de cointégration revient alors à tester la non stationnarité du résidu ε_t.

que chaque série de taux d'intérêt est caractérisée par la présence d'une racine unitaire, le système regroupant les cinq séries étudiées possède une composante fractionnaire. Ainsi, les chocs sur ce système ont des conséquences durables, mais non permanentes, suggérant ainsi l'existence d'une relation d'équilibre à long terme. Notons que ces résultats sont cohérents avec ceux de Baillie et Bollerslev (1994) mettant en avant l'existence d'une relation de cointégration fractionnaire entre les taux de change quotidiens des pays du G7. On pourra également se reporter à Robinson et Marinucci (1998) pour une application de la cointégration fractionnaire à diverses séries financières et macroéconomiques.

b/ Retour à la moyenne

Le phénomène de retour à la moyenne des prix à été mis en avant par Summers (1986). Ce dernier montre que, si le modèle habituel de formation des cours boursiers est vérifié, alors les excès de rentabilité doivent présenter une corrélation sérielle négative. Un tel schéma reflète la tendance supposée des prix à retourner vers leur valeur fondamentale. Ce comportement de retour à la moyenne des prix témoigne de deux éléments :

- d'une part, il signifie que les prix tendent à retourner vers leur valeur fondamentale : à la suite d'un choc, le prix s'écarte de sa valeur fondamentale, mais finit toujours par y revenir. Il existe alors une relation stable à long terme entre les cours et les dividendes et l'hypothèse d'efficience est vérifiée.
- d'autre part, l'existence du phénomène de retour à la moyenne induit bien évidemment l'existence d'un écart entre le cours et sa valeur fondamentale.

Alors que l'absence de cointégration standard avait pour conséquence l'existence d'un écart infiniment durable entre le cours et sa valeur fondamentale, la présence de *mean reversion* implique quant à elle un (lent) retour des prix vers leur valeur fondamentale. Il est important de souligner ici que l'écart entre les deux valeurs peut être très durable, mais les prix finiront toujours par retourner vers leur valeur fondamentale. Le phénomène de mean reversion peut alors selon nous être interprété en termes de mémoire longue. Ainsi, plus la durée du retour des prix vers leur valeur fondamentale est longue, plus la série exhibe une mémoire de long terme et inversement. On peut alors penser que plus la série est persistante, donc plus l'écart est durable, plus il sera possible d'établir une stratégie rémunératrice sur les marchés financiers. Effectivement, dans ces conditions, l'agent peut spéculer sur la différence entre le cours et sa valeur fondamentale.

En résumé, la présence de mémoire longue dans une série de rentabilités boursières a pour conséquence essentielle l'existence d'un écart *durable* entre le cours et sa valeur fondamentale. Ce phénomène témoigne d'un lent délai d'ajustement des prix à l'information, ce qui entre en contradiction avec l'hypothèse d'efficience. On comprend dès lors tout l'intérêt d'une étude sur les phénomènes de mémoire longue sur les marchés financiers, puisque les tests traditionnels d'autocorrélation et de marche aléatoire ne peuvent conduire à de telles conclusions (sur ce point, voir Lardic et Mignon, 1996 et 1997; Mignon, 1996 et 1998).

5.2. Applications en macroéconomie

Nous nous proposons à présent de mener une revue de la littérature des résultats empiriques obtenus sur diverses séries macroéconomiques. Comme précédemment, nous décomposerons cet aperçu des résultats en fonction de la nature des séries étudiées.

5.2.1. Les séries de PNB

La question concernant la persistance des chocs occupe une place centrale en macroéconomie, même si les travaux empiriques portant sur la mémoire longue sont nettement moins nombreux que dans le domaine financier. La principale controverse est centrée sur le fait de savoir si les séries macroéconomiques sont mieux approximées par des fluctuations autour d'une tendance déterministe ou par une marche aléatoire incluant une composante stationnaire. Les recherches d'une structure de dépendance de long terme dans les séries macroéconomiques portent ainsi principalement sur les séries de PNB. Les travaux empiriques dans ce domaine ont dès lors pour objet d'apporter des éléments de réponse au débat relatif à la stationnarité autour d'une tendance déterministe *versus* stationnarité en différence du PNB réel. Dans ce contexte, Diebold et Rudebusch (1989) ont appliqué la procédure de Geweke et Porter-Hudak (1983) sur la série américaine du PNB trimestriel d'après guerre. Ils mettent en avant la présence d'une structure de dépendance de long terme dans cette série (la valeur estimée du paramètre d'intégration fractionnaire est de l'ordre de 0,9). Un tel résultat est confirmé par Sowell (1992) qui, utilisant la méthode du maximum de vraisemblance exact, met en avant la supériorité des processus ARFIMA sur les processus de type ARIMA et ARMA usuels. Mentionnons également dans ce cadre l'étude de Koop *et al.* (1995) étudiant les propriétés de persistance du PNB réel américain au moyen d'une analyse bayésienne des processus ARFIMA.

5.2.2. Les séries d'indices des prix

Les travaux portant sur les séries de prix font également généralement ressortir la présence d'une structure de dépendance de long terme. L'étude de Baillie (1996) montre ainsi que la série de l'indice des prix du blé de Beveridge est caractérisée par une mémoire longue dans la mesure où une telle série est non stationnaire en niveau alors qu'elle apparaît sur-différenciée en différence première. Cet aspect est par ailleurs confirmé par l'estimation de processus ARFIMA caractérisés par un paramètre d'intégration fractionnaire significatif et proche de 0,5. Baillie, Chung et Tieslau (1996) étudient quant à eux diverses séries mensuelles d'indices des prix à la consommation et montrent que les processus ARFIMA semblent bien saisir l'évolution de ces séries. Notons en outre que ces mêmes auteurs ont estimé des processus ARFIMA avec erreurs GARCH pour les séries de taux d'inflation et ont mis en avant l'existence d'un paramètre d'intégration fractionnaire significatif pour tous les pays du

G7 à l'exception du Japon. Les études réalisées par Tieslau (1992), Delgado et Robinson (1994) et Hassler et Wolters (1995) font également ressortir ce phénomène de mémoire longue dans diverses séries de taux d'inflation. Signalons en outre qu'une telle conclusion est validée par Bos, Franses et Ooms (1998) sur les séries de taux d'inflation des pays du G7. Enfin, l'étude de Baum, Barkoulas et Caglayan (1997) met en avant un phénomène de persistance dans les séries de taux d'inflation à la fois des pays industrialisés et des pays en développement. Ils confirment par ailleurs ce résultat sur diverses séries d'indices des prix de gros par le biais d'estimations de processus ARFIMA.

5.2.3. Les séries de consommation

L'intérêt des études de la mémoire longue dans le domaine de la consommation est particulièrement notable en référence aux travaux de Brown (1952). Ce dernier, prolongeant les travaux de Duesenberry, met en effet en avant l'existence d'un effet d'inertie de la consommation, dû à la formation d'habitudes : le passé intervient ainsi de façon continue par le biais de la consommation passée. On retrouve par ailleurs ces effets de mémoire dans la théorie du revenu permanent de Friedman (1957). On sait en effet que le revenu permanent peut être estimé à l'aide des revenus observés de la période courante et des périodes passées. On a donc bien ici un phénomène de mémoire dans la mesure où le revenu courant dépend de l'ensemble des revenus passés. Dans un tel contexte, Diebold et Rudebusch (1991) et Haubrich (1992) ont étudié la relation entre la consommation et le revenu ainsi que le paradoxe de Deaton selon lequel la consommation est trop lisse par rapport à ce que suggère la théorie du revenu permanent. Haubrich (1992) montre ainsi qu'en modélisant le revenu par un processus ARFIMA, alors la variance de la consommation et celle du revenu deviennent compatibles avec l'hypothèse de revenu permanent.

5.2.4. Les séries d'investissement

La recherche d'une structure de dépendance de long terme a également fait l'objet de travaux appliqués aux séries d'investissement. Lardic et Mignon (1997) ont ainsi estimé des processus ARFIMA, par la méthode de Geweke et Porter-Hudak et par la technique du maximum de vraisemblance exact, sur les séries trimestrielles d'investissement des pays du G7 (voir tableau 3 en annexe). Les auteurs ont ainsi mis en avant la présence d'un phénomène de persistance dans les séries allemande, japonaise, française et italienne et un phénomène d'anti-persistance dans la série canadienne à l'aide de la technique du maximum de vraisemblance exact. La recherche d'une structure de dépendance de long terme dans de telles séries peut être clairement reliée aux théories explicatives de l'investissement et en particulier au modèle d'accélérateur flexible de Koyck (1954). Ce dernier relie en effet le stock de capital à une moyenne pondérée des productions des années précédentes, le poids attribué à chaque année décroissant au cours du temps. Les fluctuations de l'investissement sont dès lors moins brutales que dans le cas de l'accélérateur simple, en raison de l'effet d'inertie exercé par le stock de capital de la période précédente. La présence d'une mémoire longue dans les séries d'investissement témoigne ainsi d'un lent délai d'ajustement du stock de capital à la demande, ce qui est compatible avec le modèle d'accélérateur flexible.

5.2.5. Les séries de masse monétaire et études globales

Pour clore cette brève revue de la littérature macroéconomique, citons l'étude de Porter-Hudak (1990). L'auteur met en avant que les autocorrélations, saisonnières et non saisonnières, des séries de masse monétaire (M1, M2 et M3) aux Etats Unis décroissent à un taux hyperbolique suggérant la présence d'une structure de dépendance de long terme. Une telle hypothèse est confirmée par l'auteur par le biais de l'estimation de processus SARFIMA. En effet sur la période janvier 1959 à avril 1988, Porter-Hudak (1990) obtient une valeur de d égale à 0,661 pour M1, 0,461 pour M2 et 0,402 pour M3, ces valeurs étant statistiquement significatives. On trouvera également une application des processus SARFIMA sur les séries de revenus dans Ray (1993). La présence d'une mémoire de long terme est par ailleurs confirmée par l'étude de Lardic et Mignon (1997) sur les séries de masse monétaire de l'Allemagne, du Royaume Uni et de la France. Baum, Barkoulas et Caglayan (1998) ont également réalisé une étude sur diverses séries monétaires américaines. Ces travaux, utilisant la procédure de Geweke et Porter-Hudak, mettent en avant la présence d'une structure de dépendance de long terme dans les séries de masse monétaires et suggèrent que de telles séries seraient prévisibles par le biais de processus ARFIMA. On peut enfin signaler l'étude relativement complète de Crato et Rothman (1994) qui estiment les paramètres des processus ARFIMA par la méthode du maximum de vraisemblance exact sur un grand nombre de séries macroéconomiques (PNB, production industrielle, taux de chômage, indice des prix à la consommation, salaires, masse monétaire...) initialement étudiées par Nelson et Plosser (1982). A l'exception de certaines séries relatives au marché du travail, les auteurs mettent en avant la pertinence de l'hypothèse de mémoire longue. On trouvera également une application des processus ARFIMA et de l'analyse R/S sur diverses séries macroéconomiques dans Lardic et Mignon (1997).

6. CONCLUSION

Les processus ARFIMA apparaissent comme particulièrement pertinents pour la modélisation des séries temporelles comportant une structure de dépendance de long terme. Rappelons en effet qu'ils permettent de modéliser à la fois les dynamiques de court et long termes d'une série. Le comportement de court terme est pris en compte par le biais des paramètres autorégressifs et moyenne mobile usuels alors que la dynamique de long terme est représentée par un unique paramètre : le paramètre d'intégration fractionnaire d. On comprend dès lors aisément que la modélisation ARFIMA présente l'avantage d'être nettement plus parcimonieuse que la modélisation ARMA. Cette dernière nécessiterait, pour approximer le comportement de long terme de la série, un grand nombre de paramètres autorégressifs et moyenne mobile. La revue de la littérature que nous avons menée ici témoigne en outre de l'intérêt pratique des processus ARFIMA. De ce point de vue, les travaux récents portant sur la modélisation de la volatilité au moyen de processus FIGARCH semblent particulièrement prometteurs et peuvent apporter des éléments nouveaux à la compréhension du processus d'évolution du risque des séries financières.

REMERCIEMENTS

Nous tenons à remercier les deux rapporteurs anonymes pour leurs remarques et commentaires qui ont largement contribué à l'amélioration de la version initiale du papier. Nous restons seules responsables des éventuelles erreurs ou omissions.

ANNEXE

TABLEAU 1. — Mémoire longue et séries de rentabilités boursières.

	Séries nominales			Séries réelles
Série	Valeur de H Analyse R/S	Valeur de H R/S modifiée	MVE	MVE
TOPIX (Japon)	0,577	0,569*	(2; 0,052; 0) $t_d = 2,575$	(2; 0,192; 2) $t_d = 3,589$
SBF 250 (France)	0,532	0,524	(0, 0,065; 0) $t_d = 2,484$	(0, 0,020; 1) $t_d = 0,350$
TSE 300 (Canada)	0,531	0,512	(0, 0,102, 0) $t_d = 5,910$	(3, 0,874; 2) $t_d = -13,219$
BCI (Italie)	0,575	0,566*	(0; 0,060; 0) $t_d = 3,547$	(1, 0,107; 1) $t_d = 1,999$

L'application porte sur des séries quotidiennes sur la période 2/01/1986 – 2/12/1994 (sauf pour le SBF 250 qui débute le 28/12/1990). Les séries réelles ont été déflatées par l'indice des prix à la consommation

* · valeur significativement différente de 1/2

t_d = valeur de la statistique de Student associée au coefficient d estimé. MVE · estimation des paramètres du processus ARFIMA(p, d, q) selon la procédure du maximum de vraisemblance exact.

Source : Lardic et Mignon (1996, 1997).

TABLEAU 2 — Mémoire longue et séries de taux de change.

Série	Valeur de d	t_d	Modèle
\$/FF	0,2275	3,9275	ARFIMA$(0, d, 0)$
\$/\$ canadien	0,1376	2,6010	ARFIMA$(0, d, 0)$
\$/Lire	0,2195	4,0306	ARFIMA$(0, d, 0)$
\$/DM	0,0466	0,6719	ARFIMA$(0, d, 1)$
\$/Yen	0,0034	0,0522	ARFIMA$(0, d, 1)$
\$/£	0,0009	0,0191	ARFIMA$(0, d, 1)$

L'application porte sur les séries des variations logarithmiques des taux de change mensuels sur la période janvier 1974 à juillet 1995. t_d est la valeur de la statistique de Student associée au coefficient d estimé

Source . Lardic et Mignon (1999a)

TABLEAU 3. — Mémoire longue et investissement.

Série	Période d'estimation	GPH	MVE
Allemagne	1960.1 1994.4	$d = -0,1473$ $t_d = -0,67$	ARFIMA$(3, d, 1)$ d = 0,4181 $t_d = 4,7801$
Canada	1947 1 1995.3	$d = 0,0753$ $t_d = 0,23$	ARFIMA$(1, d, 0)$ $d = -0,6977$ $t_d = -5,8753$
USA	1947.1 1995.3	$d = 0,3222$ $t_d = 1,93$	ARFIMA$(2, d, 0)$ $d = -0,0947$ $t_d = -0,5571$
Italie	1970.1 1995.3	$d = -0,0679$ $t_d = -0,22$	ARFIMA$(1, d, 0)$ $d = -0,4033$ $t_d = -2,3717$
Japon	1955.2 1995 3	$d = 0,4917$ $t_d = 1,33$	ARFIMA$(0, d, 0)$ $d = 0,1592$ $t_d = 4,0221$
France	1970.1 1995.3	$d = 0,2128$ $t_d = 0,67$	ARFIMA$(0, d, 0)$ $d = 0,3557$ $t_d = 4,7966$
Royaume Uni	1955.1 1995.3	d = 0,1529 $t_d = 0,76$	ARFIMA$(2, d, 2)$ $d = 0,0319$ $t_d = 0,5063$

L'application porte sur des séries trimestrielles. GPH est la méthode d'estimation de Geweke et Porter Hudak, MVE est la méthode du maximum de vraisemblance exact t_d est la valeur de la statistique de Student associée au coefficient d estimé.

Source : Lardic et Mignon (1997).

BIBLIOGRAPHIE

AKONOM J. (1988) "Processus transformés d'un ARMA ou d'un processus de Wiener. Problèmes d'estimation", Thèse, Université de Lille.

AZAÏS J.M. et LANG G. (1993) "Estimation de l'exposant de longue dépendance dans un cadre semi-paramétrique", CRAS, série I math, n° 6, pp. 611-614.

BACKUS D.K. et ZIN S.E. (1993) "Long-Memory Inflation Uncertainty : Evidence from the Term Structure of Interest Rates", *Journal of Money, Credit and Banking*, pp. 681-700.

BAILLIE R.T. (1996), "Long Memory Processes and Fractional Integration in Econometrics", *Journal of Econometrics*, Vol. 73, n° 1, pp. 5-59.

BAILLIE R.T. et BOLLERSLEV T. (1994) "Cointegration, Fractional Cointegration, and Exchange Rate Dynamics", *Journal of Finance*, Vol. 49, pp. 737-745.

BAILLIE R.T., BOLLERSLEV T., et MIKKELSEN H.O. (1996), "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, Vol. 74, n° 1, pp. 3-30.

BAILLIE R.T., CHUNG C.F. et TIESLAU M.A. (1996) "Analyzing Inflation by the Fractionally Integrated ARFIMA-GARCH Model", *Journal of Applied Econometrics*, 11, pp. 23-40.

BAUM C.F., BARKOULAS J.T. et TRAVLOS N. (1996) "Long Memory in the Greek Stock Market", *Working Paper*, Boston College.

BAUM C.F., BARKOULAS J.T. et CAGLAYAN M. (1997) "Persistence in International Inflation Rates", *Working Paper*, Boston College.

BAUM C.F., BARKOULAS J.T. et CAGLAYAN M. (1998) "Fractional Monetary Dynamics", *Working Paper*, Boston College.

BAUM C.F., BARKOULAS J.T. et OGUZ G.S. (1998) "Fractional Dynamics in a System of Long Term International Interest Rates", *Working Paper*, Boston College.

BAUM C.F., BARKOULAS J.T. et CAGLAYAN M. (1999a) "Long Memory or Structural Breaks : Can Either Explain Nonstationarity Real Exchange Rates under the Current Float ?", *Working Paper*, Boston College.

BAUM C.F., BARKOULAS J.T. et CAGLAYAN M. (1999b) "A Reexamination of the Long-Memory Evidence in the Foreign Currency Market", *Working Paper*, Boston College.

BERG L. et LYHAGEN J. (1996) "Short and Long Run Dependence in Swedish Stock Returns", *Working Paper*, Uppsala University, n° 1996 :19.

BOLLERSLEV T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, Vol. 31, pp. 307-327.

BOLLERSLEV T., CHOU R.Y., JAYARAMAN N. et KRONER K.F. (1991), "Les modèles ARCH en finance : un point sur la théorie et les résultats empiriques", *Annales d'économie et de statistique*, n° 24, pp. 1-59.

BOLLERSLEV T. et ENGLE R.F. (1986), "Modelling the Persistence of Conditional Variances", *Econometric Reviews*, 5, pp. 1-50.

BOLLERSLEV T. et MIKKELSEN H.O. (1996), "Modeling and Pricing Long Memory in Stock Market Volatility", *Journal of Econometrics*, Vol. 73, n° 1, pp. 151-184.

BOOTH G.G., KAEN F.R. et KOVEOS P.E. "R/S Analysis of Foreign Exchange Rates under Two International Monetary Regimes", *Journal of Monetary Economics*, pp. 407-415.

BOS C.S., FRANSES P.H. et OOMS M. (1998) "Long Memory and Level shifts : Re-Analyzing Inflation rates", *Tinbergen Institute Discussion Paper*, TI98-039/4.

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

BROWN T.M. (1952) "Habit Persistence and Lags in Consumer Behavior", *Journal of Econometrics.*

CHEUNG Y-W. (1993) "Tests for Fractional Integration : A Monte Carlo Investigation", *Journal of Time Series Analysis,* Vol. 14, n° 4. pp. 331-345.

CHEUNG Y-W. (1993) "Long-Memory in Foreign Exchange Rates", *Journal of Business and Economic Statistics,* 11, pp. 93-101.

CHEUNG Y-W. et DIEBOLD F.X. (1994) "On Maximum Likelihood Estimation of the Differencing Parameter of Fractionally-Integrated Noise with Unknown Mean", *Journal of Econometrics,* Vol. 62, pp.301-316.

CHEUNG Y-W et LAI (1993) "A Fractional Cointegration Analysis of Purchasing Power Parity", *Journal of Business and Economic Statistics,* 11, pp. 103-112

CHUNG C.F. (1996) "Estimating a Generalized Long Memory Process", *Journal of Econometrics,* Vol. 73, pp. 237-259.

CRATO N. et DE LIMA P.J.F. (1994) "Long-Range Dependence in the Conditional Variance of Stock Returns, *Economics Letters,* 45, pp. 281-285.

CRATO N. et ROTHMAN P. (1994) "Fractional Integration Analysis of Long Run Behavior for US Macroeconomic Time Series", *Economics Letters,* 45, pp. 287-291.

DELGADO M.A. et ROBINSON P.M. (1994) "New Methods for the Analysis of Long-Memory Time Series : Application to Spanish Inflation", *Journal of Forecasting,* Vol. 13, pp. 97-107.

DIEBOLD F.X., HUSTED S. et RUSH M. (1991) "Real Exchange Rates Under the Gold Standard", *Journal of Political Economy,* 99, pp. 1252-1271.

DIEBOLD F.X. et RUDEBUSCH G.D. (1989) "Long Memory and Persistence in Aggregate Output", *Journal of Monetary Economics,* Vol. 24, n° 2, pp. 189-209.

DIEBOLD F.X. et RUDEBUSCH G.D. (1991) "Is Consumption too Smooth? Long Memory and the Deaton Paradox", *Review of Economics and Statistics,* 71, pp. 1-9.

DING Z. et GRANGER C.W.J. (1996), "Modeling Volatility Persistence of Speculative Returns : A New Approach", *Journal of Econometrics,* Vol. 73, n° 1, pp. 185-215.

FAMA E.F. (1970) "Efficient Capital Markets : A Review of Theory and Empirical Work", *Journal of Finance,* n° 2, pp. 383-417.

FAMA E.F. (1991) "Efficient Capital Markets : II", *Journal of Finance,* n° 5, pp. 1575-1617.

FOX R. et TAQQU M.S. (1986) "Large-Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series", *Annals of Statistics,* Vol. 14, pp. 517-532.

FRIEDMAN M. (1957) *A Theory of the Consumption Function,* New York.

GEWEKE J. et PORTER-HUDAK S. (1983) "The Estimation and Application of Long Memory Time Series Models", *Journal of Time Series Analysis,* Vol. 4, n° 4, pp. 221-238.

GRAF (1983) *Long Range Correlations and the Estimation of the Self Similarity Parameter,* Ph. D. Thesis, ETH Zrich, n° 7357.

GRANGER C.W.J. (1986) "Developments in the Study of Cointegrated Economic Variables", *Oxford Bulletin of Economics and Statistics,* Vol. 48, n° 3, pp. 213-228.

GRANGER C.W.J. et JOYEUX R. (1980) "An Introduction to Long-Memory Time Series Models and Fractional Differencing", *Journal of Time Series Analysis,* Vol. 1, n° 1, pp. 15-29.

GRAY H.L., ZHANG N.F. et WOODWARD W.A. (1989) "On Generalized Fractional Processes", *Journal of Time Series Analysis*, Vol. 10, n° 3, pp. 233-257.

GREENE M.T. et FIELITZ B.D. (1977) "Long-Term Dependence in Common Stock Returns", *Journal of Financial Economics*, Vol. 4, pp. 339-349.

HANNAN E.J. (1970) *Multiple Time Series*, John Wiley.

HANNAN E.J. (1973) "The Asymptotic Theory of Linear Time Series Models", *Journal of Applied Probabilities*, Vol. 10, pp. 130-145.

HARDOUIN C. (1993) "Estimation du paramètre fractionnaire par la méthode des contrastes", Thèse, Université de Paris I.

HARVEY A.C. (1993) "Long Memory in Stochastic Volatility", *Working Paper*, London School of Economics.

HASSETT (1990) "Is the Aggregate Labor Market Exploding?", Manuscript, Graduate School of Business, Columbia University, New-York.

HASSLER U. et WOLTERS J. (1995) "Long Memory in Inflation Rates : International Evidence", *Journal of Business and Economic Statistics*, 13, pp. 37-45.

HAUBRICH J.G. (1992) "Consumption and Fractional Differencing : Old and New Anomalies", *Working Paper*, University of Pennsylvania.

HAUSER M.A. (1998) "Maximum Likelihood Estimators for ARMA and ARFIMA Models : A Monte Carlo Study", à paraître dans *Journal of Statistical Planning and Inference.*

HIGUCHI T. (1988) "Approach to an Irregular Time Series on the Basis of the Fractal Theory", Physica, 31D, pp. 277-283.

HOSKING J.R.M. (1981) "Fractional Differencing", *Biometrika*, Vol. 68, n° 1.

HOSKING J.R.M. (1984) "Modeling Persistence in Hydrological Time Series Using Fractional Differencing", *Water Resources Research*, Vol. 20, pp. 1898-1908.

HURST H.E. (1951) "Long-Term Storage Capacity of Reservoirs", *Transactions of the American Society of Civil Engineers*, Vol. 116, pp. 770-799.

JANACEK G.J. (1982) "Determining the Degree for Time Series via the Log Spectrum", *Journal of Time Series Analysis*, Vol. 3, pp. 177-183.

JEFFREYS (1939, 1948, 1961) *Theory of Probability*, Clarendon Press, Oxford.

KASHYAP R.L. et EOM K.B. (1988) "Estimation in Long-Memory Time Series Model", *Journal of Time Series Analysis*, Vol. 9, n° 1, pp. 35-41.

KOLMOGOROV A.N. (1940) "Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum", *Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS*, XXVI, n° 2, pp. 115-118.

KOOP G., LEY E., OSIEWALSKI J. et STEEL M.F.J. (1995) "Bayesian Analysis of Long Memory and Persistence using ARFIMA Models", à paraître dans *Journal of Econometrics.*

KOYCK L.M. (1954) *Distributed Lags and Investment Analysis*, Amsterdam.

LARDIC S. (1992) *La non-stationnarité des séries macro économiques*, Mémoire de DEA, Université Paris X-Nanterre.

LARDIC S. (1996) "Non stationnarité, mémoire des séries et hystérésis", *Revue d'Economie Politique*, Vol. 106, n° 3, pp. 417-450.

LARDIC S. (1997) *L'hystérésis en économie : théorie et mesure*, Thèse pour le doctorat de Sciences Economiques, Université Paris X – Nanterre.

LARDIC S. et MIGNON V. (1996) "Les tests de mémoire longue appartiennent-ils au "camp du démon" ?", *Revue Economique*, Vol. 47, n° 3, pp. 531-540.

LARDIC S. et MIGNON V. (1997) "Essai de mesure du degré de mémoire longue des séries. L'exemple de la modélisation ARFIMA", *Economie Appliquée*, n° 2, pp. 161-195.

LARDIC S. et MIGNON V. (1999a) "Prévision ARFIMA des taux de change : les modélisateurs doivent-ils encore exhorter à la naïveté des prévisions ?", *Annales d'Economie et de Statistique*, n° 54, pp. 47-68.

LARDIC S. et MIGNON V. (1999b) "Modélisation FIGARCH appliquée à l'analyse de la structure par terme des taux d'intérêt", *Finance*, Vol. 20, pp. 91-114.

LI W.K. et MCLEOD A.I. (1986) "Fractional Time Series Modelling", *Biometrika*, Vol. 73, n° 1, pp. 217-221.

LO A.W. (1991) "Long-Term Memory in Stock Market Prices", *Econometrica*, Vol. 59, n° 5, pp. 1279-1313.

LOBATO I.N. et SAVIN N.E. (1996) "Real and Spurious Long Memory Properties of Stock Market Data", *Working Paper*, University of Iowa.

MCLEOD et HIPEL (1978) "Preservation of the Rescaled Adjusted Range, 1, A Reassessment of the Hurst Phenomenon", *Water Resources Research*, 14, pp. 491-508.

MANDELBROT B.B. (1965) "Une classe de processus stochastiques homothétiques à soi; application à la loi climatologique de H. E. Hurst", *Comptes rendus de l'Académie des Sciences de Paris*, 260, pp. 3274-3277.

MANDELBROT B.B. (1972) "Statistical Methodology for Nonperiodic Cycles : From the Covariance to R/S Analysis", *Annals of Economic and Social Measurement*, Vol. 1, n° 3, pp. 259-290.

MANDELBROT B.B. (1973) "Le problème de la réalité des cycles lents et le syndrome de Joseph", *Economie Appliquée*, Vol. 26, pp. 349-365.

MANDELBROT B.B. et TAQQU M. (1979) "Robust R/S Analysis of Long Run Serial Correlation", *Bulletin of the International Statistical Institute*, 48, n° 2, pp. 69-104.

MANDELBROT B.B. et VAN NESS J.W. (1968) "Fractional Brownian Motions, Fractional Noises and Applications", *SIAM Review*, Vol. 10, n° 4, pp. 422-437.

MANDELBROT B.B. et WALLIS, (1968) "Noah, Joseph and Operational Hydrology", *Water Resources Research*, 4, n° 5, pp. 909-918.

MANDELBROT B.B. et WALLIS, (1969a) "Computer Experiments with Fractional Gaussian Noises", *Water Resources Research*, 5, pp. 228-267.

MANDELBROT B.B. et WALLIS, (1969b) "Some Long Run Properties of Geophysical Records", *Water Resources Research*, 5, n° 2, pp. 321-340.

MANDELBROT BB. et WALLIS, (1969c) "Robustness of the Rescalled Range R/S in the Measurement of Noncyclic Long Run Statistical Dependance", *Water Resources Research*, 5, n° 5, pp. 967-988.

MEESE R. et ROGOFF K. (1983) "Empirical Exchange Rate Models of the Seventies : Do They Fit Out of Sample ?", *Journal of International Economics*, Vol. 14, pp. 3-24.

MIGNON V. (1996) "Les implications de la mémoire longue et de la non linéarité sur l'efficience du marché des changes", *Journal de la Société de Statistique de Paris*, n° 1, pp. 51-72.

MIGNON V. (1998) *Marchés financiers et modélisation des rentabilités boursières*, Economica, collection Approfondissement de la Science Economique.

MIGNON V. (1998) "Méthodes d'estimation de l'exposant de Hurst. Application aux rentabilités boursières", *Economie et Prévision*, n° 132-133, pp. 193-214.

NELSON C.R. et PLOSSER C.I. (1982) "Trends and Random Walks in Macroeconomic Time Series : Some Evidence and Implications", *Journal of Monetary Economics*, 10, pp. 139-162.

NEWCOMB (1886) "A Generalized Theory of the Combination of Observations so as to Obtain the Best Result", *American Journal of Mathematics*, 8, pp. 343-366.

OLSHEN R. (1967) "Asymptotic Properties of the Periodogram of a Discrete Stationary Process", *Journal of Applied Probability*, 5, pp. 508-528.

PEITGEN H.O. et SAUPE D. (1988) *The Science of Fractal Images*, Springer Verlag.

PETERS E.E. (1991) *Chaos and Order in the Capital Markets*, John Wiley & Sons.

PETERS E.E. (1994) *Fractal Market Analysis*, John Wiley & Sons.

PORTER-HUDAK S. (1990) "An Application of the Seasonal Fractionally Differenced Model to the Monetary Agregates", *Journal of the American Statistical Association*, Vol. 85, n° 410, pp. 338-344.

RAY B.K. (1993) "Long-Range Forecasting of IBM Product Revenues Using a Seasonal Fractionally Differenced ARMA Model", *International Journal of Forecasting*, 9, pp. 255-269.

ROBINSON P.M. (1994) "Semiparametric Analysis of Long Memory Time Series", *Annals of Statistics*, 22, pp. 515-539.

ROBINSON P.M. (1995a) "Gaussian Semiparametric Estimation of Long Range Dependence", *Annals of Statistics*, 23, pp. 1630-1661.

ROBINSON P.M. (1995b) "Log Periodogram Regression of Time Series with Long Range Dependence", *Annals of Statistics*, 23, pp. 1048-1072.

ROBINSON P.M. et MARINUCCI D. (1998) "Semiparametric Frequency Domain Analysis of Fractional Cointegration", *London School of Economics*, Discussion Paper n° EM/98/348.

SADIQUE S. et SILVAPULLE P. (1998) "Long-Term Memory in Stock Market Prices : International Evidence", *Working Paper*, La Trobe, Department of Economics.

SAMORODNITSKY G. et TAQQU M.S. (1994) Stable Non-Gaussian Processes : Stochastic Models with Infinite Variance, Chapman & Hall.

SHEA G.S. (1991) "Uncertainty and Implied Variance Bounds in Long-Memory Models of the Interest Rate Term Structure", *Empirical Economics*, Vol. 16, pp. 287-312.

SMITH (1938) "An Empirical Law Describing Heterogeneity in the Yields of Agricultural Crops", Journ. Agric. Sci., 28, pp. 1-23.

SOWELL F. (1992a) "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models", *Journal of Econometrics*, Vol. 53, pp. 165-188.

SOWELL F. (1992b) "Modeling Long-Run Behavior with the Fractional ARIMA Model", *Journal of Monetary Economics*, Vol. 29, pp. 277-302.

STUDENT (1927) "Errors of Routine Analysis", *Biometrika*, 19, pp. 151-164.

SUMMERS L.H. (1986) "Does the Stock Market Rationally Reflect Fundamental Values ?", *Journal of Finance*, n° 3, pp. 591-601.

TAQQU M.S., TEVEROVSKY V. et WILLINGER W. (1995) "Estimators of Long-Range Dependence : An Empirical Study", Preprint, Boston University, 18 pages.

TIESLAU M.A. (1992) "Strongly Dependent Economic Time Series : Theory and Applications", Ph.D., Michigan State University.

WHITTLE (1956) "On the Variation of Yield Variance with Plot Size", *Biometrika*, 43, pp. 337-343.

WILLINGER W., TAQQU M.S. et TEVEROVSKY V. (1999) "Stock Market Prices and Long Range Dependence", *Finance and Stochastics*, n° 1, pp. 1-14.

YAJIMA Y. (1985) "On Estimation of Long Memory Time Series Models", *Australian Journal of Statistics*, Vol.27, pp. 303-320.