

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Abderrahmane Mira de Béjaïa



Faculté des Sciences exactes
Département Mathématiques

Mémoire de fin de cycle

En vue de l'obtention d'un diplôme de Master en mathématiques

Option : Statistique et Analyse Décisionnelle

Réalisé par

DJAFRI Tifithene & IDRES Sonia

THÈME

PROCESSUS DE LONGUE MEMOIRE

Soutenu publiquement, le 23/06/2013 devant le jury composé de :

Mr.	A. AIT SAIDI	Professeur	U.A.M. Béjaïa	Président.
Mme.	K. TIMERIDJINE	MAA	U.A.M. Béjaïa	Rapporteur.
Mme.	L. BOURAINE	MCB	U.A.M. Béjaïa	Examineur.

2012/2013

Table des matières

Table des Matières	i
Table des Figures	iv
Liste des tableaux	v
Introduction	1
1 Préliminaires	4
1.1 Opérateur de retard	4
1.2 Opérateur de différenciation	4
1.3 Processus aléatoire	4
1.4 Processus stationnaire	5
1.5 Série chronologique	5
1.6 Bruit blanc	7
1.7 Les processus stationnaires à courtes mémoires	7
1.7.1 Processus autoregressif	7
1.7.2 Processus moyenne mobile	8
1.7.3 Processus autoregressif, moyenne mobile	9
1.7.4 Processus ARIMA	10
1.7.5 Processus SARIMA	10
1.8 Les processus stationnaires de longue mémoire	11
2 Analyse temporelle et spectrale	12
2.1 Définitions	12
2.2 Quelques rappels d'analyse	14
2.2.1 Série de Fourier	14
2.2.2 Intégrale de Fourier	15
2.3 Analyse spectrale des processus stationnaire	15
2.3.1 Densité spectrale d'un processus aléatoire stationnaire	15

2.3.2	Distribution spectrale d'un processus aléatoire stationnaire	16
2.3.3	Densité spectrale du processus bruit blanc	16
2.3.4	Densité spectrale du processus ARMA(p, q)	16
2.3.5	Représentation de Wold	17
2.4	Périodogramme	17
3	Processus de longue mémoire	19
3.1	L'opérateur de dérivation fractionnaire	19
3.2	Processus de longue mémoire	20
3.2.1	Le modèle autoregressif moyenne mobile fractionnaire (FARIMA)	21
3.3	Analyse d'une série temporelle	22
3.3.1	Dickey-Fuller simple	23
3.3.2	Dickey-Fuller augmenté	23
3.4	les méthodes de détection de mémoire longue	24
3.4.1	La statistique R/S (exposant de Hurst)	24
3.4.2	L'analyse Lo	25
3.5	Test de bruit blanc	26
3.6	Test de Student	27
4	Caractéristiques d'un processus à mémoire longue pure	28
4.1	Processus fractionnaire pure FI(d)	28
4.1.1	Conditions de causalité, d'inversibilité et stationnarité	28
4.1.1.1	Notion de filtre de moyenne mobile	29
4.1.2	Fonction d'autocovariance	30
4.1.3	Caractéristiques de la fonction d'autocovariance :	32
4.1.4	Fonction d'autocorrélation	33
4.1.5	Fonction d'autocorrélation partielle	34
4.1.6	Densité spectrale	34
4.2	La relation entre l'exposant de Hurst et le paramètre d'intégration fractionnaire	36
5	Estimation du paramètre de longue mémoire	37
5.1	Méthodes d'estimation du processus FARIMA(p, d, q)	37
5.1.1	Méthodes à deux étapes	37
5.1.1.1	Méthode de Geweke et Porter Hudak	37
5.1.1.2	Méthodes de Robinson	40
5.1.2	Méthodes à une étape (Méthodes du maximum de vraisemblance)	41
5.1.2.1	Méthodes du maximum de vraisemblance exacte	41
5.1.2.2	Méthodes du maximum de vraisemblance approximé par Whittle	42
5.2	Simulation	43

Conclusion	45
A Annexe	46
A.1 Quelques formules concernant la fonction hypergéométrique	46
A.2 Méthode de Beran	47
Bibliographie	48

Table des figures

1.1	Nombre annuel de tâches solaires observées à la surface du soleil de 1700 à 1980	5
1.2	Nombre de passagers par mois (en milliers) dans les transports aériens, de 1949 à 1960.	6
1.3	Nombre de morts accidentelles aux Etats-Unis de 1973 à 1978.	6
1.4	Représentation graphique d'un bruit blanc.	7
1.5	Représentation graphique d'un processus AR(3), sa fonction d'autocorrélation et sa fonction d'autocorrélation partielle.	8
1.6	Représentation graphique d'un processus MA(1), sa fonction d'autocorrélation et sa fonction d'autocorrélation partielle.	9
1.7	Représentation graphique d'un processus ARMA(1, 1), sa fonction d'autocorrélation et sa fonction d'autocorrélation partielle.	10
1.8	Représentation graphique d'un processus SARIMA.	11

Liste des tableaux

5.1	Estimation du paramètre " d " pour différentes tailles avec la méthode de Geweke et Porter Hudak (GPH).	43
5.2	Estimation du paramètre " d " pour différentes tailles avec la méthode de Valderio A. Reisen.	44
5.3	Estimation du paramètre " d " pour différentes tailles par la méthode du maximum de vraisemblance approximée.	44

Introduction

Une série chronologique est une succession d'observations d'une grandeur effectuée à un intervalle régulier au cours du temps. La spécificité de l'analyse d'une série chronologique, qui la distingue des autres analyses statistiques, est précisément l'importance accordée à l'ordre dans lequel sont effectuées les observations.

L'étude des séries chronologiques (ou temporelles) a connu un développement très important depuis de très longues dates, aussi bien en pratique qu'en théorie, telle que l'économie, " On the periodicity of sunspots", (1906), en météorologie, " Time series regression of sea level on weather", (1968), en théorie de signal, "Noise in FM receivers ", (1963), en biologie, "The auto-correlation curves of schizophrenic brain waves and the power spectrum", (1960), en économie, "Time series analysis of imports, exports and other economic variables", (1971).

Les astronomes ont été les premiers à utiliser les séries de Fourier pour les séries chronologiques. Toutefois, il a fallu attendre (1889) pour que Sir Arthur Shuster introduise le périodogramme, qui constitue la base des méthodes spectrales d'analyse des séries chronologiques. Si cette théorie a donné de bons résultats en astronomie, son application en économie a conduit à des résultats nettement moins concluants. En (1921) et (1922) Beveridge a utilisé cette théorie sur le prix du blé ("wheat prices and rainfall in western europe"). La série présentait tellement de pics qu'au moins 20 périodicités étaient possibles, et plus encore, si en tenait compte des facteurs économiques ou météorologiques. Cette méthode s'est réveillée à mettre en œuvre en sciences humaines.

En (1884), Poynting a introduit pour la première fois, l'approche temporelle qui se repose sur l'autocorrélogramme ou plus généralement sur l'utilisation de la corrélation sérielle, en étudiant la corrélation entre le mouvement du prix du blé et les importations du coton et de la soie. Le coefficient de corrélation sérielle a été défini par Hooker en (1901), dans une étude sur le taux de mariage en Angleterre et l'indice du commerce.

En séries chronologiques, la dépendance temporelle entre les variables constitue la source principale de l'information. Celle-ci peut être entièrement contenue dans la valeur moyenne des variables, dont les erreurs sont supposées stochastiquement indépendantes, mais bien souvent cette hypothèse n'est qu'une approximation de la réalité.

L'analyse des séries chronologiques a connu un grand développement depuis l'apparition du livre de Box et Jenkins (1970)¹ qui ont proposé un algorithme d'ajustement d'une chronique temporelle par les modèles autorégressifs moyenne mobile (ARMA) ; en offrant une explication de la valeur présente de la série étudiée selon ses propres valeurs passées et/ou les valeurs des

¹Modèles à courte mémoire

résidus, où les principales propriétés des processus stationnaires autorégressif moyenne mobile (ARMA) ont été décrites avec les méthodes d'identification, d'estimation et de validation.

Ces modèles ont un aspect régulier et des corrélations qui décroissent exponentiellement. Ils ont permis d'améliorer les résultats obtenus sous l'hypothèse d'indépendance. Citons comme exemple, le calcul de l'écart type $\bar{\sigma}$, de la moyenne empirique \bar{X} de n observations indépendantes $(X_1; \dots; X_n)$ supposées identiquement distribuées, centrées et de variance σ^2 .

Soit $\bar{\sigma} = \frac{\sigma}{\sqrt{n}} = \sqrt{\text{var}(\bar{X})}$. Dans le cas d'un modèle ARMA, $\bar{\sigma}$ est proportionnel à $\frac{\sigma}{\sqrt{n}}$, mais pas forcément égale à $\frac{\sigma}{\sqrt{n}}$; cependant, les modèles à mémoire courte se sont avérés insuffisants dans le domaine de la modélisation. En effet il est arrivé que l'on modélise à l'aide de variables aléatoires indépendantes des séries qui présentaient des corrélations $\rho(h)$ décroissant en $|h|^\alpha$, avec $\alpha \in [0, 1]$ donc des corrélations non sommables. Or, même si les corrélations sont faibles, l'effet de ce comportement dit de longue mémoire peut être désastreux si l'on suppose l'indépendance des variables. Dans l'exemple précédent, pour un processus à longue mémoire, $\bar{\sigma}$ décroît avec une vitesse $n^{-\frac{\alpha}{2}}$. Ce comportement est nettement différent de ce que l'on rencontre dans le cas d'indépendance des variables, mais correspond à certaines situations observées en pratique. L'idée de développer les modèles à mémoire longue s'impose naturellement.

Exemple Le phénomène de longue mémoire n'est pas sans effet sur \bar{X} et S_n^2 . Soient,

$$\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t \text{ et } S_n^2 = \frac{1}{n-1} \sum_{t=1}^n (X_t - \bar{X})^2$$

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n} \left(1 + 2 \sum_{h=1}^{n-1} \left(1 - \frac{h}{n} \right) \rho(h) \right)$$

$\rho(h)$ est le coefficient d'autocorrélation d'ordre h .

Dans un modèle ARMA classique $\rho(h)$ est telle que $\rho(h) < cr^h$, où c et r sont des constantes vérifiant, $0 < c < \infty$ et $0 < r < 1$. Alors que dans un modèle autorégressif fractionnaire à mémoire longue, $\rho(h)$ décroît de la même façon que h^{2d-1} . ($\rho(h) \sim h^{2d-1}$), quand h tend vers ∞ . Il est clair que $\text{Var}(\bar{X})$ décroît plus lentement dans le cas de mémoire longue ou courte que dans le cas d'un modèle ARMA classique.

$$\text{Var}(\bar{X}) = \sigma^2 \left(1 + 2(n-1) \sum_{h=1}^{n-1} \left(1 - \frac{h}{n} \right) \rho(h) \right)$$

Le biais dépend alors de $\rho(h)$.

Suite à ces problèmes, Granger et Joyeux [25], Hosking [30] ont introduit le concept de différenciation fractionnaire, des modèles ARIMA à différenciation fractionnaire (FARIMA) (Auto Regressive Fractionally Integrated Moving Average), pour lequel sa structure d'autocorrélation décroît lentement jusqu'à zéro lorsque le décalage augmente, ou sa densité spectrale est fortement concentrée aux fréquences proche de zéro. Cette structure d'autocorrélation indique que l'état actuel du processus doit être fortement dépendant des valeurs passées depuis un temps très long et que par conséquent, pour modéliser ces processus, l'information passée doit être

intégrée dans la description du processus.

L'étude de ces modèles et leurs applications constituent actuellement un des domaines populaires de la recherche statistique. Depuis une bonne quinzaine d'années la littérature est abondante et riche en résultats. Le lecteur peut se référer aux ouvrages généraux de Beran [5], Deniau et al [15], Doukhan et al [18] et Robinson [49]. Les domaines d'applications utilisant les modèles à longue mémoire n'ont cessé de se développer ces dernières années, touchant des domaines variés. Pour illustrer cette affirmation citons quelques travaux couvrant un large spectre : Hurst [32], Hosking [31] en hydrologie, Mandelbrot [36], Lo [34], Willinger et al [57] en finance, Hassler et Wolters [29] en macroéconomie, Graf [25], Taquq et al [53].

Pourquoi s'intéresser à des processus FARIMA(p, d, q) ?

Un avantage des modèles FARIMA sur les autres modèles est que leurs structures d'autocorrélation sont capable de mettre en évidence un comportement " mémoire à court terme" semblable à ceux des modèles ARMA mais aussi un comportement de "mémoire à long terme".

Dans ce mémoire nous étudions un modèle de longue mémoire donné par l'équation stochastique suivante :

$$(1 - L)^d y_t = \varepsilon_t$$

$(y_t, t \in \mathbb{T})$ est un processus aléatoire ;

ε_t est un bruit blanc ;

d est le paramètre d'intégration fractionnaire.

Notre mémoire est organisé comme suit :

Le chapitre 1 de ce mémoire est consacré aux préliminaires. Nous rappelons quelques notions de bases utilisées dans ce travail (les processus aléatoire stationnaire, les processus à courte mémoire et leurs propriétés ...etc.)

Dans le chapitre 2, nous présentons l'analyse temporelle et spectrale des processus.

Le chapitre 3 porte sur les processus à mémoire longue FARIMA(p, d, q).

Dans le chapitre 4 nous présentons une étude détaillée sur le processus à mémoire longue pure FI(d), nous avons démontré que les fonctions, autocovariance, génératrices des autocovariance et autocorrélation sont des fractions rationnelles qui est une propriété importante des processus.

Dans le chapitre 5, nous présentons quelques méthodes d'estimation du paramètre de mémoire longue d , en se basant essentiellement sur la fonction de vraisemblance et la densité spectrale.

Dans le chapitre 6 une étude de simulation est conçue pour éclaircir les notions présentées dans les chapitres 4 et 5.

On termine par une conclusion et perspectives.

1

Préliminaires

Le but de ce chapitre est de se familiariser avec les principaux aspects des processus aléatoires. On commence par des rappels et définitions supposés déjà assimilés. On y donne ensuite la définition des séries chronologiques, des éléments de classification des séries chronologiques selon plusieurs critères, des exemples pratiques, quelques notions utiles pour la suite (processus à courte mémoire), ainsi que les raisons qui motivent l'étude des séries chronologiques.

1.1 Opérateur de retard

On note par y_1, y_2, \dots, y_t une série de longueur t . L'opérateur de retard notée L tel que

$$Ly_t = y_{t-1} \quad (1.1)$$

1.2 Opérateur de différenciation

La différence première élimine la tendance linéaire notée ∇ est donnée par

$$\nabla y_t = y_t - y_{t-1}$$

Avec $\nabla = I - L$ et I est l'opérateur d'identité $Iy_t = y_t$

Si on applique k fois l'opérateur de différenciation ∇ , on décale le processus de k unités de temps :

$$L^k y_t = y_{t-k}$$

1.3 Processus aléatoire

On utilise le terme de processus aléatoire pour décrire une variable dont le comportement ne peut pas être exprimé entièrement par une relation déterministe.

Un processus aléatoire est une suite de variables aléatoires indexées dans le temps et définies sur un espace des états de la nature. Ainsi, pour chaque instant du temps, la valeur de la quantité étudiée y_t est appelée variable aléatoire et l'ensemble des valeurs y_t quand t varie est appelé processus aléatoire. L'indice t ($t \in \mathbb{T}$) s'interprète comme la date à laquelle est faite l'observation ou comme la période sur laquelle elle porte. Les observations étant en nombre fini, il pourrait paraître naturel de choisir pour \mathbb{T} un ensemble fini. On préfère, cependant retenir $\mathbb{T} = \mathbb{N}$ ou $\mathbb{T} = \mathbb{Z}$. Prolonger l'ensemble des indices vers $+\infty$, permet en effet de prendre en compte la possibilité d'observations nouvelles et d'étudier les propriétés asymptotiques des diverses procédures statistiques. L'intérêt du prolongement vers $-\infty$ est principalement mathématique : il permet en effet dans certains cas des écritures et des résultats plus simples [47].

1.4 Processus stationnaire

Définition 1.1. Un processus $(y_t, t \in \mathbb{Z})$ est **fortement stationnaire** ou **strictement stationnaire** si la loi de $(y_{t_1}, \dots, y_{t_n})$ est la même que la loi de $(y_{t_1+h}, \dots, y_{t_n+h})$ pour tout (t_1, \dots, t_n) avec $t_i \in \mathbb{Z}$, pour tout $i = 1$ et pour tout $h \in \mathbb{Z}$ avec $t_i+h \in \mathbb{Z}$

Ainsi, un processus aléatoire est strictement stationnaire si toutes ses caractéristiques, c'est-à-dire tous ses moments sont invariants pour tout changement de l'origine du temps.

Mais la stationnarité au sens strict est trop restrictive et on assouplit cette condition en définissant la stationnarité du second ordre.

Définition 1.2. Un processus $(y_t, t \in \mathbb{Z})$ du second ordre est **faiblement stationnaire** ou **stationnaire à l'ordre 2**, si son espérance $\mathbb{E}(y_t)$ et ses autocovariances $cov(y_t, y_s)$ sont invariantes par translation dans le temps :

- $\mathbb{E}(y_t) = \mu$,
- $var(y_t) = \sigma^2$,
- $\forall (s, t) \in \mathbb{Z}^2, \forall h, (s+h, t+h) \in T^2, cov(y_s, y_t) = cov(y_{s+h}, y_{t+h})$

Exemple :

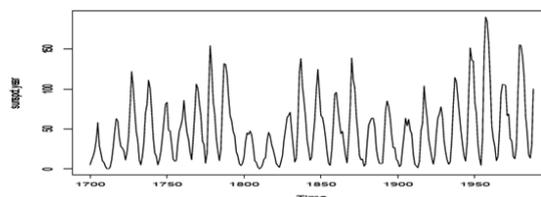


FIG. 1.1 – Nombre annuel de tâches solaires observées à la surface du soleil de 1700 à 1980

1.5 Série chronologique

Une série chronologique ou encore série temporelle est un ensemble de données mesurant un phénomène et repéré dans l'ordre croissant du temps [11]. Son analyse repose sur :

- L'observation des valeurs dans le passé à des dates fixées et équidistantes ;
- La suggestion d'hypothèses de travail permettant de justifier l'emploi de certaines méthodes de prévisions.

On distingue deux types de séries chronologiques

- Série continue : C'est une série où l'observation est faite d'une manière continue ($T \in \mathbb{R}$).
- Série discrète : C'est une série où l'observation est faite sur des intervalles de temps fixés à priori : heure, jour, mois, année,...($T \in \mathbb{Z}$).

Définition 1.3. Une série chronologique $(y_1, y_2, y_3, \dots, y_t)$ est dite stationnaire si elle est la réalisation d'un processus stationnaire. C'est à dire qu'elle n'admet ni tendance ni saisonnalité (elle est homogène par rapport au temps).

Les composantes d'une série chronologique sont :

- La tendance qui représente l'évolution à long terme de la série y_t étudiée, elle traduit le comportement moyen de la série.

Exemple :

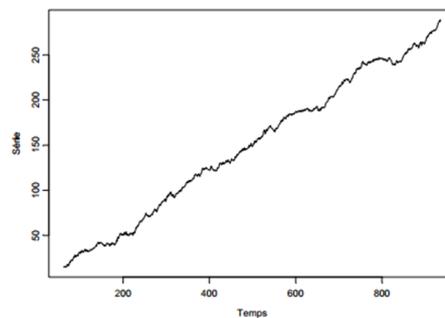


FIG. 1.2 – Nombre de passagers par mois (en milliers) dans les transports aériens, de 1949 à 1960.

- La composante saisonnière (périodique) notée S_t , elle correspond à un phénomène qui se répète à l'intervalle de temps régulier . En général c'est un phénomène saisonnier d'où le terme de variation saisonnier.

Exemple :

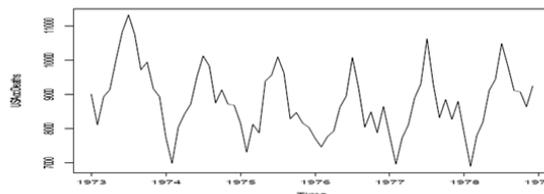


FIG. 1.3 – Nombre de morts accidentelles aux Etats-Unis de 1973 à 1978.

- La composante régulière (composante non contrôlable aléatoire) appelée aussi Bruit Blanc ou résidu. Elle comporte tout ce qui n'est pas expliqué par les autres composantes, on la note ε_t .

Exemple : Trajectoire d'un bruit blanc Gaussien $\mathcal{N}(0,1)$

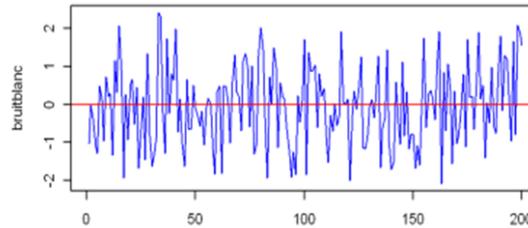


FIG. 1.4 – Représentation graphique d'un bruit blanc.

1.6 Bruit blanc

Définition 1.4. Un processus $(\varepsilon_t, t \in \mathbb{Z})$ est dit bruit blanc, si ε_t est une suite de variables aléatoires indépendantes, identiquement distribuées, d'espérance nulle et de variance constante. Généralement en pratique, on considère que $\varepsilon_t \sim \mathcal{N}(0, \sigma)$

Remarque 1.1. Si les variables aléatoires sont gaussiennes, le bruit blanc est dit gaussien.

1.7 Les processus stationnaires à courtes mémoires

Les travaux de G. Box et G. Jenkins (dans les années 1970) ont beaucoup contribué dans la théorie et la pratique des modèles des séries temporelles. L'objectif auquel ils se proposent de répondre dans leur ouvrage, "Time Series Analysis ; Forecasting and Control ", est de construire un modèle aléatoire de type ARMA permettant de reproduire au mieux les réalisations d'une série temporelle.

1.7.1 Processus autoregressif

Un processus $(y_t, t \in \mathbb{Z})$ est dit autoregressif d'ordre p noté $AR(p)$ s'il vérifie l'équation stochastique suivante [17] :

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \iff \Phi(L)y_t = \varepsilon_t \quad (1.2)$$

où ε_t est un bruit blanc centré et de variance σ^2 ,

et $\phi_1, \phi_2, \dots, \phi_p$ sont des paramètres réels, et Φ est un polynôme de degré p .

Le polynôme retard associé à un processus autoregressif $AR(p)$ est défini par :

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad (1.3)$$

Le polynôme caractéristique associé à un processus autoregressif $AR(p)$ est défini par :

$$y(z) = z^p - \phi_1 z^{p-1} - \phi_2 z^{p-2} - \dots - \phi_{p-1} z - \phi_p \quad (1.4)$$

Proposition 1. Si z une racine du polynôme retard (1.3) alors $\frac{1}{z}$ est une racine du polynôme caractéristique (pour la démonstration voir [1]).

Proposition 2. Le processus $AR(p)$ est stationnaire si seulement si les racines du polynôme retard (1.3) sont à l'extérieur du cercle unité c'est à dire plus grande que l'unité en module $|\phi_i| > 1, i = 1, \dots, p$.

Proposition 3. Le processus $AR(p)$ est stationnaire si seulement si les racines du polynôme caractéristique (1.4) sont à l'intérieur du cercle unité c'est à dire plus petite que l'unité en module $|\frac{1}{z}| < 1$.

Exemple : Série d'un processus $AR(3)$ simulé [12], $y_t = 0.3y_{t-1} - 0.7y_{t-2} + 0.5y_{t-3} + \varepsilon_t$

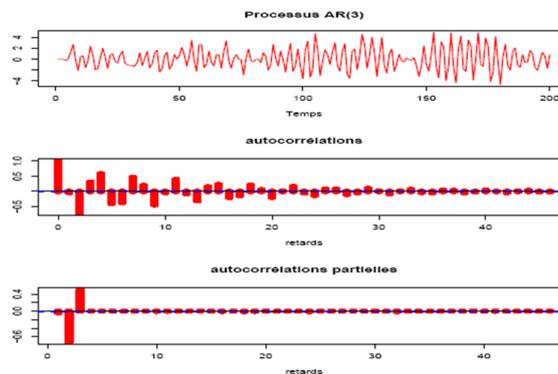


FIG. 1.5 – Représentation graphique d'un processus $AR(3)$, sa fonction d'autocorrélation et sa fonction d'autocorrélation partielle.

1.7.2 Processus moyenne mobile

Un processus $(y_t, t \in \mathbb{Z})$ est dit moyenne mobile d'ordre q noté $MA(q)$ (moving average en anglais) s'il vérifie l'équation stochastique suivante [17] :

$$\begin{aligned} y_t &= \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q} \\ &= \Theta(L)\varepsilon_t \end{aligned} \tag{1.5}$$

où Θ est un polynôme de degré q dont les coefficients sont $\{1, \theta_1, \theta_2, \dots, \theta_q\}$ et ε_t est un processus bruit blanc.

Le plynôme retard associé à un processus moyenne mobile $MA(q)$ est défini par :

$$\Theta(L) = 1 - \theta_1L - \theta_2L^2 - \dots - \theta_pL^p \tag{1.6}$$

Proposition 4. [37] Le polynôme retard Θ est inversible si et seulement si $|\theta_i| < 1, i = 1, \dots, q$.

Exemple : Série d'un processus MA(1) simulé [12], $y_t = \varepsilon_t - \varepsilon_{t-1}$

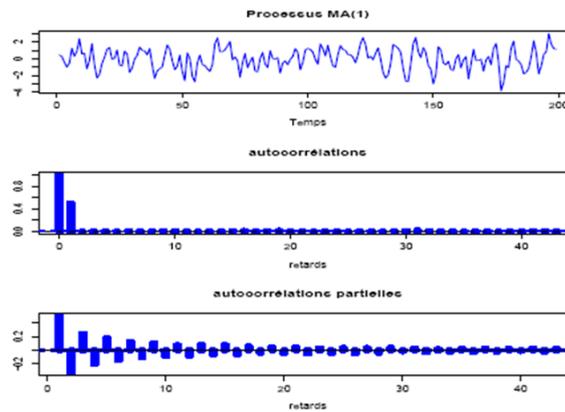


FIG. 1.6 – Représentation graphique d'un processus MA(1), sa fonction d'autocorrélation et sa fonction d'autocorrélation partielle.

1.7.3 Processus autoregressif, moyenne mobile

Un processus $(y_t, t \in \mathbb{Z})$ est dit autoregressif, moyenne mobile d'ordre (p, q) noté ARMA(p, q) s'il vérifie l'équation stochastique suivante [17] :

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1.7)$$

On peut l'écrire aussi

$$\Phi(L)y_t = \Theta(L)\varepsilon_t$$

Proposition 5. *stationnarité : le processus ARMA(p, q) est stationnaire si toutes les racines de l'équation (1.2) sont à l'extérieur du disque unité.*

Proposition 6. *invertibilité : le processus ARMA(p, q) est inversible si toutes les racines de l'équation (1.5) sont à l'extérieur du disque unité.*

Exemple : Série d'un processus ARMA(1,1) simulé [12], $y_t = 0.4y_{t-1} + \varepsilon_t - 0.8\varepsilon_{t-1}$.

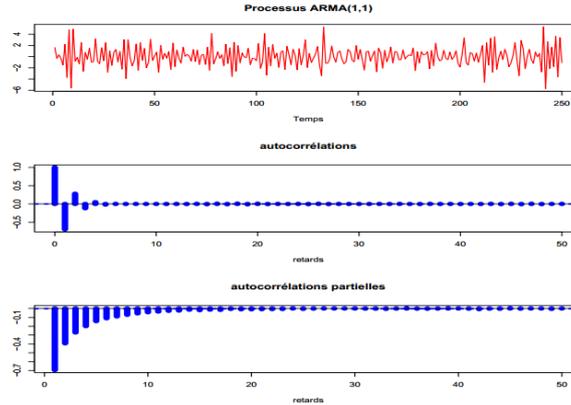


FIG. 1.7 – Représentation graphique d'un processus ARMA(1,1), sa fonction d'autocorrélation et sa fonction d'autocorrélation partielle.

1.7.4 Processus ARIMA

On appelle processus ARIMA(p, d, q) un processus non stationnaire ($y_t, t \in \mathbb{Z}$), pour le quel le processus différencié d fois, vérifie la relation suivante :

$$\Phi(L)\nabla^d y_t = \Theta(L)\varepsilon_t \quad (1.8)$$

Où Θ et Φ sont deux polynômes de degré respectifs p et q .

Remarque 1.2. Le rapport des polynômes $\frac{\Theta(z)}{\Phi(z)}$, pour $z \in \mathbb{Z}$ est dit fonction de transfert.

1.7.5 Processus SARIMA

Ces processus sont une généralisation des modèles ARIMA(p, d, q), contenant une partie saisonnière. Soit ($y_t, t \in \mathbb{Z}$) suit un processus SARIMA(p, d, q)(P, D, Q) $_s$, si cette série a une saisonnalité de période s qu'on peut écrire comme suit :

$$\Phi_1(L)\Phi_2(L^s)(1-L)^d(1-L^s)^D y_t = \Theta_1(L)\Theta_2(L^s)\varepsilon_t \quad (1.9)$$

Avec

- Φ_1 : un polynôme de degré p ;
- Φ_2 : un polynôme de degré P ;
- Θ_1 : un polynôme de degré q ;
- Θ_2 : un polynôme de degré Q .

Exemple : Marché total des anti-inflammatoires [54] :Différence saisonnière (s = 12).

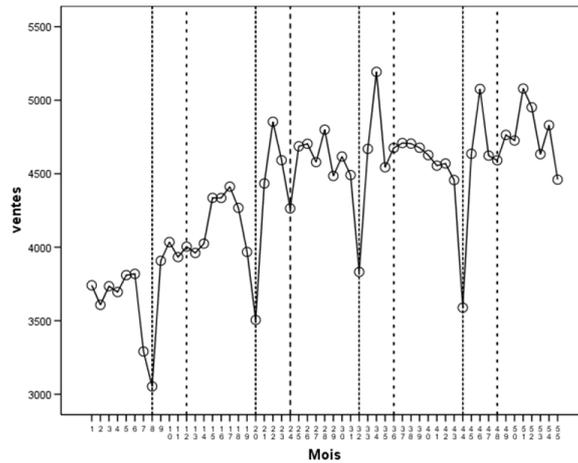


FIG. 1.8 – Représentation graphique d'un processus SARIMA.

1.8 Les processus stationnaires de longue mémoire

Les modèles de longue mémoire les plus étudiés sont les processus stationnaires qui ont des corrélations qui décroissent hyperboliquement. Les deux classes les plus connues sont d'une part les accroissements du mouvement brownien fractionnaire¹, et d'autre part les processus ARIMA fractionnaire (ou FARIMA) que nous allons détailler dans les chapitres 3, 4.

¹Un mouvement brownien Fractionnaire de paramètre H où $H \in (0, 1)$ est un processus gaussien réel centré noté $B^H(t), t \in \mathbb{R}$ défini sur un espace probabilisé $(\Omega, \mathbb{F}, \mathbf{P})$ et vérifiant :

1. $B_0^H = 0$
2. $\mathbf{E}[(B_t^H)^2] = |t|^{2H} \quad \forall t \in \mathbb{R}$
3. B^H a des accroissements stationnaires

2

Analyse temporelle et spectrale

Dans l'analyse temporelle, on s'intéresse essentiellement aux corrélations inter-temporelles entre les observations construisant ainsi la suite des corrélations de la même façon que nous disposons de la suite des observations. Plus précisément, nous essaierons de montrer quelle mesure il est possible de passer de l'étude d'une série dans le domaine des temps à l'étude dans le domaine des fréquences [33]. L'objectif premier de la théorie spectrale est l'identification des fluctuations d'une série temporelle aux principales fréquences. L'analyse fréquentielle d'une série temporelle trouve sa justification dans le fait que chaque composante d'une série peut être identifiée comme une fonction périodique. Celle-ci pouvant être décomposée en un nombre fini d'oscillations¹ sinusoïdales. Ce principe découle de l'analyse harmonique de Fourier (1822) qui précise que toute oscillation est caractérisée par son amplitude, sa phase et sa période. Le point de départ de toute discussion sur le sujet est donc la théorie des séries et intégrales de Fourier [8].

Ce chapitre décrit les notions de densité spectrale des processus aléatoires stationnaires, ainsi que l'estimation de la densité spectrale, en particulier le périodogramme.

2.1 Définitions

Définition 2.1. *Fonction d'autocovariance*

La fonction d'autocovariance d'un processus aléatoire complexe est donnée par :

$$\text{cov}(t, s) = \text{cov}(y_t, y_s) = E([y_t - E(y_t)][y_s - E(y_s)]^*), \quad (2.1)$$

Où $'^*$ désigne le complexe conjugué. Pour un processus stationnaire, on peut réécrire la fonction de covariance comme une fonction d'une seule variable :

$$\forall t, s \in \mathbb{Z} \quad \text{et} \quad s = t + \tau, \quad \tau \in \mathbb{T} \quad \text{cov}(y_t, y_s) = \text{cov}(y_t, y_{t+\tau}) = \gamma(\tau) \quad (2.2)$$

¹Une oscillation est un mouvement ou une fluctuation périodique. Les oscillations sont soit à amplitude constante soit amorties.

Remarque 2.1. *Un processus aléatoire stationnaire dont la fonction de covariance est identiquement nulle sauf pour $\tau = 0$ est qualifié de bruit blanc faible.*

Définition 2.2. Fonction d'autocorrélation

La fonction d'autocorrélation d'un processus $(y_t, t \in \mathbb{Z})$, de moyenne $\mathbb{E}(y_t) = \mu$, notée $\rho(h)$ est définie $\forall h \in \mathbb{Z}$ [47] :

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} \quad (2.3)$$

avec $\rho(h) \in [-1, 1]$ et γ_h désigne la fonction d'autocovariance

$$\text{cov}(y_t, y_{t-h}) = \mathbb{E}[(y_t - \mu)(y_{t-h} - \mu)]$$

Remarque 2.2. *Les fonctions $\gamma(h)$, $\rho(h)$ sont symétriques $\forall h \in \mathbb{R}$:*

$$\gamma(h) = \gamma(-h).$$

et

$$\rho(h) = \rho(-h).$$

Les fonctions $\gamma(h)$, $\rho(h)$ sont symétriques $\forall h \in \mathbb{C}$:

$$\overline{\gamma(h)} = \gamma(-h).$$

et

$$\overline{\rho(h)} = \rho(-h).$$

Où $\bar{\gamma}$ est le conjugué de γ .

Où $\bar{\rho}$ est le conjugué de ρ .

Définition 2.3. Fonction d'autocorrélation partielle

L'autocorrélation partielle d'ordre h désigne la corrélation entre y_t et y_{t-h} obtenue lorsque l'influence des variables y_{t-h-i} (avec $i < h$) a été retirée [47]. La fonction d'autocorrélation partielle d'un processus $(y_t, t \in \mathbb{Z})$, de moyenne $\mathbb{E}(y_t) = \mu$, notée $\alpha(h)$ est définie par :

$$\alpha(h) = \frac{|P_h^*|}{|P_h|}, \quad \forall h \in \mathbb{N} \quad (2.4)$$

Avec

$$P_h^* = \begin{pmatrix} 1 & \rho_1 & \cdot & \rho_1 \\ \rho_1 & 1 & \cdot & \rho_2 \\ \cdot & \cdot & \cdot & \cdot \\ \rho_{h-1} & \cdot & \cdot & \rho_h \end{pmatrix}$$

Et

$$P_h = \begin{pmatrix} 1 & \rho_1 & \cdot & \rho_{h-1} \\ \rho_1 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \rho_{h-1} & \cdot & \cdot & 1 \end{pmatrix}$$

2.2 Quelques rappels d'analyse

Définition 2.4. *Fonction localement intégrable*

Soit f une fonction $:\mathbb{R} \rightarrow \mathbb{R}$, f est dite localement intégrable, si $\int_a^b |f(x)|dx$ converge pour tout couple de réels (a, b) avec $a < b$.

Définition 2.5. *Fonction périodique*

Une fonction $f : \mathbb{I} \rightarrow \mathbb{R}$ est périodique s'il existe $T \in \mathbb{N}^*$, pour tout x tel que $x + T \in \mathbb{I}$,

$$f(x) = f(x + T) \quad (2.5)$$

Le plus petit entier naturel $T > 0$, qui vérifie l'équation (2.5) est appelé période.

Définition 2.6. *Discontinuité de première espèce*

Une fonction f admet une discontinuité de première espèce en un point x_0 si les limites à droite et à gauche de x_0 existent. (Celles-ci ne sont pas forcément égales sauf en cas de continuité.)

2.2.1 Série de Fourier

Soit $f : \mathbb{R} \rightarrow \mathbb{C}$ une fonction 2π -périodique et intégrable sur tout compact² de \mathbb{R} .

Définition 2.7. On appelle série de Fourier associée à f , la série trigonométrique :

$$f(t) = \frac{a_0}{2} + \sum_{j=1}^{\infty} [a_j \cos(\omega_j t) + b_j \sin(\omega_j t)] \quad (2.6)$$

Avec

$$a_j = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(\omega_j t) dt$$

$$b_j = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(\omega_j t) dt$$

où a_j et b_j sont appelés les coefficients de Fourier de f .

²Un ensemble $K \subset \mathbb{E}$ est compact si toute suite d'éléments de K , on peut extraire une sous suite convergente vers un élément de K

Théorème 2.1. *Dirichlet (2.6)*

Soit $f : \mathbb{R} \rightarrow \mathbb{C}$ une fonction périodique vérifiant :

1. Les discontinuités de f (si elles existent) sont de première espèce et sont en nombre fini dans tout intervalle fini.
2. f admet en tout point une dérivée à droite et une dérivée à gauche.

Alors la série de Fourier associée à f est simplement convergente sur \mathbb{R} , et sa somme vaut pour tout $x \in \mathbb{R}$:

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)] = \begin{cases} f(x), & \text{si } f \text{ est continue en } x; \\ \frac{f(x+0)+f(x-0)}{2}, & \text{si } f \text{ est discontinue en } x. \end{cases} \quad (2.7)$$

Où $f(x+0)$ et $f(x-0)$ représentent respectivement les limites à droite et à gauche de f au point x .

2.2.2 Intégrale de Fourier

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction localement intégrable sur \mathbb{R} et telle que

$$I = \int_{-\infty}^{+\infty} |f(t)| dt$$

converge.

On suppose que f satisfait aux conditions de Dirichlet et admet un développement en série de Fourier.

L'intégrale de Fourier (voir la démonstration [2]) s'écrit :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(t) e^{-2\pi i \nu t} dt \right] e^{2\pi i \nu x} d\nu \quad (2.8)$$

Soit

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\nu) e^{2\pi i \nu t} d\nu \quad \text{avec} \quad f(\nu) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i \nu t} dt \quad (2.9)$$

La fonction $f(\nu)$ est appelé transformée de Fourier et $f(t)$ est la représentation intégrale de Fourier de la série.

2.3 Analyse spectrale des processus stationnaire**2.3.1 Densité spectrale d'un processus aléatoire stationnaire**

L'analyse spectrale s'agit de la vision "fréquence" des séries temporelles pour un processus aléatoire, le spectre désigne l'inverse de la transformée de Fourier des coefficients d'autocorrélations.

Définition 2.8. Soit $(y_t, t \in \mathbb{Z})$ un processus stationnaire de fonction d'autocovariance $\gamma(h)$. On appelle densité spectrale de y la transformée de Fourier discrète, f , de la suite des autocovariances (lorsqu'elle existe) :

$$f(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \gamma(h) e^{-ih\omega} \quad (2.10)$$

Cette densité spectrale existe lorsque :

$$\sum_{h=-\infty}^{+\infty} |\gamma(h)| < +\infty \quad (2.11)$$

Si elle existe elle est unique, continue, positive, paire et 2π -périodique.

Théorème 2.2. (théorème spectral) Si f est la densité spectrale de y alors :

$$\gamma(h) = \int_{-\pi}^{\pi} f(\omega) e^{ih\omega} d\omega \quad (2.12)$$

2.3.2 Distribution spectrale d'un processus aléatoire stationnaire

La distribution spectrale d'un processus aléatoire stationnaire est la fonction $D(\omega)$ définie sur l'intervalle $[-\pi, \pi]$, [8] telle que :

$$F(\omega) = \int_{-\pi}^{\omega} f(u) du \quad (2.13)$$

(La densité spectrale est la dérivée de la distribution spectrale).

2.3.3 Densité spectrale du processus bruit blanc

La densité spectrale d'un processus bruit blanc est constante et vaut [19] :

$$\forall \omega \in [-\pi, \pi], \quad f_{\varepsilon}(\omega) = \frac{\sigma^2}{2\pi} \quad (2.14)$$

Cette propriété est à l'origine de la terminologie bruit blanc qui provient de l'analogie avec le spectre de la lumière blanche constant dans toute la bande de fréquences visibles.

2.3.4 Densité spectrale du processus ARMA(p, q)

De façon plus générale, la densité spectrale d'un processus ARMA(p, q) est égale au rapport de deux polynômes trigonométriques, on la qualifie de densité spectrale rationnelle.

La densité spectrale d'un processus ARMA(p, q) est donné par [19] :

$$\forall \omega \in [-\pi, \pi], \quad f_y(\omega) = \frac{\sigma^2 |\Theta(e^{-i\omega})|^2}{2\pi |\Phi(e^{-i\omega})|^2} \quad (2.15)$$

Bien que la densité spectrale soit définie pour des processus aléatoires stationnaires, ses estimateurs sont utilisables pour l'analyse des séries chronologiques en général.

2.3.5 Représentation de Wold

Nous avons vu que la densité spectrale d'un processus stationnaire est la transformée de Fourier de la fonction de covariance. Or, cette fonction possède des caractéristiques spécifiques selon la nature des processus ARMA(p, q). Il en est donc de même pour la densité spectrale des processus ARMA(p, q) [47].

Théorème 2.3. *Théorème de Wold*

Tout processus stationnaire du second ordre peut être représenté sous la forme :

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + m. \quad (2.16)$$

Où les paramètres ψ_j satisfont $\psi_0 = 1, \psi_j \in \mathbb{R}, \forall j \in \mathbb{N}^*, \sum_{j=0}^{\infty} \psi_j^2 < \infty$ et ε_t est un bruit blanc i.i.d. ($0, \sigma_\varepsilon^2$)

La somme des chocs passés correspond à la composante linéaire stochastique de y_t et le terme m désigne la moyenne du processus.

Ainsi, tout processus stationnaire peut s'écrire comme une somme pondérée infinie de chocs passés, ces chocs étant représentés par un bruit blanc de variance finie.

La condition $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ est très importante. Elle assure l'existence des moments d'ordre 2.

2.4 Périodogramme

Le périodogramme a été introduit par Sir Arthur Schuster (1898) pour étudier les périodes cachées apparaissant dans la série de tâches solaires. L'analyse spectrale des séries temporelles s'est ensuite considérablement développée avec l'apparition de moyens de calculs performants, et la découverte d'algorithmes de transformée de Fourier rapides (voir Brillinger, 1981). Ce dernier joue un rôle fondamental pour l'estimation de la densité spectrale.

Définition 2.9. Soit $(y_t, t \in \mathbb{Z})$ un processus stationnaire. Si on dispose de (y_1, \dots, y_T) , on appelle périodogramme, la fonction I_T

$$I_T(\omega) = \frac{1}{T} \left| \sum_{t=1}^T y_t e^{-it\omega} \right|^2 \quad (2.17)$$

Si le processus $(y_t, t \in \mathbb{Z})$ admet une densité spectrale alors $\frac{1}{2\pi} I_T(\omega)$ est un estimateur sans biais de la densité spectrale (mais non convergent). On peut résoudre le problème de convergence en lissant le périodogramme et en considérant l'estimateur pondéré suivant :

$$\hat{f}(\omega) = \frac{1}{2\pi} \sum_{j \leq m_T} W_T(j) I_T(g(T, \omega) + \frac{2\pi j}{T}) \quad (2.18)$$

Où

ω_T représente une fonction poids tel que : $\forall j \in \mathbb{Z}, W_T(j) \geq 0, W_T(-j) = W_T(j)$,

et $\sum_{j \leq m_T} W_T(j) = 1$ et $\sum_{j \leq m_T} W_T^2(j) \rightarrow 0$;

$g(T, \omega)$ est le multiple de $\frac{2\pi}{T}$ le plus proche de ω ;

On a alors pour $T \rightarrow +\infty$, $m_T \rightarrow +\infty$ et $\frac{m_T}{T} \rightarrow 0$.

3

Processus de longue mémoire

La présence de «mémoire longue» dans les séries économiques et financières constitue un enjeu tant théorique qu'empirique, qui a donné naissance à de nombreux travaux¹. La mémoire longue décrit une série possédant des autocorrélations non nulles à long terme : dans ce cas, des observations très espacées dans le temps afficheront tout de même une dépendance ; autrement dit un choc aura des répercussions à long terme. Les processus fractionnaires exhibent une telle caractéristique puisque leurs autocorrélations décroissent seulement à une vitesse hyperbolique.

Les processus de longues mémoires stationnaires ont été largement étudiés au cours des dernières décennies. Nous abordons dans ce chapitre, une extension importante dans l'analyse des processus aléatoires. La recherche de mémoire longue concerne les modèles ARIMA(p, d, q) fractionnaire, noté FARIMA(p, d, q) pour lequel d , le degré de différenciation n'est pas un entier.

3.1 L'opérateur de dérivation fractionnaire

En convenant que l'opérateur $\nabla = 1 - L$ correspond à la différenciation, nous assimilerons aisément l'opérateur ∇^d à une différenciation fractionnaire pour $d \in \mathbb{R}$. Ainsi,

Définition 3.1. (*OPERATEUR DE DERIVATION FRACTIONNAIRE*)

$$\begin{aligned}(1 - L)^d &= \sum_{j=0}^{\infty} C_j^d (-L)^j \\ &= \sum_{j=0}^{\infty} \frac{d(d-1)\dots(d-j+1)}{j!} (-1)^j (L)^j \\ &= \sum_{j=0}^{\infty} \pi_j L^j\end{aligned}\tag{3.1}$$

¹Hurst(1951) en hydrologie ; Mandelbrot(1962), Lo(1991), Willinger, Taqqu et teverovsky(1999) en finance ; Wagennakers, Farrel et Ratcliff(2004) en psychologie...etc.

avec

$$\frac{d(d-1)\dots(d-j+1)}{j!}(-1)^j = \frac{(j-d-1)\dots(1-d)(-d)}{j!}$$

et

$$\pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} \quad \forall j \in \mathbb{Z} \quad (3.2)$$

où $\Gamma(\cdot)$ est la fonction gamma définie sur \mathbb{R} (voir [27]) par :

$$\Gamma(x) = \begin{cases} \int_0^\infty t^{x-1} e^{-t} dt, & \text{si } x > 0; \\ \infty, & \text{si } x = 0; \\ \frac{\Gamma(1+x)}{x}, & \text{si } x < 0. \end{cases} \quad (3.3)$$

On a : $\forall x \neq 0, \quad x\Gamma(x) = \Gamma(x+1)$

et $x = n, \Gamma(n) = (n-1)!$

Remarque 3.1. Si $d \in \mathbb{N}$, $(1-L)^d = \sum_{j=0}^d C_j^d (-L)^j$ est dit binôme de Newton.

3.2 Processus de longue mémoire

Les processus de longue mémoire dépendent des valeurs passées depuis un temps très long et que par conséquent, pour modéliser ces processus, l'information passée doit être intégrée dans la description du processus. La présence de mémoire longue dans une série temporelle, renvoie à des conséquences durables mais non permanentes d'un choc sur la dite série. Le processus de longue mémoire peut être défini de façon équivalente dans le domaine temporel et dans le domaine spectral.

a) Le domaine temporel [26]

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 exponentiellement.

Un processus à mémoire longue répond à la définition suivante :

Un processus stationnaire $(y_t, t \in \mathbb{Z})$ est à mémoire longue s'il existe un nombre réel α , avec $0 < \alpha < 1$ et une constante $c, c > 0$, vérifiant :

$$\lim_{h \rightarrow \infty} \frac{\rho_h}{ch^\alpha} = 1 \quad (3.4)$$

Où h est le nombre de retard Par conséquent, les autocorrélations² d'un processus à mémoire longue vérifient la relation asymptotique suivante :

$$\rho_h \simeq ch^{-\alpha}, \quad h \rightarrow \infty \quad (3.5)$$

²Les autocorrélations ϕ_h décroissent très lentement, c'est à dire a un taux hyperbolique.

b) Le domaine fréquentiel [26]

Dans le domaine fréquentiel les processus de longue mémoire sont caractérisés par une densité spectrale s'accroissant sans limite quand la fréquence tend vers zéro, formellement on retiendra la définition suivante :

Un processus stationnaire $(y_t, t \in \mathbb{Z})$ est à mémoire longue s'il existe un nombre réel β , avec $0 < \beta < 1$, et une constante $c', c' > 0$, vérifiant :

$$\lim_{\omega \rightarrow 0} \frac{f(\omega)}{c'|\omega|^{-\beta}} = 1 \tag{3.6}$$

Où $f(\omega)$ est la densité spectrale du processus $(y_t, t \in \mathbb{Z})$ à la fréquence (ω) , voir (2.10)

On en déduit immédiatement que $f(\omega) \sim c'|\omega|^{-\beta}$ quand $\omega \rightarrow 0$. Ainsi la densité spectrale exhibe un pôle à la fréquence zéro.

De nombreuses études ont montré l'importance des phénomènes à mémoire longue dans des domaines très variés (économie, météorologie,ect...). Dans la littérature statistique, ce type de phénomène se trouve dans la classe des processus FARIMA(p, d, q).

3.2.1 Le modèle autoregressif moyenne mobile fractionnaire (FARIMA)

Les modèles autoregressif, moyenne mobile fractionnairement intégrés (FARIMA) sont des processus qui ont été développés par GRANGER et JOYEUX [1980] et HOSKING [1981] et constituent une généralisation des processus ARIMA de BOX et JENKINS dont lequell'exposant de différenciation d n'est pas forcément entier. Une série fractionnairement intégrée a pour caractéristique une dépendance entre des observations éloignées comme on peut le voir dans la fonction d'autocovariance ou dans la fonction de densité spectrale.

Définition 3.2. *On dit que $(y_t, t \in \mathbb{Z})$ est un processus FARIMA(p, d, q) s'il vérifie l'équation stochastique suivante :*

$$\Phi(L)_p(1 - L)^d y_t = \Theta_q(L)\varepsilon_t \tag{3.7}$$

où

1. Φ et Θ sont respectivement des polynômes de degré p, q ;
2. ε_t est une suite de variable aléatoire i.i.d $(0, \sigma_\varepsilon^2)$;
3. d est un nombre fractionnaire ;
4. $(1 - L)^d$ porte le nom d'opérateur de dérivation fractionnaire et s'écrit à partir du développement en série donnée dans la formule (3.1) :

$$(1 - L)^d = \sum_{j=0}^{\infty} \pi_j L^j$$

Définition 3.3. : Inversibilité

Un processus stochastique $(y_t, t \in \mathbb{Z})$ est inversible si les erreurs s'écrivent en fonction des innovations passées [51] .

$$\varepsilon_t = \sum_{j \geq 0} \pi_j y_{t-j} \tag{3.8}$$

Définition 3.4. : Causalité

Un processus stochastique est $(y_t, t \in \mathbb{Z})$ causal si les innovations s'écrivent en fonction des erreurs passées [51] :

$$y_t = \sum_{j \geq 0} \psi_j \varepsilon_{t-j}, \quad (3.9)$$

Proposition 7. : En notant par $f(\omega)$, $\gamma(h)$, $\rho(h)$, la densité spectrale, la fonction de covariance et la fonction d'autocorrélation respectivement, alors on a :

1. La densité spectrale [44]

La fonction de densité spectrale du processus FARIMA(p, d, q) est donnée d'une manière générale par :

$$f(\omega) = \frac{\sigma_\varepsilon^2 |\Theta(e^{-i\omega})|^2}{2\pi |\Phi(e^{-i\omega})|^2} |1 - e^{i\omega}|^{2d} \text{ pour } -\pi < \omega < \pi \quad (3.10)$$

Cette fonction est infiniment différentiable pour toutes les fréquences sur l'intervalle $[-\pi, \pi]$, sauf à la fréquence nulle et pour $d > 0$, elle est discontinue et non bornée.

2. La fonction d'autocovariance

La fonction de covariance $\gamma(h)$ du processus FARIMA(p, d, q) est telle que :

$$\gamma(h) = c_\gamma(d, \phi, \theta) |h|^{2d-1}, \quad \text{lorsque } |h| \rightarrow \infty \quad (3.11)$$

où

$$c_\gamma(d, \phi, \theta) = \frac{\sigma_\varepsilon^2 |\theta(1)|^2}{\pi |\phi(1)|^2} \Gamma(1 - 2d) \sin \pi d \quad (3.12)$$

3. La fonction d'autocorrélation

La fonction d'autocorrélation de FARIMA(p, d, q) notée $\rho(h)$ est telle que :

$$\rho(h) = c_\rho(d, \phi, \theta) |h|^{2d-1}, \quad |h| \rightarrow \infty \quad (3.13)$$

avec

$$c_\rho(d, \phi, \theta) = \frac{c_\gamma(d, \phi, \theta)}{\int f(\omega) d\omega} \quad (3.14)$$

3.3 Analyse d'une série temporelle

Avant d'appliquer une méthode d'estimation, une analyse approfondie des propriétés des séries univariées est indispensable, cela a pour principale objectif de révéler la (non) stationnarité des séries. il existe plusieurs tests de stationnarité, on peut citer : tests de Dickey-fuller simple et Dickey-fuller augmenté, test de Phillippe-Perron, test de Schmidlt et Shin (KPSS)...ect [8].

Dans notre étude nous nous limiterons au test le plus courant, à savoir celui de Dickey-fuller (Dickey-fuller simple et Dickey-fuller augmenté) .

3.3.1 Dickey-Fuller simple

Le test de Dickey-Fuller permet de détecter la nature de la série (stationnaire ou non stationnaire) [28]. Les hypothèses du test sont données par :

– H_0 : processus non stationnaire, il correspond à une de ces formes de non stationnarité ;

1.

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \quad (3.15)$$

2.

$$y_t = \phi_1 y_{t-1} + c + \varepsilon_t \quad (3.16)$$

3.

$$y_t = \phi_1 y_{t-1} + bt + c + \varepsilon_t \quad (3.17)$$

– H_1 : processus stationnaire, $|\phi_1| < 1$.

Où $\phi_1 = 1$, $\phi_1 - 1 = 0$ et $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$

La statistique de décision est donnée par :

$$t_{\hat{\phi}_1} = \frac{\hat{\phi}_1 - 1}{\hat{\sigma}_{\hat{\phi}_1}} \quad (3.18)$$

On commence par étudier le modèle général (3.17), on regarde si le paramètre b est significativement différent de 0 ou non, si b n'est pas différent de 0, on passe à l'étude du modèle (3.16), et on cherche à savoir si c est significativement différent de 0 ou non, si c n'est pas différent de 0, on passe à l'étude du modèle (3.15).

Les valeurs calculées sont à comparer avec les valeurs tabulées de la table de Dickey-Fuller.

La région critique est :

$$t_{\hat{\phi}_1} = \frac{\hat{\phi}_1 - 1}{\hat{\sigma}_{\hat{\phi}_1}} > t_{DF}$$

Où t_{DF} désigne la valeur tabulée.

3.3.2 Dickey-Fuller augmenté

Dans le test de Dickey-Fuller simple que nous venons d'étudier, le processus ε_t est par hypothèse un bruit blanc. Or il n'y a aucune raison pour que, a priori, les erreurs soient corrélées. Le test de Dickey-Fuller augmenté ne suppose pas que ε_t soit un bruit blanc [17].

Les hypothèses du test de Dickey-Fuller augmenté se définissent de la façon suivante :

1. H_0 : processus non stationnaire, il correspond à une de ces formes de non stationnarité ;

$$\mathbf{A} : \Delta y_t = \rho_1 y_{t-1} - \sum_{k=2}^p \gamma_k y_{t-k+1} + \zeta_t$$

$$\mathbf{B} : \Delta y_t = \rho_1 y_{t-1} - \sum_{k=2}^p \gamma_k y_{t-k+1} + c + \zeta_t$$

$$\mathbf{C} : \Delta y_t = \rho_1 y_{t-1} - \sum_{k=2}^p \gamma_k y_{t-k+1} + bt + c + \zeta_t$$

2. H_1 : processus stationnaire, $\rho < 0$.

$$\text{Avec } \rho = (\phi_1 - 1)(1 - \theta_1 - \theta_2 - \dots - \theta_{p-1})$$

Et ρ le nombre de retards à ajouter dans la régression afin de prendre en compte l'autocorrélation et donc de blanchir³ les résidus. On parle de corrélation paramétrique de l'autocorrélation. Pour le choix de p à introduire dans la régression, on peut se baser sur l'autocorrélogramme partiel $(1 - L)y_t$.

La statistique du test est donnée par :

$$t_{\hat{\rho}} = \frac{\hat{\rho}_n - 1}{\widehat{\sigma}_n} \quad (3.19)$$

- Si la valeur empirique de la statistique du test est inférieure à la valeur critique de la table, alors on rejette l'hypothèse nulle de non stationnarité,
- Si la valeur empirique de la statistique du test est supérieure à la valeur critique de la table, alors on accepte l'hypothèse nulle de non stationnarité c'est-à-dire qu'il y'a présence d'une racine unitaire.

Après avoir appliqué les tests de non stationnarité sur chaque modèle (A, B, C) , on peut passé a l'analyse de détection de mémoire longue, les deux approches les plus utilisées sont : la statistique de R/S (Rescaled Range) et la statistique de Lo.

3.4 les méthodes de détection de mémoire longue

Pour pouvoir détecter la dépendance de long terme Mandelbrot suggère l'utilisation de la statistique R/S appelée aussi (Rescaled Range) qui a été développée initialement par Hurst (1951) et a été amélioré par Lo (1991) [3].

Notons que l'objectif initial de l'auteur était de modéliser la série temporelle de la hauteur des crues du Nil, de l'antiquité à nos jours. Selon cette méthode, une série chaotique peut être caractérisée par un exposant (noté H), qui représente la probabilité pour qu'un événement soit suivi par un événement similaire. C'est donc les aspects de persistance qui sont principalement visés par cette analyse. Le mode de calcul de l'exposant de Hurst n'est pas encore fixé, et certaines différences peuvent apparaître en fonction des détails des algorithmes utilisés [14].

3.4.1 La statistique R/S (exposant de Hurst)

La statistique R/S a été présenté en 1951 dans une étude sur les débits du Nil par l'hydrologue Harold Edwin Hurst. Son but est de rechercher l'intensité d'une composante cyclique apériodique dans une chronique qui est un des aspects de la dépendance à long terme (longue mémoire) développé par Mandelbrot. [6]

Soit une réalisation (y_1, \dots, y_T) d'un processus aléatoire stationnaire $(y_t, t \in \mathbb{Z})$, avec $t = 1, \dots, T$ de moyenne

$$\bar{y}_t = \frac{1}{T} \sum_{i=1}^T y_i \quad (3.20)$$

³Les residus suivent un bruit blanc $(0, \sigma^2)$

et de variance

$$S^2(t) = \frac{1}{T} \sum_{i=1}^T (y_i - \bar{y}_t)^2 \quad (3.21)$$

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) notée ici Q_T , est donnée par :

$$Q_T = \frac{R}{S_t} \quad (3.22)$$

Avec, l'étendue R qui s'écrit comme suit :

$$R = \max \sum_{j=1}^k (y_j - \bar{y}_T) - \min \sum_{j=1}^k (y_j - \bar{y}_T)$$

Et l'écart type de la série de type

$$S_t = \sqrt{\frac{1}{T} \sum_{j=1}^k (y_j - \bar{y}_T)^2}$$

En pratique cette méthode se fait en plusieurs étapes :

- Tout d'abord, on détermine une suite des entiers $(k_i)_{i=1, \dots, m}$ de longueur m, choisie arbitrairement, telle que $1 < k_m < \dots < k_1 < n$, pour laquelle on utilise la suite définie par Davies et Harte(1987) telle que pour $i = 1, 2, 3, \dots, 6$, $k_i = \lfloor \frac{n}{i} \rfloor$ et pour $i = 7, 8, 9, \dots, m$, $k_i = \lfloor \frac{k_{i-1}}{1.15^i} \rfloor$,
- Ensuite, pour chaque k_i , on détermine la statistique Q_{k_i} ,
- Puis, on détermine les logarithmes de Q_{k_i} ainsi que les logarithmes de k_i , et on trace $\log Q_{k_i}$ contre $\log k_i$,
- Enfin, on trace une droite d'équation :

$$\log Q_{k_i} = a + b \log k_i + \mu$$

et on applique la méthode des moindres carré ordinaires ce qui donne les estimateurs \hat{a} et \hat{b} et par la suite on peut déterminer l'estimation du coefficient de Hurst tel que $\hat{H} = \hat{b}$.

3.4.2 L'analyse Lo

Parmi les inconvénients de la statistique R/S proposé par Hurst, on peut citer sa sensibilité à la présence de mémoire courte. Pour régler ce problème, Lo (1991) a proposé une autre statistique, appelé "la statistique R/S modifiée".

Sa distribution asymptotique est invariante aux différentes formes des processus à mémoire courte. Cette méthode proposée par Lo permet de tester :

- H_0 : Absence de dépendance de long terme.
- H_1 : Dépendance de long terme.

La statistique R/S modifiée, notée $\bar{Q}_q(n)$, possède la forme suivante :

$$\bar{Q}_q(n) = \frac{1}{\sqrt{n}} \frac{R(n)}{S_q(n)} \quad (3.23)$$

Où $S_q(n)$ est exprimé par :

$$S_q(n) = s_n^2 + \frac{2}{n} \sum_{j=1}^q \omega_j(q) \left[\sum_{i=j+1}^n (x_i - \bar{x}_n)(x_{i-j} - \bar{x}_n) \right]^{1/2} \quad (3.24)$$

\bar{x}_n et s_n^2 sont respectivement la moyenne empirique et la variance qui sont définies par les relations (3.20) et (3.21).

$\omega_j(q) = 1 - \frac{j}{q+1}$, sont les poids qui sont proposés par Newey et West (1987), avec $j = 1, \dots, q$

Remarque 3.2. *Phillips (1987) a montré sous forme d'un théorème la convergence de l'estimateur $S_q(n)$ sous les deux conditions suivantes [45] :*

1. $\sup_t \mathbf{E}[|\varepsilon_t|^{2B}] < \infty$, pour $B > 2$
2. Lorsque $n \rightarrow \infty$, alors $q \rightarrow \infty$ tel que $q \sim 0(n^{1/4})$

Remarque 3.3. *Dans la pratique, le choix de l'entier q représente un vrai problème. Lo, MacKinlay (1989) et Andrews (1991) ont montré à partir des études faites par monté carlo que lorsque q est relativement grand par rapport à la taille de l'échantillon ; l'estimateur est biaisé et donc q doit être choisi comme un petit entier, alors que d'autres études faites par monté carlo ont montré que $q = 1$ est un choix acceptable.*

Alors, selon les valeurs de H , on peut établir une classification des séries temporelles en fonction de leurs structure de dépendances [40] :

1. Si $H = 1/2$, les autocorrélations sont nulles et la densité spectrale est constante et positive, le processus ne présente donc aucune dépendance à long terme.
2. Si $1/2 < H < 1$, les autocorrélations sont toutes positives et diminuent hyperboliquement vers zéro, la densité spectrale exhibe un pôle à la fréquence zéro. La série présente des cycles non périodiques de tous ordres. Les basses fréquences sont très importantes et les cycles (non périodiques) lent deviennent de plus en plus marqués. Le processus présente une forme persistante de mémoire longue.
3. Si $0 < H < 1/2$, les autocorrélation alternent de signe et la densité spectrale, nulle en zéro, est dominée par les composantes de haute fréquence. Le processus est anti-persistant ; des phases de hausse ont tendance à être suivies par des phases de baisse.

3.5 Test de bruit blanc

Le test de Box-Pierce permet d'identifier les processus de bruit blanc (i.e. les processus aléatoires i.i.d. et non autocorrélé. Autrement dit, il convient de tester que $cov(\varepsilon_t, \varepsilon_{t-h}) = 0$ pour tout h , soit $\rho(h) = 0$ pour tout h . ce test s'écrit [12], [17] :

$H_0 : \rho(1) = \rho(2) = \dots = \rho(h) = 0.$

$H_1 : \exists i \text{ tel que } \rho(i) \neq 0.$

Pour effectuer ce test, on utilise la statistique de Box et Pierce (1970) Q , donnée par :

$$Q = T \sum_{i=1}^h \widehat{\rho}_i^2 \quad (3.25)$$

Où T est le nombre d'observations et $\widehat{\rho}_i$ l'autocorrélation empirique. Asymptotiquement, sous H_0 , Q suit une Khi-deux à $(k - p - q)$ degrés de liberté. il est donc possible de construire un test qui consiste à rejeter l'hypothèse nulle (la série est un bruit blanc) si $Q > X_{k-p-q, \alpha}^2$.

Ljung et Box (1978) ont amélioré ce test pour les échantillons de grandes tailles.

Considérant la statistique de décision suivante :

$$Q' = T(T + 2) \sum_{i=1}^h \frac{\widehat{\rho}_i^2}{T - h} \quad (3.26)$$

Dont la distribution est mieux approximée que la précédente par une loi de Khi-deux à k degré de liberté. Ces tests sont appelés par les anglo-saxons 'portmanteau tests', soit littéralement tests 'fourre-tout'.

3.6 Test de Student

Le test de student nous permet de tester si les paramètres du modèle (3.7) estimés sont significativement différent de 0. Si un coefficient est significativement égale à 0, on élimine le terme correspondant dans la formulation du modèle. On test alors les hypothèses suivantes :

$H_0 : \theta = 0.$

$H_1 : \theta \neq 0.$

La statistique de décision du test de student est donnée par :

$$T_\theta = \frac{|\widehat{\theta}|}{\widehat{\sigma}(\widehat{\theta})} \quad (3.27)$$

T_θ suit une loi de Student à $(T - k)$ degré de liberté.

Où T représente la taille de la série considérée et k le nombre de paramètre estimés .

On note par $\theta \in \mathbb{R}^k$ le paramètre du modèle à estimé et $\widehat{\theta}$ son estimateur et $\widehat{\sigma}(\widehat{\theta})$ est l'estimateur de l'écart type associé à l'estimateur $\widehat{\theta}$ de θ .

- Si $T_\theta < t_{T-k, \alpha/2}$ on accepte l'hypothèse H_0 , le coefficient n'est pas significatif ;
- Si $T_\theta > t_{T-k, \alpha/2}$ on rejette l'hypothèse H_0 , le coefficient est significatif.

4

Caractéristiques d'un processus à mémoire longue pure

Dans ce chapitre nous étudions un cas particulier des modèles FARIMA(p, d, q), le modèle FARIMA($0, d, 0$) appelé processus à mémoire longue pure, on va s'intéresser aux différentes caractéristiques de ce processus.

4.1 Processus fractionnaire pure FI(d)

Le processus FI(d) est le processus le plus simple des processus FARIMA(p, d, q). C'est un bruit gaussien fractionnaire d'ordre $p = 0$ et $q = 0$ (FARIMA($0, d, 0$)) défini par l'équation stochastique suivante :

$$(1 - L)^d y_t = \varepsilon_t \quad \text{avec} \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \quad (4.1)$$

$(1 - z)^d$ est un polynôme qu'on peut développer selon la formule (3.1).

4.1.1 Conditions de causalité, d'inversibilité et stationnarité

Définition 4.1. Causalité [23]

Un processus stochastique $(y_t, t \in \mathbb{Z})$ est dit causal si les innovations s'écrivent en fonction des erreurs passées :

$$y_t = \sum_{j \geq 0} \psi_j \varepsilon_{t-j}, \quad (4.2)$$

Avec

$$\psi_j = \frac{\Gamma(d + j)}{\Gamma(d)\Gamma(j + 1)}$$

Propriété 1.

$d < 0$, alors $\sum_{j \geq 0} |\psi_j| < \infty$
 $0 < d < 1/2$, alors $\sum_{j \geq 0} |\psi_j| = \infty$ et $\sum_{j \geq 0} |\psi_j|^2 < \infty$
 $d \geq 1/2$, alors $\sum_{j \geq 0} |\psi_j| = \infty$ et $\sum_{j \geq 0} |\psi_j|^2 = \infty$

Proposition 8. (Gonçalves 1987) [23]

Le processus stationnaire satisfaisant l'équation (4.1) est causal si et seulement si $d < 1/2$.

4.1.1.1 Notion de filtre de moyenne mobile

$$y_t = \sum_{j \geq 0} \psi_j \varepsilon_{t-j}, \quad (4.3)$$

sur le processus $(y_t, t \in \mathbb{Z})$ au cours du temps se représente commodément à l'aide de l'opérateur retard donné dans la formule (1.1) $Ly_t = y_{t-1}$

$$y_t = \sum_{j \geq 0} \psi_j L^j \varepsilon_t = A(L) \varepsilon_t \quad (4.4)$$

Où A désigne le polynôme $A(L) = \sum_{j \geq 0} \psi_j L^j$

Notons $U = [-\pi, \pi[$, la fonction $\mathbf{R} : U \rightarrow \mathbb{C}$, définie par :

$$\mathbf{R}(\omega) = \sum_{j \geq 0} \psi_j e^{-ij\omega} = A(e^{-i\omega}) \quad (4.5)$$

est appelée fonction réponse du filtre moyenne mobile.

Définition 4.2. Inversibilité [23]

Un processus stochastique $(y_t, t \in \mathbb{Z})$ est dit inversible si les erreurs peuvent s'écrire en fonction des innovations passées :

$$\varepsilon_t = \sum_{j \geq 0} \pi_j y_{t-j} \quad (4.6)$$

Avec

$$\pi_j = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}$$

Voir les formules (3.1) et (3.2).

Propriété 2.

$d > 0$, alors $\sum_{j \geq 0} |\pi_j| < \infty$
 $-1/2 < d < 0$, alors $\sum_{j \geq 0} |\pi_j| = \infty$ et $\sum_{j \geq 0} |\pi_j|^2 < \infty$
 $d \leq 1/2$, alors $\sum_{j \geq 0} |\pi_j| = \infty$ et $\sum_{j \geq 0} |\pi_j|^2 = \infty$

Proposition 9. (Gonçalves 1987)

Le processus défini satisfaisant l'équation (4.1) est inversible si et seulement si $d > -1/2$.

Remarque 4.1. En utilisant la formule de shepard suivante :

$$\frac{\Gamma(j+a)}{\Gamma(j+b)} \approx j^{a-b}, \quad \text{si } j \text{ tend vers } \infty, \quad (4.7)$$

de la fonction gamma on obtient les valeurs asymptotiques des coefficients ψ_j et π_j sont :

$$\lim_{j \rightarrow +\infty} \psi_j \approx \frac{j^{d-1}}{\Gamma(d)}$$

et

$$\lim_{j \rightarrow +\infty} \pi_j \approx \frac{j^{-d-1}}{\Gamma(-d)}$$

Les coefficients ψ_j et π_j sont décroissants mais à une vitesse hyperbolique qui est plus faible que la vitesse exponentielle des processus ARMA. La fonction d'autocorrélation possède ce même type de comportement ce qui permet de caractériser les processus FI(d).

Proposition 10. Le processus défini par l'équation (4.1) est inversible et causal si et seulement si $-\frac{1}{2} < d < \frac{1}{2}$

4.1.2 Fonction d'autocovariance

Soit un processus $(y_t, t \in \mathbb{Z})$, sa fonction d'autocovariance notée $\gamma(h)$ pour $h \in \mathbb{N}$ est :

$$\begin{aligned} \gamma(h) &= \mathbb{E}[y_{t+h} - \mathbb{E}(y_{t+h})][y_t - \mathbb{E}(y_t)] \\ &= \mathbb{E}(y_{t+h} \cdot y_t) \\ &= \mathbb{E}\left(\left[\sum_{j \geq 0} \psi_j \varepsilon_{t-j}\right] \left[\sum_{j \geq 0} \psi_j \varepsilon_{t+h-j}\right]\right) \\ &= \mathbb{E}\left(\sum_{j \geq 0} \psi_j^2 \varepsilon_{t-j} \varepsilon_{t+h-j}\right) + \mathbb{E}\left(\sum_{j \neq i} \psi_j \psi_i \varepsilon_{t-j} \varepsilon_{t+h-i}\right) \\ &= \sum_{j \geq 0} \psi_j^2 \mathbb{E}(\varepsilon_{t-j} \varepsilon_{t+h-j}) + \sum_{j \neq i} \psi_j \psi_i \mathbb{E}(\varepsilon_{t-j} \varepsilon_{t+h-i}) \end{aligned} \quad (4.8)$$

Pour $h \neq 0$

$$\gamma(h) = \sigma^2 \sum_{j \geq 0} \psi_j \psi_{j+h} \quad (4.9)$$

Pour $h = 0$

$$\begin{aligned} \gamma(h) &= \mathbb{E}\left(\left[\sum_{j > 0} \psi_j \varepsilon_{t-j}\right] \left[\sum_{j \geq 0} \psi_j \varepsilon_{t-j}\right]\right) \\ &= \mathbb{E}\left(\sum_{j > 0} \psi_j^2 \varepsilon_{t-j}^2\right) + \sum_{j \neq i} \psi_j \psi_i \mathbb{E}(\varepsilon_{t-j} \varepsilon_{t-i}) \end{aligned} \quad (4.10)$$

Donc

$$\gamma(0) = \sigma^2 \sum_{j \geq 0} \psi_j^2 \quad (4.11)$$

La forme explicite des autocovariances se déduit directement de celle des coefficients ψ_j , tel que

$$\psi_j = \frac{\Gamma(d+j)}{\Gamma(d)j!}$$

En remplaçant ψ_j par sa valeur, dans la formule (4.9), on aura :

$$\begin{aligned} \gamma(h) &= \sigma^2 \sum_{j \geq 0} \frac{\Gamma(d+j)}{\Gamma(d)j!} \frac{\Gamma(d+j+h)}{\Gamma(d)(j+h)!} \\ &= \left[\frac{\sigma^2 \Gamma(h+d)}{\Gamma(h+1)\Gamma(d)} \right] \cdot \left[\frac{\Gamma(h+1)}{\Gamma(h+d)\Gamma(d)} \sum_{j \geq 0} \frac{\Gamma(d+h)\Gamma(d+h+j)}{\Gamma(h+1+j)j!} \right] \end{aligned} \quad (4.12)$$

On reconnaît dans le second terme entre crochet une valeur de la fonction hypergéométrique donnée par :

$$F(a, b, c, z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{j \geq 0} \frac{\Gamma(a+j)\Gamma(b+j)z^j}{\Gamma(c+j)j!} \quad (4.13)$$

Alors la fonction d'autocovariance peut s'écrire comme suit :

$$\gamma(h) = \sigma^2 \sum_{j \geq 0} \frac{\Gamma(d+h)}{\Gamma(h+1)\Gamma(d)} F(d, h+d, h+1, 1) \quad \forall h \geq 0 \quad (4.14)$$

En utilisant la propriété de la fonction hypergéométrique suivante :

$$F(a, b, c, 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \quad (4.15)$$

On obtient alors

$$F(d, h+d, h+1, 1) = \frac{\Gamma(h+1)\Gamma(1-2d)}{\Gamma(h+1-d)\Gamma(1-d)}$$

La formule (4.14) s'écrit alors comme suit :

$$\gamma(h) = \sigma^2 \sum_{j \geq 0} \frac{\Gamma(d+h)\Gamma(1-2d)}{\Gamma(d)\Gamma(h+1-d)} \quad (4.16)$$

Pour $h = 0$

$$\gamma(0) = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma^2(1-d)} \quad (4.17)$$

En utilisant la formule de Sheppard (4.7), quand h tend vers ∞ , on a :

$$\gamma(h) \approx |h|^{2d-1} \quad (4.18)$$

Remarque 4.2. Le processus $FI(d)$ est à mémoire longue si $0 < d < \frac{1}{2}$, car dans ce cas la fonction d'autocorrélation décroît lentement vers zéro.

4.1.3 Caractéristiques de la fonction d'autocovariance :

La fonction de transfert Soit un processus $(y_t, t \in \mathbb{Z})$, admet une représentation moyenne mobile définie par la relation (4.2)

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = A(L)\varepsilon_t = (1-L)^{-d} \varepsilon_t$$

Pour $z \in \mathbb{Z}$ la fonction de transfert du processus $(y_t, t \in \mathbb{Z})$ est donnée par :

$$A(z) = \frac{\theta(z)}{\psi(z)} \quad (4.19)$$

Où $\theta(z) = 1$ et $\psi(z) = (1-z)^d$, $z \in \mathbb{C}$

Et donc

$$A(z) = \frac{1}{(1-z)^d} = (1-z)^{-d} = \sum_{j=0}^{\infty} \psi_j z^j \quad (4.20)$$

La dérivé logarithmique de la fonction de transfert

$$\begin{aligned} (\log(A(z)))' &= \frac{d(A(z))}{A(z)dz} = \frac{d(\frac{1}{\psi(z)})}{\psi(z)^{-1}dz} \\ &= \frac{-d\psi(z)}{\psi(z)^2} \frac{\psi(z)}{dz} \\ &= \frac{-d\psi(z)}{\psi(z)dz} \end{aligned} \quad (4.21)$$

On a, $\psi(z) = (1-z)^d$, alors

$$\begin{aligned} (\log(\psi(z)))' &= \frac{d(\psi(z))}{dz} \\ &= \frac{d}{dz} e^{d \log(1-z)} = \frac{-d}{1-z} (1-z)^d \end{aligned} \quad (4.22)$$

En remplaçant dans la formule du $(\log(A(z)))'$, on obtient :

$$(\log(A(z)))' = \frac{d}{1-z} \frac{(1-z)^d}{(1-z)^d} = \frac{d}{(1-z)} \quad (4.23)$$

Fonction génératrice des autocovariances On définit la fonction génératrice des autocovariances par,

$$G(z) = \sum_{-\infty}^{+\infty} \gamma(h) z^h, \quad z \in \mathbb{C} \quad (4.24)$$

D'après l'expression $\gamma(h)$ donnée par la formule (4.9), on obtient :

$$\begin{aligned} G(z) &= \sigma^2 \sum_{-\infty}^{+\infty} \sum_{j \geq 0} \psi_j \psi_{j+h} z^h \\ &= \sigma^2 \sum_{j \geq 0} \psi_j z^j \left(\sum_{-\infty}^{+\infty} \psi_{j+h} z^{h-j} \right) \end{aligned} \tag{4.25}$$

On pose $k = h + j$, donc :

$$G(z) = \sigma^2 \sum_{j \geq 0} \psi_j z^{-j} \left(\sum_{k \geq 0} \psi_k z^k \right) \tag{4.26}$$

D'après la formule (4.20), on peut écrire $G(z)$ comme suit :

$$G(z) = \sigma^2 \frac{1}{\psi(\frac{1}{z})\psi(z)} = \frac{\sigma^2}{(1 - \frac{1}{z})^d (1 - z)^d} \tag{4.27}$$

Remarque 4.3. *Le processus à longue mémoire pure est caractérisé par sa fonction de transfert, la dérivé logarithmique de la fonction de transfert et la fonction génératrice des autocovariances qui s'écrivent sous forme d'une fraction rationnelle.*

4.1.4 Fonction d'autocorrélation

La fonction d'autocorrélation pour un retard $h \in \mathbb{Z}$ est donnée par :

$$\begin{aligned} \rho(h) &= \frac{\gamma(h)}{\gamma(0)} \\ &= \frac{\sigma^2 \frac{\Gamma(d+h)}{\Gamma(h+1)\Gamma(d)} F(d, h+d, h+1, 1)}{\sigma^2 \frac{\Gamma(d)}{\Gamma(1)\Gamma(d)} F(d, d, 1, 1)} \end{aligned} \tag{4.28}$$

Donc

$$\rho(h) = \frac{\Gamma(d+h)F(d, h+d, h+1, 1)}{\Gamma(d)\Gamma(h+1)F(d, d, 1, 1)} \tag{4.29}$$

On peut écrire $F(d, d, 1, 1)$ en utilisant la formule (4.15) :

$$F(d, d, 1, 1) = \frac{\Gamma(1-2d)}{\Gamma(1+d)^2}$$

Alors la fonction d'autocorrélation donnée dans la formule (4.28), peut s'écrire comme suit :

$$\rho(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(h-d+1)\Gamma(d)} \tag{4.30}$$

Notons grâce a la formule de Sheppard (4.7), on obtient :

$$\rho(h) \sim h^{2d-1} \frac{\Gamma(1-d)}{\Gamma(d)}$$

Remarque 4.4. Comme $0 < d < 1/2$, quand $h \rightarrow \infty$ alors $h^{2d-1} \rightarrow 0$, ce qui fait que le processus $FI(d)$ est caractérisé par sa fonction d'autocorrélation qui décroît hyperboliquement vers 0 ($\rho(h) \searrow 0$).

4.1.5 Fonction d'autocorrélation partielle

L'application du logarithme de Durbin-Levinson donne pour l'autocorrélation d'ordre h notée $\alpha(h)$, le résultat suivant [56] :

$$\begin{aligned}\alpha(h) &= -\frac{\Gamma(h-d)\Gamma(1-d)}{\Gamma(d)\Gamma(h-d+1)} \\ &= \frac{d}{h-d}, \quad \text{pour } h \in \mathbb{N}^*\end{aligned}\tag{4.31}$$

4.1.6 Densité spectrale

La densité spectrale d'un $FI(d)$ est donnée par [10] :

$$f(\omega) = \frac{\sigma^2}{2\pi} |1 - e^{i\omega}|^{-2d} = \frac{\sigma^2}{2\pi} |2 \sin(\frac{\omega}{2})|^{-2d} \quad -\pi < \omega < \pi\tag{4.32}$$

Démonstration 1. On démontre que

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

Alors, on a

$$f(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \gamma(h) e^{-ih\omega}$$

donnée dans la formule (2.10)

Avec $\gamma(h) = \sigma^2 \sum_{j \geq 0} \psi_j \psi_{j+h}$, quand $h \neq 0$

On peut écrire $f(\omega)$ comme suit

$$\begin{aligned}f(\omega) &= \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \gamma(h) e^{-ih\omega} \\ &= f(\omega) = \frac{\sigma^2}{2\pi} \sum_{h=-\infty}^{+\infty} \sum_{j \geq 0} \psi_j \psi_{j+h} e^{-ih\omega}\end{aligned}\tag{4.33}$$

posons $j + h = m$

$$\begin{aligned}
f(\omega) &= \frac{\sigma^2}{2\pi} \sum_{m=-\infty}^{+\infty} \sum_{j \geq 0} \psi_j \psi_m e^{-(m-j)i\omega} \\
f(\omega) &= \frac{\sigma^2}{2\pi} \sum_{m=-\infty}^{+\infty} \sum_{j \geq 0} \psi_j \psi_m e^{-(m-j)i\omega} \\
&= \frac{\sigma^2}{2\pi} \left| \sum_{j \geq 0} \psi_j e^{-ij\omega} \right|^{-2d}
\end{aligned} \tag{4.34}$$

Avec $\sum_{j \geq 0} \psi_j e^{-ij\omega} = A(e^{-i\omega}) = (1 - e^{-i\omega})^{-d}$ dit le filtre de la moyenne mobile, et donc

$$\begin{aligned}
f(\omega) &= \frac{\sigma^2}{2\pi} (1 - e^{i\omega})^{-d} (1 - e^{-i\omega})^{-d} \\
&= \frac{\sigma^2}{2\pi} |1 - e^{i\omega}|^{-2d}
\end{aligned} \tag{4.35}$$

Par suite, on démontre que

$$f(\omega) = \frac{\sigma^2}{2\pi} \left| 2 \sin\left(\frac{\omega}{2}\right) \right|^{-2d} \quad -\pi < \omega < \pi$$

$$\begin{aligned}
f(\omega) &= \frac{\sigma^2}{2\pi} |1 - e^{i\omega}|^{-2d} \\
&= \frac{\sigma^2}{2\pi} [(1 - e^{i\omega})(1 - e^{-i\omega})]^{-d} \\
&= \frac{\sigma^2}{2\pi} (1 - e^{i\omega} - e^{-i\omega} + 1)^{-d} \\
&= \frac{\sigma^2}{2\pi} (2 - (\cos \omega + i \sin \omega + \cos \omega - i \sin \omega))^{-d} \\
&= \frac{\sigma^2}{2\pi} (2 - 2 \cos \omega)^{-d} \\
&= \frac{\sigma^2}{2\pi} (2(1 - \cos^2(\frac{\omega}{2}) + \sin^2(\frac{\omega}{2})))^{-d} \\
&= \frac{\sigma^2}{2\pi} (4 \sin^2(\frac{\omega}{2}))^{-d} \\
&= \frac{\sigma^2}{2\pi} (2 \sin(\frac{\omega}{2}))^{-2d}
\end{aligned} \tag{4.36}$$

Remarque 4.5. Lorsque $-1/2 < d < 1/2$, le processus $(y_t, t \in \mathbb{Z})$, est stationnaire et inversible. Les autocorrélations décroissent hyperboliquement vers zéro et la densité spectrale n'est pas limitée à la fréquence zéro. Ces processus vérifient donc la définition générale des processus à mémoire longue donnée dans le paragraphe précédent. En outre, d décrit le comportement de long terme de la série alors que les paramètres auto-regressifs et moyenne mobile prennent en compte les fluctuations de court terme.

4.2 La relation entre l'exposant de Hurst et le paramètre d'intégration fractionnaire

En reprenant les formulations¹ des fonctions d'autocovariance du bruit gaussien fractionnaire (4.10) et des processus FARIMA (4.18), on établit une relation entre le paramètre d des processus FARIMA(p, d, q) et l'exposant de Hurst H . Ainsi, on liant les deux exposants, on peut obtenir la relation suivante [39] :

$$d = H - \frac{1}{2} \quad (4.37)$$

Dès lors, à partir de cette relation, il est possible de classifier les séries temporelles en fonction de leur structure de dépendance selon le même schéma que celui retenu dans le cas du bruit gaussien fractionnaire [56] :

1. Si $d = 0$, le processus FARIMA(p, d, q) se réduit au processus ARMA(p, q) standard ;
2. Si $0 < d < 1/2$, le processus est persistant ;
3. Si $-1/2 < d < 0$, le processus est anti-persistant.

Ainsi, tout comme l'exposant de Hurst H , le paramètre d'intégration fractionnaire d nous fournit une mesure de l'intensité de la dépendance présente dans une série temporelle.

¹GEWEKE et Parter-HUDAK (1983) ont démontré la relation théorique entre le paramètre d et l'exposant de Hurst H .

5

Estimation du paramètre de longue mémoire

Plusieurs méthodes d'estimations du paramètre de longue mémoire ont été proposées. Nous citons dans ce chapitre, la famille des méthodes dites "en une étape" et celle dite "en deux étapes". La première estime simultanément le coefficient d'intégration fractionnaire et les paramètres de la composante de mémoire courte. et la deuxième estime d'abord le coefficient d'intégration fractionnaire sur la série brute, ensuite les paramètres de la composante de court terme sur la série différenciée fractionnaire par les techniques classiques d'estimation.

Ces méthodes sont faciles à mettre en oeuvre mais pèchent par leur imprécision, particulièrement s'il existe une composante de court terme importante.

5.1 Méthodes d'estimation du processus FARIMA(p, d, q)

Nous présentons d'abord quelques méthodes de la famille "à deux étapes". En suite, la méthode de vraisemblance comme méthode "à une étape".

5.1.1 Méthodes à deux étapes

5.1.1.1 Méthode de Geweke et Porter Hudak

La technique la plus utilisée est celle de l'estimation de Geweke et Porter Hudak(1983) notée GPH. Elle se base sur l'expression de la densité spectrale du processus FARIMA(p, d, q) lorsque les fréquences tendent vers 0.

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

$$f(\omega) = 4 \sin^2(\omega/2)^{-d} * f_y(\omega) \quad (5.1)$$

où $f_y(\omega) = f_\varepsilon(\omega) \frac{|\theta(e^{-i\omega})|^2}{|\phi(e^{-i\omega})|^2} = \frac{\sigma^2}{2\pi} \frac{|\theta(e^{-i\omega})|^2}{|\phi(e^{-i\omega})|^2}$ est la densité spectrale du processus ARMA déjà définie dans la formule voir (2.15).

On a,

$$\frac{f(\omega)}{f_\varepsilon(0)} = [4 \sin^2(\omega/2)]^{-d} * \frac{f_y(\omega)}{f_\varepsilon(0)} \quad (5.2)$$

En appliquant le logarithme à la formule (5.2), on obtient :

$$\ln \frac{f(\omega)}{f_\varepsilon(0)} = \ln [4 \sin^2(\omega/2)]^{-d} * \frac{f_y(\omega)}{f_\varepsilon(0)}$$

$$\ln f(\omega) = \ln f_\varepsilon(0) - 2d \ln 4 \sin(\omega/2) + \ln \frac{f_y(\omega)}{f_\varepsilon(0)}$$

$$\ln f(\omega_j) = \ln f_\varepsilon(0) - 2d \ln 4 \sin(\omega_j/2) + \ln \frac{f_y(\omega_j)}{f_\varepsilon(0)}$$

Considérons l'estimateur de la densité spectrale du processus $(y_t, t \in \mathbb{Z})$ donné par le périodogramme

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

Ce périodogramme est évalué en fréquences de Fourier $\omega_j = \frac{2\pi j}{T}$

où $1 < j < m_T$, m_T le paramètre de la troncature¹.

On pose :

$$Y_j = \ln f(\omega_j)$$

$$X_j = \ln 4 \sin(\omega_j/2) = \ln 4 \sin(2\pi j/2T) = \log 4 \sin(\pi j/T)$$

$$e_j = \ln \frac{f_y(\omega_j)}{f_\varepsilon(0)} + \gamma \text{ où } \gamma \text{ est la constante d'euler qui est égale à } \gamma = 0.5772$$

$$a = \ln f_\varepsilon(0) - \gamma$$

$$b = -2d$$

On obtient donc l'équation linéaire suivante :

$$Y_j = a + bX_j + e_j \quad (5.3)$$

¹La troncature consiste à remplacer la somme allant à l'infini par une somme allant à $n-1+h$ (Cf. BROCKWELL et DAVIS [1991, p. 533])

Par la méthode des moindres carrés ordinaire, l'estimateur de GPH s'écrit :

$$d^{GPH} = \frac{cov(Y, X)}{\sigma_X^2} = \frac{\sum_{j=1}^{m_T} (Y_j - \bar{Y}_{m_T})(X_j - \bar{X}_{m_T})}{\sum_{j=1}^{m_T} (Y_j - \bar{Y}_{m_T})^2} \quad (5.4)$$

On peut calculer sa variance :

$$\begin{aligned} var(d^{GPH}) &= \frac{1}{\sum_{j=1}^{m_T} (Y_j - \bar{Y}_{m_T})^2} var\left(\sum_{j=1}^{m_T} (Y_j - \bar{Y}_{m_T})\alpha_j\right) \quad \text{avec } \alpha_j = (X_j - \bar{X}_{m_T}) \\ &= \frac{1}{\sum_{j=1}^{m_T} (Y_j - \bar{Y}_{m_T})^2} \sum_{j=1}^{m_T} (Y_j - \bar{Y}_{m_T}) var(\alpha_j) \end{aligned} \quad (5.5)$$

Où $\sum_{j=1}^{m_T} (Y_j - \bar{Y}_{m_T}) var(\alpha_j) = \pi^2/6$ [6],

Et

$$\begin{aligned} var(\alpha_j) &= var(\log f(\omega_j) - \bar{y}_{m_T}) \\ &= var(\log f(\omega_j)) \end{aligned} \quad (5.6)$$

Donc sa variance est donnée par :

$$var(d^{GPH}) = \frac{\pi^2}{6} \frac{1}{\sum_{j=1}^{m_T} (Y_j - \bar{Y}_{m_T})^2} \quad (5.7)$$

GPH(1983) ont montré la convergence et la normalité asymptotique pour $d < 0$ et pour $m_T = g(n)$ avec $\lim_{n \rightarrow \infty} g(n) = \infty$ et $\lim_{n \rightarrow \infty} \frac{g(n)}{n} = 0$. Porter-Hudak (1990), Crato et Lima (1994), ont montré que les fréquences m_T doivent être choisies de telle sorte que $m_T = n^\nu$, avec $\nu = 0.5, 0.6, 0.7$. [27].

$$\frac{d^{GPH} - d}{\sqrt{var(d^{GPH})}} \sim \mathcal{N}(0, 1)$$

Cet estimateur peut être calculé en utilisant des fenêtres spectrales (Chen, Abraham, Peiris(1994)) dans le périodogramme.

Malgré les bonnes propriétés de l'estimateur proposé par GPH, Newbold et Al (1991) ont montré qu'il possède un inconvénient important qui correspond à la présence d'un biais asymptotique d'un périodogramme estimé. C'est la raison qui a amené Robinson (1995) a proposé un autre estimateur sans biais pour d .

Remarque 5.1. Valderi A. Reisen (1994) a proposé un autre estimateur qui est une extension de celui de GPH. Pour déterminer son estimateur, il a suivi les mêmes étapes que celles proposées par GPH mais au lieu d'utiliser le périodogramme qui est un estimateur biaisé de la fonction de la densité spectrale, il a utilisé le périodogramme lissé².

²Voir chapitre 2

5.1.1.2 Méthodes de Robinson

Robinson a proposé deux méthodes simples (1995a,1995b) pour estimer uniquement le paramètre de mémoire longue d , sans donner aucune information sur l'estimation des autres paramètres.

a) Première méthode

Robinson (1995a) a proposé un estimateur du log périodogramme, qui est en fait l'extension de l'estimateur GPH. Pour déterminer l'expression de cet estimateur, on procède de la même façon que précédemment mais on ne prend pas les m premières fréquences qui correspondent aux basses fréquences. Cet estimateur noté par \hat{d}_{Rm} est donné par :

$$\hat{d}_{Rm} = \frac{\sum_{j=1+m}^l (Y_j - \bar{Y}) a_j}{\sum_{j=1+m}^l (Y_j - \bar{Y})^2}, \quad 0 \leq m < l \quad (5.8)$$

Avec

$$a_j = X_j - \bar{X}_{mT}$$

La non existence d'un critère de choix optimal des paramètres l et m est la cause du manque de robustesse³ de cet estimateur.

b) Deuxième méthode

Robinson(1995b) a proposé un autre estimateur semi-paramétrique gaussien pour le paramètre H^4 ; il est asymptotiquement plus efficace que l'estimateur proposé par la première méthode et celui de GPH (1983). C'est un estimateur qui ne suppose pas la normalité du processus. Pour déterminer l'expression de \hat{H} , on va tout d'abord donner quelques définitions [44] :

Soient :

$R(H)$, une fonction définit par l'expression suivante :

$$R(H) = \ln \hat{G}(H) - (2H - 1) \frac{1}{m} \sum_{j=1}^m \ln \omega_j \quad (5.9)$$

$\hat{G}(H)$ l'estimateur de $G(H)$ qui est donné par :

$$\hat{G}(H) = \frac{1}{m} \sum_{j=1}^m \ln \omega_j^{2H-1} I_j \quad (5.10)$$

Où

I_j représente le périodogramme,

$m = \frac{1}{2}n$, pour n pair et $m = \frac{1}{2}(n + 1)$, pour n impair. Dans la théorie asymptotique, ce paramètre m converge plus lentement vers ∞ que n .

L'estimateur \hat{H} est donné par :

$$\hat{H} = \arg \min_{H \in \Theta} R(H) \quad (5.11)$$

Avec $\Theta \in [\Delta_1, \Delta_2]$, Δ_1 et Δ_2 sont choisis arbitrairement entre 0 et 1 telle que $0 < \Delta_1 < \Delta_2 < 1$

³Le test reste valable alors que les hypothèses d'application ne sont pas toutes réunies

⁴Paramètre de Hurst

Dans son article, Robinson (1995b) a montré dans deux théorèmes la convergence ainsi que la normalité asymptotique de cet estimateur et sous des conditions de régularités assez forte.

5.1.2 Méthodes à une étape (Méthodes du maximum de vraisemblance)

Les méthodes du maximum de vraisemblance sont considérées parmi les méthodes d'estimation les plus importantes pour estimer le paramètre de mémoire longue d . Ce sont les méthodes les plus efficaces qui permettent d'estimer simultanément tous les paramètres. On considère deux approches la méthode du maximum de vraisemblance exacte ainsi que la méthode du maximum de vraisemblance approximée de Whittle (1951). Chacune de ces deux méthodes donne un bon estimateur. Les propriétés ont été étudiées par différents auteurs tels que Fox et Taquq (1986), Dahlhas (1989), Fallow Sowell (1992), voir [20].

5.1.2.1 Méthodes du maximum de vraisemblance exacte

Cette procédure, développée par Sowell (1992), a été utilisée par Lardic et Mignon (1995, 1996a et 1996b) sur diverses séries financières. L'intérêt de la méthode du maximum de vraisemblance exacte est qu'elle utilise toute l'information concernant le comportement de la série, a la fois à court terme et à long terme puisque sont estimés simultanément tous les paramètres de la représentation de FARIMA(0, d , 0) ou FI(d).

Soit $(y_t, t \in \mathbb{Z})$ un processus FARIMA(0, d , 0) qui vérifie les conditions suivantes :

- La moyenne μ du processus $(y_t, t \in \mathbb{Z})$ est supposée généralement connu et égale a zéro,
- Le paramètre de mémoire longue est strictement inférieur à 1/2.

La fonction log-vraisemblance est donnée par :

$$\ln(y_t, \theta) = -\frac{t}{2} \ln(2\pi) - \frac{1}{2} \ln |\sum \theta| - \frac{1}{2} y_t' \sum^{-1} \theta y_t \quad (5.12)$$

Où

\sum : représente la matrice de variance-covariance de y_t

$|\sum|$ représente le déterminant de la matrice de variance-covariance de y_t

$\theta = (\sigma_\varepsilon^2, d, \theta_1, \dots, \theta_q, \phi_1, \dots, \phi_p)$ est le vecteur des paramètres inconnus contenant la variance du bruit blanc, le paramètre d'intégration fractionnaire, les paramètres d'un processus moyenne mobile et les paramètres d'un processus autoregressif.

Dans notre étude $\theta = (\sigma_\varepsilon^2, d)$

L'estimateur du maximum de vraisemblance $\hat{\theta}$ de θ est obtenu en maximisant la fonction de log-vraisemblance conditionnellement au vecteur des paramètres θ

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \ln(y, \theta) \quad (5.13)$$

Autrement dit, l'estimateur $\hat{\theta}$ sera la solution de l'équation suivante :

$$\begin{aligned} \ln'(y_t, \theta) &= \frac{\partial}{\partial \theta_j} \ln(y_t, \theta) \\ &= -\frac{\delta}{\delta \theta_j} \frac{1}{2} \ln \left| \sum \theta \right| - \frac{1}{2} y_t' \left[\frac{\delta}{\delta \theta_j} \sum^{-1} \theta \right] y_t \end{aligned} \quad (5.14)$$

Avec $j = 1, \dots, m$

L'estimateur du maximum de vraisemblance $\hat{\theta}$ de θ est alors la solution de :

$$\ln'(y_t, \hat{\theta}) = 0$$

De nombreux auteurs se sont intéressés à l'étude des propriétés de cet estimateur dont Yajima (1985) qui a montré la convergence et la normalité asymptotique et ceci sous les conditions de régularité forte.

Par la suite Dahlhaus (1989) montre que l'estimateur du maximum de vraisemblance exacte est asymptotiquement efficace au sens de Fisher.

5.1.2.2 Méthodes du maximum de vraisemblance approximé par Whittle

La méthode proposée par Whittle permet d'estimer le vecteur des paramètres d'intérêt θ . Elle est construite à partir du périodogramme et elle nécessite la connaissance de la fonction de densité spectrale [46]. L'approximation du déterminant de la matrice variance-covariance $|\sum \theta|$ est donnée par :

$$\lim \frac{1}{T} \ln \left| \sum \theta \right| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\omega, \theta) d\omega \quad (5.15)$$

L'approximation de l'inverse de la matrice variance-covariance $\sum^{-1} \theta$ est donnée par :

$$y_t' \sum^{-1} \theta y_t = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I(\omega)}{f(\omega, \theta)} d\omega \quad (5.16)$$

La fonction log-vraisemblance peut alors s'écrire :

$$L(y_t, \theta) = -\frac{T}{2} \log(2\pi) - \frac{T}{4\pi} \int_{-\pi}^{\pi} \log f(\omega, \theta) d\omega - \frac{T}{4\pi} \int_{-\pi}^{\pi} \frac{I(\omega, \theta)}{f(\omega, \theta)} d\omega \quad (5.17)$$

L'estimateur du maximum de vraisemblance approximé de Whittle est un bon estimateur puisqu'il est convergent et asymptotiquement normal ; ces deux bonnes propriétés ont été montrées par Fox et Taqu (1986) et Dalhlaus (1989). De plus Dalhlaus (1989) a établi l'efficacité de cet estimateur.

5.2 Simulation

1. Expérience de simulation

Une étude de simulation est conçue pour étudier la consistance des méthodes d'estimation du paramètre de mémoire longue d citées dans le chapitre précédent.

Les expériences de simulation ont aussi pour but d'implémenter quelques résultats théoriques de ces estimateurs.

Nous avons générer des séries de différentes taille ($T = 30$, $T = 200$ et $T = 1000$) avec la méthode de Beran [4] des processus de longue mémoire pure pour différentes valeurs de d ($d = 0.1$, $d = 0.15$, $d = 0.2$, $d = 0.25$, $d = 0.3$, $d = 0.35$ et $d = 0.4$). Les méthodes d'estimation utilisées sont d'une part la méthode de Geweke et Porter Hudak et d'autre part la méthode de maximum de vraisemblance approximée. Les résultats obtenus sont résumés dans les tableaux suivants :

d	T=30		T=200		T=1000	
	\hat{d}	$S_{\hat{E}d}$	\hat{d}	$S_{\hat{E}d}$	\hat{d}	$S_{\hat{E}d}$
0.1	0,670	0,517	0,266	0,230	0,09	0,137
0.15	-0,123	0,517	0,031	0,230	0,179	0,137
0.20	0,007	0,517	0,173	0,230	0,224	0,137
0.25	-0.153	0,517	0,157	0,230	0,253	0,137
0.30	0,172	0,517	0,219	0,230	0,318	0,137
0.35	0,137	0,517	0,290	0,230	0,336	0,137
0.40	0,694	0,517	0,632	0,230	0,407	0,137

TAB. 5.1 – Estimation du paramètre "d" pour différentes tailles avec la méthode de Geweke et Porter Hudak (GPH).

d	T=30		T=200		T=1000	
	\hat{d}	$S_{\hat{E}d}$	\hat{d}	$S_{\hat{E}d}$	\hat{d}	$S_{\hat{E}d}$
0.1	-0.203	0,248	0,056	0,1	0,104	0,055
0.15	0,081	0,248	0,133	0,1	0,161	0,055
0.20	0,115	0,248	0,192	0,1	0,209	0,055
0.25	-0.151	0,248	0,226	0,1	0,253	0,055
0.30	0,202	0,248	0,281	0,1	0,304	0,055
0.35	0,144	0,248	0,293	0,1	0,341	0,055
0.40	0,006	0,248	0,275	0,1	0,420	0,055

TAB. 5.2 – Estimation du paramètre "d" pour différentes tailles avec la méthode de Valderio A. Reisen.

d	T=30		T=200		T=1000	
	\hat{d}	$S_{\hat{E}d}$	\hat{d}	$S_{\hat{E}d}$	\hat{d}	$S_{\hat{E}d}$
0.1	0.135	0,118	0,064	0,0450	0,123	0.0204
0.15	0,003	0,114	0,077	0,0451	0,164	0.0206
0.20	0,376	0,122	0,128	0,457	0,203	0.027
0.25	0.082	0,116	0,228	0,0466	0,255	0.0209
0.30	0,18	0,119	0,243	0,0466	0,302	0.0211
0.35	0,286	0,121	0,223	0,0467	0,309	0,0211
0.40	0,278	0,121	0,371	0,0475	0,424	0,0213

TAB. 5.3 – Estimation du paramètre "d" pour différentes tailles par la méthode du maximum de vraisemblance approximée.

2. Analyse des résultats de simulation

Afin d'appliquer quelques méthodes d'estimation du paramètre de mémoire longue, on a utilisé comme processus générateur des données, le processus FARIMA(0, d, 0) ou FI(d) suivant :

$$(1 - L)^d y_t = \varepsilon_t, \quad \text{avec } \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{5.18}$$

Le tableau 1 présente les résultats de l'application de la méthode de GPH. Les résultats obtenus montrent l'influence de la taille de l'échantillon sur l'application de cette méthode. Les meilleurs résultats sont obtenus pour un (T= 1000) car les valeurs de \hat{d} sont plus proches des vraies valeurs, de plus, ils possèdent de faibles écarts types estimés.

Comme on le remarque dans le tableau 2 Les résultats obtenus montrent aussi l'influence de la taille de l'échantillon ; mais aussi ces estimateurs sont meilleurs que ceux de la méthode du GPH et cela peut s'expliquer par le fait que la méthode de Valderio A. Reisen utilise le périodogramme lissé comme estimateur de la densité spectrale qui est sans biais.

Le tableau 3 présente les résultats de l'application de la méthode du maximum de vraisemblance approximée. Les résultats obtenus montrent l'influence de la taille de l'échantillon sur l'application de cette méthode. Les meilleurs résultats sont obtenus pour un (T= 1000) car les valeurs de \hat{d} sont plus proches des vraies valeurs, de plus, ils possèdent de faibles écarts types estimés.

Cette simulation prouve ce qui a été démontré en théorie que les méthodes d'estimation de Valderio A. Reisen est meilleure que celle de GPH qui est un bon estimateur, la méthode de vraisemblance approximée est aussi un bon estimateur.

Conclusion

Nous avons présenté une classe de modèles à longue mémoire fractionnaire FARIMA(p, d, q), en particulier FARIMA($0, d, 0$), qui est une extension de la classe ARMA(p, q).

Les processus à longue mémoire permettent de modéliser un grand nombre de phénomènes rencontrés en pratique. Ils sont utilisés dans divers domaines d'applications tel que l'hydrologie, l'économie, la météorologie...

Nous avons étudié en détails les aspects temporel et spectral du processus fractionnaire de longue mémoire pure FI(d), ce dernier est caractérisé par ses fonctions autocovariance, autocorrélation, génératrice des autocovariance qui se présentent sous forme de fractions rationnelles.

La fonction d'autocorrélation décroît hyperboliquement (donc lentement) vers 0. Contrairement au processus à courte mémoire (ARMA, ARIMA) qui ont une fonction d'autocorrélation qui décroît exponentiellement (plus rapidement) vers 0. La fonction de densité spectrale exhibe un pôle en 0.

Les méthodes d'estimation du paramètre d du processus FI(d) sont nombreuses, nous avons présenté dans ce mémoire quelques une d'entres elles. Nous avons appliqué les méthodes de Geweke et Peter Hudak (GPH), Valderi A. Reisen et celle du maximum de vraisemblance approximée (Whittle) pour estimer le paramètre d . Ces méthodes nous ont donné de bons estimateurs pour un choix de taille d'échantillon assez grande.

Perspectives de recherches

Modélisation des phénomènes naturels, nous citons :

- Les températures (journalière, mensuelle, trimestrielle) de la ville Béjaia ;
- Le problème du crue de la Soumam, Béjaia ;
- Prix du pétrole ;
- Taux de change.

A

Annexe

A.1 Quelques formules concernant la fonction hypergéométrique

1. Définition

$$F(a, b, c, z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{j \geq 0} \frac{\Gamma(a+j)\Gamma(b+j)z^j}{\Gamma(c+j)j!}$$

2. Relation de contiguïté de Gauss

$$(c-b-1)F(a, b, c, z) + bF(a, b+1, c, z) - (c-1)F(a, b, c-1, z) = 0$$
$$c(1-z)F(a, b, c, z) - cF(a, b-1, c, z) + (c-a)F(a, b, c+1, z) = 0$$

On déduit de ces deux relation, la nouvelle égalité :

$$(c-1)F(a, b-1, c-1, z) + (az - bz - c + 1)F(a, b, c, z) + \frac{b(c-a)}{c}zF(a, b+1, c+1, z) = 0$$

preuve

3. Fonctions s'exprimant en terme de fonctions hypergéométriques

$$(1 - \alpha z)^\delta F(-\delta, b, b, \alpha z)$$
$$\cos(2az) = F(-a, a, \frac{1}{2}, \sin^2 z)$$
$$\pi \sec(\pi z) = \Gamma(\frac{1}{2} - z)\Gamma(\frac{1}{2} + z)$$

4. Propriétés diverses de la fonction hypergéométrique

$$F(a, b, c, 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$$

$$\frac{\Gamma(d+1)}{\Gamma(d-h+1)} = (-1)^h \frac{\Gamma(-d+h)}{\Gamma(-d)} \quad (\text{par calcul direct})$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

5. Formule de Sheppard

$$\frac{\Gamma(j+a)}{\Gamma(j+b)} \approx j^{a-b}, \quad \text{si } j \text{ tend vers } \infty,$$

(déduite de la formule de stirling)

A.2 Méthode de Beran

Le générateur nommé "Beran" utilise la transformée de Fourier rapide pour générer des processus FARIMA(0,d,0) qui est écrit à l'origine par J. Beran (1994) [4].

Bibliographie

- [1] S. ADJEMIAN. Le processus AR(p). Université de Maine, GAINS AND CEPREMAP. 24 avril 2008. pp. 1-19.
- [2] N. AMROUNE. Série de fourier [http ://www.univ-sba.dz/fsi/lmd/Math3/Cours-seriesFourier.pdf](http://www.univ-sba.dz/fsi/lmd/Math3/Cours-seriesFourier.pdf)
- [3] G. ASMA et A.CHAKER(1974). Efficience informationnelle, dynamique des rendements Boursiers et apport des processus de mémoire longue :fondements théoriques et validations internationales, Tunis Tunisie.
- [4] D.L. BANKS, C. B. READ, S. KOTZ. Encyclopédia of Statistical Sciences : Update. University of Minnesota. Vol. 1. Illustré. 1997.
- [5] J.BERAN, (1994). Statistics for long memory processes. New York : (Chapman Hall).
- [6] R. BOURBONNAIS, M. TERRAZA. Analyse des séries temporelles, Applications à l'économie, 2^e edition DUNOD, 2004.
- [7] G. BOX, G. JENKINS, (1970). Time Series Analysis : Forecasting and Control, Holden-Day, San Fransisco.
- [8] G. BRESSON et A. PIROTTE(1998).Econométrie et series temporelles, Théorie et applications , Presse universitaire de France.
- [9] P. BRIAND. Espace probabilisé Septembre 2003.
- [10] P.J. BROCKWELL, R.A. DAVIS. Times Series : Theory and Methods. 1991. Springer-Verlag.
- [11] N. CHABANE, S. ZIDOUNI. *Méthodes statistique pour l'audit interne des banques*, mémoire de fin d'étude de l'université A/Mira de Béjaia, 2002.
- [12] A. CHARPENTIER. Econométrie de la finance, introduction aux modèles dynamiques. ENSAI-CREST .Master Gestion des Risques, IGR, 2006/2007.
- [13] R. DAHLHAUS,(1989). Efficient parameter estimation for self-similar processes. The Annals of Statistics. 17. 1749-1766.
- [14] D. DELIGNIERES. L'analyse des processus stochastiques. EA 2991 " Sport performance, santé", Université de Montpellier 1. Janvier 2001. pp. 1-12.
- [15] C. DENIAU, P. DOUKHAN, G. OPPENHEIM, E. RENAULTE, (2000).Théorèmes limites et longue mémoire en statistique, Statistical Inference For stochastic Processes 3, Special issue.
- [16] D. A. DICKEY, W. A. FULLER. "Likelihood ratio statistics for autoregressive time series with a unit root", *Econometrica* 49, 1981, p. 1057-1072.
- [17] T. DJAFRI et S. IDRES. Etude statistique des produits d'importation les plus dominants au port et l'indépendance entre l'import et l'export. 2011.

- [18] P. DOUKHAN, G. OPPENHEIM, M. S. TAQQU, (2003), Theory and applications of long-range dependence, Basel (Suisse), Birkhäuser.
- [19] J.M. DUFOUR. Histoire de l'analyse des séries chronologiques. 19 janvier 2003.
- [20] R. FOX et M.S. TAQQU (1968). Large-sample properties of parameter for strongly dependent stationary gaussian time series. Ann statist. 1968, 1-12.
- [21] O. FRANÇOIS. Notes de cours de processus aléatoire 2004-2005.
- [22] J. GEWEKE, S. PORTER-HUDAK, (1983). The Estimation and Application of Long Memory Time Series Models. Journal of Time Series Analysis. 4, 221-238.
- [23] E. GONÇALVES. Une généralisation des processus ARMA. In : Annales d'économie et de statistique.- Amiens. Vol. 5.1987. pp. 109-145.
- [24] W. L. GRAF,(1983). The arroyo problem : paleohydrology and paleohydraulics in the short term, Gregory K. G. (ed.), Background to paleohydrology, New York, wiley, 279-302.
- [25] C.W.J. GRANGER, R. JOYEUX (1980), An Introduction to Long Memory Time Series and Fractional Differencing, Journal of Time Series Analysis, Vol. 1.
- [26] A. GRIRA, C. ALOUI. Efficence informationnelle, dynamique des rendements boursiers et apport des processus de mémoire longue : fondements théoriques et validation internationale.
- [27] V. GUIRAUD (1997). Intrégration fractionnaire et etude des chroniques économiques. université de Montpellier 1. 1996-1997.
- [28] H. HAMISULTANE. Econométrie des séries temporelles.
- [29] U. Hassler, J. Wolters, (1995). Long memory in inflation rates : international evidence, Journal of Business and Economic statistics 13, 37-45.
- [30] J. R. M. HOSKING (1981), Fractional Differencing ; Biometrika, Vol. 68.
- [31] J. R. M. HOSKING (1984), Modelling persistence in hydrological time series using fractional differencing. Water Resources Res. 20, 1898-908.
- [32] H. HURST(1951). Long Term Storage Capacity of Reservoirs. Transaction of the American society of civil engineers, Vol. 116.
- [33] P. A. JAYET. Quelques notions sur l'analyse spectrale des séries temporelles. In : Histoire et Mesure, 1991 Volume 6 - n° 1-2. pp. 7-2.
- [34] A. W. Lo, (1991). Long-term memory in stock market prices, Econometrica, 59, 1279-1313.
- [35] M. LUBRANO. Introduction à la modélisation des séries temporelles univariées. Septembre 2008. pp. 60-61
- [36] B. MANDELBROT, (1962). Sur certain prix spéculatifs : faits empiriques et modèle basé sur les processus stables additifs de Paul Lévy. Comptes rendus(Paris) : 254, 3968-3970.
- [37] B. MANDELBROT, Van Ness J.W. (1968).Fractional Brownian Motion, Fractional Noises and Application. SIAM Review. Vol. 10.
- [38] V. MIGNON. Méthode d'estimation de l'exposant de hurst.
- [39] V. MIGNON. Méthode d'estimation de l'exposant de hurst, Application aux rentabilités boursières. Economie et prévision. Numéro 132-133, 1998-1-2. pp. 193-214.
- [40] V. MIGNON (1996). les tests de mémoire longue appartiennent-ils au camp du démon ? (version longue), document de travail MODEM.Université Paris Nanterre (en collaboration avec S. Lardic).

- [41] V. MIGNON (1996). les tests de mémoire longue appartiennent-ils au camp du démon ?. Revue économiques, vol 47,n°3,pp 531-540 (en collaboration avec S. Lardic).
- [42] V. MIGNON (1996). les implications de la mémoire longue de la non linéarité sur l'efficacité du marché des change. Journal de la société statistique de Paris, Tom 137, PP 51-72.
- [43] V. MIGNON. Méthodes d'estimation de l'exposant de Hurst. Application aux rentabilités boursières. In : économie et prévision. Numéro 132-133. 1998-1-2. pp.193-194.
- [44] L. NOUIRA, M. BOUTAHAR et V. MARIMOUTOU, (2003). Les méthodes d'estimation du paramètre de mémoire longue : Evidence Numérique. GREQAM, université de la méditerranée. June 30.
- [45] L. NOUIRA, M. BOUTAHAR et V. MARIMOUTOU, (2003). Les méthodes d'estimation du paramètre de mémoire longue : Evidence Numérique. GREQAM, université de la méditerranée. June 30.pp. 10.
- [46] L. NOUIRA, M. BOUTAHAR et V. MARIMOUTOU (2003). Les méthodes d'estimation du paramètre de mémoire longue : Evidence Numérique. GREQAM, université de la méditerranée. June 30. pp. 19-21.
- [47] C. PERRAUDIN, (2004-2005). SERIES CHRONOLOGIQUES. Quelques éléments du cours. Université Paris I.
- [48] V. A. REISEN,(1994). Estimation of fractional difference parameter in the AR-FIMA(p,d,q) model using the smoothed periodogram. Journal of Time Series Analysis. 15, 335-350.
- [49] P. M. ROBINSON, (2003). Time series with long memory. Oxford, Oxford University Press.
- [50] A . SCHUSTER, (1897) . "Sur périodicités lunaires et solaires des tremblements de terre", Actes de la Royal Society of London , vol. 61, pp 455-465.
- [51] A. SERROUKII, (1995-1996). inférence asymptotique paramétrique et non paramétrique pour les modèles ARMA fractionnaires. Université libre de Bruxelles, Institut de Statistique. .
- [52] F. SOWELL,(1992). Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models. Journal of Econometrics. 53, 165-188.
- [53] M. S. TAQQU, W. WILLINGER, R. SHERMAN, (1997).Proof of a fundamental result in Self-similar traffic modeling. Comp. Comm. Rev, self-similar.
- [54] M. TENENHAUS. La Méthodologie de Box-Jenkins. site web.
- [55] P. WHITLLE,(1983). Prediction and regulation by linear least-square methods. University of Minnesota Press, Minneapolis, MN, second edition.With a foreword by Thomas J. Sargent.
- [56] S. ZETILI. Inférence Statistique dans les processus ARMA fractionnaires. Université Mentouri Constantine.
- [57] W. Willinger , M. S. Taqqu , V. Teverovsky, (1999). Stock market prices and Long memory ependence, Finance and stochastic 3, 1-13.