

# UNIVERSITÉ ASSANE SECK DE ZIGUINCHOR



## U. F. R DES SCIENCES ET TECHNOLOGIES DÉPARTEMENT DE MATHÉMATIQUES

### Mémoire de Master

DOMAINE : SCIENCES ET TECHNOLOGIES  
MENTION : MATHÉMATIQUES ET APPLICATIONS  
SPÉCIALITÉ : MATH-APPLIQUÉE  
OPTION : STATISTIQUE

### Estimation des paramètres de mémoire longue d'un processus ARFISMA $(p, d, q)(P, D, Q)_s$

Présenté par :

SEYDI DIAMIL DIOUF

Sous la direction de : Dr Mor NDONGO

Sous la supervision de : Pr Alassane DIEDHIOU

Soutenu publiquement le 09 Aout 2022

Devant le jury ci-après :

Prénom(s) et Nom	Grade	Qualité	Établissement
Alassane DIEDHIOU	Professeur titulaire	Président du jury	UASZ
Clement MANGA	Professeur assimilé	Examineur	UASZ
Emmanuel Nicolas CABRAL	Maitre de conférence Titulaire	Examineur	UASZ
Mor NDONGO	Maitre de conférence Titulaire	Directeur	UASZ

Année universitaire : 2020–2021



---

## Remerciements

Je voudrais tout d'abord exprimer ma profonde reconnaissance à mon encadreur Monsieur Mor NDONGO, maître conférence titulaire à l'Université Assane SECK de Ziguinchor pour avoir accepté de diriger ce travail. Je tiens à le remercier de m'avoir accordé sa confiance en faisant toujours preuve d'une très grande disponibilité à mon égard. Ses conseils, encouragements et remarques minutieuses ont été déterminants et je réalise aujourd'hui à quel point travailler avec lui a été enrichissant.

Je voudrais exprimer ma gratitude à Monsieur Alassane DIEDHIOU, professeur titulaire à l'Université Assane SECK de Ziguinchor pour avoir accepté de présider mon jury.

Mes remerciements vont également à Messieurs Emmanuel Nicolas CABRAL, maître de conférence titulaire à l'Université Assane SECK de Ziguinchor ; Clément MANGA, professeur assimilé à l'Université Assane SECK de Ziguinchor, pour avoir accepté de faire partie de mon jury.

Je tiens ensuite à remercier docteur souhaibou sambou pour sa disponibilité, ses conseils, ses remarques scientifiques m'ont beaucoup apporté aussi bien sur le plan mathématique et sur le plan humain. Je remercie aussi l'ensemble du personnel de l'institut polytechnique de Ziguinchor (ZIP), lieu qui a abrité une grande partie de la rédaction de mon mémoire.

Je souhaite aussi remercier l'ensemble des enseignants du département de mathématiques ainsi que toute l'équipe du laboratoire de mathématiques de l'Université Assane SECK de Ziguinchor.

Mes remerciements vont aussi à tous ceux qui ont de près ou de loin contribué à la réalisation de ce travail.

Je profite de cette occasion pour remercier mes prédécesseurs docteurs : Marcel BADIANE, Ablaye DIOUF, Gorgui GACKOU, ainsi que les doctorants Papa BADIANE, Moustapha CAMARA pour leurs conseils.

J'adresse mes vives remerciements à tous mes camarades de promotion de master de maths appliquées et maths pures de l'Université Assane Seck de Ziguinchor en particuliers : Saliou Diaw, Tidiane BA, Aliou BA, Yaya COULIBALY, Daouda DIACK, Diang yoro BA, et aussi mes camarades de promotion licence de l'Université Iba Der Thiam de Thiès en particuliers : Omar Sarr, Ibrahima Traoré, Abdou ngom.

Mes remerciements vont aussi à tous mes camarades de l'Université Iba Der Thiam de Thiès en particuliers " wa chambre 5" : Daouda Mbathie, El hadj Mballo, Ablaye Ndiaye, Matar Ndiaye et " wa chambre 18" : Moustapha KA, Amadou Diouldé Diallo, Mbagnick Ndiaye.

Toute ma reconnaissance à la famille SOW au quartier Mbour 4 à Thiès qui m'a considéré comme leur propre fils et la famille YOUM en particuliers : ma tante Ngoné DIOUF et son époux Saliou YOUM que je veux remercier pour divers services rendus.

Je souhaite très sincèrement remercier Monsieur DIOUF qui était mon professeur de maths au lycée et qui m'a beaucoup aidé sur tous les plans et qui a su me transmettre la

---

passion des mathématiques. Recevez par ces quelques mots toute ma profonde reconnaissance.

J'exprime ma gratitude à mes amis : Mouhamadou DIEBATE, Cherif KABA, Baba TOURE avec qui nous avons passé de bons moments à Kédougou sans oublier ma grande soeur Fatou Traoré à qui je leur suis très reconnaissant toute ma vie pour tout ce qui l'a fait pour moi .

Mes remerciements vont aussi à mes amis et voisins à Ziguinchor : Ndiaga MENGUE, Mbacké DIOP, Ibrahima MBENGUE, Ibrahima CISSOKHO, Mamadou Saliou DIALLO pour divers services rendus.

Je ne saurais terminé sans remercier ma famille plus particulièrement à mon Père Ibou DIOUF, à ma mère Diouma DIOUF, à mon oncle Gorgui DIOUF qui m'ont toujours soutenus dans mes études. Je veux remercier aussi mes frères et sœurs pour leurs soutiens permanents, tant financier que moral.

---

## Résumé

Dans ce travail, nous nous sommes intéressé par l'estimation des paramètres de mémoire longue d'un processus  $ARFISMA(p, d, q)(P, D, Q)_s$ .

Dans un premier temps, nous aborderons la Généralité sur les séries chronologiques. Dans cette partie, nous présenterons les notions fondamentales sur les séries chronologiques et quelques modèles linéaires à savoir :  $AR$ ,  $MA$ ,  $ARMA$ ,  $ARIMA$ ,  $SARIMA$ ,  $ARFIMA$ . Dans la seconde partie de ce travail, nous présenterons le modèle  $ARFISMA(p, d, q)(P, D, Q)_s$  et ses propriétés probabilistiques. Nous ferons une simulation numérique pour illustrer le comportement du processus  $ARFISMA$ .

Dans la dernière partie, nous présenterons les méthodes d'estimations à savoir la méthode semiparamétrique ( $GPH$ ) et la méthode paramétrique ( whittle ) et nous terminerons ce travail par l'estimation des paramètres des modèle  $ARFIMA(0, d, 0)$ ,  $ARFISMA(0, D, 0)$  et  $ARFISMA(0, d, 0)(0, D, 0)_s$  par les méthodes présentés dans cette partie avec le logiciel R.

# Table des matières

<b>Table des matières</b>	<b>4</b>
<b>Liste des tableaux</b>	<b>6</b>
<b>Table des figures</b>	<b>7</b>
Introduction générale . . . . .	9
<b>1 Généralité sur les Séries chronologiques</b>	<b>10</b>
1.1 Description d'une série chronologique . . . . .	11
1.1.1 Processus stochastique . . . . .	11
1.1.2 Opérateur de retard . . . . .	15
1.1.3 Les autocorrélations et la densité spectrale . . . . .	15
1.2 Les processus ARMA . . . . .	16
1.2.1 Processus autorégressif AR(p) . . . . .	16
1.2.2 Processus moyenne mobile MA(q) . . . . .	17
1.2.3 Les processus ARMA(p, q) . . . . .	18
1.2.4 Présentation du modèle . . . . .	18
1.2.5 Propriétés probabilistiques . . . . .	18
1.3 Les processus ARIMA et SARIMA . . . . .	19
1.3.1 Les processus ARIMA . . . . .	19
1.3.2 Les processus SARIMA . . . . .	20
1.4 Processus ARFIMA . . . . .	20
1.4.1 Présentation du modèle . . . . .	20
1.4.2 Propriétés probalistiques . . . . .	20
<b>2 Processus ARFISMA <math>(p, d, q)(P, D, Q)_s</math></b>	<b>22</b>
2.1 Présentation du modèle . . . . .	22
2.2 Propriétés probabilistes . . . . .	24
2.3 Quelques simulations . . . . .	26
2.4 conclusion . . . . .	28
<b>3 Estimation des paramètres de mémoire longue d'un processus ARFISMA</b>	<b>29</b>
3.1 Les méthodes d'estimations . . . . .	29
3.1.1 Méthodes Sémiparamétriques . . . . .	30
3.1.2 Méthodes de Whittle . . . . .	31
3.2 Simulation de Monté carlo . . . . .	31
3.2.1 Estimation du modèle $ARFIMA(0, d, 0)$ . . . . .	31

TABLE DES MATIÈRES

---

3.2.2	Estimation du modèle $ARFISMA(0, D, 0)_s$ . . . . .	32
3.2.3	Estimation du modèle $ARFISMA(0, d, 0)(0, D, 0)_s$ . . . . .	34
	Conclusion générale . . . . .	36
	<b>Bibliographie</b>	<b>37</b>

# Liste des tableaux

3.1	Estimation du paramètre $d$ d'un processus ARFIMA(0, $d$ , 0) pour différente valeur $\alpha$ : 0.5, 0.6, et 0.7 et pour $d = 0.15$ . . . . .	32
3.2	Estimation du paramètre $D$ d'un processus $ARFISMA(0, D, 0)_s$ pour $D = 0.20$ et $s = 4$ . . . . .	33
3.3	Estimation du paramètre $D$ d'un processus $ARFISMA(0, D, 0)_s$ pour $D = 0.20$ et $s = 6$ . . . . .	33
3.4	Estimation du paramètre $D$ d'un processus $ARFISMA(0, D, 0)_s$ pour $D = 0.20$ et $s = 12$ . . . . .	34
3.5	Estimation des paramètres $d$ et $D$ d'un processus $ARFISMA(0, d, 0)(0, D, 0)_s$ pour $d = 0.15$ , $D = 0.20$ et $s = 4$ . . . . .	35
3.6	Estimation des paramètres $d$ et $D$ pour d'un $ARFISMA(0, d, 0)(0, D, 0)_s$ pour $d = 0.15$ , $D = 0.20$ et $s = 6$ . . . . .	35
3.7	Estimation des paramètres $d$ et $D$ d'un processus $ARFISMA(0, d, 0)(0, D, 0)_s$ pour $d = 0.15$ , $D = 0.20$ et $s = 12$ . . . . .	36



# Table des figures

2.1	Périodogramme (a) et ACF (b) d'un processus ARFISMA $(0, d, 0)(0, D, 0)_s$ pour une saisonnalité $s=4$ . . . . .	27
2.2	Périodogramme (a) et ACF (b) d'un processus ARFISMA $(0, d, 0)(0, D, 0)_s$ pour une saisonnalité $s= 6$ . . . . .	27
2.3	Périodogramme (a) et ACF (b) d'un processus ARFISMA $(0, d, 0)(0, D, 0)_s$ pour une saisonnalité $s= 12$ . . . . .	27

## Notations

**ft** : La tendance

**S** : La saisonalité

$\epsilon_t$  : Le résidu

**bb** : Bruit blanc

**i.i.d** : Indépendant identiquement distribué

**E** : Espérance

**Var** : Variance

**Cov** : Covariance

**Corr** : Corrélation

**ACF** : Auto correlation function "Fonction d'autocorrélation"

**PACF** : Partial auto correlation function "Fonction d'autocorrélation partiel"

**ARMA** : Auto regressive moving average "Autorégressif moyenne mobile"

**AR** : Auto regressive "Autorégressif"

**MA** : Moving average "Moyenne mobile"

**ARIMA** : Autorégressif intégré moyenne mobile

**SARIMA** : Autorégressif moyenne mobile intégré saisonier

**ARFIMA** : Autorégressif fractionnée intégré moyenne mobile

**GPH** : La méthode de Geweke et Porter Hudak

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

**RMSE** : La racine de l'erreur quadratique moyenne

**MAE** : L'erreur absolue moyenne

## Introduction générale

Une série chronologique est un ensemble d'observations qui est obtenu de façon séquentielle à des intervalles de temps réguliers ou non. La modélisation et la prévision des séries chronologiques sont des enjeux importants dans de nombreux domaines : l'économétrie financière, l'économétrie classique, les sciences sociales et médicales, la démographie, et les sciences du climat.

Les objectifs pour l'étude des séries chronologiques sont la description et la compréhension du mécanisme de production de la série d'une part, et la prévision des valeurs futures d'autre part.

Les séries chronologiques linéaires peuvent être divisées en deux principales catégories. Les processus à mémoire courte et les processus à mémoire longue. Nous disons qu'un processus linéaire est à mémoire courte si sa fonction d'autocovariance est absolument sommable i.e  $\sum_{h \in \mathbb{Z}} |\gamma(h)| < \infty$ . On dit que le processus est à mémoire longue si  $\sum_{h \in \mathbb{Z}} |\gamma(h)| = \infty$ . Les processus à mémoire courte sont largement utilisés dans de nombreux domaines et de multiples modèles ont été proposés : AR, MA, ARMA. En revanche, les modèles permettant de représenter les processus à mémoire longue ne sont apparus qu'au début des années 1950, historiquement pour l'étude du comportement inhabituel des niveaux du fleuve Nil en Égypte.

Les processus à mémoire longue s'avèrent plus adaptés à l'étude des séries chronologiques issues par exemple de l'économie, la climatologie et l'économétrie financière. On peut, dans cette situation, adapter des modèles plus larges appelés modèle Autorégressifs Moyenne mobile fractionnairement Intégrés ARFIMA et modèle Autorégressifs Moyenne mobile fractionnairement Intégrés Saisonnier ARFISMA qui présentent « des mémoires longues ». L'objectif de notre travail porte sur l'estimation des paramètres de mémoire longue d'un processus  $ARFISMA(p, d, q)(P, D, Q)_s$ .

Ce mémoire est composé de trois chapitres.

Dans le premier chapitre intitulé «Généralité sur les séries chronologiques», nous rappellerons les notions fondamentales sur les séries chronologiques. Ce chapitre est consacré aussi à l'étude des processus stationnaires de mémoire courte : AR, MA et ARMA, les processus non stationnaires de mémoire courte : ARIMA et SARIMA et le processus de mémoire longue ARFIMA.

Le deuxième chapitre intitulé «Les processus  $ARFISMA(p, d, q)(P, D, Q)_s$ » est consacré à la présentation du modèle  $ARFISMA(p, d, q)(P, D, Q)_s$  et la simulation du modèle  $ARFISMA(0, d, 0)(0, D, Q)_s$ .

Enfin, le dernier chapitre intitulé

«Estimation des paramètres de mémoire longue d'un processus  $ARFISMA(p, d, q)(P, D, Q)_s$ ».

Dans ce chapitre nous avons présenté deux méthodes d'estimations : la méthode semiparamétrique ( $GPH_T$  et  $GPH_P$ ) et la méthode paramétrique (Whittle). Nous allons faire l'estimation de paramètre de mémoire longue non saisonnier  $d$  du modèle  $ARFISMA(0, d, 0)$  ; du paramètre de mémoire longue saisonnier  $D$  du modèle  $ARFISMA(0, D, 0)_s$  et des paramètres de mémoire longue non saisonnier et saisonnier  $d$  et  $D$  du modèle  $ARFISMA(0, d, 0)(0, D, 0)_s$  par les deux méthodes : semiparamétrique et paramétrique avec le logiciel R.

A la fin, nous donnons une conclusion générale.

# Chapitre 1

## Généralité sur les Séries chronologiques

### Introduction

Depuis toujours, l'Homme a voulu prédire l'avenir, que ce soit pour prendre de meilleures décisions ou simplement pour satisfaire sa curiosité. Les premières tentatives étaient basées sur l'astrologie ou autres superstitions. Toutefois, ces méthodes de prédiction ne peuvent pas être considérées comme rigoureuses aux yeux de la science. Heureusement, la science a énormément évolué au cours des dernières décennies ; les méthodes de prévision qui en découlent ne font pas exception. Une méthode de prévision très populaire est basée sur l'étude rigoureuse de séries chronologiques. Cette approche permet de prédire, par exemple, de nombreux phénomènes naturels et financiers.

Une série chronologique est constituée de valeurs observées à des intervalles de temps réguliers. Par exemple, les débits annuels sur un cours d'eau ou encore les valeurs mensuelles de titres boursiers sont des séries chronologiques. À la base, l'étude formelle des séries chronologiques consiste à trouver un modèle mathématique qui explique le mieux possible les données observées. À partir de ce modèle, il est possible de faire de la prévision. Cependant, la justesse des prévisions dépend fortement de la qualité du modèle choisi. Il est donc primordial de trouver des modèles qui reflètent le mieux possible la réalité afin de minimiser les erreurs de prévision.

Dans ce chapitre nous allons présenter brièvement les notions classiques des séries chronologiques, les processus ARMA, ARIMA, SARIMA et ARFIMA et leurs caractéristiques statistiques en terme de stationnarité et d'inversibilité.

**Définition 1.0.1** (voir [10])

*On appelle série chronologique ou chronique une suite  $(X_t)_{t \in T}$  d'observation chiffrées d'un même phénomène, ordonnée dans le temps.*

- Une série chronologique est aussi appelée série temporelle ou chronique ;
- $t$  est le numéro de l'observation et représente la date à la quelle est faite l'observation ;
- $T$  est appelé espace de temps qui peut être discret ou continu.

### Remarque

Les dates d'observations sont généralement ordonnées de manière régulière dans le temps. On distingue deux types de séries chronologiques :

- **série continue** : c'est une série où l'observation est faite d'une manière continue dans le temps.

A titre d'exemple : résultat d'un électrocardiogramme.

- **série discret** : C'est une série où l'observation est faite sur des intervalles de temps fixés à priori. Dans ce cas  $T \in \mathbb{Z}$  et les dates d'observations sont plus souvent équidistantes.

A titre d'exemple : relevés mensuels, trimestriels, . . . .

Les objectifs des séries chronologiques sont :

- comprendre le passé : expliquer les variations observées ;
- prédire les valeurs futures (proche) ;
- étudier l'influence du passé sur le futur (proche ou lointain).

## 1.1 Description d'une série chronologique

On considère qu'une série chronologique  $(X_t)_{t \in T}$  est la résultante de trois (3) composantes fondamentales.

- La tendance ou retard notée  $f_t$ , représente l'évolution à long termes de la série étudiée. Elle traduit le comportement moyen de la série.

- La composante saisonnière notée  $s$  correspond à un phénomène qui se répète à intervalle de temps réguliers (périodique). En général c'est un phénomène saisonnier d'où le terme de variations saisonnières.

- La composante résiduelle (bruit ou résidu) notée  $\epsilon_t$  correspond à des fluctuations irrégulières, en général de faible intensité mais de nature aléatoire. (voir [10] )

### 1.1.1 Processus stochastique

On parle de processus stochastique, dit aléatoire, lorsque l'évolution d'une variable dans le temps est imprévisible c'est à dire qu'il est impossible, connaissant la position de la variable au temps  $t$  de prédire avec exactitude sa position au temps  $t + \Delta t$ .

#### Définition 1.1.1

Un processus stochastique est une suite de variable aléatoire  $(X_t)_{t \in T}$  définie sur un espace de probabilité  $(\Omega, A, \mathbf{P})$  à valeurs dans un espace  $\mathbb{E}$ .

\* Si  $t$  est fixé dans  $T$  :  $\omega \rightarrow X_t(\omega)$  est une variable aléatoire.

\* Si  $\omega$  est fixé dans  $\Omega$  :  $t \rightarrow X_t(\omega)$  est une trajectoire.

De façon générale, nous allons donc supposer qu'une série chronologique univariée  $(X_t)$  (une seule variable observée) peut être considérée comme réalisation d'une fonction aléatoire ; plus précisément, la valeur observée  $x_t$  est envisagée comme la valeur particulière d'une variable aléatoire  $X_t$ , dont la distribution de probabilité décrit les valeurs possibles à l'instant  $t_0$ . La série chronologique est donc considérée comme une réalisation finie d'un processus stochastique.

- Si  $T \subset \mathbb{Z}$ , le processus est dit discret.
- Si  $T$  est un intervalle de  $\mathbb{R}$ , le processus est dit continu.

### Bruit blanc

On appelle bruit toute variation imprévisible d'une quantité dans le temps. Le bruit blanc est le plus aléatoire des bruits. Dans ce cas, il existe aucune corrélation entre les accroissements successifs de la quantité. Autrement dit, la fonction d'autocorrélation est donc nulle.

#### Définition 1.1.2

Un processus  $(\varepsilon_t, t \in \mathbb{Z})$  est dit bruit blanc si  $(\varepsilon_t)$  est une suite de variable aléatoire indépendante et identiquement distribuée, d'espérance nulle et de variance constante c'est à dire vérifiant les conditions suivantes :

- $E(\varepsilon_t) = 0, \forall t,$
- $Var(\varepsilon_t) = \sigma_\varepsilon^2, \forall t,$
- $Cov(\varepsilon_t, \varepsilon_s) = 0, \forall t \neq s,$
- $(\varepsilon_t, \varepsilon_s) : \text{indépendant } \forall t \neq s.$

Un tel processus n'a ni tendance, ni mémoire par conséquent la connaissance de la valeur du processus à une date donnée n'apporte aucune information pour la prédiction de sa valeur à une date ultérieure.

### Processus stationnaire

Un processus stochastique peut être stationnaire ou non stationnaire. Pour chaque instant du temps, le processus  $X_t, t \in T$  a une distribution de probabilité. Si on ne fait aucune hypothèse particulière sur la nature du processus aléatoire, chaque  $X_t$  est une variable aléatoire avec son espérance (ou moyenne) et sa variance propre. Ainsi, la moyenne  $E(X_t) = \mu_t$  et la variance  $Var(X_t) = \sigma_t^2$  varient également en fonction du temps, sachant que l'on dispose d'une seule observation  $x_t$  de cette variable. Les estimations de son espérance et de sa variance devraient donc se fonder sur une seule observation, ce qui est impossible. On est donc conduit à formuler des contraintes sur les processus aléatoires à prendre en considération.

On étudie donc une classe particulière de processus aléatoire appelée processus aléatoire stationnaire.

#### Définition 1.1.3

Le processus  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire au sens strict si pour tout  $(t_1, t_2, \dots, t_n)$  avec  $t_i \in T, i = 1, \dots, n$  et si pour tout  $\tau \in T$  avec  $t_i + \tau \in T, \{X_{t_1}, \dots, X_{t_n}\}$  a la même distribution de probabilité jointe que  $\{X_{t_1+\tau}, \dots, X_{t_n+\tau}\}$ .

Ainsi, la stationnarité dite forte exprime qu'il y a invariance dans le temps de toutes les caractéristiques du processus. Mais la stationnarité au sens strict est trop restrictive et on assouplit cette condition en définissant la stationnarité faible ou la stationnarité du second ordre.

#### Définition 1.1.4

Un processus  $(X_t)_{t \in \mathbb{Z}}$  est stationnaire au second ordre si :

- $E(X_t) = m < +\infty \quad \forall t \in \mathbb{Z};$
- $E(X_t^2) < +\infty \quad \forall t \in \mathbb{Z};$
- $\text{Cov}(X_t, X_{t+h}) = \gamma(h)$  (ne dépend pas de  $t$ )  $\forall t$  et  $h \in \mathbb{Z}$ ,

où  $\gamma(h)$  est la fonction d'autocovariance du processus.

Cette hypothèse signifie que la covariance entre les valeurs prises en deux dates distinctes par le processus ne dépend que de la longueur de la période qui les sépare (et pas de la date initiale).

**Exemple 1.1.1**

\* Par définition un processus bruit blanc faible  $(\varepsilon_t) \sim \text{bb}(0, \sigma_\varepsilon^2)$  est faiblement stationnaire.

\* Un processus linéaire  $X_t = \sum_{i=-\infty}^{+\infty} \psi_i \varepsilon_{t-i}$ , où  $\sum_{i=-\infty}^{+\infty} \psi_i^2 < +\infty$  est stationnaire au second ordre car :

$$1. E(X_t) = E\left(\sum_{i=-\infty}^{+\infty} \psi_i \varepsilon_{t-i}\right) = \sum_{i=-\infty}^{+\infty} \psi_i E(\varepsilon_{t-i}) = 0 < +\infty.$$

2.

$$\begin{aligned} E(X_t^2) &= \text{Var}(X_t) = E\left[\left(\sum_{i=-\infty}^{+\infty} \psi_i \varepsilon_{t-i}\right)^2\right] \\ &= E\left(\sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t-j}\right) \\ &= E\left(\sum_{i=-\infty, i=j}^{+\infty} \psi_i^2 \varepsilon_{t-i}^2 + \sum_{i=-\infty, i \neq j}^{+\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t-j}\right) \\ &= E\left(\sum_{i=-\infty}^{+\infty} \psi_i^2 \varepsilon_{t-i}^2\right) = \sum_{i=-\infty}^{+\infty} \psi_i^2 E(\varepsilon_{t-i}^2) = \sigma_\varepsilon^2 \sum_{i=-\infty}^{+\infty} \psi_i^2 < +\infty \end{aligned}$$

3.

$$\begin{aligned}
 \text{Cov}(X_t, X_{t+h}) &= E(X_t X_{t+h}) - E(X_t)E(X_{t+h}) = E(X_t X_{t+h}) \\
 &= E \left[ \left( \sum_{i=-\infty}^{+\infty} \psi_i \varepsilon_{t-i} \right) \left( \sum_{j=-\infty}^{+\infty} \psi_j \varepsilon_{t+h-j} \right) \right] \\
 &= E \left( \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t+h-j} \right) \\
 &= E \left( \sum_{i=-\infty, i=j-h}^{+\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t-(j-h)} + \sum_{i=-\infty, i \neq j-h}^{+\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t-(j-h)} \right) \\
 &= E \left( \sum_{i=-\infty}^{+\infty} \psi_i \psi_{i+h} \varepsilon_{t-i}^2 \right) = \sum_{i=-\infty}^{+\infty} \psi_i \psi_{i+h} E(\varepsilon_{t-i}^2) \\
 &= \sigma_\varepsilon^2 \sum_{i=-\infty}^{+\infty} \psi_i \psi_{i+h}.
 \end{aligned}$$

### Causalité

#### Définition 1.1.5

On dit qu'un processus linéaire  $(X_t)_{t \in \mathbb{Z}}$  est causal s'il peut être représenté sous la forme :

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad (1.1)$$

où  $\varepsilon_t$  est un bruit blanc et

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty.$$

### Inversibilité

#### Définition 1.1.6

Un processus  $(X_t)_{t \in \mathbb{Z}}$  est dit inversible s'il existe une suite constante  $\pi_i$  telle que :

$$\varepsilon_t = \sum_{i=0}^{\infty} \pi_i X_{t-i}, \quad (1.2)$$

où,

$$\sum_{i=0}^{\infty} \pi_i^2 < \infty.$$



### 1.1.2 Opérateur de retard

On aura souvent à considérer une variable en fonction de son passé. Il est donc commode de définir un opérateur qui transforme une variable  $X_t$  en sa valeur passée. C'est l'opérateur retard désigné par la lettre B tel que :

$$\left\{ \begin{array}{l} BX_t = X_{t-1} \text{ ( l'opérateur B décale le processus d'une unité de temps vers le passé),} \\ B^2 X_t = B(BX_t) = BX_{t-1} = X_{t-2}, \\ \vdots \\ B^d X_t = X_{t-d} \text{ (le processus est décalé de d unité de temps).} \end{array} \right. \quad (1.3)$$

- On suppose également que  $B^0 = 1$  de sorte que  $1X_t = X_t$ .
- Les opérations usuelles telles que l'addition, multiplication, division et inverse sont possibles sur l'ensemble des polynômes de retard avec les mêmes propriétés que sur les séries entières.

### Opérateurs de différenciation

L'opérateur  $\Delta$  fait la différence entre le processus et sa version décalée d'une unité de temps. Cet opérateur se construit en utilisant l'opérateur précédent

$$\left\{ \begin{array}{l} \Delta X_t = X_t - X_{t-1} = X_t - BX_t = (1 - B)X_t \Leftrightarrow \Delta = 1 - B, \\ \Delta^2 X_t = (1 - B)^2 X_t = (1 - 2B + B^2)X_t, \\ \vdots \\ \Delta^d X_t = (1 - B)^d X_t. \end{array} \right. \quad (1.4)$$

### 1.1.3 Les autocorrélations et la densité spectrale

#### Fonction d'autocovariance

##### Définition 1.1.7

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire de variance finie. On appelle fonction d'autocovariance  $\gamma(h)$  de  $X_t$  la fonction

$$\gamma(h) = Cov(X_t, X_{t-h}) = \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-h} - \mathbb{E}(X_{t-h}))], \forall h \in \mathbb{Z}. \quad (1.5)$$

#### Fonction d'autocorrélation

##### Définition 1.1.8

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire. On appelle fonction d'autocorrélation  $\rho(h)$  de  $X_t$  la fonction

$$\rho(h) = Corr(X_t, X_{t-h}) = \frac{Cov(X_t, X_{t-h})}{\sqrt{(Var(X_t)(Var(X_{t-h}))}} = \frac{\gamma(h)}{\gamma(0)}, \forall h \in \mathbb{Z}. \quad (1.6)$$

## Densité spectrale

### Définition 1.1.9

La densité spectrale d'un processus aléatoire stationnaire est donnée par :

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \gamma(h) \exp(-i\lambda h) \quad \forall \lambda \in \mathbb{R},$$

où  $\gamma(h)$  est la fonction d'autocovariance du processus.

### Propriété 1.1.1

La densité spectrale est une fonction paire, positive, continue et périodique de période  $2\pi$ .

### Remarque 1.1.1

la fonction d'autocovariance  $\gamma(h)$  s'écrit :

$$\begin{aligned} \gamma(h) &= \text{Cov}(X_t, X_{t-h}) \\ &= \int_{-\pi}^{\pi} f(\lambda) \cos(\lambda h) d\lambda \\ &= \int_{-\pi}^{\pi} f(\lambda) \exp(i\lambda h) d\lambda. \end{aligned}$$

### Exemple :

Un bruit blanc  $(\varepsilon_t)$  est caractérisé par :

$$\begin{cases} \gamma(0) = \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, \text{ pour } h = 0 \\ \gamma(h) = 0, \text{ pour } h \neq 0 \end{cases},$$

alors sa densité spectrale est donnée par :

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi}.$$

## 1.2 Les processus ARMA

### 1.2.1 Processus autorégressif AR(p)

Les processus autorégressifs ont été introduits par Yule en 1927. (voir [30])

#### Définition 1.2.1 (voir [6])

On appelle processus autorégressif d'ordre  $p$  noté  $AR(p)$ , un processus stationnaire au sens faible  $(X_t, t \in \mathbb{Z})$  vérifiant une relation du type :

$$X_t - \sum_{i=1}^p \Phi_i X_{t-i} = \varepsilon_t, \quad \forall t \in \mathbb{Z}, \tag{1.7}$$

où  $\Phi_i$  sont des réels,  $\Phi_p \neq 0$  et  $\{\epsilon_t, t \in \mathbb{Z}\}$  est un bruit blanc de variance  $\sigma_\epsilon^2$ .  
On utilise généralement la notation suivante :

$$\Phi(B)X_t = \epsilon_t, \quad (1.8)$$

avec

$$\Phi(B) = 1 - \sum_{i=1}^p \Phi_i B^i, \quad (1.9)$$

où  $\Phi$  est un polynôme de degré  $p$ .

- Les processus autorégressifs possèdent une «mémoire» au sens où chaque valeur est corrélée à l'ensemble des valeurs qui la précède.
- L'équation caractéristique associée à ce polynôme se note :

$$\Phi(z) = 1 - \Phi_1 z - \Phi_2 z^2 - \dots - \Phi_p z^p. \quad (1.10)$$

- Les processus AR(p) sont stationnaires si toutes les racines de cette équation caractéristique sont à l'extérieur du cercle unité, c'est à dire plus grand que l'unité en module ;

$$|\Phi_i| > 1, \quad i = 1, \dots, p. \quad (1.11)$$

- La représentation AR(p) est inversible par définition.

### 1.2.2 Processus moyenne mobile MA(q)

Les processus moyennes mobiles ont été introduits par Slutsky en 1927. ( voir [30] )

#### Définition 1.2.2

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

$$X_t = \epsilon_t + \sum_{i=1}^q \Theta_i \epsilon_{t-i}. \quad (1.12)$$

On utilise généralement la notation suivante :

$$X_t = \Theta(B)\epsilon_t, \quad (1.13)$$

avec

$$\Theta(B) = 1 + \sum_{i=1}^q \Theta_i B^i, \quad (1.14)$$

où  $\Theta$  est un polynôme de degré  $q$  et  $\epsilon_t$  est un processus de bruit blanc.

- Un tel modèle est appelé moyenne mobile d'ordre  $q$  car  $X_t$  est une moyenne mobile appliquée aux variables aléatoires  $\epsilon_t, \epsilon_{t-1}, \dots, \epsilon_{t-q}$ . Le terme moyenne est à prendre dans un sens très large dans la mesure où la somme des coefficients  $\Theta_k$  n'est pas nécessairement égale à 1. (voir [19])
- La représentation du processus MA(q) est causale par définition.
- Le polynôme retard  $\Theta$  est inversible si et seulement si

$$|\Theta_i| < 1, \quad i = 1, \dots, q. \quad (\text{voir [3]})$$

### 1.2.3 Les processus ARMA(p, q)

l'analyse des séries chronologiques a connu un grand développement depuis la parution du livre de Box et Jenkins (1970), où les principales propriétés des processus stationnaires autoregressifs moyenne mobile (ARMA) ont été décrites avec les méthodes d'identifications, d'estimations et de validations.

Les processus ARMA(p, q) généralisent les modèles autorégressifs et moyennes mobiles. Ces modèles sont très utiles en pratique pour modéliser des séries réelles en nécessitant moins de paramètres que les modèles AR ou MA simples (on le comprend en regardant la décomposition AR ou MA infinie).

### 1.2.4 Présentation du modèle

**Définition 1.2.3** (voir [6])

Un processus stationnaire  $(X_t, t \in \mathbb{Z})$  est un processus ARMA (Autorégressif moyenne mobile) noté ARMA(p, q) s'il existe un  $\Phi = (\Phi_1, \dots, \Phi_p) \in \mathbb{R}^p$  avec  $\Phi_p \neq 0$  et  $\Theta = (\Theta_1, \dots, \Theta_q) \in \mathbb{R}^q$  avec  $\Theta_q \neq 0$  tel que :

$$X_t - \sum_{i=1}^p \Phi_i X_{t-i} = \epsilon_t - \sum_{i=1}^q \Theta_i \epsilon_{t-i}, \quad (1.15)$$

qu'on note généralement par :

$$\Phi(B)X_t = \Theta(B)\epsilon_t, \quad (1.16)$$

où

$$\Phi(B) = 1 - \sum_{i=1}^p \Phi_i B^i \text{ et } \Theta(B) = 1 + \sum_{i=1}^q \Theta_i B^i, \quad (1.17)$$

et  $\epsilon_t$  est un bruit blanc.

- Ce modèle utilise des valeurs retardées de la variable  $X_t$  (d'où le terme autoregressif) et des chocs aléatoires  $\epsilon_t$  qui sont généralement de moyenne nulle, de variance constante et non autocorrélés (bruit blanc); lorsque la variable qui représente ces chocs est retardée, on parle de moyenne mobile. (voir [25])
- Un tel modèle est caractérisé par le paramètre  $p$  de la partie autoregressive et le paramètre  $q$  de la moyenne mobile.

### 1.2.5 Propriétés probabilistiques

Dans cette section, on rappelle les résultats relatifs à la stationnarité et l'inversibilité du processus ARMA(p, q).

**Définition 1.2.4**

Le processus ARMA(p, q)  $(X_t)_{t \in \mathbb{Z}}$  est dit causal s'il existe une suite  $(\psi_j)_{j \geq 0}$  telle que

$$X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \quad (1.18)$$

et

$$\sum_{j=0}^{\infty} \psi_j < \infty.$$

**Proposition 1.2.1**

Supposons que les polynômes  $\Phi(z)$  et  $\Theta(z)$  n'ont pas de racines communes dans le disque unité. Alors nous avons :

*i . Si le polynôme  $\Phi(z)$  n'a pas de racines dans le disque unité alors le processus ARMA( $p, q$ ) est causal, stationnaire et admet une représentation moyenne mobile infinie donnée par l'équation (1.18), où les coefficients  $(\psi_j)_{j \geq 0}$  sont définis par :*

$$\forall |z| \leq 1, \frac{\Phi(z)}{\Theta(z)} = \sum_{j=0}^{\infty} \psi_j z^j.$$

*ii . Si le polynôme  $\Theta(z)$  n'a pas de racines dans le disque unité, alors le processus ARMA( $p, q$ )  $(X_t)_{t \in \mathbb{Z}}$  est inversible et admet une représentation autorégressive infinie définie par :*

$$\epsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \tag{1.19}$$

où les coefficients  $\pi_j$  sont déterminés par :

$$\forall |z| \leq 1, \frac{\Phi(z)}{\Theta(z)} = \sum_{j=0}^{\infty} \pi_j z^j.$$

**Preuve 1.2.1** (Voir [4])

## 1.3 Les processus ARIMA et SARIMA

### 1.3.1 Les processus ARIMA

La plupart des séries chronologiques présente une tendance et / ou une saisonnalité et n'est donc pas modélisable par un processus stationnaire. (voir [19])

**Définition 1.3.1**

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

$$\Phi(B)\nabla^d X_t = \Theta(B)\epsilon_t, \tag{1.20}$$

où  $\Phi$  et  $\Theta$  sont deux polynômes de degré respectifs  $p$  et  $q$ .

- Le modèle ARMA ne sert à traiter que les séries dites stationnaires alors que le modèle ARIMA permet de traiter les séries non stationnaires après avoir déterminé le niveau d'intégration (le nombre de fois qu'il faut différencier la série avant de la rendre stationnaire). Autrement dit, le modèle ARIMA revient à appliquer un modèle ARMA sur le processus différencié.

**Remarque 1.3.1**

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

### 1.3.2 Les processus SARIMA

Le processus SARIMA est une généralisation du processus ARIMA. Autrement dit c'est un processus ARIMA contenant une partie saisonnière.

**Définition 1.3.2**

On dit qu'un processus  $(X_t, t \in \mathbb{Z})$  est un processus SARIMA  $(p, d, q)(P, D, Q)_s$ , s'il vérifie l'équation suivante :

$$\Phi(B)\Phi_s(B^s)(1 - B)^d(1 - B^s)^D x_t = \Theta(B)\Theta_s(B^s)\epsilon_t, \quad (1.21)$$

où  $\Phi, \Phi_s, \Theta, \Theta_s$  sont des polynômes de degré respectifs  $p, P, q, Q$ .

## 1.4 Processus ARFIMA

Les modèles autorégressifs intégrés fractionnaires et à moyenne mobile ARFIMA ont été développés par Granger et Joyeux, (1980) et Hosking (1981) (voir [29]) et constituent une généralisation des processus ARIMA de Box et Jenkins dans lesquels l'exposant de différenciation  $d$  était un entier. Dans le cas des processus ARFIMA,  $d$  peut prendre des valeurs réelles, et non plus seulement des valeurs entières. (Lardic, Mignon, 1999)(voir [1])

### 1.4.1 Présentation du modèle

**Définition 1.4.1**

On dit que le processus  $(X_t, t \in \mathbb{Z})$  est un processus ARFIMA  $(p, d, q)$  s'il vérifie l'équation suivante :

$$\Phi(B)X_t = (1 - B)^{-d}\Theta(B)\epsilon_t, \quad (1.22)$$

où  $d$  est un nombre fractionnaire et  $B$  est l'opérateur retard. Les polynômes  $\Phi(B)$  et  $\Theta(B)$  sont respectivement les polynômes autorégressifs d'ordre  $p$  et moyenne mobile d'ordre  $q$  définis précédemment.

Le filtre dans l'équation (1.22) peut s'écrire de la manière suivante :

$$(1 - B)^{-d}\epsilon_t = \sum_{j=0}^{\infty} b_j \epsilon_{t-j}, \quad (1.23)$$

où les  $(b_j), j \geq 0$  sont les coefficients du développement en série  $(1 - z)^{-d}$  pour  $|z| < 1$  c'est à dire

$$b_0 = 1, b_j = \frac{\Gamma(j + d)}{\Gamma(d)\Gamma(j + 1)}, j \geq 1. \quad (1.24)$$

### 1.4.2 Propriétés probalistiques

**Proposition 1.4.1**

Supposons que les polynômes  $\Phi(z)$  et  $\Theta(z)$  n'ont pas de racines communes dans le disque unité. Alors :

*i* Si le polynôme  $\Phi(z)$  n'admet pas de racines dans le disque unité alors le processus ARFIMA( $p, d, q$ ) est causal, stationnaire et admet une représentation moyenne mobile infinie donnée par :

$$X_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j}, \quad (1.25)$$

où les coefficients  $(c_j)_{j \geq 0}$  sont obtenus par l'équation suivante :

$$\forall |z| \leq 1, \frac{\Theta(z)}{\Phi(z)} (1-z)^{-d} = \sum_{j=0}^{\infty} c_j z^j.$$

*ii* Si de plus le polynôme  $\Theta(z)$  n'a pas de racines dans le disque unité, alors le processus ARFIMA( $p, d, q$ )  $(X_t)_{t \in \mathbb{Z}}$  est inversible et admet une représentation autorégressive infinie définie par :

$$\epsilon_t = \sum_{j=0}^{\infty} \tilde{c}_j X_{t-j}, \text{ p.s} \quad (1.26)$$

où les coefficients  $(\tilde{c}_j)_{j \geq 0}$  sont donnés par :

$$\forall |z| \leq 1, \sum_{j=0}^{\infty} \tilde{c}_j z^j = \frac{\Phi(z)}{\Theta(z)} (1-z)^d.$$

**Preuve 1.4.1** (Voir [20])

## Conclusion

Dans ce chapitre, nous avons présenté quelques notions fondamentales sur les séries chronologiques et les processus linéaires stationnaires et non stationnaires à mémoire courte. Nous avons aussi présenté le processus linéaire à mémoire longue ARFIMA avec leurs propriétés probabilistiques.

# Chapitre 2

## Processus ARFISMA $(p, d, q)(P, D, Q)_s$

### Introduction

Dans de nombreuses situations pratiques, une série chronologique peut présenter un modèle périodique. C'est une caractéristique commune dans de nombreux domaines tels que la météorologie, l'économie, l'hydrologie et l'astronomie ( voir [9]). Parfois même dans ces domaines, la périodicité peut dépendre du temps c'est à dire la structure d'autocorrélation des données varie d'une saison à l'autre. Le modèle d'un tel type de série chronologique ne peut pas être modélisé par un processus ARFIMA. Les processus ARFISMA permettent de modéliser un tel type de modèle. Ils permettent de prendre en compte non seulement des dépendances à long terme mais en plus de la saisonnalité avec parcimonie. Un avantage considérable de ces modèles est qu'ils peuvent prendre en compte à la fois des dépendances à court terme non saisonnier via les paramètres  $p$  et  $q$ , des dépendances à long terme non saisonnier via le paramètre  $d$  mais aussi ils peuvent en même temps prendre en compte des dépendances à court terme saisonnier via les paramètres  $P$  et  $Q$ , et des dépendances à long terme saisonnier via le paramètre  $D$ . Autrement dit le processus ARFISMA est un processus ARFIMA saisonnier.

### 2.1 Présentation du modèle

#### Définition 2.1.1

Soit  $(\epsilon_t)_{t \in \mathbb{Z}}$  un processus suivant un bruit blanc gaussien de variance constante finie  $\sigma_\epsilon^2$ . On dit qu'un processus centré, noté  $(X_t)_{t \in \mathbb{Z}}$ , est un processus ARFISMA  $(p, d, q)(P, D, Q)_s$  s'il vérifie l'équation suivante :

$$\varphi(B)\Phi(B^s)X_t = (1 - B)^{-d}(1 - B^s)^{-D}\theta(B)\Theta(B^s)\epsilon_t, \quad (2.1)$$

où les réels  $d$  et  $D$  sont les paramètres de mémoire longue,  $B$  est l'opérateur de retard et  $s$  désigne la période de saisonnalité.

Les polynômes  $\varphi(B)$  et  $\theta(B)$  sont de degré respectifs  $p$  et  $q$  sont les polynômes autorégressif et moyenne mobile de la partie non saisonnière et sont donnés par :

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p = 1 - \sum_{i=1}^p \varphi_i B^i, \quad (2.2)$$



$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q = 1 + \sum_{i=1}^q \theta_i B^i. \quad (2.3)$$

Les polynômes  $\Phi(B^s)$  et  $\Theta(B^s)$  de degrés  $P$  et  $Q$  sont respectivement les polynômes autorégressif et moyenne mobile de la partie saisonnière et sont définis respectivement par :

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps} = 1 - \sum_{j=1}^P \Phi_j B^{js}, \quad (2.4)$$

$$\Theta(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \cdots + \Theta_Q B^{Qs} = 1 + \sum_{j=1}^Q \Theta_j B^{js}. \quad (2.5)$$

### Remarque 2.1.1

1. Dans l'équation (2.1), le terme  $(1 - B)^{-d}$  contrôle le comportement de la mémoire longue, la saisonnalité est captée par la quantité  $(1 - B^s)^{-D}$ .
2. Dans le cas, où  $P = Q = 0$  et  $D = 0$ , le processus défini en (2.1) devient un processus ARFIMA introduit par Granger et Joyeux. (voir [14])

Ainsi les processus ARFISMA constituent une généralisation des processus ARFIMA.

Par un développement en série, nous avons :

$$(1 - z)^{-d} = \sum_j C_j^d (-z)^j \text{ pour } |z| = 1, \quad d \in \mathbb{R}, \quad (2.6)$$

où

$$C_j^d = \frac{d(1+d)(2+d)\cdots(d+j-1)}{j!} = \frac{(d+j-1)!}{j!(d-1)!} = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}. \quad (2.7)$$

De même,

$$(1 - z^s)^{-D} = \sum_i C_i^{-D} (-z^s)^i \text{ pour } |z| = 1, \quad D \in \mathbb{R}, \quad (2.8)$$

où

$$C_i^{-D} = \frac{D(1+D)(2+D)\cdots(D+i-1)}{i!} = \frac{(D+i-1)!}{i!(D-1)!} = \frac{\Gamma(i+D)}{\Gamma(i+1)\Gamma(D)}. \quad (2.9)$$

En se basant sur les équations (2.6), (2.7), (2.8) et (2.9), le filtre  $(1 - B)^{-d}(1 - B^s)^{-D}$  dans l'équation (2.1) peut s'écrire sous la forme :

$$(1 - B)^{-d}(1 - B^s)^{-D} = \sum_{j=0}^{\infty} \psi_j B^j, \quad (2.10)$$

où les coefficients  $(\psi_j)_{j \geq 0}$  sont définis par :

$$\psi_0 = 1 \text{ et } \psi_j = \pi_j + \sum_{i=0}^{\infty} \pi_i^s \pi_{j-is}, \quad \forall j \geq 1. \quad (2.11)$$

Les poids  $(\pi_j)_{j \geq 0}$  et  $(\pi_j^s)_{j \geq 0}$  sont donnés par :

$$\pi_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(D)}, \quad j = 0, 1, \dots, \quad (2.12)$$

et

$$\pi_i^s = \frac{\Gamma(i+1)}{\Gamma(i+1)\Gamma(D)}, \quad i = 0, 1, \dots; \quad (2.13)$$

où  $\Gamma(\cdot)$  désigne la fonction gamma et définie pour tout réel  $a$  positif ou nul, par :

$$\Gamma(a) = \int_0^{+\infty} t^{a-1} e^{-t} dt \quad \text{et} \quad \Gamma(n+1) = n!. \quad (2.14)$$

(voir [22])

## 2.2 Propriétés probabilistes

Dans cette section, on rappelle les résultats relatifs à la stationnarité et à l'inversibilité des processus ARFISMA définis par l'équation (2.1). Les propriétés, établies par Giraitis et Leipus (1995) et Woodward et al (1998), sont résumées dans la proposition suivante.

### Proposition 2.2.1

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus ARFISMA  $(p, d, q)(P, D, Q)_s$  donné par l'expression (2.1). Supposons que les polynômes  $\varphi(z)\Phi(z^s)$  et  $\theta(z)\Theta(z^s)$  n'ont aucune racine en commun, et les paramètres de mémoire longue  $d$  et  $D$  vérifient les conditions suivantes :

$$|D+d| < \frac{1}{2} \quad \text{et} \quad |D| < \frac{1}{2}. \quad (2.15)$$

Alors nous avons les résultats suivants :

1. Si  $\varphi(z)\Phi(z^s) \neq 0$  et pour  $|z| = 1$ , alors  $X_t$  est stationnaire (au sens faible) et admet une unique représentation moyenne mobile infinie donnée par :

$$X_t = \sum_{j=0}^{\infty} \psi_j \frac{\theta(z)\Theta(z^s)}{\varphi(z)\Phi(z^s)} \epsilon_{t-j}, \quad (2.16)$$

où les coefficients  $(\psi_j)_{j \geq 0}$  sont données par (2.11).

2. Si  $\theta(z)\Theta(z^s) \neq 0$ , pour  $|z| = 1$ , alors  $X_t$  est inversible et sa représentation autorégressif infinie est donnée par :

$$\epsilon_t = \sum_{j=0}^{\infty} \psi_j^* \frac{\varphi(z)\Phi(z^s)}{\theta(z)\Theta(z^s)} X_{t-j}, \quad (2.17)$$

où les coefficients  $\psi_j^*$  sont donnés par :

$$\psi_0^* = 1 \quad \text{et} \quad \psi_j^* = \pi_j^* + \sum_{i=0}^{\infty} \pi_i^{*(s)} \pi_{j-is}^*, \quad \forall j \geq 1, \quad (2.18)$$

avec les  $(\pi_j^*)$  et  $(\pi_i^{*(s)})$  sont obtenus par les expressions suivantes :

$$\pi_j^* = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} \quad , \quad j = 0, 1, \dots, \quad (2.19)$$

et

$$\pi_i^{*s} = \frac{\Gamma(i-D)}{\Gamma(i+1)\Gamma(-D)} \quad , \quad i = 0, 1, \dots \quad (2.20)$$

**Preuve 2.2.1** : la preuve de cette proposition suit les grandes lignes de la preuve du théorème 1 de Giraitis et Leipus (1995). (voir [13])

Ainsi sous les conditions de stationnarités et d'inversibilités, les représentations  $MA(\infty)$  et  $AR(\infty)$  peuvent s'écrire respectivement par :

$$X_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j} \quad \text{et} \quad \epsilon_t = \sum_{j=0}^{\infty} \tilde{c}_j X_{t-j}, \quad (2.21)$$

où les coefficients  $(c_j)_{j \geq 0}$  et  $(\tilde{c}_j)_{j \geq 0}$  sont déterminés respectivement par :

$$\theta(z)\Theta(z^s) \sum_{j=0}^{\infty} \psi_j z^j = \varphi(z)\Phi(z^s) \sum_{j=0}^{\infty} c_j z^j, \quad |z| \leq 1 \quad (2.22)$$

et

$$\varphi(z)\Phi(z^s) \sum_{j=0}^{\infty} \psi_j^* z^j = \theta(z)\Theta(z^s) \sum_{j=0}^{\infty} \tilde{c}_j z^j, \quad |z| \leq 1. \quad (2.23)$$

Dans le cas particulier où  $P = Q = 0$ , on vérifie que les coefficients  $(c_j)_{j \geq 0}$  et  $(\tilde{c}_j)_{j \geq 0}$  peuvent être calculé par les formules récurrentes suivantes :

$$c_0 = 1 \quad \text{et} \quad c_j = \psi_j + \sum_{i=1}^{\min(j,p)} \varphi_i c_{j-i} - \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}, \quad \forall j \geq 1 \quad (2.24)$$

$$\tilde{c}_0 = 1 \quad \text{et} \quad \tilde{c}_j = \psi_j^* + \sum_{i=1}^{\min(j,q)} \theta_i \tilde{c}_{j-i} - \sum_{i=1}^{\min(j,p)} \varphi_i \psi_{j-i}^*, \quad \forall j \geq 1. \quad (2.25)$$

La densité spectrale du processus (2.1) est définie par :

$$f(\lambda) = \frac{|\Theta(e^{i\lambda s})\theta(e^{i\lambda})|^2 \sigma_\epsilon^2}{|\Phi(e^{i\lambda s})\varphi(e^{i\lambda})|^2} \frac{\sigma_\epsilon^2}{2\pi} [2 \sin(\frac{\lambda s}{2})]^{-2D} [2 \sin(\frac{\lambda}{2})]^{-2d}, \quad -\pi \leq \lambda \leq \pi, \quad (2.26)$$

avec  $\sigma_\epsilon^2$  la variance du bruit.

Soit  $(X_1, \dots, X_n)$  un échantillon extrait d'un processus  $(X_t)$  défini par (2.1). Le périodogramme de  $X_t$  est donné par :

$$I(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^n X_t e^{i\lambda t} \right|^2, \quad -\pi \leq \lambda \leq \pi. \quad (2.27)$$

## 2.3 Quelques simulations

Nous présentons dans cette partie un ensemble de simulation pour illustrer le comportement asymptotique d'un processus ARFISMA( $p, d, q$ )( $P, D, Q$ ) $_s$ . Sans perte de généralité, nous simulons le modèle ARFISMA( $0, d, 0$ )( $0, D, 0$ ) $_s$ , c'est à dire  $p = q = P = Q = 0$ . Plus précisément, nous simulons, en utilisant la méthode décrit par Stoev et Taqu pour simuler un processus ARFIMA( $p, d, q$ ). Cette méthode consiste à générer une trajectoire  $X(t)$ ,  $t = 0, \dots, N - 1$  avec  $N \in \mathbb{N}$ , issu d'un processus ARFISMA( $p, d, q$ )( $P, D, Q$ ) $_s$ , en utilisant sa représentation moyenne mobile infinie donnée par l'équation (2.21). En d'autres termes, elle approxime cette trajectoire de taille  $N$  par sa version tronquée de la manière suivante :

$$X(t) = X_M(t) := \sum_{j=0}^{M-1} c_j \epsilon_{t-j}, \quad t = 0, \dots, N - 1 \quad (2.28)$$

où les coefficients sont définies par l'équation (2.24),  $\epsilon_t$  est une suite de variable aléatoire indépendante et identiquement distribuée de variance  $\sigma_\epsilon^2$  et  $M$  est le paramètre de troncature. Pour plus de détail, (voire la section (2.4) de la thèse de NDONGO [24]). Le processus est donné par :

$$X_t = (1 - B)^{-d} (1 - B^s)^{-D} \epsilon_t, \quad (2.29)$$

d'après (2.1).

Nous choisissons le paramètre de troncature  $M = 5000$ , la taille d'échantillon  $N = 2000$ , les paramètres de mémoire longue non saisonnier et de mémoire longue saisonnier respectifs  $d = 0.15$  et  $D = 0.20$  pour différentes périodes de saisonnalité  $s = 4$ ,  $s = 6$ ,  $s = 12$ .

Les outils de bases utilisés pour analyser une série chronologique par des processus de mémoire longue sont la fonction d'autocorrélation empirique (ACF) et la densité spectrale empirique. La densité spectrale d'un processus ARFISMA gaussien est largement étudiée dans la littérature (voir Reisen et al [28]). Ainsi, d'après (2.30), la densité spectrale de ce modèle ARFISMA pour  $p = q = P = Q = 0$ , est définie par :

$$f(\lambda) = \frac{\sigma_\epsilon^2}{2\pi} [2 \sin(\frac{\lambda s}{2})]^{-2D} [2 \sin(\frac{\lambda}{2})]^{-2d}, \quad -\pi \leq \lambda \leq \pi, \quad (2.30)$$

où  $\sigma_\epsilon^2$  désigne la variance du bruit. Cette densité spectrale est non bornée aux fréquences saisonnières  $\lambda_j = \frac{2\pi j}{s}$ , pour  $j = 0, 1, 2, \dots, [\frac{s}{2}]$  (voir Reisen et al. [28] pour plus de détails). A échantillon fini, la densité spectrale est estimée par le périodogramme et qui est visualisé respectivement pour  $s = 4$ ,  $s = 6$  et  $s = 12$  sur les Figures 2.1( a), 2.2( a) et 2.3( a) et la fonction d'autocorrélation ( ACF) est visualisée sur les Figures 2.1( b), 2.2( b) et 2.3( b)

### 2.3. QUELQUES SIMULATIONS

---

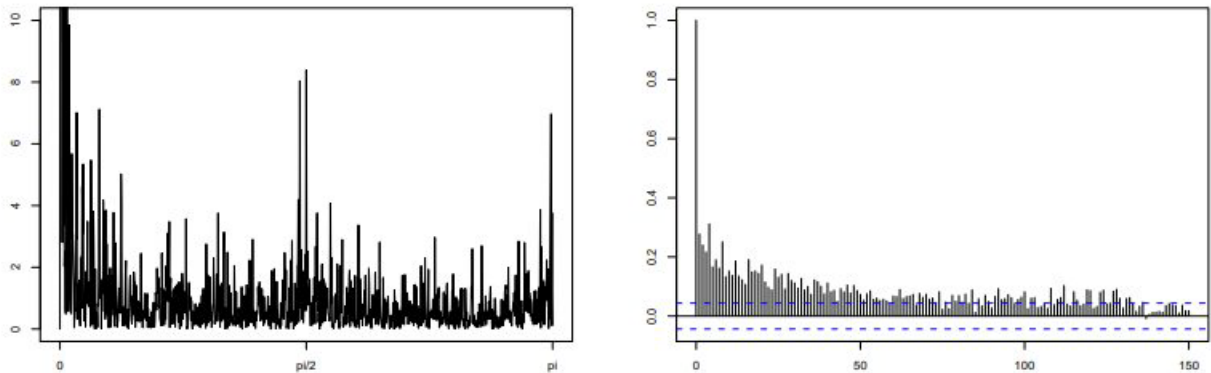


FIGURE 2.1 – Périodogramme (a) et ACF (b) d'un processus ARFISMA  $(0, d, 0)(0, D, 0)_s$  pour une saisonnalité  $s=4$

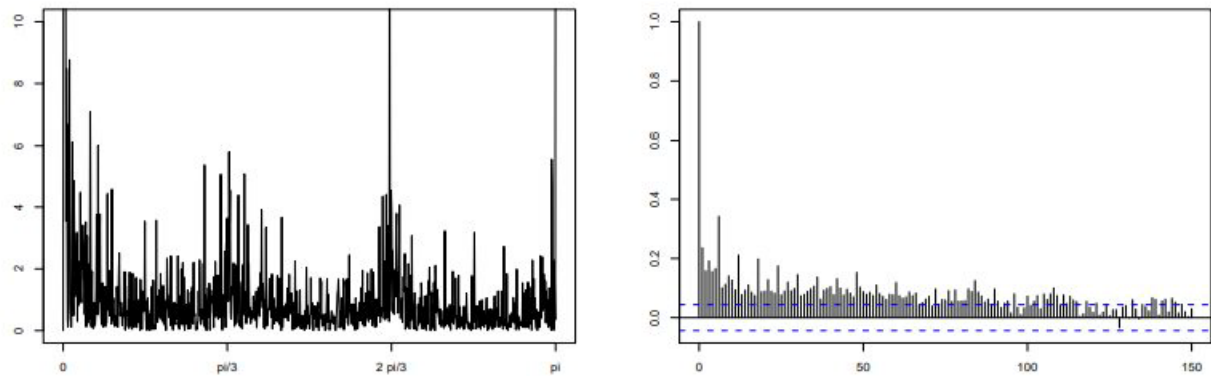


FIGURE 2.2 – Périodogramme (a) et ACF (b) d'un processus ARFISMA  $(0, d, 0)(0, D, 0)_s$  pour une saisonnalité  $s=6$

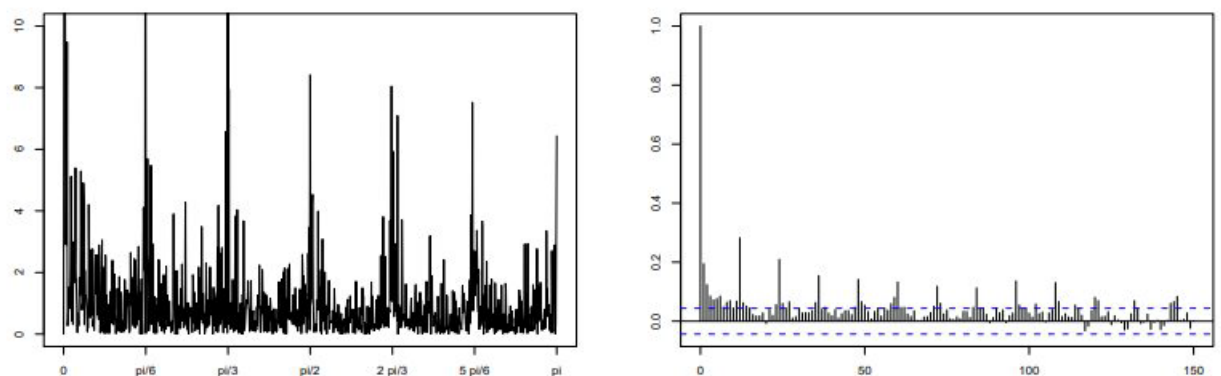


FIGURE 2.3 – Périodogramme (a) et ACF (b) d'un processus ARFISMA  $(0, d, 0)(0, D, 0)_s$  pour une saisonnalité  $s=12$

Les graphiques (périodogrammes) montrent un nombre fréquence dépendant de la saisonnalité. Pour une saisonnalité  $s = 4$ , nous avons 3 fréquences, 4 fréquences pour une saisonnalité  $s = 6$  et 7 fréquences pour une saisonnalité  $s = 12$  car les fréquences sont définies par un  $\lambda_j = \frac{2\pi j}{s}$ ,  $j = 0, 1, 2, \dots, \lfloor \frac{s}{2} \rfloor$  ce qui explique le nombre de fréquence pour chaque saisonnalité. Les périodogrammes montrent aussi des pics sur chaque fréquence saisonnière et un pic à zéro, ce qui est en accord avec le fait que dans le domaine spectral un pic sur la densité spectrale à une fréquence donnée  $\lambda$  ou simplement un pic en zéro indique un cycle périodique de période  $\frac{2\pi}{\lambda}$  sur le processus. D'où la présence de la saisonnalité. Nous constatons sur les graphiques (ACF) une décroissance lente de la fonction d'auto-corrélation lorsque les retards augmentent, assurant ainsi une dépendance à long terme du processus  $(X_t)_{t \in \mathbb{Z}}$  pour différente période de saisonnalité  $s = 4$ ,  $s = 6$  et  $s = 12$ . D'où la présence de mémoire longue.

## 2.4 conclusion

Dans ce chapitre, nous avons présenté le modèle  $ARFISMA(p, d, q)(P, D, Q)_s$  en étudiant les conditions de stationnarité et d'inversibilité et nous avons fait la simulation de ce modèle en prenant compte seulement les paramètres de mémoire longues. c'est à dire  $p = P = 0$  et  $q = Q = 0$  pour des saisonnalités  $s = 4$ ,  $s = 6$  et  $s = 12$ . Nous avons obtenu des graphiques (ACF) qui montrent la présence de mémoire longue et des graphiques (périodogrammes) qui montrent la présence de la saisonnalité.

# Chapitre 3

## Estimation des paramètres de mémoire longue d'un processus ARFISMA

### Introduction

Depuis les années 1980, un certain nombre d'estimateurs a été proposé pour estimer les paramètres d'un processus *ARFIMA* stationnaire et non stationnaire.

Cette méthodologie de modélisation des séries chronologiques avec un comportement de longue mémoire a été récemment étendue aux séries chronologiques à longue mémoire avec des composantes saisonnières liées au modèle *ARFIMA* noté *ARFISMA*.

Dans ce chapitre, nous allons présenter deux méthodes d'estimations : la méthode semiparamétrique appelée GPH proposée par Geweke et Porter-Hudak pour estimer le paramètre de mémoire longue  $d$  d'un processus ARFIMA que Reisen et al. étendent dans le cas saisonnier pour estimer les paramètres de mémoire longue non saisonnier  $d$  et le paramètre de mémoire longue saisonnier  $D$  d'un processus *ARFISMA* (voire [28]) et la méthode Whittle qui est basée sur le périodogramme et la fonction de densité spectrale.

### 3.1 Les méthodes d'estimations

Dans ce paragraphe, nous nous intéressons à l'estimation des paramètres de mémoire longue d'un processus *ARFISMA*. Nous présentons deux méthodes d'estimations : la méthode semiparamétrique et la méthode whittle.

Soit  $(X_1, \dots, X_n)$  une trajectoire de taille finie  $n$  issue d'un processus *ARFISMA* $(0, d, 0)(0, D, 0)_s$ , noté  $(X_t)_{t \in \mathbb{Z}}$  centré, stationnaire, inversible, introduit dans la définition (2.1.1).

On s'intéresse à l'estimation du paramètre vectoriel  $\theta = (d, D)$ . Supposons que  $\theta_0 = (d_0, D_0)$  est la vraie valeur du paramètre  $\theta$  et se trouve à l'intérieur de l'ensemble compact  $\Theta$  donné par :

$$\Theta = \{\theta \in \mathbf{R}^2 : |d + D| < \frac{1}{2} \text{ et } |D| < \frac{1}{2}\}$$

### 3.1.1 Méthodes Sémiparamétriques

Dans cette section, nous présentons la méthode semiparamétrique (GPH) pour estimer les paramètres de mémoire longue d'un processus *ARFISMA*. C'est une méthode qui a été proposée par Geweke et Porter-Hudak d'où le nom GPH pour estimer le paramètre de mémoire longue d'un processus *ARFIMA*( $p, d, q$ ) gaussien (voir [12]). Reisen et al. étendent la méthode GPH dans le cas saisonnier (voir [28]). Cette méthode d'estimation, notée par la suite GPH, est dite semiparamétrique, car elle est fondée sur l'expression locale de la densité spectrale du processus lorsque les fréquences tendent vers zéro. Une équation de régression multilinéaire est obtenue en prenant des logarithmes dans l'expression (2.30),

$$\log f(\lambda) = \log \frac{\sigma_\epsilon^2}{2\pi} \left[ 2 \sin \left( \frac{\lambda s}{2} \right) \right]^{-2D} \left[ 2 \sin \left( \frac{\lambda}{2} \right) \right]^{-2d}. \quad (3.1)$$

Pour  $0 \leq \lambda \leq \pi$ , les estimateurs de  $d$  et  $D$  peuvent être obtenues en remplaçant  $f(\lambda)$  par  $I(\lambda)$  puis en rapprochant la régression (3.1) par :

$$\log I(\lambda) \cong a_0 - D \log \left[ 2 \sin \left( \frac{\lambda s}{2} \right) \right]^2 - d \log \left[ 2 \sin \left( \frac{\lambda}{2} \right) \right]^2 + U_\lambda, \quad (3.2)$$

où  $a_0$  est une constante et

$$U_\lambda = \frac{\ln I(\lambda)}{f(\lambda)} - \mathbb{E} \frac{\ln I(\lambda)}{f(\lambda)}.$$

On considère les fréquences suivantes :

$$\lambda_{\nu,j} = \frac{2\pi\nu}{s} + \frac{2\pi j}{n}, \nu = 0, 1, \dots, \left[ \frac{s}{2} \right] - 1, j = 1, 2, \dots, m;$$

où  $m$  désigne la largeur de bande satisfaisant la condition  $\left(\frac{1}{m}\right) + \left(\frac{m}{n}\right) \rightarrow 0$  lorsque  $n \rightarrow \infty$ . Puisque  $\lambda$  doit se trouver dans l'intervalle  $(0, \pi)$ , nous posons :

$$\lambda_{\frac{s}{2},j} = \pi - \left( \frac{2\pi j}{n} \right). \quad (3.3)$$

Différentes méthodes d'estimations pour  $D$  et  $d$  peuvent être obtenues par des choix appropriés de la fréquence harmonique  $\lambda_j = \frac{2\pi j}{n}$ .

Les méthodes de régressions proposées sont :

1. L'estimateur  $GPH_T$ ,  $T$  pour la totale, produit les estimations en utilisant toute les fréquences harmoniques dans l'équation de régression c'est à dire la régression est construite à partir de :

$$\lambda_{\nu,j} = \frac{2\pi\nu}{s} + \frac{2\pi j}{n}, \nu = 0, 1, \dots, \left[ \frac{s}{2} \right] - 1, j = 1, 2, \dots, m \quad (3.4)$$

avec

$$m = \left[ \frac{n}{s} \right] - 1 \quad (3.5)$$

2. L'estimateur est le  $GPH_p$ ,  $p$  pour le partiel, considère une collection de fréquence harmonique choisies autour au coté droite de chaque fréquences saisonnières ( coté gauche lorsque  $\nu = \left[ \frac{s}{2} \right]$  et  $m = \left[ \frac{n}{2s} \right] - 1$ ).



### 3.1.2 Méthodes de Whittle

Cet estimateur, désigné ci après par  $\hat{\theta}_w$  est une procédure paramétrique due à Fox, Taquq (voir [11]) et whittle (voir [32]) pour les processus gaussiens à mémoire longue. Il est basé sur le périodogramme et les fonctions de densité spectrale. L'indice w signifie whittle. L'estimateur  $\hat{\theta}_w(d, D)$  par la méthode de whittle est obtenu en minimisant la fonction du vraisemblance de whittle

$$\iota_\lambda(\theta) = \frac{1}{2n} \sum_j \left\{ \frac{I(\lambda_j)}{f_\theta(\lambda_j)} \right\} \quad (3.6)$$

où  $f_\theta$  est la densité spectrale , avec

$$\lambda_j = \frac{2\pi j}{n}, \forall j = 1, \dots, n$$

## 3.2 Simulation de Monté carlo

Dans cette partie, nous étudions la performance en échantillon fini des méthodes discutées dans la partie précédente à travers des expériences de Monté carlo. Nous avons effectué plusieurs simulations pour différentes combinaisons des paramètres, de la période saisonnière et de la taille des échantillons.

La moyenne de l'échantillon , l'erreur quadratique moyenne (Mae) et la racine quadratique moyenne (Rmse) sont présentées dans les tableaux 1 à 7.

Nous avons considéré pour  $n = 100$ ,  $n = 1000$ ,  $n = 2000$  et les périodes saisonnières  $s = 4$ ,  $s = 6$  et  $s = 12$ . Les modèles et les valeurs des paramètres de mémoire longue D et d sont spécifiés dans chaque tableau.

### 3.2.1 Estimation du modèle $ARFIMA(0, d, 0)$

Dans cette partie, nous faisons une estimation du paramètre d du modèle  $ARFIMA(0, d, 0)$  qui est un cas particulier du modèle  $ARFISMA$  en utilisant les méthodes sémi-paramétriques ( $GPH_T$  et  $GPH_P$ ) et la méthode de whittle. Pour ce modèle, la méthode  $GPH_T$  utilise toutes les fréquences de Fourier  $\lambda_j = \frac{2\pi j}{n}$ ,  $j = 1, \dots, n-1$ . La méthode  $GPH_P$  sélectionne partiellement les fréquences de Fourier avec une largeur de bande classique m, telle  $m = n^\alpha$  et  $\alpha$  est compris entre 0 et 1. (voir [26]). Dans cette partie nous avons choisi :  $\alpha = 0.5, 0.6, 0.7$ .

### 3.2. SIMULATION DE MONTÉ CARLO

Dans le tableau 3.1, nous remarquons que la méthode du  $GPH_P$  dépend du choix de  $\alpha$  et plus  $\alpha$  est grand, plus nous avons la meilleure estimation mais il semble que les meilleures estimations ont été obtenues avec la méthode  $GPH_T$  car la méthode  $GPH_T$  utilise toutes les fréquences de Fourier  $\lambda_j = \frac{2\pi j}{n}$ ,  $j = 1, \dots, n - 1$ , alors que la méthode  $GPH_P$  sélectionne partiellement les fréquences de Fourier avec une largeur de bande classique  $m$ , telle que  $m = n^\alpha$  et  $\alpha$  est compris entre 0 et 1. Par contre la méthode de whittle est meilleure que la méthode du  $GPH_T$ . Nous remarquons aussi, plus la taille d'échantillon est grande, plus nous avons une meilleur estimation.

n	Statistics	$GPH_P$			$GPH_T$	Whittle
		$\hat{d}$ $\alpha = 0.5$	$\hat{d}$ $\alpha = 0.6$	$\hat{d}$ $\alpha = 0.7$	$\hat{d}$	$\hat{d}$
100	Mean	0.1448	0.1424	0.1441	0.1509	0.1454
	Mae	0.0123	0.0055	0.0024	0.0005	0.0003
	Rmse	0.1066	0.0744	0.0493	0.0236	0.0175
1000	Mean	0.1508	0.1563	0.1506	0.1501	0.1476
	Mae	0.0114	0.0054	0.0022	0.0005	0.0003
	Rmse	0.1059	0.0739	0.0469	0.0227	0.0168
2000	Mean	0.1486	0.1488	0.1490	0.1511	0.1477
	Mae	0.0111	0.0051	0.0022	0.0005	0.0003
	Rmse	0.1055	0.0718	0.0467	0.0226	0.0162

TABLE 3.1 – Estimation du paramètre  $d$  d'un processus ARFIMA(0,  $d$ , 0) pour différente valeur  $\alpha$  : 0.5, 0.6, et 0.7 et pour  $d = 0.15$

#### 3.2.2 Estimation du modèle $ARFISMA(0, D, 0)_s$

Dans cette partie, nous estimons le paramètre de mémoire longue saisonnier  $D$  du modèle  $ARFISMA(0, D, 0)_s$  avec les méthodes semiparamétrique ( $GPH_T$  et  $GPH_P$ ) et la méthode whittle que nous avons décrit dans les sections précédentes avec des saisonnalités  $s = 4$ ,  $s = 6$  et  $s = 12$

### 3.2. SIMULATION DE MONTÉ CARLO

---

		$GPH_T$	$GPH_P$	Whittle
n	Statistics	$\hat{D}$	$\hat{D}$	$\hat{D}$
100	Mean	0.1989	0.2015	0.1933
	Mae	0.0006	0.0006	0.0004
	Rmse	0.0245	0.0262	0.0217
1000	Mean	0.1996	0.2010	0.1923
	Mae	0.0005	0.0006	0.0004
	Rmse	0.0240	0.0250	0.0215
2000	Mean	0.2008	0.2009	0.1920
	Mae	0.0005	0.0006	0.0004
	Rmse	0.0239	0.0249	0.0208

TABLE 3.2 – Estimation du paramètre  $D$  d'un processus  $ARFISMA(0, D, 0)_s$  pour  $D = 0.20$  et  $s = 4$

		$GPH_T$	$GPH_P$	Whittle
n	Statistics	$\hat{D}$	$\hat{D}$	$\hat{D}$
100	Mean	0.1993	0.2023	0.1952
	Mae	0.0005	0.0005	0.0004
	Rmse	0.0236	0.0244	0.0207
1000	Mean	0.2007	0.2016	0.1971
	Mae	0.0004	0.0004	0.0003
	Rmse	0.0215	0.0217	0.0190
2000	Mean	0.2018	0.2011	0.1974
	Mae	0.0004	0.0004	0.0003
	Rmse	0.0213	0.0216	0.0189

TABLE 3.3 – Estimation du paramètre  $D$  d'un processus  $ARFISMA(0, D, 0)_s$  pour  $D = 0.20$  et  $s = 6$

		$GPH_T$	$GPH_P$	Whittle
n	Statistics	$\hat{D}$	$\hat{D}$	$\hat{D}$
100	Mean	0.2026	0.2002	0.1935
	Mae	0.0005	0.0005	0.0004
	Rmse	0.0233	0.0243	0.0208
1000	Mean	0.2010	0.2017	0.1947
	Mae	0.0005	0.0005	0.0004
	Rmse	0.0229	0.0234	0.0201
2000	Mean	0.2016	0.2012	0.1956
	Mae	0.0005	0.0005	0.0003
	Rmse	0.0223	0.0233	0.0196

TABLE 3.4 – Estimation du paramètre  $D$  d'un processus  $ARFISMA(0, D, 0)_s$  pour  $D = 0.20$  et  $s = 12$

Dans les tableaux 3.2 à 3.7, nous remarquons que les méthodes  $GPH_T$  et  $GPH_p$  ont de bonnes performances (Mae et Rmse), même pour un échantillon de petite taille, mais il semble que de meilleures estimations soient obtenues à partir de  $GPH_T$  que de  $GPH_p$  car la première implique toutes les fréquences harmoniques dans l'équation de régression alors que  $GPH_p$  considère une collection de fréquence harmonique choisies autour au coté du droite de chaque fréquences saisonnières. Pour les échantillons de petites tailles ,(Mae) de toute les méthodes de régressions diminue avec  $s$  en raison du fait que plus de fréquences harmoniques sont utilisées. Cet effet devient insignifiant lorsque  $n$  augmente. Par contre l'estimateur de Whittle est meilleur que  $GPH_T$ . Nous remarquons aussi, plus la taille d'échantillon est grande plus nous avons une meilleur estimation.

### 3.2.3 Estimation du modèle $ARFISMA(0, d, 0)(0, D, 0)_s$

Dans cette partie, nous estimons le paramètre de mémoire longues non saisonnier  $d$  et le paramètre de mémoire longue saisonnier  $D$  du modèle  $ARFISMA(0, d, 0)(0, D, 0)_s$  avec la méthodes semiparametrique ( $GPH_T$  et  $GPH_P$ ) et la méthode de whittle que nous avons décrit dans les sections précédentes le paramètre avec des saisonnalités  $s = 4$ ,  $s = 6$  et  $s = 12$ .

### 3.2. SIMULATION DE MONTÉ CARLO

n	Statistics	$GPH_T$		$GPH_P$		Whittle	
		$\hat{d}$	$\hat{D}$	$\hat{d}$	$\hat{D}$	$\hat{d}$	$\hat{D}$
100	Mean	0.1487	0.2005	0.1516	0.2037	0.1483	0.1937
	Mae	0.0005	0.0005	0.0005	0.0006	0.0003	0.0004
	Rmse	0.0235	0.0255	0.0243	0.0260	0.0181	0.0199
1000	Mean	0.1515	0.2007	0.1510	0.2018	0.1509	0.1921
	Mae	0.0005	0.0005	0.0005	0.0006	0.0003	0.0004
	Rmse	0.0224	0.0244	0.0231	0.0247	0.0180	0.0198
2000	Mean	0.1501	0.2016	0.1520	0.2004	0.1501	0.1918
	Mae	0.0005	0.0005	0.0005	0.0006	0.0003	0.0004
	Rmse	0.0217	0.0237	0.0222	0.0241	0.0178	0.0190

TABLE 3.5 – Estimation des paramètres  $d$  et  $D$  d'un processus  $ARFISMA(0, d, 0)(0, D, 0)_s$  pour  $d = 0.15$ ,  $D = 0.20$  et  $s = 4$

n	Statistics	$GPH_T$		$GPH_P$		Whittle	
		$\hat{d}$	$\hat{D}$	$\hat{d}$	$\hat{D}$	$\hat{d}$	$\hat{D}$
100	Mean	0.1492	0.2031	0.1516	0.1998	0.1471	0.1968
	Mae	0.0004	0.0004	0.0006	0.0006	0.0003	0.0003
	Rmse	0.0214	0.0221	0.0250	0.0252	0.0179	0.0189
1000	Mean	0.1502	0.2024	0.1511	0.2012	0.1485	0.1974
	Mae	0.0004	0.0004	0.0004	0.0005	0.0003	0.0003
	Rmse	0.0210	0.0220	0.0218	0.0240	0.0175	0.0189
2000	Mean	0.1519	0.2004	0.1499	0.2020	0.1482	0.1983
	Mae	0.0004	0.0004	0.0004	0.0005	0.0003	0.0003
	Rmse	0.0209	0.0217	0.0215	0.0230	0.0164	0.0188

TABLE 3.6 – Estimation des paramètres  $d$  et  $D$  pour d'un  $ARFISMA(0, d, 0)(0, D, 0)_s$  pour  $d = 0.15$ ,  $D = 0.20$  et  $s = 6$

n	Statistics	$GPH_T$		$GPH_P$		Whittle	
		$\hat{d}$	$\hat{D}$	$\hat{d}$	$\hat{D}$	$\hat{d}$	$\hat{D}$
100	Mean	0.1476	0.2042	0.1488	0.2022	0.1466	0.1954
	Mae	0.0004	0.0004	0.0004	0.0005	0.0003	0.0003
	Rmse	0.0193	0.0232	0.0211	0.0233	0.0174	0.0174
1000	Mean	0.1506	0.2025	0.1507	0.2025	0.1492	0.1957
	Mae	0.0004	0.0004	0.0004	0.0005	0.0003	0.0003
	Rmse	0.0185	0.0230	0.0210	0.0230	0.0173	0.0161
2000	Mean	0.1502	0.2019	0.1512	0.2021	0.1484	0.1946
	Mae	0.0004	0.0004	0.0004	0.0005	0.0003	0.0003
	Rmse	0.0175	0.0225	0.0208	0.0227	0.0163	0.0154

TABLE 3.7 – Estimation des paramètres  $d$  et  $D$  d'un processus  $ARFISMA(0, d, 0)(0, D, 0)_s$  pour  $d = 0.15$ ,  $D = 0.20$  et  $s = 12$

## Conclusion générale

Ce mémoire s'intéresse à l'estimation des paramètres de mémoire longue d'un processus  $ARFISMA(p, d, q)(P, D, Q)_s$ .

Dans la première partie du travail, nous avons rappelé les notions fondamentales sur les séries chronologiques avec la présentation des modèles linéaires à mémoire courte stationnaire à savoir  $AR$ ,  $MA$ ,  $ARMA$ ; les modèles à mémoire courte non stationnaire  $ARIMA$ ,  $SARIMA$ ; le modèle à mémoire longue non stationnaire  $ARFIMA$ .

Dans la seconde partie, nous avons présenté le modèle  $ARFISMA(p, d, q)(P, D, Q)_s$  et faire la simulation du modèle  $ARFISMA(0, d, 0)(0, D, 0)_s$ .

En fin, nous avons fait l'estimation des paramètres de mémoire longue d'un processus  $ARFISMA$ . Dans cette partie nous avons présenté la méthodes sémicparamétriques ( $GPH_T$  et  $GPH_P$ ) et la méthode paramétrique (whittle) sur le modèle  $ARFISMA(0, d, 0)$ , le modèle  $ARFISMA(0, D, 0)_s$  et le modèles  $ARFISMA(p, d, q)(P, D, Q)_s$ .

L'estimation avec la méthode semiparamétrique montre bien que les estimateurs des méthode  $GPH_T$  et  $GPH_P$  fonctionnent bien, même pour les échantillons de petits tailles. Par conséquent, ils sont des méthodes raisonnables pour traiter les données saisonnières dépendantes à mémoire longue. Ces méthodes sont basées sur des périodogrammes et ont l'avantage qu'ils ne sont pas seulement limités au boitier stationnaire mais ils peuvent également être utilisées pour le modèle non stationnaire. Ce pendant la méthode de whittle nécessite la mise en œuvre d'un procédure numérique de précision afin de minimiser de quasi vraisemblance.

# Bibliographie

- [1] A. Charpentier. *Cours de séries temporelles, Théorie et applications*. Université Paris Dauphine.
- [2] Arteche, J. and Robinson, P.M., 2000, *Semiparametric inference in seasonal and cyclical long memory processes*. *Journal of Time Series Analysis*, 21(1), 1–25.
- [3] B. Mandelbrot et J.W. Van Ness. 1968 *Fractional Brownian motions, fractional noises and applications*. *SIAM Review*, Vol 10, 422-437.
- [4] Brockwell, P. J. et Davis, R. A. 1991 *Time Series : Theory and Methods*. Springer Series in statistics. Springer-Verlag, second éd, 1991. 7, 25, 26, 27, 36, 37, 96, 97, 114
- [5] C. Keribin et M. Prenat et R. Rossignol. *Séries Chronologiques, cours et exercices*. Université Paris-Sud.
- [6] C. Perraudin. 2004-2005 *Séries Chronologiques*. Quelques éléments du cours, Université Paris 1, .
- [7] Dahlhaus, R., 1989, *Efficient parameter estimation of self-similar processes*. *Annals of Statistics*, 17, 1749 – 1766
- [8] D. Delignières, *L'analyse des processus stochastiques*. EA 2991 "Sport performance, santé",
- [9] Doukhan, P., Oppenheim, G. and Taqqu, M.S., 2003, *Theory and Applications of Long-Range Dependence* (Boston : Birkhäuser Boston Inc.).
- [10] F. Nicolau. 2005 - 2006 *Séries Chronologiques*. Cour, D'épartement STID, université Nice d'Azur, .
- [11] Fox, R. and Taqqu, M.S., 1986, *Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series*. *Annals of Statistics*, 14, 517–532

- [12] Geweke, J. et Porter-Hudak, S. 1983 *The estimation and application of long memory time series models*. Journal of Time Series Analysis, 4 : 221-228, 2, 3, 59, 64
- [13] Giraitis, L. et Leipus, R. A 1995 *generalized fractionally differencing approach in long memory modelling*. Lithuanian Mathematical Journal, 35 : 65-81, 38, 39
- [14] Granger C.W.J et Joyeux, R. 1980 *An introduction to long memory time serie and fractionel differencing*. Journal of Time Serie Analysis ,1 : 15-30, 29, 35,36
- [15] Hosking, J.R.M. 1981 *Fractional differencing*. *Biometrika*, 68(1) : 165-176, 29,35
- [16] J.Beran , .Chapmann and Hall 1994 *Statistics for Long Memory Processus*
- [17] J-D. Hamilton, 2016 *Time Séries Analysis*. Princeton University Press.
- [18] J.Hosking , 1981 *Fractional Differencing*. *Biometrika* 165-167
- [19] J-J. Daudin, C. Duby, S. Robin et P. Trécourt. 1996 *Analyse des Séries Chronologiques*. INA-PG , Mathématiques.
- [20] Kokoszka, P. S. et Taqqu, M. S, 1995. *Fractional ARIMA with stable innovations*. *Stochastic Processes and their Applications*, 60 : 19-47. iv, v, 1, 3, 7, 29, 30, 35, 37, 41, 42, 44, 64
- [21] Magnus, W., Oberhettinger, F. et Soni, R. P. 1996 *Formulas and Theorems for the special function of mathematical physics*. 3rd edn. Springer-Verlag, New York, 37
- [22] Modeling and forecasting daily average PM10 concentrations by a seasonal ARFIMA model with volatility
- [23] Mr DERGAOUI ramdane, 2016 *Processus de longue mémoire* ,Université A.Mira de Béjaia - Algérie , mémoire de master
- [24] M. Ndongo, 2011, *les processus à mémoire longue saisonniers avec variance infinie des innovations et applications*, thèse de doctorat Université Gaston Berger de saint louis Sénégal
- [25] P-A. Jayet, 1991. *Quelques notion sur l'analyse spectrale des séries temporelles*. In Histoire et mesure, Volume 6, n 1-2, PP 7-2.
- [26] Porter-Hudak, S. 1990 *An application to seasonal fractionally differenced model to the monetary aggregates*. Journal of American Statistical Association, 85 : 338-344.



- [27] Rainville, E. D. 1960 *Special Functions*. Mac Millan, New York. 37
- [28] Reisen, V. A., Rodrigues, A. L. et Palma, W. 2006, *Estimation os seasonal fractionally integrated processes*. Computational Statistics and Data Analysis, 50 : 568-582, iv, v, 1, 2, 3, 35, 37, 38, 48, 58, 59, 60, 65, 66.
- [29] Sandrine LARDIC, Valérie MIGNON (1999) *La mémoire longue en économie*, revue de la littérature.
- [30] Steven Fortier, *Les modèles AR, MA et ARMA multidimensionnels : estimation et causalité*. Département de mathématique, Université de SHERBROOK
- [31] Szegö, G. 1959, *Othogonal polynomials*. AMS Colloquium Publications, vol. 23, Amer. Math. Soc. New York.
- [32] Whittle, P. 1953, *Estimation and information in stationary time series*. Arkiv for Matematik, 2, 423-434.