

UNIVERSITÉ GASTON BERGER DE SAINT-LOUIS

U.F.R DES SCIENCES APPLIQUÉES ET TECHNOLOGIES



T H È S E

présentée pour obtenir

LE GRADE DE DOCTEUR EN MATHÉMATIQUES APPLIQUÉES
DE L'UNIVERSITÉ GASTON BERGER

Spécialité : Probabilités-Statistique

par

MOR NDONGO

Les processus à mémoire longue saisonniers avec variance infinie des innovations et leurs applications

soutenue publiquement le 29 juillet 2011, devant le jury ci-dessous :

M. Gane Samb LO	Professeur (UGB, Sénégal)	Président
M. Papa NGOM	Maître de Conférences (UCAD, Sénégal)	Examineur
M. Aboubakary DIAKHABY	Maître de Conférences (UGB, Sénégal)	Examineur
M. Ouagnina HILI	Maître de Conférences (INP-HB, Côte d'Ivoire)	Rapporteur
M. Zaher MOHDEB	Professeur (Université de Constantine, Algérie)	Rapporteur
M. Aliou DIOP	Maître de Conférences (UGB, Sénégal)	Co-directeur de Thèse
M. Abdou Kâ DIONGUE	Maître de Conférences (UGB, Sénégal)	Co-directeur de Thèse

Dédicace

A mon cher père feu Samba NDONGO, que la terre lui soit légère.

A ma mère, ma tante, mes frères et soeurs.

A mon amour Fatou GUEYE.

Remerciements

Je tiens tout d'abord à exprimer ma profonde reconnaissance à mes co-directeurs de thèse les professeurs Aliou DIOP et Abdou Kâ DIONGUE de l'Université Gaston Berger (UGB) Saint-Louis. Je leur adresse mes plus sincères remerciements pour la confiance qu'ils m'ont accordée, pour leur disponibilité et leur précieux conseils.

Toute ma reconnaissance aux programmes AIRES-SUD (Appuis Intégrés pour le Renforcement des Equipes Scientifiques du Sud) et FIRST (Fonds d'Impulsion pour la Recherche Scientifique et Technique) qui ont contribué financièrement à ma thèse. Je remercie également l'Université de Pau et des Pays de l'Adour (UPPA) et la communauté d'Agglomération de Pau-Pyrénées qui m'ont accordé deux fois de suite une bourse d'excellence pour des stages au sien du Laboratoire des Mathématiques et leurs Application (LMA) de l'UPPA. Merci à Simplicite DOSSOU-GBETE, professeur à l'UPPA et responsable de l'équipe de Probabilités-Statistique du LMA, pour sa disponibilité, ses conseils ainsi que la liberté qu'il a su me donner lors de mes stages à Pau.

Je suis très honoré que les professeurs Ouagnina HILI (Institut National Polytechnique Félix Houphouët-Boigny de Yamoussoukro, Côte d'Ivoire) et Zaher MOHDEB (Université de Constantine, Algérie) aient accepté de rapporter cette thèse. Je leur suis très reconnaissant pour l'intérêt qu'ils ont accordé à mon travail.

Je voudrais exprimer ma gratitude à Monsieur Gane Samb LO, professeur à l'UGB pour l'intérêt qu'il a bien voulu porter à ce travail et avoir accepté de présider mon jury. Mes remerciements vont également à Monsieur Papa NGOM, professeur à l'Université Cheikh Anta DIOP de Dakar et à Monsieur Aboubakary DIAKHABY professeur à l'UGB qui ont accepté de faire partie de mon jury.

Je remercie toute l'équipe du Laboratoire des Mathématiques et leurs Applications (LMA) de l'UPPA, pour l'accueil et l'aide scientifique indispensable à mes recherches. Je remercie aussi toute l'équipe administrative et technique de l'UFR SAT. Je remercie particulièrement Monsieur Galaye DIA, professeur à l'UGB, pour son aide précise dans tous les problèmes scientifiques que j'ai pu rencontrer et pour tous ses conseils.

Je remercie tous les doctorants et jeunes docteurs du Laboratoire d'Études et de Recherches en Statistique et Développement (LERSTAD) avec qui j'ai passé de très bons moments et ainsi que ceux du Laboratoire d'Analyse Numérique et Informatique (LANI).

Je tiens à saluer tous les résidents du Village E, en particulier ceux du "GASS" et ainsi que ceux du Village J qui m'ont soutenu durant mon travail de recherche. Mes derniers remerciements sont bien sûr pour ma famille et mes amis qui ont su me soutenir pendant ces années. Je remercie particulièrement Fatou GUEYE à qui je dois beaucoup, pour son aide, sa présence et sa patience tout au long de mes recherches.

Résumé & Abstract

Résumé : Dans ce travail, nous étudions de manière approfondie les processus à mémoire longue saisonniers avec variance infinie des innovations. Dans le premier chapitre, nous rappelons les différentes propriétés des lois α -stables univariées (stabilité, calcul des moments, simulation, ...). Nous introduisons ensuite deux modèles à variance infinie largement utilisés dans la littérature statistique : les modèles ARMA α -stables et les modèles ARFIMA α -stables développés respectivement par Mikosch et al. [57] et Kokoszka et Taqqu [45]. Constatant les limites de ces modèles, nous construisons dans le second chapitre de nouveaux modèles appelés processus ARFISMA symétriques α -stables. Ces modèles nous permettent de prendre en compte dans une modélisation la présence éventuelle des trois éléments caractéristiques suivants : *mémoire longue, saisonnalité et variance infinie*, que l'on rencontre souvent en finance, en télécommunication ou en hydrologie. Après avoir conclu le chapitre par l'étude du comportement asymptotique du modèle par des simulations, nous abordons dans le troisième chapitre, le problème d'estimation des paramètres d'un processus ARFISMA α -stable. Nous présentons plusieurs méthodes d'estimation : une méthode semiparamétrique développée par Reisen et al. [67], une méthode de Whittle classique utilisée par Mikosch et al. [57] et par Kokoszka et Taqqu [45], et une autre approche de la méthode de Whittle basée sur l'évaluation de la vraisemblance de Whittle par une méthode de Monte Carlo par chaînes de Markov (MCMC). De nombreuses simulations, effectuées avec le logiciel R [64], permettent de comparer ces méthodes d'estimation. Cependant, ces méthodes ne permettent pas d'estimer le paramètre d'innovation α . Ainsi, nous introduisons, dans le quatrième chapitre deux méthodes d'estimation : la méthode de la fonction caractéristique empirique développée par Knight et Yu [43] et la méthode des moments généralisés basée sur des moments conditionnels continus, suggérée par Carrasco et Florens [16]. De plus, afin de comparer les propriétés asymptotiques des estimateurs, des simulations de Monte Carlo sont effectuées. Enfin, dans le cinquième chapitre, nous appliquons ce modèle sur des données de débits du fleuve Sénégal à la station de Bakel. En guise de comparaison, nous considérons le modèle linéaire classique de Box et Jenkins [11], et nous comparons leurs capacités prédictives.

Mots-clés : distribution alpha-stable, variance infinie, saisonnalité, mémoire longue, estimations semiparamétriques, estimations de Whittle, Monte Carlo chaînes de Markov.

Abstract: In this work, we thoroughly study the seasonal fractionally integrated ARIMA time series with stable innovations. In the first chapter, we make an overview of the various properties of the univariate α -stable distribution (stability, calculus of moments, simulation, ...). Then we introduce two models with infinite variance which are widely used in the statistic literature: the stable ARMA model and the stable ARFIMA model developed respectively by Mikosch et al. [57] and Kokoszka and Taqqu [45]. In the second chapter, noting that these models admit some limit, we build a new model called the ARFISMA symmetric α -stable model. These models help us to take into account, in a modelisation, the following three stylized facts: long-memory, seasonality and infinite variance, which are often encountered at finance, telecommunication, or hydrology. After having concluded the chapter with the study of the asymptotic behaviour of the model by some simulations, we focus, in the third chapter, on the parameter estimation problem of the stable ARFISMA model. We present various estimation procedures: the semiparametric estimation methods proposed by Reisen et al.[67], the classical Whittle estimation method used by Mikosch et al. [57] and Kokoszka and Taqqu [45], and another approach of Whittle's method based on the approximation of the Whittle's likelihood function by Markov Chains Monte Carlo (MCMC) method. Many simulations, carried out in R [64], allow to compare these estimation methods. However, these methods do not permit to estimate the α innovation parameter. Thus, in the fourth chapter, we introduce two estimation methods: the empirical characteristic function method developed by Knight et Yu [43] and the generalized method of moments with a continuum of moment conditions, suggested by Carrasco et Florens [16]. Moreover, the asymptotic properties are validated and compared via Monte Carlo simulations. To finish, in the fifth chapter, we apply this model to monthly mean level observations in Bakel (Sénégal). For comparison purpose, we consider the classical linear model of Box et Jenkins [11], and a comparison is made between their forecasting abilities.

Keywords: alpha-stable distribution, infinite variance, seasonality, long-memory, semi-parametric estimates, Whittle estimates, Markov chains Monte Carlo.

Abréviations & Notations

Variabes aléatoires et modes de convergence

- $\mathbb{P}(A)$: La probabilité de l'événement A .
 $\mathbb{E}(X)$: L'espérance mathématique de la variable aléatoire X .
 $X \stackrel{\mathcal{D}}{=} Y$: Les deux variables aléatoires X et Y ont la même loi de probabilité.
 $X \xrightarrow{\mathcal{D}} Y$: La variable aléatoire X converge en loi vers Y .
 $X_n \xrightarrow{p.s.} Y$: La suite de variables aléatoires $(X_n)_n$ converge presque sûrement vers Y .

Notations α -stables

- $S\alpha S$: Symétrique α -stable.
 $S_\alpha(\gamma, \beta, \mu)$: La variable aléatoire X suit une loi α -stable de paramètres α , γ , β et μ .

Notations d'ordre général

- \mathbb{N} : ensemble des entiers naturels.
 \mathbb{N}^* : ensemble des entiers naturels non nuls.
 \mathbb{R} : ensemble des réels et $\mathbb{R}^d = \underbrace{\mathbb{R} \times \dots \times \mathbb{R}}_{d \text{ fois}}$.
 \mathbb{Z} : ensemble des entiers relatifs.
 \mathbb{C} : ensemble des nombres complexes.
 i : le nombre complexe solution de l'équation $i^2 = -1$.
 X' : Transposée du vecteur X .
 $\mathcal{D}_R(a)$: Transformées de Fourier Discrète (TFD) de a .
FFT : Fast Fourier Transform.
i.i.d : indépendantes et identiquement distribuées.
MCMC : Monte Carlo par Chaînes de Markov.
ECF : Empirical Characteristic Function

Remarque sur la bibliographie

Nous signalons que les numéros qui apparaissent après une référence dans la bibliographie désignent les numéros de pages où cette référence a été citée.

Table des matières

Dédicace	i
Remerciements	ii
Résumé & Abstract	iv
Abréviations & Notations	vi
Introduction Générale	1
1 Distributions à queues lourdes et processus symétriques α-stables	6
1.1 Introduction	7
1.2 Les lois stables univariées	7
1.2.1 Définitions et résultats préliminaires	8
1.2.2 Propriétés des lois stables	11
1.2.3 Simulation des lois stables	14
1.2.4 Inférence Statistique des lois stables	15
1.3 Les processus ARMA(p, q)-S α S	25
1.3.1 Présentation du modèle	25
1.3.2 Propriétés probabilistes	25
1.3.3 Estimation des paramètres	26
1.4 Les processus ARFIMA(p, d, q)-S α S	29
1.4.1 Présentation du modèle	29
1.4.2 Propriétés probabilistes	29
1.4.3 Estimation des paramètres	31
1.5 Conclusion	33
2 Les processus ARFISMA symétriques α-stables	34
2.1 Introduction	35
2.2 Présentation du modèle	36
2.3 Propriétés probabilistes	37
2.4 Méthode de simulations numériques	45
2.5 Quelques simulations	47
2.5.1 Cas gaussien	47

2.5.2	Cas stable	52
2.6	Conclusions et perspectives	56
3	Estimation des paramètres d'un processus ARFISMA symétrique α-stable	57
3.1	Introduction	58
3.2	Quelques notations	58
3.3	Estimation semiparamétrique	59
3.4	Estimation par la méthode de Whittle	61
3.4.1	Whittle classique	61
3.4.2	Whittle par la méthode MCMC	61
3.5	Simulations numériques	63
3.5.1	Modèle ARFIMA(0, d, 0)-SaS	64
3.5.2	Modèle ARFISMA(0, D, 0) _s -SaS	65
3.5.3	Modèle ARFISMA(0, d, 0)(0, D, 0) _s -SaS	65
3.6	Conclusions et perspectives	66
4	Estimation simultanée des paramètres d'un processus ARFISMA symétrique α-stable via la méthode de la fonction caractéristique empirique	68
4.1	Introduction	69
4.2	Procédure d'estimation	69
4.3	Propriétés asymptotiques	75
4.4	Simulations de Monte Carlo	77
4.4.1	Choix empirique de la taille des blocs m	79
4.4.2	Comparaison des estimateurs	85
4.5	Autre alternative	87
4.5.1	Méthode C-GMM	87
4.5.2	Calcul simple de l'estimateur C-GMM	89
4.5.3	Quelques simulations	90
4.6	Conclusions et perspectives	91
5	Illustration sur des données réelles	93
5.1	Introduction	94
5.2	Méthodologie	94
5.2.1	Tests de validation	94
5.2.2	Méthode de prévision	97
5.3	Application sur des données hydrologiques	99

5.3.1	Première approche	104
5.3.2	Deuxième approche	106
5.3.3	Comparaison des résultats	110
5.4	Conclusions et perspectives	112
Conclusion générale		113
Bibliographie		116
A	Calcul des coefficients $(c_j)_{j \geq 0}$	123
B	Démonstration du Théorème 4.3.8	126
C	Comparaison des méthodes d'estimation	141

Table des figures

1.1	Test graphique 1 : variance empirique calculée pas à pas sur l'échantillon de 5000 réalisations d'une loi SaS standard pour différentes valeurs de α	17
1.2	Test graphique 2 calculé sur l'échantillon de 5000 réalisations d'une loi SaS standard pour différentes valeurs de α	18
2.1	Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 4$ et $\alpha = 2$. . .	49
2.2	Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 6$ et $\alpha = 2$. . .	50
2.3	Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 12$ et $\alpha = 2$. . .	51
2.4	Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 4$ et $\alpha = 1.7$. . .	53
2.5	Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 6$ et $\alpha = 1.7$. . .	54
2.6	Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 12$ et $\alpha = 1.7$. . .	55
5.1	Position de la station de Bakel sur le fleuve Sénégal.	99
5.2	Evolution de la série du débit du fleuve à Bakel, de Janvier 1904 à Septembre 2010.	100
5.3	Débits mensuels moyens du fleuve à Bakel, de Janvier 1904 à Septembre 2004.	101
5.4	Densité de la série d'étude, comparée avec celle de la loi normale.	102
5.5	Test graphique de la variance empirique sur la série d'étude.	103
5.6	ACF empirique et périodogramme normalisé de la série d'étude.	103
5.7	Evolution du processus résiduel $(\xi_t)_{t=1, \dots, T}$ issu de la modélisation de (5.9)	107
5.8	ACF et PACF empirique du processus résiduel $(\xi_t)_{t=1, \dots, T}$	108
5.9	Trajectoire et ACF empirique du processus résiduel $(\varepsilon_t)_{t=1, \dots, T}$	109
5.10	Comparaison de l'évolution du critère RMSE en fonction des mois, pour les prévisions des deux modèles	111
C.1	Boxplot des paramètres estimés d du modèle 1, avec $\alpha = 1.3$ et $\alpha = 2$ par les méthodes GPH_T , GPH_P , Whittle classique et Whittle par MCMC, pour $n = 100$, $n = 500$ et $n = 1500$, basé sur 1500 réplifications.	143
C.2	Boxplot des paramètres estimés de D du modèle 2, avec $\alpha = 1.3$ et $\alpha = 2$ par les méthodes GPH_T , GPH_P , Whittle classique et Whittle par MCMC, pour $n = 100$, $n = 500$ et $n = 1500$, basé sur 1500 réplifications.	145

C.3	Boxplot des paramètres estimés de d du modèle 3, avec $\alpha = 1.3$ et $\alpha = 2$ par les méthodes GPH_T , GPH_P , Whittle classique et Whittle par MCMC, pour $n = 100$, $n = 500$ et $n = 1500$, basé sur 1500 réplifications.	147
C.4	Boxplot des paramètres estimés de D du modèle 3, avec $\alpha = 1.3$ et $\alpha = 2$ par les méthodes GPH_T , GPH_P , Whittle classique et Whittle par MCMC, pour $n = 100$, $n = 500$ et $n = 1500$, basé sur 1500 réplifications.	148

Liste des tableaux

1.1	Moyenne et variance empirique calculées sur 5000 réalisations.	15
3.1	Processus générateur de données.	63
4.1	Processus générateur de données.	78
4.2	Paramètres estimés du Modèle 1, pour $m = 1, 2, \dots, 8$.	81
4.3	Paramètres estimés du Modèle 2, pour $m = 1, 2, \dots, 8$.	82
4.4	Paramètres estimés du Modèle 3, pour $m = 1, 2, \dots, 8$.	83
4.5	Paramètres estimés du Modèle 4, pour $m = 1, 2, \dots, 8$.	84
4.6	Comparaison entre les méthodes ECF et Whittle-MCMC (Modèle 1)	85
4.7	Comparaison entre les méthodes ECF et Whittle-MCMC (Modèle 2)	86
4.8	Comparaison entre les méthodes ECF et Whittle-MCMC (Modèle 3)	86
4.9	Comparaison entre les méthodes ECF et Whittle-MCMC (Modèle 4)	86
4.10	Comparaison entre les méthodes ECF et C-GMM	91
5.1	Statistiques descriptives de la série du débit du fleuve à Bakel, de Janvier 1904 à Septembre 2004.	101
5.2	Paramètres estimés pour le modèle de Box et Jenkins, défini par 5.8	105
5.3	Résultats en prévision sur la série mensuelle du débit du fleuve Sénégal à Bakel, par le modèle de Box et Jenkins.	106
5.4	Statistiques descriptives du processus résiduel $(\xi_t)_{t=1, \dots, T}$.	107
5.5	Statistiques descriptives du processus $(\varepsilon_t)_{t=1, \dots, T}$.	109
5.6	Estimation des paramètres de la distribution du processus $(\varepsilon_t)_{t=1, \dots, T}$. . .	110
5.7	Résultats en prévision sur la série mensuelle du débit du fleuve Sénégal à Bakel, par le modèle ARFISMA.	110
C.1	Paramètres estimés du modèle 1, pour $\alpha = 1.3$. . .	141
C.2	Paramètres estimés du modèle 1, pour $\alpha = 2$. . .	142
C.3	Paramètres estimés du modèle 2, pour $\alpha = 1.3$. . .	144
C.4	Paramètres estimés du modèle 2, pour $\alpha = 2$. . .	144
C.5	Paramètres estimés du modèle 3, pour $\alpha = 1.3$. . .	146
C.6	Paramètres estimés du modèle 3, pour $\alpha = 2$. . .	146

Introduction Générale

Une revue de la littérature

L'étude des séries temporelles ou séries chronologiques, correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Leur domaine d'application est très vaste et s'étend de l'astronomie à l'économie en passant par la biologie ou la théorie du signal. Elles ont donc suscité un très vif intérêt, ce qui a eu pour conséquence le développement de nombreux modèles : AR, ARMA, ARIMA, GARCH¹ s'appliquant particulièrement à la compréhension des processus à mémoire courte, i.e. ceux pour lesquels il n'y a pas persistance des chocs. Pour tenir compte de la persistance à long terme des chocs dans certaines séries, Granger et Joyeux [31] ainsi que Hosking [37] ont suggéré d'utiliser les modèles ARFIMA². Ils supposent dans leur théorie que les résidus sont des bruits blancs gaussiens. Cependant, l'hypothèse de gaussianité s'avère trop restrictive, en particulier dans certains domaines tels que la finance ou les télécommunications pour lesquels, il faut prendre en compte une plus grande variabilité des données qui peut se traduire par le phénomène de variance infinie. Ainsi, on assiste au développement de modèles à variance infinie, parmi lesquels on peut citer : les processus ARMA avec variance infinie des innovations introduits par Mikosch et al. [57] et les processus ARFIMA symétriques α -stables développés par Kokoszka et Taqqu [45]. Mais toujours dans un cadre longue mémoire, si on désire modéliser simultanément un comportement longue mémoire et un comportement cyclique périodique, on ne pourra malheureusement pas utiliser un modèle ARFIMA. Pour palier à ce manque, plusieurs auteurs ont développé différents modèles longue mémoire intégrant une composante de saisonnalité. Par exemple, Porter-Hudak [63] ainsi que Reisen et al. [67] ont proposé les processus ARFISMA gaussiens qui prennent en compte les phénomènes de mémoire longue et de saisonnalité. Nous pouvons également citer dans cette classe de processus, les modèles GARMA³ à k facteurs développés par Woodward et al. [78] afin de pou-

¹Respectivement AutoRegressive, AR Moving Average, Integrated ARMA, Generalized AR Conditionally Heteroscedastic

²Fractionary ARIMA

³Gegenbauer Autoregressive Moving Average

voir modéliser à la fois les phénomènes de cycles plus ou moins fixes et de mémoire longue.

De nombreuses méthodes d'estimation du paramètre de mémoire longue d'un modèle ARFIMA ont été développées depuis les années 50s. Elles peuvent être regroupées en trois grandes classes : les méthodes heuristiques (voir Hurst [38], Higuchi [36], Lo [50]), les méthodes semiparamétriques (Geweke et Porter-Hudak [29], Robinson ([69], [70]), Reisen [66], Lobato et Robinson [51]) et les méthodes d'estimation paramétriques telles que les méthodes du maximum de vraisemblance (Whittle [77], Sowell [72]), la méthode du minimum de distance (Kouamé et Hili [47]) et la méthode du minimum de distance de Hellinger (Bitty et Hili [10]). Pour les modèles ARFISMA gaussiens, Reisen et al. [67] ont utilisé une forme modifiée de la méthode de Geweke et Porter-Hudak [29] et la méthode du maximum de vraisemblance pour estimer les paramètres de mémoire longue. Dans le cas stable, la méthode du maximum de vraisemblance souffre d'un manque de forme explicite de la densité de probabilité. Ainsi, pour les modèles ARMA(p, q)-S α S et ARFIMA(p, d, q)-S α S, le problème d'estimation est résolu en utilisant la méthode de Whittle classique (voir respectivement Mikosch et al. [57] et Kokoszka et Taqqu [46], pour plus de détails).

L'utilisation de modèles à variance infinie a reçu une attention particulière dans la littérature statistique, depuis les années 1960. En effet, dans les années 60s, Mandelbrot [55] étudie les fluctuations boursières, pour lesquelles il était tout-à-fait clair que le modèle gaussien ne convenait pas. Il s'appuie alors sur les lois de Pareto pour mettre en évidence un nouveau modèle de variation des prix, appelé "loi α -stable". Le paramètre α compris entre 0 et 2, représente l'exposant caractéristique des lois stables et lorsque celui-ci est strictement inférieur à 2, la variance de la loi stable est infinie. Bien qu'une certaine réticence persistât à propos de l'utilisation de ces modèles en finance, la théorie de ces lois à variance infinie s'est développée petit à petit. Mandelbrot [55] confirme que son modèle décrit de façon réaliste la variation des prix pratiqués sur certaines bourses des valeurs. Fama [24] va valider le modèle précédent sur le prix du marché des actions. Quelques années plus tard, Stuck et Kleiner [75] analysent le signal téléphonique en utilisant des modèles basés sur des processus à variance infinie. Gallagher [27] modélise des données de température de la surface de la mer, en considérant un processus AR gaussien et un processus AR stable. Il montre que les meilleurs résultats sont obtenus par le modèle AR stable. Lewis et Ray [49] ont également analysé des données de température de la surface de la mer, en utilisant un modèle à mémoire longue combiné avec un phénomène périodique. Récemment, Stanislavsky et al. [73] décrivent l'activité de radiation solaire

et suggèrent d'utiliser un modèle ARFIMA(2, d, 0) avec un bruit de type Pareto.

Notre contribution

Contribution du Chapitre 1 : Le premier chapitre est surtout un chapitre de généralités qui permet de définir l'ensemble des notions que nous utiliserons dans la suite de ce travail. Nous commençons par les définitions et les propriétés des variables α -stables unidimensionnelles. Nous prêtons une attention particulière à la présentation de deux modèles à variance infinie : les processus ARMA et ARFIMA avec variance infinie des innovations. Dans chacune de ces étapes, nous présentons les résultats les plus importants qui nous seront utiles dans les chapitres qui suivent.

Contribution du Chapitre 2 : Le deuxième chapitre apporte une réponse à la question principale que nous nous sommes posés, à savoir : Existe-t-il un modèle plus général qui permet de prendre en compte les trois caractéristiques suivantes : mémoire longue, variance infinie et saisonnalité ? Dans ce chapitre, nous introduisons le modèle ARFISMA symétrique α -stable et précisons les conditions permettant d'assurer la stationnarité et l'inversibilité de ce modèle. Nous étudions une méthode de simulation pour ce type de processus. En particulier, nous étendons la méthode basée sur l'approche proposée par Stoev et Taquq [74] pour les processus ARFIMA α -stables. Nous montrons sur des données simulées⁴ les trois caractéristiques du modèle. Ces résultats ont fait l'objet de la publication Diongue, Diop et Ndongo [21]. Toutefois, notons que la contribution présentée dans ce chapitre est plus générale par rapport à l'article cité précédemment.

Contribution du Chapitre 3 : Nous nous intéressons à l'estimation des paramètres d'un processus ARFISMA symétrique α -stable. Nous proposons plusieurs méthodes d'estimation. La première méthode est une méthode semiparamétrique développée par Reisen et al.[67], pour estimer les paramètres de mémoire longue d'un processus ARFISMA gaussien. Cette méthode est basée sur une extension dans le cas saisonnier de la méthode de Geweke et Porter-Hudak [29]. Nous proposons d'étendre cette méthode dans le cas stable en utilisant la fonction de transfert puisque la densité spectrale n'existe pas lorsque la variance est infinie. La seconde est la méthode de Whittle classique utilisée par Mikosch et al. [57] et par Kokoszka et Taquq [45], pour estimer les paramètres des modèles ARMA et ARFIMA α -stables. Elle est basée

⁴Tous les calculs numériques établis dans ce document sont effectués avec le logiciel R [64].

sur une évaluation de la fonction de vraisemblance de Whittle par une méthode de quadrature. Nous proposons aussi une autre approche de la méthode de Whittle qui consiste à évaluer cette fonction de vraisemblance par une méthode de Monte Carlo par chaînes de Markov (ou MCMC). Afin d'apprécier les propriétés asymptotiques et la performance de ces différentes méthodes d'estimations, des simulations numériques sont effectuées. Nous faisons varier la taille de l'échantillon, ainsi que les paramètres de mémoire longue du modèle et nous avons étudié deux cas : un cas gaussien ($\alpha = 2$) et un cas stable avec $\alpha = 1.3$. Les résultats de simulations ont permis de conclure que la méthode de Whittle par MCMC est alors privilégiée par rapport aux autres méthodes. Ces travaux ont donné lieu à la publication Ndongo, Diongue, Diop et Dossou-Gbété [59].

Contribution du Chapitre 4 : Vue la limite des méthodes proposées dans le Chapitre 3, pour estimer les paramètres de la distribution du bruit, on se propose ici d'adapter la méthode basée sur la fonction caractéristique empirique (ou méthode ECF) pour estimer simultanément l'ensemble des paramètres du modèle. Nous fournissons des conditions nécessaires pour établir les propriétés asymptotiques de cet estimateur. Nous effectuons des simulations dans le but de déterminer la taille optimale des blocs, puis de comparer la performance de cette méthode avec celle de la méthode de Whittle par MCMC. Nous définissons également une autre alternative : la méthode des moments généralisés basée sur des moments conditionnels continus qui est comparée avec la méthode ECF dans un modèle simple. Les résultats de simulation ont montré que la méthode ECF est meilleure pour le modèle avec présence de parties courtes mémoires et longues mémoires, alors que pour le modèle sans partie courte mémoire la méthode de Whittle par MCMC est préférée. Nous avons également montré, sous certaines conditions de régularité, que l'estimateur ECF est convergent et asymptotiquement normal.

Ces résultats ont fait l'objet de présentations à des séminaires et Workshop :

- Séminaire LMA, 27 Mai 2008, Université de Pau et des pays de l'Adour, France.
- Séminaire LERSTAD, 27 Novembre 2008, Université Gaston Berger de Saint-Louis, Sénégal.
- Séminaire LMA, 19 Novembre 2009, Université de Pau et des pays de l'Adour, France.
- Workshop : Risques Extrêmes et Applications, 24-25 Novembre 2009, LERSTAD, Université Gaston Berger de Saint-Louis, Sénégal.
- Workshop : Modélisation et Environnement, 20-26 Mai 2010, LTI, ESP, UCAD, Sénégal.
- Workshop : Statistique Spatiale et Applications, 29 Novembre-04 Décembre 2010,

LERSTAD, Université Gaston Berger de Saint-Louis, Sénégal.

- Séminaire LERSTAD, 9 Février 2011, Université Gaston Berger de Saint-Louis, Sénégal.

- Atelier CORUS AIRES-SUD : Protection des ressources naturelles, 14-19 Février 2011, Mèknes, Maroc.

Contribution du Chapitre 5 : Le Chapitre 5 a pour objet de réviser les différentes étapes de la modélisation de Box et Jenkins [11] dans le cas gaussien et d'effectuer des applications sur des données réelles. Ainsi, nous avons considéré des données du débit du fleuve Sénégal à la station de Bakel. Nous avons mis en compétition deux types de modèles : le modèle classique SARIMA de Box et Jenkins et le modèle ARFISMA α -stable développé dans le Chapitre 2. Les résultats ne sont pas très satisfaisants pour les deux modèles. Cependant, le modèle ARFISMA α -stable semble fournir une meilleure performance, en terme des critères RMSE ("Root Mean Squared Error") et MAPE ("Mean Absolute Percentage Error").

Distributions à queues lourdes et processus symétriques α -stables

Résumé

Dans ce chapitre, nous résumons les définitions et les propriétés des distributions α -stables dans le cas univarié. Nous présentons également deux modèles à variance infinie, ainsi que leurs propriétés probabilistes et statistiques.

Sommaire

1.1	Introduction	7
1.2	Les lois stables univariées	7
1.2.1	Définitions et résultats préliminaires	8
1.2.2	Propriétés des lois stables	11
1.2.3	Simulation des lois stables	14
1.2.4	Inférence Statistique des lois stables	15
1.3	Les processus ARMA(p, q)-SαS	25
1.3.1	Présentation du modèle	25
1.3.2	Propriétés probabilistes	25
1.3.3	Estimation des paramètres	26
1.4	Les processus ARFIMA(p, d, q)-SαS	29
1.4.1	Présentation du modèle	29
1.4.2	Propriétés probabilistes	29
1.4.3	Estimation des paramètres	31
1.5	Conclusion	33

1.1 Introduction

La propriété de stabilité, le théorème de la limite centrale et la caractérisation par-faite par les moments d'ordre un (la moyenne) et d'ordre deux (la variance ou la covariance) sont des propriétés qui font de la loi gaussienne une des lois les plus utilisées en modélisation statistique. Cependant, bien que les calculs d'inférence statistique soient simples, l'hypothèse de gaussianité s'avère trop restrictive en particulier dans certains domaines pour lesquels il faut prendre en compte une plus grande variabilité des données qui se traduit par une variance infinie. Dans le cadre des distributions non-gaussiennes à variance infinie sont alors apparues les lois α -stables, dont les moments d'ordre 2 sont infinis dès que l'exposant caractéristique α est strictement inférieur à 2. Ces lois sont utilisées dans de nombreux domaines tels que les télécommunications (voir Bestavros et al.[9]), le traitement du signal (voir Nikias et Shao [60]) et la finance (voir Bassi et al.[8]) etc... Elles font partie de la classe des lois de probabilités non-gaussiennes à queue lourde¹ qui englobent d'autres modèles existant dans la littérature et qui ont attiré l'attention de beaucoup de chercheurs en statistique et en séries chronologiques (voir Mikosch et al. [57], Kokoszka et Taqqu [45], Brockwell et Davis [13], Embrechts et al. [23]). Le but de ce chapitre n'est pas de faire une description exhaustive des modèles non-gaussiens, mais d'introduire ceux qui sont particulièrement adaptés à la modélisation des phénomènes de mémoire courte et de mémoire longue. En d'autres termes, nous présentons de façon plus détaillée la famille des distributions α -stables univariées. Ainsi, après un bref rappel historique sur les lois stables, leurs distributions univariées sont définies et diverses propriétés sont présentées dans un premier temps. Puis, sont abordés le problème du test d'une variance finie ou infinie ainsi que l'estimation des quatre paramètres caractérisant une loi α -stable. Dans une deuxième et troisième section, nous présentons deux modèles non-gaussiens à queue algébrique largement utilisés en séries temporelles pour la modélisation des phénomènes de mémoire (courte et longue) : les processus ARMA symétriques α -stables et les processus ARFIMA symétriques α -stables.

1.2 Les lois stables univariées

La théorie des distributions α -stables apparaît dans le cadre de la généralisation du théorème de la limite centrale. En effet, ce théorème permet d'écrire pour une suite de variables aléatoires X_1, X_2, \dots, X_n indépendantes et de même loi, d'espérance et de

¹Formellement, une variable aléatoire réelle a une queue lourde si elle a une queue algébrique : il existent c, α tel que $\mathbb{P}(|X| > x) \sim cx^\alpha$, quand $x \rightarrow \infty$.

variance σ^2 finies, que :

$$\frac{S_n - n\mathbb{E}(X)}{\sigma\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \quad \text{quand } n \rightarrow \infty, \quad (1.1)$$

où $S_n = \sum_{i=1}^n X_i$. Dans le cas où l'espérance et la variance ne sont pas nécessairement finies, est-il possible de trouver une loi limite ? Plus précisément, étant donnée une suite de variables aléatoires indépendantes et identiquement distribuées (i.i.d.), peut-on trouver des suites $a_n \in \mathbb{R}$ et $b_n > 0$, et une loi non dégénérée² caractérisée par sa fonction de répartition H telles que :

$$\frac{S_n - a_n}{b_n} \xrightarrow{\mathcal{D}} H, \quad \text{quand } n \rightarrow \infty. \quad (1.2)$$

En 1924, Paul Lévy [53] a montré que de telles lois existaient, en introduisant pour cela la notion de loi dite *stable*.

1.2.1 Définitions et résultats préliminaires

Dans cette section, nous donnons les quatre définitions équivalentes d'une variable aléatoire stable.

Définition 1.2.1 Une variable aléatoire X est dite stable (ou a une distribution stable) si et seulement si pour tous réels positifs A et B , il existe un réel positif C et un réel D tels que :

$$AX_1 + BX_2 \stackrel{\mathcal{D}}{=} CX + D, \quad (1.3)$$

où les variables aléatoires X_1 et X_2 sont des copies aléatoires indépendantes de X .

La loi de la somme de deux variables aléatoires indépendantes est le produit de convolution des lois respectives de ces deux variables. La Définition 1.2.1 justifie la nomination "stable", car elle fait intervenir la stabilité par produit de convolution. Il est montré dans Feller ([25], pages 170-171) que pour toute variable aléatoire X stable, il existe un réel α , $0 < \alpha \leq 2$, tel que les nombres réels positifs A , B , et C dans l'équation (1.3) vérifient

$$A^\alpha + B^\alpha = C^\alpha. \quad (1.4)$$

Le nombre α est unique et ne dépend que de la loi de la variable aléatoire X , d'où la nomination α -stable. La relation (1.4) peut être généralisée pour un nombre fini de copies indépendantes d'une variable aléatoire α -stable. La définition suivante permet d'expliquer cela.

²Une variable aléatoire est dégénérée si toute la masse de probabilité est concentrée en un seul point.

Définition 1.2.2 Une variable aléatoire X a une distribution stable si et seulement si pour toute suite $(A_n)_{n \in \mathbb{N}^*}$ de nombres réels positifs, il existe un réel positif C_n et un réel D_n tels que :

$$A_1 X_1 + A_2 X_2 + \cdots + A_n X_n \stackrel{\mathcal{D}}{=} C_n X + D_n,$$

où X_1, X_2, \dots, X_n sont des copies aléatoires indépendantes de X .

En utilisant la relation (1.4) et en procédant par récurrence, il est facile de montrer qu'il existe une constante α , $0 < \alpha \leq 2$, telle que $C_n^\alpha = A_1^\alpha + \cdots + A_n^\alpha$ pour tout $n \in \mathbb{N}^*$.

La définition suivante explicite les variables stables en terme du théorème de la limite centrale.

Définition 1.2.3 Une variable aléatoire X est dite stable si et seulement si elle possède un domaine d'attraction, c'est à dire qu'il existe une suite de variables aléatoires Y_1, Y_2, \dots , i.i.d. et des suites $\{a_n\}$ de nombres réels strictement positifs et $\{b_n\}$ de nombres réels telles que

$$\frac{Y_1 + Y_2 + \cdots + Y_n}{a_n} + b_n \xrightarrow{\mathcal{D}} X.$$

La Définition 1.2.3 est équivalente aux Définitions 1.2.1 et 1.2.2. Pour plus de détails voir Feller[25] et Samorodnitsky et Taqqu[71]. Elle permet de définir une variable aléatoire stable comme somme normalisée d'une suite de variables i.i.d.. Remarquons que ce résultat est très utile dans la pratique, car il permet d'approcher une somme normalisée d'une suite de variables aléatoires i.i.d par une variable aléatoire stable. Cela généralise le cas classique du théorème de la limite centrale pour les variables aléatoires de variance finie (cas Gaussien). Ces trois dernières définitions introduisent les variables α -stables d'une manière abstraite et ne donnent aucune précision sur la distribution de probabilité de ces variables. La définition suivante donne la fonction caractéristique de ces variables.

Définition 1.2.4 Une variable aléatoire X est dite avoir une distribution stable si et seulement si, il existe quatre paramètres uniques : $0 < \alpha \leq 2$, $\gamma > 0$, $-1 \leq \beta \leq 1$, et μ réel tels que sa fonction caractéristique s'écrit sous la forme :

$$\Phi_X(t) = \begin{cases} \exp \left\{ i\mu t - \gamma |t|^\alpha \left[1 - i\beta \text{sign}(t) \tan \frac{\pi\alpha}{2} \right] \right\} & \text{si } \alpha \neq 1, \\ \exp \left\{ i\mu t - \gamma |t| \left[1 + i\beta \frac{2}{\pi} \text{sign}(t) \ln |t| \right] \right\} & \text{si } \alpha = 1, \end{cases} \quad (1.5)$$

$$\text{où } \text{sign}(t) = \begin{cases} 1 & \text{si } t > 0, \\ 0 & \text{si } t = 0, \\ -1 & \text{si } t < 0. \end{cases}$$

Une loi stable notée $S_\alpha(\gamma, \beta, \mu)$ est caractérisée par quatre paramètres. Par conséquent, il est important pour un praticien de connaître la signification statistique de chaque paramètre.

1. α : **l'exposant caractéristique**, $0 < \alpha \leq 2$. Il caractérise les queues de distribution en mesurant leurs épaisseurs. C'est pourquoi on parle de distributions α -stables à queue lourde ou à queue épaisse. Quand α est proche de 2, la probabilité d'observer des valeurs de la variable aléatoire loin de la position centrale est faible. Une valeur proche de 0 de l'indice α signifie que la masse de la queue a une probabilité considérable. Le cas $\alpha = 2$ et $\beta = 0$ correspond à la loi normale (loi de Gauss), alors que $\alpha = 1$ et $\beta = 0$ correspond à la loi de Cauchy.
2. μ : **paramètre de position**. Il mesure la tendance centrale de la distribution. Lorsque $1 < \alpha \leq 2$, μ caractérise la moyenne mais il représente la médiane dans le cas où $0 < \alpha < 1$.
3. γ : **paramètre de dispersion**. Son rôle est similaire à celui de l'écart-type dans le cas d'une variable normale.
4. β : **paramètre de symétrie**, $-1 \leq \beta \leq 1$. Si $\beta = 0$, la loi est symétrique par rapport au paramètre de position μ , de fonction caractéristique

$$\Phi_X(t) = \exp \{i\mu t - \gamma |t|^\alpha\}, \quad \forall t \in \mathbb{R}.$$

Dans ce cas, la loi de probabilité est dite α -stable symétrique ou tout simplement $S\alpha S$. Les distributions α -stables symétriques représentent une sous classe importante des distributions α -stables.

Remarque 1.2.5 Par convention, une loi α -stable est dite standard si $\mu = 0$ et $\gamma = 1$. Enfin, il reste à noter aussi qu'il est assez courant dans la littérature de remplacer la dispersion γ par σ^α et d'appeler σ paramètre d'échelle.

Exemple 1.2.6 Certaines des lois connues appartiennent à la classe des lois stables :

1. La loi normale $\mathcal{N}(m, \sigma^2)$ est une loi $S_2\left(\frac{\sigma^2}{2}, \beta, m\right)$ et réciproquement une loi $S_2(\gamma, \beta, \mu)$ est une loi normale $\mathcal{N}(\mu, 2\gamma)$.
2. La distribution symétrique de Cauchy est une loi $S_1(\gamma, 0, \mu)$.
3. La distribution de Lévy est une loi $S_{\frac{1}{2}}(\gamma, 1, \mu)$.
4. La loi de Poisson $\mathcal{P}(\lambda)$ n'est pas stable.

Preuve :

Soient X_1 et X_2 deux variables aléatoires réelles i.i.d. suivant une loi de Poisson de paramètre λ . Supposons que X_1 et X_2 sont stables, alors il existe $a > 0$ et b réel tels que :

$$X_1 + X_2 \stackrel{\mathcal{D}}{=} aX_1 + b.$$

Par égalité des moyennes et des variances, nous pouvons voir que :

$$\begin{cases} 2\lambda = a\lambda + b \\ 2\lambda = a^2\lambda. \end{cases} \quad (1.6)$$

En résolvant le système (1.6), on obtient $a = \sqrt{2}$ et $b = (2 - \sqrt{2})\lambda$. Ce qui entraîne une contradiction car la quantité $X_1 + X_2$ a ses valeurs uniquement dans \mathbb{N} alors que $\sqrt{2}X_1 + (2 - \sqrt{2})\lambda$ n'est pas forcément à valeurs dans \mathbb{N} . ■

1.2.2 Propriétés des lois stables

Dans cette partie, les propriétés les plus importantes des lois α -stables seront présentées.

1.2.2.1 Densité de probabilité

Pour la plupart des lois connues, nous avons une forme explicite de la densité de probabilité (normale, Cauchy, gamma, ...). Pour la loi α -stable, nous n'avons que la forme explicite de la fonction caractéristique. Cependant, on peut obtenir une expression de la densité, sous forme d'une intégrale à l'aide de la transformée de Fourier inverse de la fonction caractéristique Φ_X ,

$$f(x; \alpha, \gamma, \beta, \mu) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-itx) \Phi_X(t) dt.$$

Quand la distribution représentée par cette densité est symétrique ($\beta = 0$) autour de zéro ($\mu = 0$), alors la fonction caractéristique est une fonction réelle et paire, ce qui permet de simplifier l'expression de la densité de probabilité qu'on peut écrire comme suit :

$$f(x; \alpha, \gamma) = \frac{1}{\pi} \int_0^{+\infty} \exp(-\gamma |t|^\alpha) \cos(tx) dt. \quad (1.7)$$

A ce jour l'évaluation explicite de l'intégrale (1.7) n'est possible que dans les trois cas particuliers suivants : distribution de Lévy quand $\alpha = \frac{1}{2}$, distribution de Cauchy quand $\alpha = 1$ et distribution gaussienne quand $\alpha = 2$.

Proposition 1.2.7 (Propriété de la densité)

1. La densité de probabilité vérifie : $f(x; \alpha, \beta) = f(-x; \alpha, -\beta)$
2. La densité de probabilité d'une distribution α -stable est une fonction bornée.
3. La densité de probabilité d'une distribution α -stable est de classe \mathcal{C}^∞ .

Pour la démonstration de la Proposition 1.2.7, voir Zolotarev [80].

1.2.2.2 Stabilité

Proposition 1.2.8 Pour $\alpha \neq 1$, nous avons l'équivalence suivante

$$X \sim S_\alpha(\gamma, \beta, \mu) \iff Y = \frac{X - \mu}{\gamma^{1/\alpha}} \sim S_\alpha(1, \beta, 0)$$

Preuve :

Tout d'abord remarquons que $\Phi_{mX+p}(t) = \mathbb{E}[e^{it(mX+p)}] = e^{itp}\Phi_X(mt)$.

Condition nécessaire : Soit $m = \frac{1}{\gamma^{1/\alpha}}$ et $p = -\frac{\mu}{\gamma^{1/\alpha}}$, alors

$$\begin{aligned} \Phi_Y(t) &= \exp\left(-\frac{i\mu t}{\gamma^{1/\alpha}}\right) \Phi_X\left(\frac{t}{\gamma^{1/\alpha}}\right) \\ &= \exp\left(-\frac{i\mu t}{\gamma^{1/\alpha}}\right) \exp\left\{\frac{i\mu t}{\gamma^{1/\alpha}} - \gamma \left|\frac{t}{\gamma^{1/\alpha}}\right|^\alpha \left[1 - i\beta \operatorname{sign}\left(\frac{t}{\gamma^{1/\alpha}}\right) \tan \frac{\pi\alpha}{2}\right]\right\}. \end{aligned}$$

Or $\operatorname{sign}\left(\frac{t}{\gamma^{1/\alpha}}\right) = \operatorname{sign}(t)$ car $\gamma > 0$, donc

$$\Phi_Y(t) = \exp\left\{-|t|^\alpha \left[1 - i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2}\right]\right\},$$

qui est bien la forme de la fonction caractéristique d'une loi $S_\alpha(1, \beta, 0)$.

Condition suffisante : la démonstration est similaire à la condition nécessaire en prenant $m = \gamma^{1/\alpha}$ et $p = \mu$. ■

1.2.2.3 Comportement de queues lourdes

Définition 1.2.9 La loi de probabilité d'une variable aléatoire est dite à queue lourde (ou "heavy-tailed") d'indice α s'il existe un nombre $\alpha \in]0, 2[$ et une fonction h à variation lente, c'est-à-dire $\lim_{x \rightarrow +\infty} \frac{h(bx)}{h(x)} = 1$, pour tout $b \in \mathbb{R}^+$ tels que :

$$\mathbb{P}(X \geq x) = x^{-\alpha} h(x).$$

Les lois α -stables sont des distributions à queue lourde. En effet, ce résultat est résumé dans la proposition suivante :

Proposition 1.2.10 Soit X une variable aléatoire de loi $S_\alpha(\gamma, \beta, \mu)$ avec $0 < \alpha < 2$, alors on a les deux résultats suivants :

$$\begin{cases} \lim_{t \rightarrow +\infty} t^\alpha \mathbb{P}(X > t) = C_\alpha \frac{1 + \beta}{2} \gamma, \\ \lim_{t \rightarrow +\infty} t^\alpha \mathbb{P}(X < -t) = C_\alpha \frac{1 - \beta}{2} \gamma, \end{cases} \quad (1.8)$$

où C_α est une constante qui ne dépend que de α et est définie par :

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x dx \right)^{-1} = \begin{cases} \frac{1 - \alpha}{\Gamma(2 - \alpha) \cos(\pi\alpha/2)} & \text{si } \alpha \neq 1, \\ \frac{2}{\pi} & \text{si } \alpha = 1. \end{cases} \quad (1.9)$$

La démonstration de cette proposition est détaillée dans Samorodnitsky et Taqqu ([71], pages 16-18).

On rappelle que la fonction Gamma, notée $\Gamma(\cdot)$, est définie, pour tout réel a positif ou nul, par :

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx.$$

L'extension de cette fonction pour des valeurs négatives de a est basée sur le fait que : $\Gamma(a + 1) = a\Gamma(a)$.

1.2.2.4 Calcul des moments

Proposition 1.2.11 Si X suit une loi $S_\alpha(\gamma, \beta, \mu)$, alors

1. si $\alpha = 2$: $\forall p \geq 0$, $\mathbb{E}|X|^p < +\infty$,

2. Si $0 < \alpha < 2$:

$$\begin{cases} \mathbb{E}|X|^p < +\infty & \forall 0 \leq p < \alpha, \\ \mathbb{E}|X|^p = +\infty & \forall p \geq \alpha. \end{cases} \quad (1.10)$$

Preuve :

Remarquons tout d'abord que pour toute variable aléatoire réelle Y positive et intégrable, on a

$$\mathbb{E}(Y) = \int_0^{+\infty} \mathbb{P}(Y > v) dv.$$

En posant $Y = |X|^p$ et $v = u^p$, nous avons

$$\begin{aligned}\mathbb{E}(|X|^p) &= \int_0^{+\infty} \mathbb{P}(|X|^p > u^p) p u^{p-1} du \\ &= \int_0^M \mathbb{P}(|X| > u) p u^{p-1} du + \int_M^{+\infty} \mathbb{P}(|X| > u) p u^{p-1} du\end{aligned}\quad (1.11)$$

$$\text{Or } \lim_{u \rightarrow 0} \frac{u^{p-1} \mathbb{P}(|X| > u)}{u^{p-1}} = \lim_{u \rightarrow 0} \mathbb{P}(|X| > u) = 1,$$

$$\text{donc } u^{p-1} \mathbb{P}(|X| > u) \underset{0}{\sim} u^{p-1}$$

$$\text{et } \int_0^M u^{p-1} \mathbb{P}(|X|^p > u^p) du < +\infty \iff \int_0^M u^{p-1} du < +\infty \iff p > 0$$

$$\text{Remarquons que } \lim_{u \rightarrow +\infty} \frac{u^{p-1} \mathbb{P}(|X| > u)}{u^{p-1-\alpha}} = \lim_{u \rightarrow +\infty} u^\alpha \mathbb{P}(|X| > u) = C_\alpha \gamma,$$

$$\text{donc } u^{p-1} \mathbb{P}(|X| > u) \underset{+\infty}{\sim} u^{p-1-\alpha}$$

$$\text{et } \int_M^{+\infty} u^{p-1} \mathbb{P}(|X|^p > u^p) du < +\infty \iff \int_M^{+\infty} u^{p-1-\alpha} du < +\infty \iff p < \alpha.$$

D'après la relation (1.11), nous avons $\mathbb{E}(|X|^p) < +\infty$ lorsque

$$\left\{ \begin{array}{l} \int_0^M u^{p-1} \mathbb{P}(|X| > u) du < +\infty \\ \int_M^{+\infty} u^{p-1} \mathbb{P}(|X| > u) du < +\infty, \end{array} \right.$$

c'est à dire lorsque $0 < p < \alpha$. ■

1.2.3 Simulation des lois stables

1.2.3.1 Sources des codes

Pour simuler les lois stables, il existe un algorithme développé par Chambers et al. [17]. Celui-ci permet de générer une loi $S_\alpha(1, \beta, 0)$. Pour obtenir une loi $S_\alpha(\gamma, \beta, \mu)$, il suffit de faire un changement de variables, en utilisant la Proposition 1.2.8. Un code en langage **FORTRAN** a été donnée dans Chambers et al. [17] et est ensuite amélioré et publié dans le livre de Samorodnitsky et Taqqu [71]. Il existe aussi une fonction **rstab** dans la bibliothèque de **S-Plus**. Pour générer des lois α -stables sous le logiciel R, on peut utiliser la fonction **rstable** disponible dans le package **fBasics**.

1.2.3.2 Quelques exemples

Nous avons simulé 5000 réalisations de lois $S\alpha S$ standard pour différentes valeurs de α . Le Tableau 1.1 suivant représente la moyenne et la variance empirique des 5000 réalisations.

Valeurs de α	2	1.7	1.5	1.2	1	0.9
Moyenne	-0.0098	-0.0005	0.0150	0.1109	0.4174	1.0815
Variance	1.9732	5.6939	16.838	204.00	807.29	16974

TAB. 1.1 – Moyenne et variance empirique calculées sur 5000 réalisations.

Ces résultats confirment l'équation sur le calcul des moments. En effet, lorsque α décroît vers 1, la variance diverge et lorsque α devient plus petit que 1, c'est la moyenne qui commence à exploser.

1.2.4 Inférence Statistique des lois stables

Dans cette partie, nous allons mettre l'accent sur les différentes manières de tester et d'estimer les paramètres d'une loi α -stable. Pour les problèmes de test de variance ou d'estimation des paramètres, nous considérons un échantillon (X_1, \dots, X_n) extrait de la variable aléatoire réelle X définie par sa fonction caractéristique (1.5).

1.2.4.1 Test graphique de la variance

Nous allons présenter deux méthodes graphiques pour tester si la distribution de nos observations est à variance finie ou infinie.

a)- Test graphique de la convergence de la variance empirique.

Ce premier test est très simple et se décompose en trois étapes :

Étape 1 : Calcul de la moyenne empirique $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.

Étape 2 : Calcul de la variance empirique $s_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$.

Étape 3 : Visualisation de la courbe (n, s_n^2) pour des valeurs de n assez grandes.

Intuitivement, si le Processus Générateur des Données (PGD) a une variance finie, alors peu d'observations proviennent de la queue et donc la variance empirique s_n^2 converge vers une valeur finie, lorsque la taille des observations n devient assez grande. Par contre, si le PGD est à variance infinie, alors beaucoup d'observations viennent de la queue et par conséquent $s_n^2 \rightarrow +\infty$, lorsque $n \rightarrow +\infty$. En d'autres termes, lorsque n augmente et lorsque la variance est finie, le tracé doit converger et si on est en présence d'une variance infinie, le tracé diverge.

b)- Test graphique de la queue

L'idée principale de ce deuxième test est basée sur le fait que :

$$\lim_{t \rightarrow +\infty} t^\alpha \mathbb{P}(|X| > t) = C_\alpha \gamma,$$

où la constante C_α est donnée par l'équation (1.9). Alors, cela implique :

$$\frac{d \mathbb{P}(|X| > t)}{d \log(t)} \sim -\alpha, \text{ quand } t \rightarrow +\infty.$$

A ce niveau également, deux étapes sont nécessaires :

Etape 1 : Fixer x et calculer le logarithme de la queue

$$q(x) = \log \left(\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{|X_i| > x} \right).$$

Etape 2 : Tracer la courbe $(\log(x), q(x))$ et voir si la pente est finie à partir d'une certaine valeur de x .

c)- Quelques exemples

Pour cela, nous avons repris les simulations de la section précédente. Les graphiques de la Figure 1.1 représentent le test graphique de la convergence de la variance empirique (Test graphique 1). Ceux de la Figure 1.2 résument le test graphique de la queue (Test graphique 2). Nous pouvons remarquer que seul le cas α égal à 2 fait converger la variance. Pour tous les autres cas, il y a divergence. Nous constatons également, sur la Figure 1.1, que plus on s'éloigne du cas gaussien ($\alpha = 2$), plus la variance explose. Ce résultat confirme celui observé sur le Tableau 1.1.

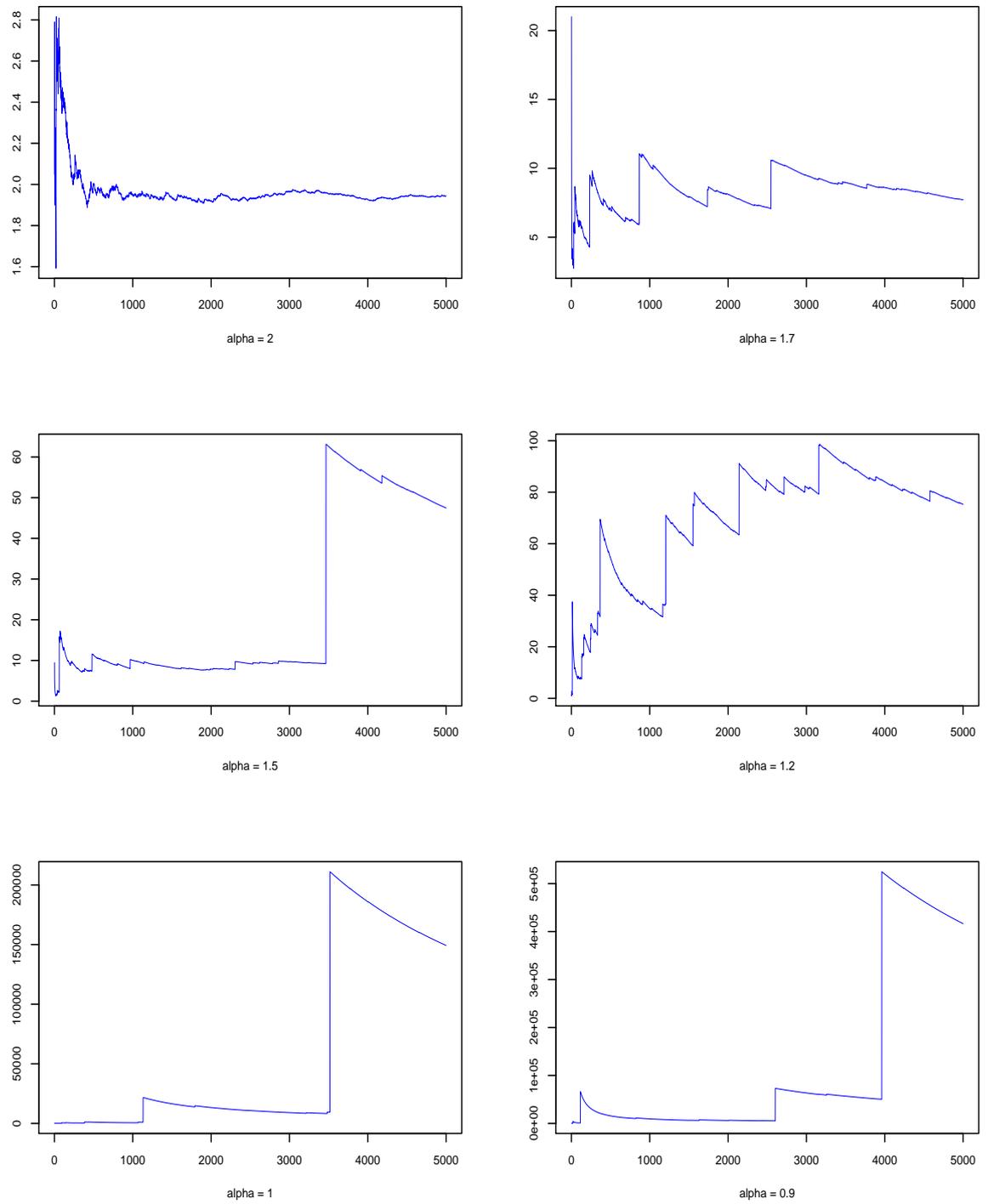


FIG. 1.1 – Test graphique 1 : variance empirique calculée pas à pas sur l'échantillon de 5000 réalisations d'une loi SaS standard pour différentes valeurs de α .

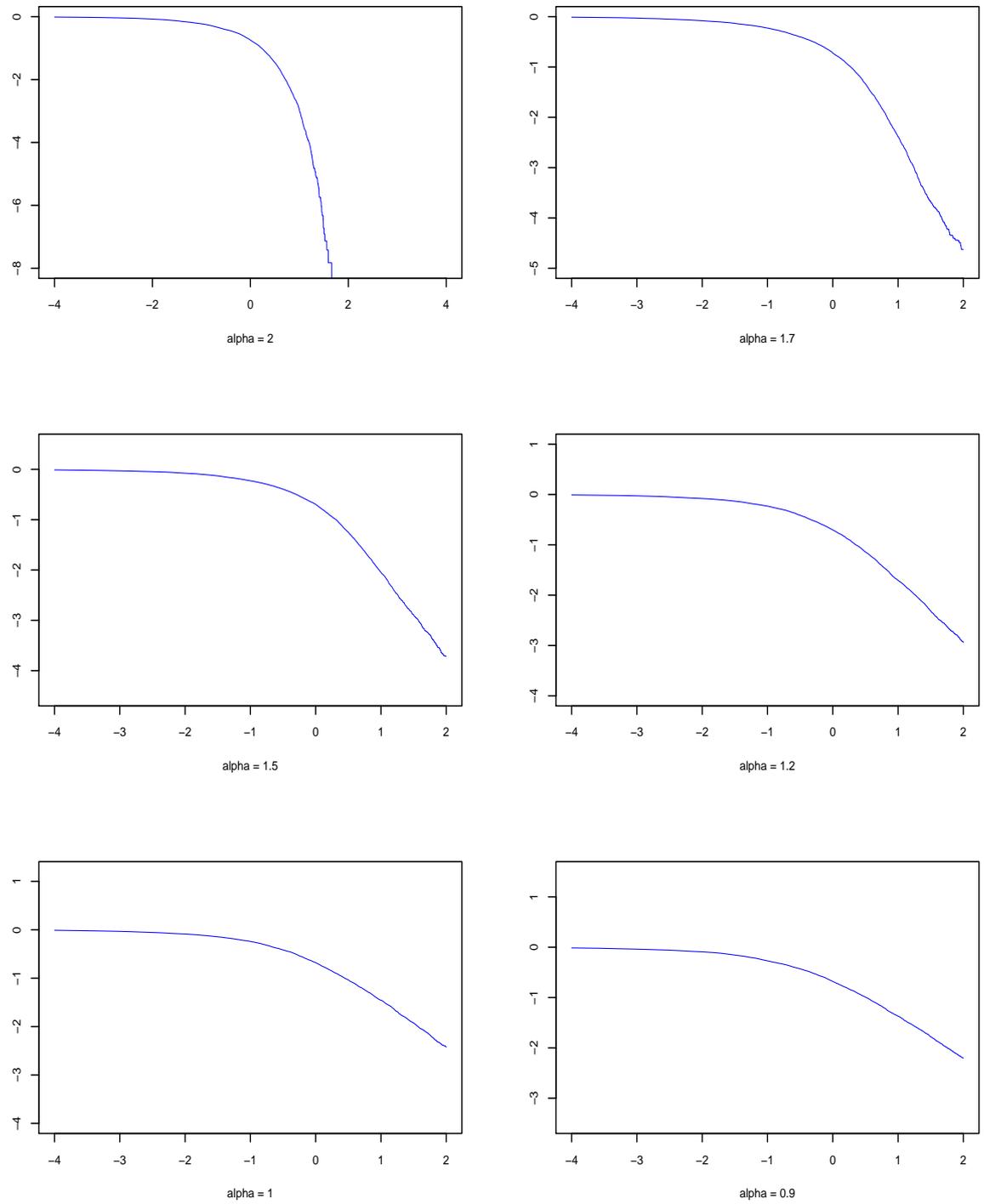


FIG. 1.2 – Test graphique 2 calculé sur l'échantillon de 5000 réalisations d'une loi $S\alpha S$ standard pour différentes valeurs de α .

1.2.4.2 Estimation des paramètres des lois α -stables

La plupart des algorithmes de traitement de données utilisant des lois α -stables exigent l'estimation a priori des paramètres de la distribution α -stable et en particulier l'exposant caractéristique α . C'est le paramètre le plus controversé car c'est lui qui indique si la variance est infinie ou non. Il semble donc primordial de pouvoir l'estimer de la manière la plus correcte qui soit. Il existe plusieurs méthodes d'estimation des paramètres d'une distribution α -stable. Alvarez et Olivares [4] proposent une revue des différentes méthodes d'estimation dont les méthodes de queue, quantiles, moments, maximum de vraisemblance et de fonction caractéristique et comparent leur performance par simulations de Monte Carlo. Après l'analyse des résultats de l'étude des simulations, ils arrivent à la conclusion que le meilleur compromis entre précision de l'estimateur et vitesse est soit la méthode de McCulloch soit la méthode de régression itérative; avec comme préférence la première. Ainsi, dans cette partie nous décrivons la méthode de McCulloch et nous renvoyons le lecteur à l'article original de McCulloch [56] ou à Alvarez et Olivares [4], pour plus de détails.

On suppose que l'on dispose d'un échantillon X_1, X_2, \dots, X_n de variables aléatoires stables, indépendantes et identiquement distribuées, avec $X_i \sim S_\alpha(\gamma, \beta, \mu)$. Soit F la fonction de répartition de la variable aléatoire stable X et soit x_p le quantile d'ordre p , c'est-à-dire, $F(x_p) = p$, et \hat{x}_p le quantile empirique correspondant. Pour éviter une fausse asymétrie des petits échantillons, une correction est nécessaire : si les x_i sont ordonnées de façon croissante cette correction se fait en posant $\hat{x}_{q(i)} = x_i$ où $q(i) = \frac{2i-1}{2n}$. Puis, on effectue une interpolation linéaire pour obtenir \hat{x}_p à partir de $\hat{x}_{q(i)}$ et $\hat{x}_{q(i+1)}$, où $q(i) \leq p \leq q(i+1)$. L'estimateur obtenu \hat{x}_p est un estimateur convergent de x_p .

En considérant les quatre tableaux suivants³, l'algorithme d'estimation des quatre paramètres peut être défini comme ci-dessous.

Procédure concrète de l'algorithme

1. Ordonner l'échantillon.
2. Calculer les quantiles empiriques $\hat{x}_{0.05}$, $\hat{x}_{0.25}$, $\hat{x}_{0.5}$, $\hat{x}_{0.75}$, $\hat{x}_{0.95}$.
3. Calculer les indices $\hat{\nu}_\alpha$ et $\hat{\nu}_\beta$:

$$\hat{\nu}_\alpha = \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}} \quad \hat{\nu}_\beta = \frac{\hat{x}_{0.95} + \hat{x}_{0.05} - 2\hat{x}_{0.5}}{\hat{x}_{0.95} - \hat{x}_{0.05}}$$

McCulloch montre que les indices ν_α et ν_β ne dépendent pas de γ ni de μ . De plus, ils sont respectivement des fonctions décroissante et croissante de α et de β .

³Les quatre tableaux (TABLE III, TABLE IV, TABLE V, TABLE VII) sont scannés et proviennent de l'article de McCulloch [56].

4. Estimer α et β :

$$\hat{\alpha} = \psi_1(\hat{\nu}_\alpha, \hat{\nu}_\beta) \quad \hat{\beta} = \psi_2(\hat{\nu}_\alpha, \hat{\nu}_\beta),$$

où les fonctions ψ_1 et ψ_2 sont respectivement déterminées par les tableaux TABLE III et TABLE IV, pour $\hat{\nu}_\alpha$ et $\hat{\nu}_\beta$ données.

5. Estimer γ

$$\hat{\gamma} = \frac{\hat{x}_{0.75} - \hat{x}_{0.25}}{\phi_3(\hat{\alpha}, \hat{\beta})},$$

avec ϕ_3 déterminée par le tableau TABLE V.

6. Estimer ζ

$$\hat{\zeta} = \hat{x}_{0.5} + \hat{\gamma}\phi_5(\hat{\alpha}, \hat{\beta}),$$

ϕ_5 est donnée par le tableau TABLE VII.

7. Estimer μ

$$\hat{\mu} = \hat{\zeta} - \hat{\beta} \hat{\gamma} \tan \frac{\pi \hat{\alpha}}{2}.$$

Caractéristiques de la méthode

Comme \hat{x}_p est un estimateur convergent et asymptotiquement de loi normale de x_p et les fonctions ψ_i et ϕ_i sont continues, alors les estimateurs des paramètres sont convergents et asymptotiquement de loi normale (voir Garcia et al.[28] pour l'analyse de l'efficacité de cette méthode). Il est à noter que, sans considérer le travail préliminaire de génération des tableaux (TABLE III, TABLE IV, TABLE V, TABLE VII), la méthode de McCulloch est facile à implémenter et sa complexité est de l'ordre de $O(n \log(n))$.

Quelques exemples

Nous allons illustrer dans cet exemple les différentes étapes de la méthode de McCulloch sur des données simulées. Pour cela, nous générons un échantillon de taille $n = 1000$ d'une variable aléatoire $X \sim S_\alpha(\gamma, \beta, \mu)$, avec $\alpha = 1.5$, $\beta = 0$, $\gamma = 1$ et $\mu = 0$.

Le calcul des quantiles donne :

$$\hat{x}_{0.05} = -3.3564, \quad \hat{x}_{0.25} = -0.9150, \quad \hat{x}_{0.5} = -0.0341, \quad \hat{x}_{0.75} = 0.9804, \quad \hat{x}_{0.95} = 3.091.$$

On en déduit : $\hat{\nu}_\alpha = 3.4016$ et $\hat{\nu}_\beta = -0.0304$.

En utilisant les tableaux III et IV et en faisant une interpolation, on obtient :

$$\hat{\alpha} = 1.4478 \quad \hat{\beta} = 0.0509.$$

Pour calculer la valeur de $\hat{\gamma}$, on détermine d'abord la valeur de $\phi_3(\hat{\alpha}, \hat{\beta})$ qui s'obtient par interpolation linéaire dans le tableau V. Ainsi, on a $\phi_3(\hat{\alpha}, \hat{\beta}) = 1.9457$ et par conséquent $\hat{\gamma} = 0.9741$. De manière similaire, on obtient la valeur de $\phi_5(\hat{\alpha}, \hat{\beta}) = 0.0143$ à partir du tableau VII. Par suite, $\hat{\zeta} = -0.0201$ et $\hat{\mu} = 0.0383$.

TABLE III

$$\alpha = \psi_1(v_\alpha, v_\beta)$$

v_α	v_β						
	0.0	0.1	0.2	0.3	0.5	0.7	1.0
2.439	2.0	2.0	2.0	2.0	2.0	2.0	2.0
2.5	1.916	1.924	1.924	1.924	1.924	1.924	1.924
2.6	1.808	1.813	1.829	1.829	1.829	1.829	1.829
2.7	1.729	1.730	1.737	1.745	1.745	1.745	1.745
2.8	1.664	1.663	1.663	1.668	1.676	1.676	1.676
3.0	1.563	1.560	1.553	1.548	1.547	1.547	1.547
3.2	1.484	1.480	1.471	1.460	1.448	1.438	1.438
3.5	1.391	1.386	1.378	1.364	1.337	1.318	1.318
4.0	1.279	1.273	1.266	1.250	1.210	1.184	1.150
5.0	1.128	1.121	1.114	1.101	1.067	1.027	0.973
6.0	1.029	1.021	1.014	1.004	0.974	0.935	0.874
8.0	0.896	0.892	0.887	0.883	0.855	0.823	0.769
10.0	0.818	0.812	0.806	0.801	0.780	0.756	0.691
15.0	0.698	0.695	0.692	0.689	0.676	0.656	0.595
25.0	0.593	0.590	0.588	0.586	0.579	0.563	0.513

Note that $\psi_1(v_\alpha, -v_\beta) = \psi_1(v_\alpha, v_\beta)$.

TABLE IV

$$\beta = \psi_2(v_\alpha, v_\beta)$$

v_α	v_β						
	0.0	0.1	0.2	0.3	0.5	0.7	1.0
2.439	0.0	2.160	1.0	1.0	1.0	1.0	1.0
2.5	0.0	1.592	3.390	1.0	1.0	1.0	1.0
2.6	0.0	0.759	1.800	1.0	1.0	1.0	1.0
2.7	0.0	0.482	1.048	1.694	1.0	1.0	1.0
2.8	0.0	0.360	0.760	1.232	2.229	1.0	1.0
3.0	0.0	0.253	0.518	0.823	1.575	1.0	1.0
3.2	0.0	0.203	0.410	0.632	1.244	1.906	1.0
3.5	0.0	0.165	0.332	0.499	0.943	1.560	1.0
4.0	0.0	0.136	0.271	0.404	0.689	1.230	2.195
5.0	0.0	0.109	0.216	0.323	0.539	0.827	1.917
6.0	0.0	0.096	0.190	0.284	0.472	0.693	1.759
8.0	0.0	0.082	0.163	0.243	0.412	0.601	1.596
10.0	0.0	0.074	0.147	0.220	0.377	0.546	1.482
15.0	0.0	0.064	0.128	0.191	0.330	0.478	1.362
25.0	0.0	0.056	0.112	0.167	0.285	0.428	1.274

Note that $\psi_2(v_\alpha, -v_\beta) = -\psi_2(v_\alpha, v_\beta)$. Entries in this table greater than 1.0 are required in order to permit accurate bivariate linear interpolation as β approaches 1.0 from below. As a result, sampling error in finite samples may yield an interpolated estimate of β greater than 1.0. In this case, the estimate should be truncated back to 1.0.

TABLE V

$$v_c = \phi_3(\alpha, \beta).$$

α	β				
	0.0	0.25	0.50	0.75	1.00
2.00	1.908	1.908	1.908	1.908	1.908
1.90	1.914	1.915	1.916	1.918	1.921
1.80	1.921	1.922	1.927	1.936	1.947
1.70	1.927	1.930	1.943	1.961	1.987
1.60	1.933	1.940	1.962	1.997	2.043
1.50	1.939	1.952	1.988	2.045	2.116
1.40	1.946	1.967	2.022	2.106	2.211
1.30	1.955	1.984	2.067	2.188	2.333
1.20	1.965	2.007	2.125	2.294	2.491
1.10	1.980	2.040	2.205	2.435	2.696
1.00	2.000	2.085	2.311	2.624	2.973
0.90	2.040	2.149	2.461	2.886	3.356
0.80	2.098	2.244	2.676	3.265	3.912
0.70	2.189	2.392	3.004	3.844	4.775
0.60	2.337	2.635	3.542	4.808	6.247
0.50	2.588	3.073	4.534	6.636	9.144

Note that $\phi_3(\alpha, -\beta) = \phi_3(\alpha, \beta)$.

TABLE VII

$$v_z = \phi_5(\alpha, \beta).$$

α	β				
	0.0	0.25	0.50	0.75	1.00
2.00	0.0	0.0	0.0	0.0	0.0
1.90	0.0	-0.017	-0.032	-0.049	-0.064
1.80	0.0	-0.030	-0.061	-0.092	-0.123
1.70	0.0	-0.043	-0.088	-0.132	-0.179
1.60	0.0	-0.056	-0.111	-0.170	-0.232
1.50	0.0	-0.066	-0.134	-0.206	-0.283
1.40	0.0	-0.075	-0.154	-0.241	-0.335
1.30	0.0	-0.084	-0.173	-0.276	-0.390
1.20	0.0	-0.090	-0.192	-0.310	-0.447
1.10	0.0	-0.095	-0.208	-0.346	-0.508
1.00	0.0	-0.098	-0.223	-0.383	-0.576
0.90	0.0	-0.099	-0.237	-0.424	-0.652
0.80	0.0	-0.096	-0.250	-0.469	-0.742
0.70	0.0	-0.089	-0.262	-0.520	-0.853
0.60	0.0	-0.078	-0.272	-0.581	-0.997
0.50	0.0	-0.061	-0.279	-0.659	-1.198

Note that $\phi_5(\alpha, -\beta) = -\phi_5(\alpha, \beta)$.

1.3 Les processus ARMA(p, q)-SαS

Dans ce paragraphe, on définit les processus ARMA(p, q) symétriques α -stables introduits par Brockwell et Davis [13], dans le but de prendre en compte dans une modélisation la présence de composantes à mémoire courte et de variance infinie. On s'intéresse ensuite aux propriétés probabilistes et statistiques de ce modèle.

1.3.1 Présentation du modèle

Définition 1.3.1 Le processus $(X_t)_{t \in \mathbb{Z}}$ est appelé processus ARMA(p, q)-SαS, s'il vérifie l'équation suivante :

$$\Phi(B) X_t = \Theta(B) Z_t, \quad (1.12)$$

où les $(Z_t)_{t \in \mathbb{Z}}$ sont indépendants et de même loi symétrique α -stable de paramètre de dispersion γ pris égal à 1 sans perte de généralités et B est l'opérateur retard, c'est à dire $BX_t = X_{t-1}$. Les polynômes $\Phi(B)$ et $\Theta(B)$ sont respectivement les polynômes autorégressifs d'ordre p et moyenne mobile d'ordre q et sont définis par :

$$\Phi(B) = 1 - \sum_{i=1}^p \phi_i B^i \quad \text{et} \quad \Theta(B) = 1 + \sum_{i=1}^q \theta_i B^i. \quad (1.13)$$

On note que si $\alpha = 2$, alors le processus défini en (1.12) est un processus ARMA(p, q) gaussien tel qu'il est décrit par Box et Jenkins [11].

1.3.2 Propriétés probabilistes

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

Définition 1.3.2 Le processus ARMA(p, q)-SαS $(X_t)_{t \in \mathbb{Z}}$ est dit causal s'il existe une suite $(\psi_j)_{j=0, \dots, +\infty}$ tel que $\sum_{j=0}^{+\infty} \psi_j < \infty$ et

$$X_t = \sum_{j=0}^{+\infty} \psi_j Z_{t-j} \quad (1.14)$$

Proposition 1.3.3 Supposons que les polynômes $\Phi(z)$ et $\Theta(z)$ n'ont pas de racines communes dans le disque unité $\{z : |z| \leq 1\}$. Alors :

i) si le polynôme $\Phi(z)$ n'a pas de racines dans le disque unité, alors le processus

ARMA(p, q)-SαS $(X_t)_{t \in \mathbb{Z}}$ est causal et stationnaire, et admet une unique représentation moyenne mobile infinie donnée par l'équation (1.14), où les coefficients $(\psi_j)_{j \geq 0}$ sont définis par :

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

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

$$Z_t = \sum_{j=0}^{+\infty} \pi_j X_{t-j}, \quad (1.15)$$

où les coefficients π_j sont déterminés par

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

La démonstration est détaillée dans Brockwell et Davis ([13], page 537).

1.3.3 Estimation des paramètres

Dans ce paragraphe, on s'intéresse à l'estimation des paramètres d'un processus ARMA(p, q)-SαS. On présente la méthode de Whittle introduite par Mickosch et al [57] pour ce type de processus.

On considère un échantillon (X_1, \dots, X_n) de taille finie n extrait d'un processus ARMA(p, q)-SαS, centré, stationnaire et inversible. On suppose que les paramètres sur la distribution des innovations (α et γ) sont connus et on s'intéresse à l'estimation du paramètre vectoriel $\beta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ composé des paramètres autorégressifs et moyenne mobile. L'espace des paramètres à estimer est alors :

$$C = \{\beta \in \mathbb{R}^{p+q} : \phi_p \neq 0, \theta_q \neq 0, \Theta(z) \text{ et } \Phi(z) \text{ n'ont pas de racines}$$

$$\text{communes, } \Phi(z)\Theta(z) \neq 0 \text{ pour } |z| \leq 1\}.$$

Supposons que β_0 est la vraie valeur du paramètre β et se trouve à l'intérieur du compact C . Ainsi on appelle estimateur de Whittle, noté β_n , la valeur de β qui minimise la vraisemblance de Whittle :

$$\sigma_n^2(\beta) = \int_{-\pi}^{\pi} \frac{\tilde{I}_n(\lambda)}{g(\lambda, \beta)} d\lambda, \quad (1.16)$$

où $\tilde{I}_n(\lambda)$ représente le périodogramme normalisé du processus $(X_t)_{t \in \mathbb{Z}}$ et est défini par :

$$\tilde{I}_n(\lambda) = \left(\sum_{t=1}^n X_t^2 \right)^{-1} \left| \sum_{t=1}^n X_t e^{-i\lambda t} \right|^2, \quad -\pi < \lambda \leq \pi. \quad (1.17)$$

La fonction $g(\lambda, \beta)$ est la fonction de transfert définie par :

$$g(\lambda, \psi) = \frac{|\Theta(e^{-i\lambda})|^2}{|\Phi(e^{-i\lambda})|^2} = \frac{\left| 1 + \sum_{k=1}^q \theta_k e^{-i\lambda k} \right|^2}{\left| 1 - \sum_{k=1}^p \phi_k e^{-i\lambda k} \right|^2}, \quad -\pi < \lambda \leq \pi. \quad (1.18)$$

Le choix de l'estimateur de Whittle est motivé par deux faits. Le premier est que la fonction $\int_{-\pi}^{\pi} \frac{g(\lambda, \beta_0)}{g(\lambda, \beta)} d\lambda$ a un minimum absolu au point $\beta = \beta_0$ dans le compact C (voir Brockwell et Davis[13], Proposition 10.8.1). En plus, d'après Klüppelberg et Mikosch[40], $\tilde{I}_n(\lambda)$ est un estimateur de $g(\lambda, \beta_0)/\Psi^2(\beta_0)$, avec $\Psi^2(\beta_0)$ est la quantité donnée par :

$$\Psi^2(\beta_0) = \sum_{j=0}^{+\infty} \psi_j,$$

où les coefficients ψ_j sont définis dans la Proposition 1.3.3.

En pratique, l'intégrale (1.16) est souvent remplacée par une somme. Ceci donne un estimateur, noté $\bar{\beta}_n$, asymptotiquement équivalent à β_n et qui est obtenu en minimisant sur le compact C la fonction objective suivante :

$$\bar{\sigma}_n^2(\beta) = \frac{2\pi}{n} \sum_{j=0}^{n-1} \frac{\tilde{I}_n(\lambda_j)}{g(\lambda_j, \beta)}, \quad (1.19)$$

où les λ_j sont des fréquences de Fourier définies par $\lambda_j = \frac{2\pi j}{n}$, $j = 0, \dots, n-1$.

Ainsi, si on suppose que $\beta_0 \in C$ est la vraie valeur des paramètres à estimer, alors les deux estimateurs naturels de β_0 sont donnés par :

$$\beta_n = \arg \min_{\beta \in C} \sigma_n^2(\beta), \quad \bar{\beta}_n = \arg \min_{\beta \in C} \bar{\sigma}_n^2(\beta).$$

Nous résumons dans le Théorème 1.3.8 suivant les propriétés asymptotiques de l'estimateur de Whittle. Nous précisons tout d'abord quelques hypothèses faites sur le processus d'innovation $(Z_t)_{t \in \mathbb{Z}}$, pour tout $b > 0$ et $a_n > 0$ tel que $a_n \uparrow \infty$:

Hypothèse 1.3.4 : $\mathbb{E}|Z_1|^b < \infty$

Hypothèse 1.3.5 : $\frac{n}{a_n^{2\delta}} \rightarrow 0$, quand $n \rightarrow +\infty$, et pour $\delta = 1 \wedge b$,

Hypothèse 1.3.6 : $\lim_{x \rightarrow 0} \limsup_{n \rightarrow +\infty} \mathbb{P} \left(a_n^{-2} \sum_{t=1}^n Z_t^2 \leq x \right) = 0$.

Ces trois hypothèses permettent d'établir la consistance de l'estimateur de Whittle. Cependant, pour obtenir des informations sur sa distribution asymptotique, nous avons besoin de l'hypothèse supplémentaire suivante :

Hypothèse 1.3.7 : Nous supposons que Z_1 appartient au domaine d'attraction d'une loi symétrique α -stable; c'est à dire :

$$n^{-1/\alpha} \sum_{t=1}^n Z_t^2 \xrightarrow{\mathcal{D}} Y, \quad (1.20)$$

où Y suit une loi symétrique α -stable.

Théorème 1.3.8 Supposons que $(X_t)_{t \in \mathbb{Z}}$ est un processus ARMA(p, q) α -stable causal et inversible introduit par la Définition 1.3.1. Alors :

i) si les Hypothèses "**Hypothèse 1.3.4**"-"**Hypothèse 1.3.6**" sont vérifiées, alors :

$$\beta_n \xrightarrow{\mathbb{P}} \beta_0 \quad \text{et} \quad \sigma_n^2(\beta_n) \xrightarrow{\mathbb{P}} 2\pi \Psi^{-2}(\beta_0), \quad \text{quand } n \rightarrow \infty. \quad (1.21)$$

ii) si de plus, l'**Hypothèse 1.3.7** est satisfaite, alors :

$$\left(\frac{n}{\ln n} \right)^{1/\alpha} (\beta_n - \beta_0) \xrightarrow{\mathcal{D}} 4\pi W^{-1}(\beta_0) \sum_{k=1}^{\infty} b_k \frac{Y_k}{Y_0}, \quad \text{quand } n \rightarrow \infty, \quad (1.22)$$

où Y_0, Y_1, Y_2, \dots sont des variables aléatoires indépendantes; Y_0 est positive $\frac{\alpha}{2}$ -stable de paramètre d'échelle $C_{\alpha/2}^{-2/\alpha}$ et $(Y_k)_{k \geq 1}$ est une suite de variables aléatoires indépendantes et de même loi symétrique α -stable de paramètre de dispersion $C_\alpha^{1/\alpha}$. La constante C_α est donnée par l'équation (1.9). La matrice $W^{-1}(\beta_0)$ est l'inverse de la matrice $W(\beta_0)$ définie par :

$$W(\beta_0) = \int_{-\pi}^{\pi} \left[\frac{\partial \ln g(\lambda, \beta_0)}{\partial \beta} \right] \left[\frac{\partial \ln g(\lambda, \beta_0)}{\partial \beta} \right]^T d\lambda,$$

et pour tout $k \in \mathbb{N}$, b_k est le vecteur défini par :

$$b_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} g(\lambda, \beta_0) \frac{\partial g^{-1}(\lambda, \beta_0)}{\partial \beta} d\lambda. \quad \blacksquare$$

La preuve de ce théorème est détaillée dans Mikosch et al.[57].

Remarque 1.3.9 Les relations (1.21) et (1.22) restent encore vraies si on remplace β_n et σ_n^2 par $\bar{\beta}_n$ et $\bar{\sigma}_n^2$.

1.4 Les processus ARFIMA(p, d, q)-S α S

Les processus ARFIMA(p, d, q) symétriques α -stables ($0 < \alpha \leq 2$) ont été introduits dans la littérature statistique par Kokoszka et Taqqu [45] afin de prendre en compte, dans une modélisation, la présence éventuelle d'une composante longue mémoire et d'une composante de variance infinie dans la série étudiée. Ces processus constituent une généralisation des processus ARFIMA gaussiens (cas $\alpha = 2$) proposés par Granger et Joyeux [31] et Hosking [37].

Dans ce paragraphe, nous allons rappeler quelques résultats sur les propriétés probabilistes et statistiques des processus ARFIMA(p, d, q)-S α S. On rappelle premièrement la définition d'un processus ARFIMA(p, d, q)-S α S et deuxièmement nous posons notre regard sur les problèmes d'estimation des paramètres de ce processus.

1.4.1 Présentation du modèle

Définition 1.4.1 Soit $(Z_t)_{t \in \mathbb{Z}}$ une suite de variables aléatoires i.i.d suivant une loi symétrique α -stable, de moyenne nulle et de paramètre de dispersion $\gamma = 1$, sans perte de généralité. On dit que le processus $(X_t)_{t \in \mathbb{Z}}$ est un processus ARFIMA(p, d, q)-S α S s'il vérifie l'équation suivante :

$$\Phi(B) X_t = \Theta(B) (1 - B)^{-d} Z_t, \quad (1.23)$$

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 linéaire dans l'équation (1.23) peut s'écrire de la manière suivante :

$$(1 - B)^{-d} Z_t = \sum_{j=0}^{\infty} b_j Z_{t-j}, \quad (1.24)$$

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

$$b_0 = 1 \quad b_j = \frac{\Gamma(j + d)}{\Gamma(d) \Gamma(j + 1)}, \quad j \geq 1. \quad (1.25)$$

1.4.2 Propriétés probabilistes

Nous nous intéressons maintenant aux conditions d'existence et d'inversibilité d'un processus ARFIMA(p, d, q)-S α S. Nous rappelons d'abord quelques hypothèses.

Hypothèse 1.4.2 : Le polynôme $\Phi(z)$ n'admet pas de racines dans le disque unité.

Hypothèse 1.4.3 : $0 < \alpha \leq 2$ et $\alpha(d-1) < -1$.

Hypothèse 1.4.4 : Le polynôme $\Theta(z)$ ne s'annule pas dans le disque unité.

Hypothèse 1.4.5 : $1 < \alpha \leq 2$ et $|d| < 1 - \frac{1}{\alpha}$.

Les résultats relatifs à la stationnarité et à l'inversibilité d'un processus ARFIMA(p, d, q)-SαS sont résumés par la Proposition 1.4.6 suivante :

Proposition 1.4.6 Supposons que les polynômes $\Phi(z)$ et $\Theta(z)$ n'ont pas de racines communes dans le disque unité $\{z : |z| \leq 1\}$. Alors :

i) si les Hypothèses "**Hypothèse 1.4.2**" et "**Hypothèse 1.4.3**" sont vraies, alors le processus ARFIMA(p, d, q)-SαS $(X_t)_{t \in \mathbb{Z}}$ est stationnaire, causal et admet une unique représentation moyenne mobile infinie donnée par :

$$X_t = \sum_{j=0}^{\infty} c_j Z_{t-j}, \quad (1.26)$$

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

$$\frac{\Theta(z)}{\Phi(z)} (1-z)^{-d} = \sum_{j=0}^{\infty} c_j z^j, \quad \forall |z| < 1, \quad (1.27)$$

ii) si de plus, les Hypothèses "**Hypothèse 1.4.4**" et "**Hypothèse 1.4.5**" sont vérifiées, alors le processus ARFIMA(p, d, q)-SαS $(X_t)_{t \in \mathbb{Z}}$ est inversible et admet une représentation autorégressive infinie donnée par

$$Z_t = \sum_{j=0}^{\infty} \tilde{c}_j X_{t-j}, \quad \text{p.s.}, \quad (1.28)$$

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

$$\sum_{j=0}^{\infty} \tilde{c}_j z^j = (1-z)^d \frac{\Phi(z)}{\Theta(z)}, \quad \forall |z| < 1. \quad (1.29)$$

■

La preuve de cette proposition se trouve dans Kokoszka et Taqqu [45].

Remarque 1.4.7 Un développement de l'équation (1.27) permet de calculer les coefficients $(c_j)_{j \geq 0}$ par la formule récursive suivante :

$$c_0 = 1 \quad \text{et} \quad c_j = b_j + \sum_{i=1}^{\min(j,q)} \theta_i b_{j-i} + \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i}, \quad j \geq 1. \quad (1.30)$$

1.4.3 Estimation des paramètres

Dans cette section, nous présentons l'estimation des paramètres d'un processus ARFIMA(p, d, q)- SαS par la méthode de Whittle proposée par Kokoszka et Taqqu [46]. Nous donnons ensuite les propriétés asymptotiques de cet estimateur. On considère une suite (X_1, \dots, X_n) de valeurs issue d'un processus ARFIMA(p, d, q)- SαS, $(X_t)_{t \in \mathbb{Z}}$ stationnaire, inversible et de moyenne nulle. La moyenne étant fixée, on suppose les paramètres de la distribution des innovations connus et on s'intéresse à l'estimation du paramètre $\psi = (d, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$. L'espace des paramètres d'intérêt, noté Ψ , est défini comme suit

$$\Psi = \{ \psi \in \mathbb{R}^{p+q+1} : \phi_p \neq 0, \theta_q \neq 0, \Theta(z) \text{ et } \Phi(z) \text{ n'ont pas de racines communes, } \Phi(z)\Theta(z) \neq 0 \text{ pour } |z| \leq 1, d \in]0, 1 - \frac{1}{\alpha}[, 1 < \alpha \leq 2 \}.$$

Supposons que ψ_0 est la vraie valeur du paramètre ψ et se trouve à l'intérieur du compact Ψ . L'estimateur par la méthode de Whittle ψ_n de ψ sur Ψ minimise la fonction objective donnée par :

$$\sigma_n^2(\psi) = \int_{-\pi}^{\pi} \frac{\tilde{I}_n(\lambda)}{f(\lambda, \psi)} d\lambda, \quad (1.31)$$

où la fonction $\tilde{I}_n(\lambda)$ représente le périodogramme normalisé défini par l'équation (1.17) et $f(\lambda, \psi)$ est la fonction de transfert définie de la manière suivante :

$$f(\lambda, \psi) = \left| \frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})(1 - e^{-i\lambda})^d} \right|^2 = \left| \sum_{j=0}^{\infty} c_j e^{-i\lambda j} \right|^2, \quad (1.32)$$

où les coefficients $(c_j)_{j \geq 0}$ dans la dernière égalité sont donnés par l'équation (1.30).

On aborde maintenant les propriétés asymptotiques de l'estimateur de Whittle ψ_n . Le théorème suivant donne la consistance de ψ_n .

Théorème 1.4.8 Si ψ_0 est la vraie valeur des paramètres à estimer et ψ_n la valeur de ψ qui minimise $\sigma_n^2(\psi)$, alors

$$\psi_n \xrightarrow{\mathbb{P}} \psi_0 \quad \text{et} \quad \sigma_n^2(\psi_n) \xrightarrow{\mathbb{P}} 2\pi \left(\sum_{j=0}^{+\infty} c_j^2 \right)^{-1}, \quad (1.33)$$

où les coefficients $(c_j)_{j \geq 0}$ sont définis par l'équation (1.30).

La preuve de ce théorème est détaillée dans la Section 2 de Kokoszka et Taqqu [46]. ■

Pour établir la distribution asymptotique de cet estimateur, nous avons besoin de l'hypothèse suivante, relative au processus d'innovations $(Z_t)_{t \in \mathbb{Z}}$.

Hypothèse 1.4.9 : Supposons que les Z_t appartiennent au domaine d'attraction d'une loi S α S (voir équation (1.20) pour la définition) ou de manière équivalente vérifient la relation suivante :

$$\lim_{\lambda \rightarrow +\infty} \lambda^\alpha P(|Z| > \lambda) = C_\alpha \gamma, \quad (1.34)$$

où la constante C_α est définie par l'équation (1.9).

Nous introduisons aussi quelques notations que nous allons utiliser dans le théorème suivant donnant la distribution asymptotique de ψ_n . Soit $W(\psi_0)$ une matrice de dimension $(p+q+1) \times (p+q+1)$ et d'éléments générateurs :

$$w_{ij} = \int_{-\pi}^{\pi} f(\lambda, \psi_0) \frac{\partial^2}{\partial \psi_i \partial \psi_j} f^{-1}(\lambda, \psi_0) \, d\lambda, \quad i, j = 1, \dots, p+q+1,$$

et soit v_k , $k \in \mathbb{Z}$ un vecteur de dimension $(p+q+1)$ dont la $j^{\text{ième}}$ coordonnée est donnée par

$$(v_k)_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda, \psi_0) \frac{\partial}{\partial \psi_j} f^{-1}(\lambda, \psi_0) \, d\lambda, \quad j = 1, \dots, p+q+1.$$

Théorème 1.4.10 Si le processus d'innovations $(Z_t)_{t \in \mathbb{Z}}$ vérifie l'**Hypothèse 1.4.9**, alors

$$\left(\frac{n}{\ln n} \right)^{1/\alpha} (\psi_n - \psi_0) \xrightarrow{\mathcal{D}} 4\pi W^{-1}(\psi_0) \sum_{k=1}^{\infty} \frac{Y_k}{Y_0} v_k, \quad (1.35)$$

où Y_0, Y_1, Y_2, \dots sont des variables aléatoires indépendantes; Y_0 est positive $\frac{\alpha}{2}$ -stable de paramètre d'échelle $C_{\alpha/2}^{-2/\alpha}$ et $(Y_k)_{k \geq 1}$ est une suite de variables aléatoires indépendantes

et de même loi symétrique α -stable de paramètre de dispersion $C_\alpha^{-1/\alpha}$. ■

La Démonstration de ce théorème se trouve dans la Section 3 de Kokoszka et Taqqu ([46]).

On peut remarquer que ce résultat est similaire à celui développé dans le Théorème 2.2 de Mikosch et al.[57] et montre ainsi que le résultat asymptotique pour $d = 0$ peut bien être étendu dans le cas $d > 0$. Il faut également noter que dans ce théorème, le paramètre de dispersion des $(Y_k)_{k \geq 1}$ est $C_\alpha^{-1/\alpha}$, au lieu de $C_\alpha^{1/\alpha}$ comme dans le cas des ARMA α -stables.

1.5 Conclusion

Comme nous venons de le remarquer, l'hypothèse de gaussianité s'avère trop restrictive dans certains domaines pour lesquels on assiste à une grande variabilité des données. Pour palier à ce manque plusieurs modèles à variance infinie ont été développés parmi lesquels on peut citer les processus ARFIMA(p, d, q)- S α S. Ils permettent de prendre en compte la présence de dépendance à long terme et de variance infinie. Mais toujours dans un cadre longue mémoire si on désire modéliser simultanément une composante longue mémoire et une composante cyclique périodique, on ne pourra malheureusement pas utiliser un modèle ARFIMA qui n'est pas assez souple pour intégrer un phénomène périodique. La question qu'on peut se poser est : est-il possible de créer un modèle plus général qui prend en compte la présence des trois éléments caractéristiques : *mémoire longue, saisonnalité et variance infinie*? C'est l'objet du chapitre suivant.

Les processus ARFISMA symétriques α -stables

Résumé

Dans ce chapitre, nous nous proposons de développer un nouveau modèle prenant en compte les phénomènes de mémoire longue, de saisonnalité et de variance infinie. Nous établissons les conditions de stationnarité et d'inversibilité. Nous proposons également une méthode de simulation et nous montrons le comportement asymptotique du modèle par des simulations.

Sommaire

2.1	Introduction	35
2.2	Présentation du modèle	36
2.3	Propriétés probabilistes	37
2.4	Méthode de simulations numériques	45
2.5	Quelques simulations	47
2.5.1	Cas gaussien	47
2.5.2	Cas stable	52
2.6	Conclusions et perspectives	56

2.1 Introduction

Afin de prendre en compte la propriété de mémoire longue dans une modélisation, de nombreux auteurs ont suggéré d'utiliser les processus ARFIMA. Granger et Joyeux [31] ainsi que Hosking [37] donnent clairement les bases de cette théorie et supposent que les résidus sont des bruits blancs gaussiens. Bien que les calculs (d'estimation, de prévisions,...) soient rapides avec cette hypothèse, on se rend compte qu'elle est trop restrictive, surtout pour des données présentant un phénomène de variance infinie, comme en finance, en télécommunication ou en hydrologie. Ainsi Kokoszka et Taqqu [45] ont développé la théorie des processus ARFIMA avec variance infinie des innovations qui prennent en compte les comportements de mémoires longues et de variance infinie. Par ailleurs, dans tous les champs d'application des statistiques, nombreuses sont les séries chronologiques qui possèdent un comportement cyclique, périodique ou non périodique. Un exemple célèbre qui a illustré de nombreux travaux sur les séries chronologiques est la série du nombre de taches solaires de Wölfer, que l'on peut trouver, par exemple dans le livre de Brockwell et Davis [12]. Ainsi, dans un cadre longue mémoire, si le statisticien désire modéliser simultanément une composante longue mémoire et une composante cyclique périodique persistante, il ne pourra malheureusement pas utiliser un modèle ARFIMA qui n'est pas assez souple pour intégrer un phénomène périodique. Pour atteindre ce but, plusieurs auteurs ont proposé différents processus longue mémoire, dont le point commun est d'avoir une densité spectrale ayant une ou plusieurs singularités sur l'intervalle $[0, \pi]$. Porter-Hudak [63] ainsi que Reisen et al.[67] ont proposé les processus ARFISMA qui prennent en compte les phénomènes de mémoire longue et de saisonnalité et dont la principale caractéristique est d'avoir une densité spectrale non bornée aux fréquences saisonnières situées dans $[0, \pi]$. Cependant ce modèle est très limité face à des données présentant une variance infinie. Il apparaît donc judicieux de proposer un modèle plus général qui prenne en compte ces trois éléments : *mémoire longue, saisonnalité et variance infinie*. On se propose dans ce chapitre de faire l'étude de ces processus que l'on qualifiera de processus ARFISMA symétrique α -stable.

Ce chapitre est organisé de la manière suivante. Dans la Section 2.2, nous commençons d'abord par définir le modèle ARFISMA symétrique α -stable et nous donnons ensuite quelques extensions de ce modèle. Dans la Section 2.3, nous discutons les conditions de stationnarité et d'inversibilité du modèle. Dans la Section 2.4, nous proposons une méthode de simulation d'un processus ARFISMA symétrique α -stable. Dans la Section 2.5, nous présentons quelques simulations, dans le but de montrer le

comportement asymptotique du modèle. La Section 2.6 est dédiée aux conclusions et perspectives.

2.2 Présentation du modèle

Dans ce paragraphe, nous introduisons d'abord le modèle avec lequel nous travaillons. Nous donnons ensuite quelques extensions de ce modèle.

Définition 2.2.1 Soit $(Z_t)_{t \in \mathbb{Z}}$ une suite de variables aléatoires indépendantes et de même loi symétrique α -stable, de moyenne nulle, de paramètre d'échelle égal à l'unité et d'indice de stabilité $0 < \alpha \leq 2$. On dit qu'un processus centré, noté $(X_t)_{t \in \mathbb{Z}}$, est un processus ARFISMA(p, d, q)(P, D, Q) $_s$ -SaS s'il vérifie l'équation suivante :

$$\Phi(B) \Phi_s(B^s) X_t = (1 - B)^{-d} (1 - B^s)^{-D} \Theta(B) \Theta_s(B^s) Z_t, \quad (2.1)$$

où les réels d et D sont les paramètres de mémoire longue, B est l'opérateur retard et s désigne la période de saisonnalité ($s = 1$ pour des données annuelles, $s = 12$ pour des données mensuelles, ...). Les polynômes $\Phi(B)$ et $\Theta(B)$ sont les polynômes autorégressifs et moyenne mobile de la partie non saisonnière donnés par l'équation (1.13). Les polynôme $\Phi_s(B^s)$ et $\Theta_s(B^s)$ sont respectivement les polynômes de degrés P et Q en B^s de la partie saisonnière et sont définis par :

$$\Phi_s(B^s) = 1 - \sum_{j=1}^P \Phi_j B^{js}, \quad \Theta_s(B^s) = 1 + \sum_{j=1}^Q \Theta_j B^{js}. \quad (2.2)$$

Remarque 2.2.2 Dans l'équation (2.1), le terme $(1 - B)^{-d}$ contrôle le comportement de mémoire longue, la saisonnalité est captée par la quantité $(1 - B^s)^{-D}$ et la variance infinie vient du fait que le processus d'innovations $(Z_t)_{t \in \mathbb{Z}}$ est symétrique α -stable avec $\alpha < 2$.

Les processus ARFISMA(p, d, q)(P, D, Q) $_s$ -SaS constituent une généralisation des processus ARFIMA(p, d, q) de Granger et Joyeux [31] et possèdent plusieurs extensions :

1. si $d = D = 0$ et $P = Q = 0$, nous avons alors les processus linéaires avec variance infinie des innovations étudiés par Fama [24], Stuck et Kleiner [75], ou Brockwell et Davis [13].
2. Si d est un entier, $D = 0$ et $P = Q = 0$, alors nous retrouvons les processus ARIMA stables de Samonodnitsky et Taqqu [71].

3. Si $D = 0$ et $P = Q = 0$, alors le modèle (2.1) devient le processus ARFIMA(p, d, q) symétrique α -stable de Kokoszka et Taqqu [45].
4. Si d et D sont des entiers et $\alpha = 2$, on obtient les processus SARIMA de Brockwell et Davis [13].
5. Si $\alpha = 2$, alors on retrouve les processus ARFISMA gaussiens introduits par Reisen et al. [67].

2.3 Propriétés probabilistes

Dans cette section, nous établissons les conditions d'existence et d'inversibilité d'un processus ARFISMA(p, d, q)(P, D, Q)_s-S α S introduit par la Définition 2.2.1. Nous définissons d'abord les polynômes de Gegenbauer. Les polynômes de Gegenbauer, notés $(C_j(d, \nu))_{j \in \mathbb{N}}$, sont des polynômes orthogonaux générés par la fonction suivante :

$$(1 - 2\nu z + z^2)^{-d} = \sum_{j \geq 0} C_j(d, \nu) z^j, \quad (2.3)$$

où $|z| \leq 1$ et $|\nu| \leq 1$. Rainville [65] montre que les coefficients $(C_j(d, \nu))_{j \in \mathbb{N}}$ peuvent être calculés très simplement de la manière suivante :

$$\forall j \geq 0, \quad C_j(d, \nu) = \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} \frac{(-1)^k \Gamma(d + j - k) (2\nu)^{j-2k}}{\Gamma(d) \Gamma(k+1) \Gamma(j-2k+1)}, \quad (2.4)$$

où $\lfloor \frac{j}{2} \rfloor$ est la partie entière de $\frac{j}{2}$. Cependant, l'expression des coefficients (2.4) est extrêmement lourde, et implique donc une grande quantité de calculs. Il est d'usage de considérer l'algorithme récursif suivant :

$$\begin{cases} C_0(d_i, \nu_i) = 1 \\ C_1(d_i, \nu_i) = 2d_i \nu_i \\ C_j(d_i, \nu_i) = 2\nu_i \left(\frac{d_i-1}{j} + 1 \right) C_{j-1}(d_i, \nu_i) - \left(2 \frac{d_i-1}{j} + 1 \right) C_{j-2}(d_i, \nu_i), \quad \forall j > 1. \end{cases}$$

Pour plus de détails sur les polynômes de Gegenbauer, on se réfère à Szegö [76], Rainville [65] ou Magnus et al.[54]. Dans la suite, nous allons transformer le modèle (2.1), afin de pouvoir l'écrire sous une forme moyenne mobile infinie. Notons d'abord par ω_j les fréquences saisonnières définies par :

$$\omega_j = \frac{2\pi j}{s}, \quad \text{pour } j = 0, 1, 2, \dots, \lfloor \frac{s}{2} \rfloor.$$

Ainsi en se basant sur Giraitis et Leipus [30] ou Reisen et al.[67], le filtre dans l'équation (2.1) peut être récrit comme suit :

$$\begin{aligned} (I - B)^{-d}(I - B^s)^{-D} &= \prod_{j=0}^{\lfloor \frac{s}{2} \rfloor} [(1 - e^{i\omega_j} B)(1 - e^{-i\omega_j} B)]^{-d_j} \\ &= \prod_{j=0}^{\lfloor \frac{s}{2} \rfloor} (1 - 2 \cos(\omega_j) B + B^2)^{-d_j}, \end{aligned} \quad (2.5)$$

avec $d_0 = \frac{d+D}{2}$, $d_i = D$, pour $i = 1, \dots, \lfloor \frac{s}{2} \rfloor - 1$, et $d_{\lfloor \frac{s}{2} \rfloor} = \frac{D}{2}$. Par conséquent, le processus défini par l'équation (2.1) se met sous la forme suivante :

$$\Phi(B) \Phi_s(B^s) X_t = \Theta(B) \Theta_s(B^s) \prod_{j=0}^{\lfloor \frac{s}{2} \rfloor} (1 - 2 \cos(\omega_j) B + B^2)^{-d_j} Z_t. \quad (2.6)$$

Par développement en série, nous avons également

$$(1 - B)^{-d}(1 - B^s)^{-D} = \sum_{j=0}^{+\infty} \psi_j(d, \nu) z^j, \quad (2.7)$$

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

$$\psi_j(d, \nu) = \sum_{\substack{0 \leq l_0, \dots, l_{\lfloor \frac{s}{2} \rfloor} \leq j, \\ l_0 + \dots + l_{\lfloor \frac{s}{2} \rfloor} = j}} C_{l_0}(d_0, \nu_0) \cdots C_{l_{\lfloor \frac{s}{2} \rfloor}}(d_{\lfloor \frac{s}{2} \rfloor}, \nu_{\lfloor \frac{s}{2} \rfloor}), \quad (2.8)$$

avec $d = (d_0, d_1, \dots, d_{\lfloor \frac{s}{2} \rfloor})$ et $\nu = (\nu_0, \nu_1, \dots, \nu_{\lfloor \frac{s}{2} \rfloor})$ où $\nu_j = \cos(\omega_j)$, pour $j = 0, \dots, \lfloor \frac{s}{2} \rfloor$. Les coefficients $(C_l(d_i, \nu_i))_{l \in \mathbb{N}}$ sont les polynômes de Gegenbauer définis précédemment. Dans toute la suite du document, on suppose que les polynômes $\Phi(z)$, $\Phi_s(z^s)$, $\Theta(z)$ et $\Theta_s(z^s)$ n'ont aucune racine commune et ont toutes leurs racines en dehors du cercle unité. Ainsi, pour montrer que l'équation (2.1) admet une unique solution qui s'écrit sous une représentation moyenne mobile infinie, nous avons besoin de définir les coefficients $(c_j)_{j \geq 0}$ de cette représentation à partir de la relation suivante :

$$\Theta_s(z^s) \Theta(z) \sum_{j=0}^{+\infty} \psi_j(d, \nu) z^j = \Phi_s(z^s) \Phi(z) \sum_{j=0}^{+\infty} c_j z^j, \quad \forall |z| < 1. \quad (2.9)$$

En développant l'équation (2.9) et par identification de polynômes, nous obtenons une formule explicite pour le calcul des coefficients $(c_j)_{j \geq 0}$ (un calcul détail est présenté en Annexe A).

Supposons sans perte de généralité que $P \leq Q$. Nous avons alors les équations suivantes :

- Pour $j \in [0, s-1] \cup [(Q+1)s, +\infty[$

$$c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} = \psi_j(d, \nu) + \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu). \quad (2.10)$$

- Pour $j \in [s, (P+1)s-1]$

$$c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} + \tilde{H}_j - \tilde{\beta}_j = \psi_j(d, \nu) + \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu) + \tilde{K}_j + \tilde{\alpha}_j. \quad (2.11)$$

- Pour $j \in [(P+1)s, (Q+1)s-1]$

$$c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} = \psi_j(d, \nu) + \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu) + \tilde{K}_j + \tilde{\alpha}_j. \quad (2.12)$$

Les quantités $\tilde{\beta}_j$, $\tilde{\alpha}_j$, \tilde{H}_j et \tilde{K}_j sont données par les expressions suivantes. En remarquant que, pour tout $j \in [s, (Q+1)s-1]$, il existe $1 \leq \lambda \leq Q$ et $k = 0, 1, \dots, s-1$ tels que $j = \lambda s + k$, nous avons :

- a) pour $1 \leq \lambda \leq P$

$$\tilde{\beta}_{\lambda s+k} = \sum_{i=1}^{\lambda} \Phi_{is} c_{(\lambda-i)s+k} \quad (2.13)$$

$$\tilde{H}_{\lambda s+k} = \Phi_{\lambda s} \sum_{i=1}^k \phi_i c_{k-i} + \sum_{l=1}^{\lambda-1} \Phi_{ls} \sum_{i=1}^p \phi_i c_{(\lambda-l)s-i+k} \quad (2.14)$$

- b) pour $1 \leq \lambda \leq Q$

$$\tilde{\alpha}_{\lambda s+k} = \sum_{i=1}^{\lambda} \Theta_{is} \psi_{(\lambda-i)s+k}(d, \nu) \quad (2.15)$$

$$\tilde{K}_{\lambda s+k} = \Theta_{\lambda s} \sum_{i=1}^k \theta_i \psi_{k-i}(d, \nu) + \sum_{l=1}^{\lambda-1} \Theta_{ls} \sum_{i=1}^q \theta_i \psi_{(\lambda-l)s-i+k}(d, \nu) \quad (2.16)$$

En utilisant l'expression (2.5) et se basant sur Giraitus et Leipus [30] ou Woodward et al.[78], nous pouvons déterminer une région de convergence des moments d'ordre α du processus (2.1). C'est à dire des conditions sur les paramètres de mémoire longue d et D telles que $\sum_{j=1}^{\infty} |c_j|^\alpha < 1$. Ce résultat est spécifié dans le théorème suivant.

Théorème 2.3.1 Sous les conditions suivantes $|d + D| < 1 - \frac{1}{\alpha}$ et $|D| < 1 - \frac{1}{\alpha}$ avec $1 < \alpha \leq 2$, nous avons :

i) si les polynômes $\Phi(z)$ et $\Phi_s(z^s)$ ne s'annulent pas dans le cercle unité, alors le processus défini en (2.1) est strictement stationnaire, causal et admet une unique représentation moyenne mobile infinie donnée par :

$$X_t = \sum_{j=0}^{+\infty} c_j Z_{t-j}, \quad (2.17)$$

où les coefficients $(c_j)_{j \geq 0}$ sont obtenus par les équations (2.10), (2.11) et (2.12).

ii) si, de plus, les polynômes $\Theta(z)$ et $\Theta_s(z^s)$ n'admettent pas de zéros dans le cercle unité, alors le processus (2.1) est inversible et a une représentation autorégressive infinie définie par :

$$Z_t = \sum_{j=0}^{+\infty} \tilde{c}_j X_{t-j}, \quad \text{p.s.}, \quad (2.18)$$

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

$$\frac{\Phi(z)\Phi_s(z^s)}{\Theta(z)\Theta_s(z^s)} \prod_{j=0}^{\lfloor \frac{s}{2} \rfloor} (1 - 2 \cos(\omega_j)z + z^2)^{d_j} = \sum_{j=0}^{\infty} \tilde{c}_j z^j. \quad (2.19)$$

■

Un calcul analogue à celui des coefficients $(c_j)_{j \geq 0}$ (voir Annexe A) permet d'établir l'expression explicite des coefficients $(\tilde{c}_j)_{j \geq 0}$ de la manière suivante :

- Pour $j \in [0, s-1] \cup [(Q+1)s, +\infty[$

$$\tilde{c}_j + \sum_{i=1}^{\min(j,q)} \theta_i \tilde{c}_{j-i} = \pi_j(d, \nu) - \sum_{i=1}^{\min(j,p)} \phi_i \pi_{j-i}(d, \nu). \quad (2.20)$$

- Pour $j \in [s, (P+1)s-1]$

$$\tilde{c}_j + \sum_{i=1}^{\min(j,q)} \theta_i \tilde{c}_{j-i} + \tilde{L}_j + \tilde{\zeta}_j = \pi_j(d, \nu) - \sum_{i=1}^{\min(j,p)} \phi_i \pi_{j-i}(d, \nu) + \tilde{G}_j - \tilde{\gamma}_j. \quad (2.21)$$

- Pour $j \in [(P+1)s, (Q+1)s-1]$

$$\tilde{c}_j + \sum_{i=1}^{\min(j,q)} \theta_i \tilde{c}_{j-i} + \tilde{L}_j + \tilde{\zeta}_j = \pi_j(d, \nu) - \sum_{i=1}^{\min(j,p)} \phi_i \pi_{j-i}(d, \nu). \quad (2.22)$$

Les poids $(\pi_j(d, \nu))_{j \geq 0}$ sont tels que $\pi_j(d, \nu) = \psi_j(-d, \nu)$, avec les coefficients $(\psi_j(-d, \nu))_{j \geq 0}$ sont donnés par l'équation (2.8). Les quantités $\tilde{L}_j, \tilde{\zeta}_j, \tilde{G}_j, \tilde{\gamma}_j$ sont définies de manière correspondantes par :

a) pour $1 \leq \lambda \leq P$

$$\tilde{\gamma}_{\lambda s+k} = \sum_{i=1}^{\lambda} \Phi_{is} \pi_{(\lambda-i)s+k}(d, \nu) \quad \tilde{G}_{\lambda s+k} = \Phi_{\lambda s} \sum_{i=1}^k \phi_i \pi_{k-i} + \sum_{l=1}^{\lambda-1} \Phi_{ls} \sum_{i=1}^p \phi_i \pi_{(\lambda-l)s-i+k}(d, \nu)$$

b) pour $1 \leq \lambda \leq Q$

$$\tilde{\zeta}_{\lambda s+k} = \sum_{i=1}^{\lambda} \Theta_{is} \tilde{c}_{(\lambda-i)s+k} \quad \tilde{L}_{\lambda s+k} = \Theta_{\lambda s} \sum_{i=1}^k \theta_i \tilde{c}_{k-i} + \sum_{l=1}^{\lambda-1} \Theta_{ls} \sum_{i=1}^q \theta_i \tilde{c}_{(\lambda-l)s-i+k}.$$

Pour établir la preuve du Théorème 2.3.1, nous avons besoin du Théorème 2.3.2 suivant que l'on doit à Kokoszka et Taqqu [45].

Théorème 2.3.2 Supposons que $\{\varepsilon_n, n \in \mathbb{Z}\}$ est une suite de variables aléatoires i.i.d selon une loi SaS, avec $0 < \alpha \leq 2$. Soit $(c_j)_{j \geq 0}$ des coefficients réels satisfaisant $\sum_{j=0}^{\infty} |c_j|^\alpha < \infty$ et soit $(\psi_j)_{j \geq 0}$ une suite de nombres réels telle que : $\sum_{j=0}^{\infty} |\psi_j| < \infty$, si $\alpha > 1$ et $\sum_{j=0}^{\infty} |\psi_j|^\alpha < \infty$, si $\alpha \leq 1$. Si

$$X_n := C(B)\varepsilon_n = \sum_{j=0}^{\infty} c_j \varepsilon_{n-j} \quad (2.23)$$

$$Y_n := \Psi(B)\varepsilon_n = \sum_{j=0}^{\infty} \psi_j \varepsilon_{n-j} \quad (2.24)$$

$$A(B)\varepsilon_n := \sum_{j=0}^{\infty} a_j \varepsilon_{n-j}, \quad (2.25)$$

où $a_j = \sum_{k=0}^j \psi_k c_{j-k}$, pour $j = 0, 1, \dots$, alors on a :

$$\lim_{m \rightarrow +\infty} \sum_{k=0}^m \psi_k X_{n-k} = \sum_{j=0}^{\infty} a_j \varepsilon_{n-j} \quad \text{et} \quad \lim_{s \rightarrow +\infty} \sum_{j=0}^s c_j Y_{n-j} = \sum_{j=0}^{\infty} a_j \varepsilon_{n-j},$$

où la convergence est dans \mathbb{L}^p , pour tout $0 < p < \alpha$. Par conséquent, on obtient :

$$\Psi(B) [C(B)\varepsilon_n] = A(B)\varepsilon_n = C(B) [\Psi(B)\varepsilon_n], \quad \text{p.s} \quad (2.26)$$

■

La preuve de ce théorème est détaillée dans Kokoszka et Taqqu [45].

Preuve du Théorème 2.3.1 :

La preuve de i) est une adaptation de la preuve du Théorème 2.1 de Kokoszka et Taqqu [45]. Sous les conditions $|d + D| < 1 - \frac{1}{\alpha}$ et $|D| < 1 - \frac{1}{\alpha}$ avec $1 < \alpha \leq 2$, les séries $\sum_{j=0}^{\infty} c_j Z_{n-j}$ et $\sum_{j=0}^{\infty} \psi_j(d, \nu) Z_{n-j}$ convergent presque sûrement.

$$\text{Posons } \Omega_0 = \left\{ \omega : \sum_{j=0}^{\infty} c_j Z_{n-j}(\omega) \text{ et } \sum_{j=0}^{\infty} \psi_j(d, \nu) Z_{n-j}(\omega) \text{ convergent} \right\}.$$

Pour $\omega_0 \in \Omega_0$ fixé, notons par :

$$Z_j = Z_j(\omega_0), \quad X_n = \sum_{j=0}^{\infty} c_j Z_{n-j}(\omega_0) \quad \text{et} \quad Y_n = \sum_{j=0}^{\infty} \psi_j(d, \nu) Z_{n-j}(\omega_0).$$

Nous devons montrer alors que

$$\Phi(B)\Phi_s(B^s)X_n = \Theta(B)\Theta_s(B^s)Y_n$$

$$\begin{aligned} \Phi(B)\Phi_s(B^s)X_n &= \left(1 - \sum_{i=1}^P \Phi_i B^{is}\right) \left(1 - \sum_{j=1}^p \phi_j B^j\right) \sum_{k=0}^{+\infty} c_k Z_{n-k} \\ &= \sum_{k=0}^{+\infty} c_k Z_{n-k} \\ &\quad - \left(\sum_{j=1}^p \phi_j B^j\right) \left(\sum_{k=0}^{+\infty} c_k Z_{n-k}\right) - \left(\sum_{i=1}^P \Phi_i B^{is}\right) \left(\sum_{k=0}^{+\infty} c_k Z_{n-k}\right) \\ &\quad + \left(\sum_{i=1}^P \Phi_i B^{is}\right) \left(\sum_{j=1}^p \phi_j B^j\right) \left(\sum_{k=0}^{+\infty} c_k Z_{n-k}\right) \end{aligned}$$

En se servant du développement de l'expression A.1 (Annexe A), nous obtenons :

$$\begin{aligned} \Phi(B)\Phi_s(B^s)X_n &= \sum_{j=0}^{s-1} \left(c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} \right) Z_{n-j} \\ &\quad + \sum_{\lambda=1}^P \sum_{k=0}^{s-1} \left(\tilde{H}_{\lambda s+k} - \tilde{\beta}_{\lambda s+k} \right) Z_{n-\lambda s-k} \\ &\quad + \sum_{\lambda=1}^{\infty} \sum_{k=0}^{s-1} \left(c_{\lambda s+k} - \sum_{i=1}^{\min(p, \lambda s+k)} \phi_i c_{\lambda s+k-i} \right) Z_{n-\lambda s-k} \end{aligned} \quad (2.27)$$

où les quantités $\tilde{H}_{\lambda s+k}$ et $\tilde{\beta}_{\lambda s+k}$ sont définies par les équations (2.14) et (2.13). En regroupant les termes qui se ressemblent, on peut récrire l'équation (2.27) comme suit :

$$\begin{aligned}
\Phi(B)\Phi_s(B^s)X_n &= \sum_{j=0}^{s-1} \left(c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} \right) Z_{n-j} \\
&+ \sum_{\lambda=1}^P \sum_{k=0}^{s-1} \left(c_{\lambda s+k} - \sum_{i=1}^{\min(p,\lambda s+k)} \phi_i c_{\lambda s+k-i} + \tilde{H}_{\lambda s+k} - \tilde{\beta}_{\lambda s+k} \right) Z_{n-\lambda s-k} \\
&+ \sum_{\lambda=P+1}^{\infty} \sum_{k=0}^{s-1} \left(c_{\lambda s+k} - \sum_{i=1}^{\min(p,\lambda s+k)} \phi_i c_{\lambda s+k-i} \right) Z_{n-\lambda s-k} \quad (2.28)
\end{aligned}$$

En utilisant les équations (2.10), (2.11) et (2.12), remplaçons chaque terme de l'équation (2.28) par son expression. Ainsi on obtient :

$$\begin{aligned}
\Phi(B)\Phi_s(B^s)X_n &= \sum_{j=0}^{s-1} \left(\psi_j(d, \nu) - \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu) \right) Z_{n-j} \\
&+ \sum_{\lambda=1}^Q \sum_{k=0}^{s-1} \left(\tilde{K}_{\lambda s+k} + \tilde{\alpha}_{\lambda s+k} \right) Z_{n-\lambda s-k} \\
&+ \sum_{\lambda=1}^{\infty} \sum_{k=0}^{s-1} \left(\psi_{\lambda s+k}(d, \nu) - \sum_{i=1}^{\min(q,\lambda s+k)} \theta_i \psi_{\lambda s+k-i}(d, \nu) \right) Z_{n-\lambda s-k}
\end{aligned}$$

$\tilde{\alpha}_{\lambda s+k}$ et $\tilde{K}_{\lambda s+k}$ sont données par les équations (2.15) et (2.16).

$$\begin{aligned}
\implies \Phi(B)\Phi_s(B^s)X_n &= \sum_{j=0}^{+\infty} \psi_j(d, \nu) Z_{n-j} + \left(\sum_{i=1}^Q \Theta_i B^{is} \right) \left(\sum_{j=0}^{+\infty} \psi_j(d, \nu) Z_{n-j} \right) \\
&+ \left(\sum_{i=1}^q \theta_i B^i \right) \left(\sum_{j=0}^{+\infty} \psi_j(d, \nu) Z_{n-j} \right) \\
&+ \left(\sum_{i=1}^Q \Theta_i B^{is} \right) \left(\sum_{k=1}^q \theta_k B^k \right) \left(\sum_{j=0}^{+\infty} \psi_j(d, \nu) Z_{n-j} \right) \\
&= \left(1 + \sum_{i=1}^Q \Phi_i B^{is} \right) \left(1 + \sum_{k=1}^q \theta_k B^k \right) \sum_{j=0}^{+\infty} \psi_j(d, \nu) Z_{n-j} \\
&= \Theta(B)\Theta_s(B^s)Y_n
\end{aligned}$$

D'où $X_n = \sum_{j=0}^{\infty} c_j Z_{n-j}$ est solution de l'équation (2.1). Pour prouver l'unicité, supposons que $X'_n = \sum_{j=0}^{\infty} c'_j Z_{n-j}$ est une autre solution du système (2.1). L'égalité $\Phi(B)\Phi_s(B^s)X_n = \Phi(B)\Phi_s(B^s)X'_n$ implique que $\sum_{j=0}^{\infty} a'_j Z_{n-j} = \sum_{j=0}^{\infty} a_j Z_{n-j}$, p.s, où les coefficients a_j sont définis par :

$$a_j = \begin{cases} c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i}, & \text{pour } j \in J_1, \\ c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} + \tilde{H}_j - \tilde{\beta}_j, & \text{pour } j \in J_2, \end{cases}$$

avec $J_1 = [0, s-1] \cup [(Q+1)s, +\infty[\cup [(P+1)s, (Q+1)s-1]$, $J_2 = [s, (P+1)s-1]$. Les coefficients a'_j sont définis de manière correspondante. Comme les variables Z_t sont i.i.d., nous avons $\sum_{j=0}^{\infty} |a_j - a'_j| = 0$, ce qui implique $a_j = a'_j$ et par conséquent $c_j = c'_j$.

La preuve de ii) suit les grandes lignes de la preuve du Théorème 2.3 de Kokoszka et Taquu [45]. Sous les conditions du Théorème 2.3.1, nous avons $\sum_{j=0}^{\infty} |c_j|^\alpha < \infty$. Or les variables aléatoires $(Z_t)_{t \in \mathbb{Z}}$ sont i.i.d suivant une loi SaS, donc le processus $(X_t)_{t \in \mathbb{Z}}$ admet également une loi SaS. Ainsi nous avons :

$$\begin{aligned} \mathbb{E} \left(\sum_{j=0}^{\infty} |\tilde{c}_j X_{t-j}| \right) &= \sum_{j=0}^{\infty} |\tilde{c}_j| \mathbb{E}(|X_{n-j}|) \\ &= \sum_{j=0}^{\infty} |\tilde{c}_j| \mathbb{E}(|X_1|), \end{aligned}$$

or $\mathbb{E}(|X_1|) < \infty$, puisque $1 < \alpha \leq 2$ (voir Proposition 1.2.11). D'où nous avons $\sum_{j=0}^{\infty} |\tilde{c}_j| < \infty$. Ainsi appliquons le Théorème 2.3.2 avec :

$$\Psi(B) = C_{d,D}(B) = \frac{\Theta(B)\Theta_s(B^s)}{\Phi(B)\Phi_s(B^s)} (1-B)^{-d} (1-B^s)^{-D}$$

$$C(B) = C_{d,D}^{-1}(B) = \frac{\Phi(B)\Phi_s(B^s)}{\Theta(B)\Theta_s(B^s)} (1-B)^d (1-B^s)^D.$$

Nous aurons alors :

$$\Psi(B) [C(B)Z_t] = A(B)Z_t = C(B) [\Psi(B)Z_t], \quad \text{p.s,}$$

or $A(B)Z_t = C_{d,D}(B)C_{d,D}^{-1}(B)Z_t = Z_t$ et

$$\begin{aligned}\Psi(B)[C(B)Z_t] &= C_{d,D}(B)[C_{d,D}^{-1}(B)Z_t] \\ &= C_{d,D}(B)X_t \\ &= \sum_{j=0}^{\infty} \tilde{c}_j X_{t-j}.\end{aligned}$$

Par suite, nous avons $Z_t = \sum_{j=0}^{\infty} \tilde{c}_j X_{t-j}$, p.s. ■

2.4 Méthode de simulations numériques

Dans ce paragraphe, on étend la méthode décrite par Stoev et Taqqu [74] pour simuler un processus ARFIMA(p, d, q)-S α S, au cas des processus ARFISMA(p, d, q)(P, D, Q)_s symétrique α -stable introduit par la Définition 2.2.1. Nous rappelons d'abord quelques définitions et résultats relatifs aux Transformées de Fourier Discrète (TFD). La raison pour laquelle la TFD est importante et d'un intérêt pratique est qu'il existe des algorithmes rapides pour effectuer cette transformation. Parmi ces algorithmes, on peut citer l'algorithme de la Transformée de Fourier rapide (sigle en anglais : FFT ou Fast Fourier Transform) qui est disponible dans plusieurs logiciels statistiques (Matlab, S-plus, R, ...).

Soit $a = (a(j), j \in \mathbb{Z})$ une suite périodique de période R telle que $a(j + R) = a(j)$, $\forall j \in \mathbb{Z}$. La TFD de la séquence a , notée $\mathcal{D}_R(a) = (\mathcal{D}_R(a)(k), k \in \mathbb{Z})$ est définie par :

$$\hat{a}(k) = \mathcal{D}_R(a)(k) = \sum_{j=0}^{R-1} a(j)e^{2\pi ijk/R}, \quad \forall k \in \mathbb{Z}.$$

Ainsi, la suite \hat{a} est périodique de période R et satisfait à la formule d'inversion

$$a(j) = \mathcal{D}_R^{-1}(a)(j) := \frac{1}{R} \sum_{k=0}^{R-1} \hat{a}(k)e^{-2\pi ijk/R}, \quad \forall j \in \mathbb{Z}.$$

En plus, si a et b sont deux suites périodiques de période R , alors nous avons par produit de convolution

$$\mathcal{D}_R(a)(k)\mathcal{D}_R(b)(k) = \mathcal{D}_R(a * b)(k), \quad k \in \mathbb{Z},$$

où $(a * b)(k) := \sum_{j=0}^{R-1} a(k-j)b(j)$, $k \in \mathbb{Z}$.

Si R est un entier tel que $R = 2^m$, $m \in \mathbb{N}$, alors la TFD de la suite a peut être calculée

de manière efficace en utilisant l'algorithme FFT.

Notre objectif est de générer une trajectoire $X(t)$, $t = 0, 1, \dots, N-1$, $N \in \mathbb{N}$, issu d'un processus ARFISMA(p, d, q)(P, D, Q)_s-SαS, en utilisant sa représentation moyenne mobile infinie donnée par l'équation (2.17). En d'autre termes, nous pouvons approximer cette trajectoire de taille N par sa version tronquée de la manière suivante :

$$X(t) \approx X_M(t) := \sum_{j=0}^{M-1} c_j Z_{t-j}, \quad t = 0, 1, \dots, N-1, \quad (2.29)$$

où les coefficients $(c_j)_{j \geq 0}$ sont définis par les équations (2.10), (2.11) et (2.12), $(Z_j)_{j \in \mathbb{Z}}$ est une suite de variables aléatoires indépendantes et de même loi symétrique α -stable standard et M est le paramètre de troncature. Nous allons décrire très rapidement une procédure de simulation et nous renvoyons à l'article de Stoev et Taquq [74] pour plus de détails.

$$\text{Posons } \tilde{b}(j) = \begin{cases} c_j & \text{pour } j = 0, 1, \dots, M-1, \\ 0 & \text{pour } j = M, \dots, M+N-1, \end{cases}$$

et soit $\tilde{b}(j + k(M+N)) := \tilde{b}(j)$, $j = 0, 1, \dots, M+N-1$, $k \in \mathbb{Z}$, de sorte que la suite $\tilde{b} = (\tilde{b}(j), j \in \mathbb{Z})$ soit périodique de période $(M+N)$. De même, définissons la séquence \tilde{Z} périodique de période $(M+N)$ comme suit $\tilde{Z}(j + k(M+N)) := \tilde{Z}(j)$, $j = 0, 1, \dots, M+N-1$, $k \in \mathbb{Z}$,

$$\text{où } \tilde{Z}(j) = \begin{cases} Z_j & \text{pour } j = 0, 1, \dots, M-1, \\ 0 & \text{pour } j = M, \dots, M+N-1, \end{cases}$$

Ainsi, d'après la relation (2.29), nous avons

$$X_M(t) \stackrel{\mathcal{D}}{=} \sum_{j=0}^{M+N-1} \tilde{b}(j) \tilde{Z}(t-j), \quad t = 0, 1, \dots, N-1. \quad (2.30)$$

Nous pouvons alors utiliser un algorithme FFT (qui n'est pas décrit dans ce document) pour calculer l'équation (2.30). Plus précisément, nous avons l'algorithme suivant.

Algorithme

1. Choisir le paramètre de troncature M et la taille de l'échantillon N tels que $(M+N) = 2^m$, $m \in \mathbb{N}$.
2. En utilisant les équations (2.10), (2.11) et (2.12), calculer les coefficients $\tilde{b}(j)$, $j = 0, 1, \dots, M+N-1$. En appliquant l'algorithme FFT, calculer la TFD $\widehat{\tilde{b}}(k) = \mathcal{D}_{(M+N)}(\tilde{b})(k)$, $k = 0, 1, \dots, M+N-1$ de la suite \tilde{b} .

3. Générer une suite de variables aléatoires Z_j , $j = 0, 1, \dots, M+N-1$, indépendantes et de même loi symétrique α -stable standard. En appliquant l'algorithme FFT, calculer la TFD $\widehat{Z}(k) = \mathcal{D}_{(M+N)}(\widetilde{Z})(k)$, $k = 0, 1, \dots, M+N-1$ de la séquence \widetilde{Z} .
4. En utilisant l'algorithme FFT, calculer l'inverse de la TFD du produit $(\widehat{b} \widehat{Z})$ et soit

$$X_M(t) := \mathcal{D}_{M+N}^{-1}(\widehat{b} \widehat{Z})(t), \quad t = 0, 1, \dots, N-1.$$

Prendre la suite $\{X_M(t), \quad t = 0, 1, \dots, N-1\}$ comme une réalisation d'un processus ARFISMA(p, d, q)(P, D, Q)_s symétrique α -stable.

2.5 Quelques simulations

Nous présentons dans cette section un ensemble de simulation pour illustrer le comportement asymptotique d'un processus ARFISMA(p, d, q)(P, D, Q)_s symétrique α -stable (*variance infinie, mémoire longue et saisonnalité*). Pour simplifier nous nous limitons sans perte de généralité au modèle avec $p = q = P = Q = 0$. Plus précisément, nous simulons, en utilisant la méthode proposée dans la section 2.4, le processus donné par :

$$X_t = (1 - B)^{-d}(1 - B^s)^{-D}Z_t, \quad (2.31)$$

avec $d = 0.15$, $D = 0.20$. Nous considérons la taille d'échantillon $N = 2000$ avec un paramètre de troncature $M = 5000$ pour différentes périodes de saisonnalité $s = 4$, $s = 6$ et $s = 12$. Nous étudions deux cas : un cas gaussien ($\alpha = 2$) et un cas stable avec $\alpha = 1.7$. Nous notons également que les valeurs des paramètres de mémoire longue sont choisies telles que $|D| < 1 - \frac{1}{\alpha}$ et $|D + d| < 1 - \frac{1}{\alpha}$, en accord avec le Théorème 2.3.1.

2.5.1 Cas gaussien

Dans cette partie, nous supposons que les bruits $(Z_t)_{t \in \mathbb{Z}}$ du modèle (2.31) suivent une loi normale standard. Nous traçons sur les Figures 2.1((a)), 2.2((a)) et 2.3((a)) les trajectoires du processus $(X_t)_{t \in \mathbb{Z}}$ pour les périodes de saisonnalité $s = 4$, $s = 6$ et $s = 12$. Dans ces trois graphiques, on peut constater que cette série semble être strictement stationnaire. De même, en considérant les deux tests graphiques (voir Section 1.2.4, page 15), visibles sur les Figures 2.1((b), (c)), 2.2((b), (c)) et 2.3((b), (c)), les résultats montrent clairement que nous sommes en présence de données à variance finie.

Les outils de bases utilisés pour analyser une série chronologique par des processus longue mémoire 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.[67]). Ainsi, elle est définie par :

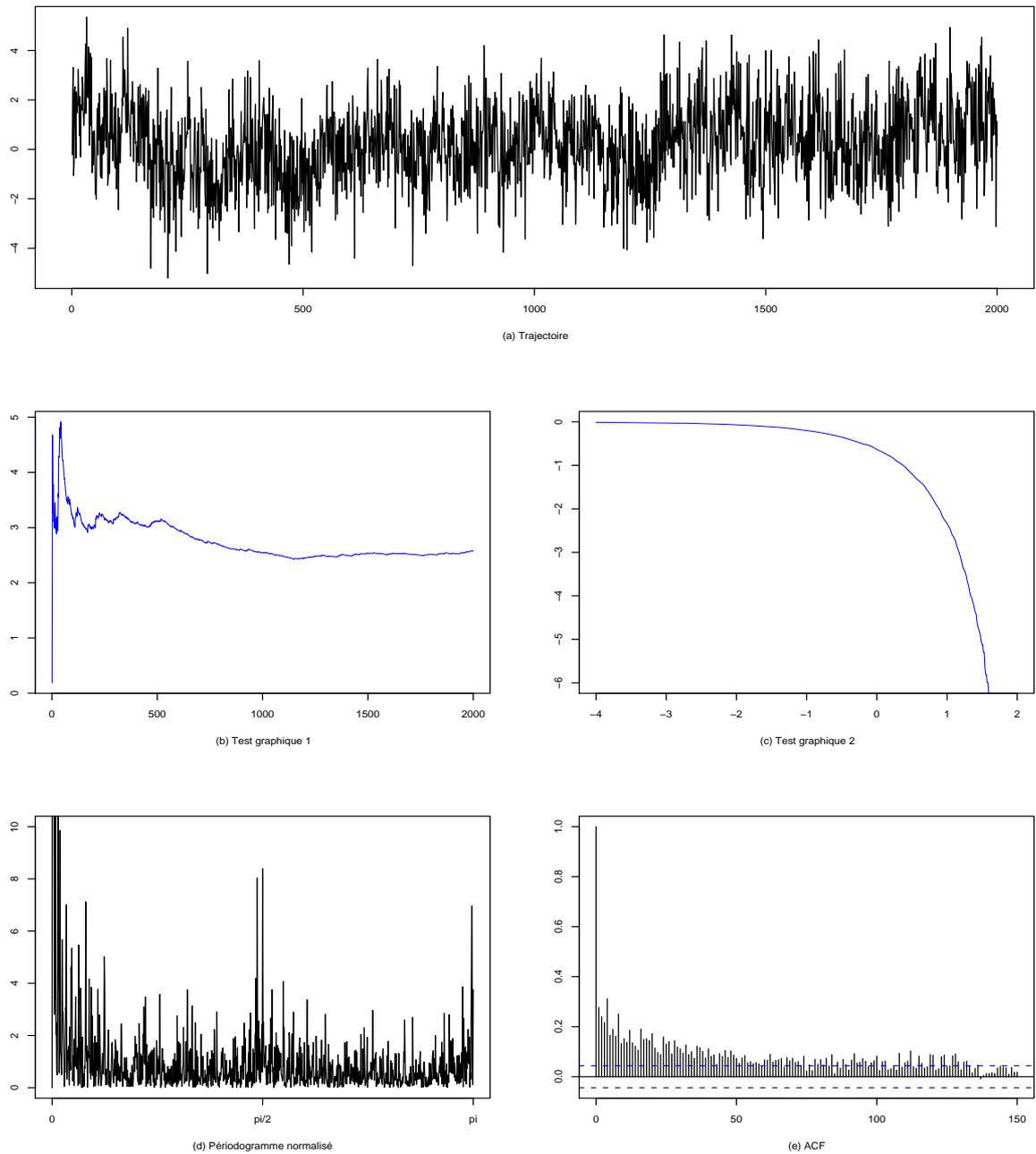
$$f(\omega) = \frac{\sigma_Z^2}{2\pi} [2 \sin(\omega s/2)]^{-2D} [2 \sin(\omega/2)]^{-2d}, \quad -\pi \leq \omega \leq \pi, \quad (2.32)$$

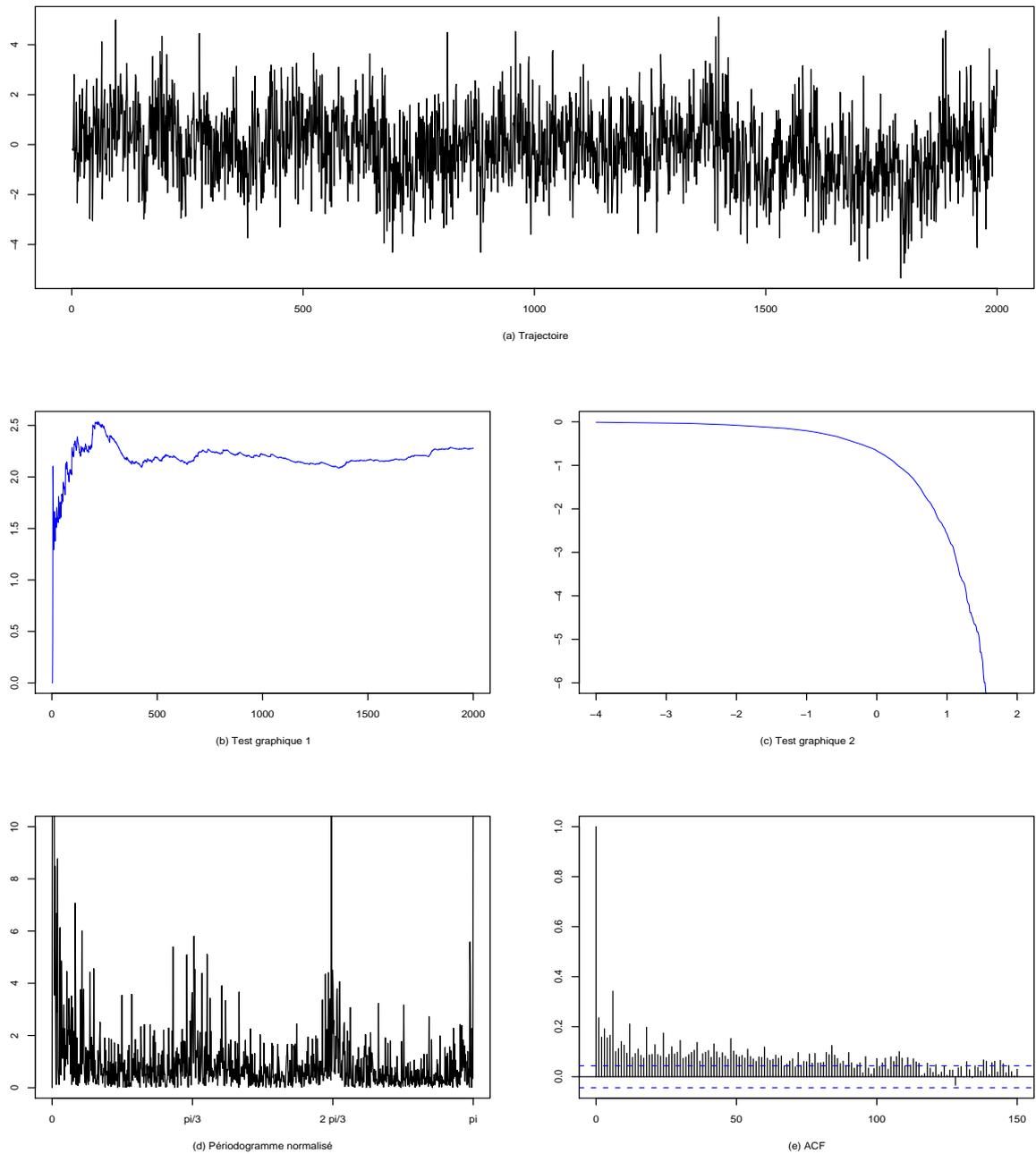
où σ_Z^2 désigne la variance du bruit. Cette densité spectrale est non bornée aux fréquences saisonnières $\omega_j = \frac{2\pi j}{s}$, pour $j = 0, 1, 2, \dots, [\frac{s}{2}]$ (voir Reisen et al.[67] pour plus de détails). A échantillon fini, la densité spectrale est estimée par le périodogramme normalisé donné par l'équation (1.17) et qui est visualisé respectivement pour $s = 4$, $s = 6$ et $s = 12$ sur les Figures 2.1((d)), 2.2((d)) et 2.3((d)). Ces graphiques montrent des pics sur les fréquences saisonnières, ce qui est en accord avec le fait que, dans le domaine spectral un pic sur la densité spectrale à une fréquence donnée ω indique un cycle périodique de période $\frac{2\pi}{\omega}$ sur le processus.

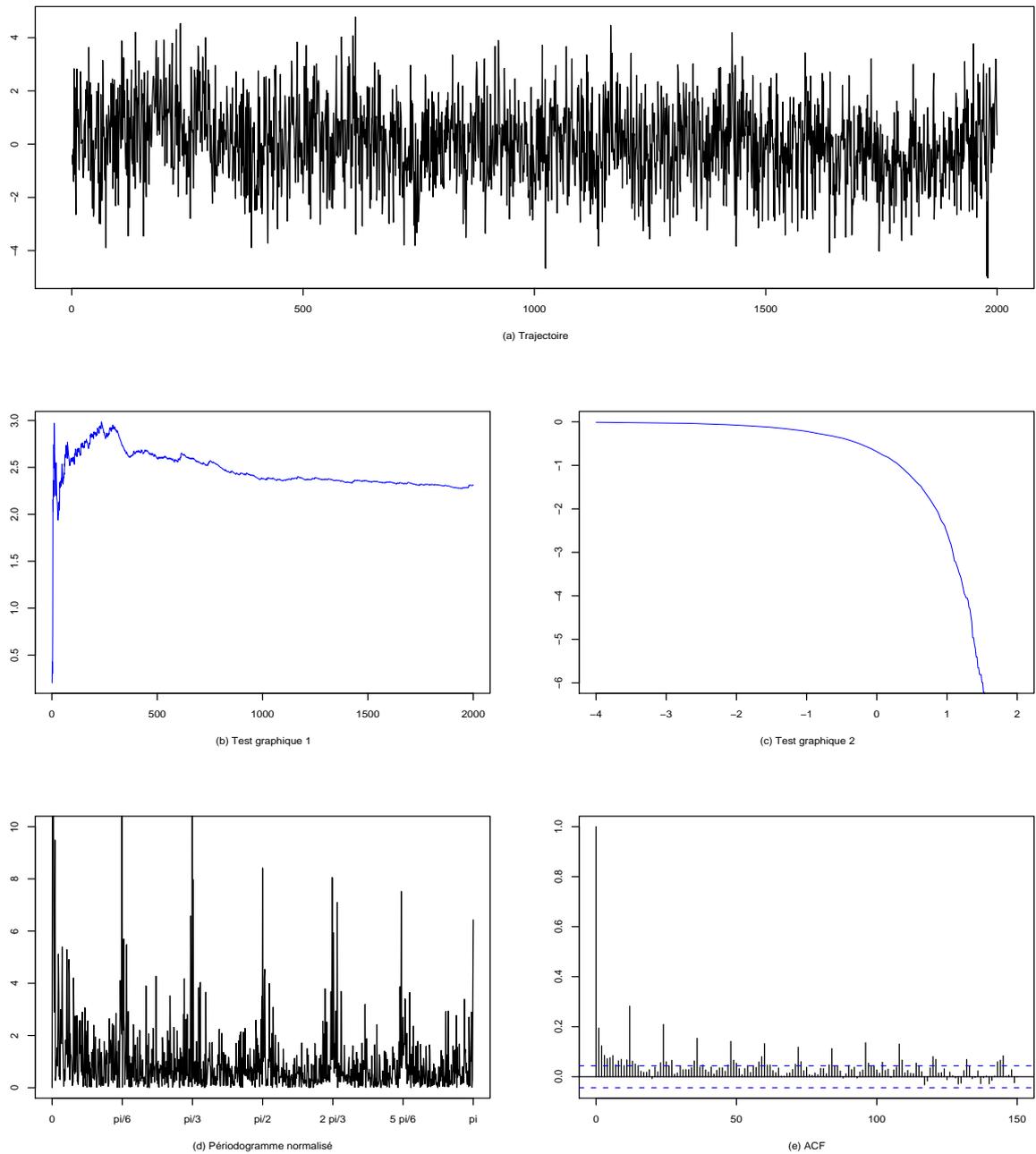
Dans le cas classique, la fonction d'autocorrélation empirique $\hat{\rho}(h)$ est définie par :

$$\hat{\rho}(h) = \frac{\sum_{i=1}^{N-|h|} (X_i - \bar{X})(X_{i+h} - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2}, \quad \text{où} \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (2.33)$$

et qui est un estimateur de la fonction d'autocorrélation théorique. Les graphiques (e) des Figures 2.1, 2.2 et 2.3 illustrent le comportement asymptotique de (2.33) pour $s = 4$, $s = 6$ et $s = 12$. Ces trois graphiques montrent une décroissance lente de la fonction d'autocorrélation lorsque les retards augmentent, assurant ainsi une dépendance à long terme du processus $(X_t)_{t \in \mathbb{Z}}$.

FIG. 2.1 – Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 4$ et $\alpha = 2$.

FIG. 2.2 – Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 6$ et $\alpha = 2$.

FIG. 2.3 – Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 12$ et $\alpha = 2$.

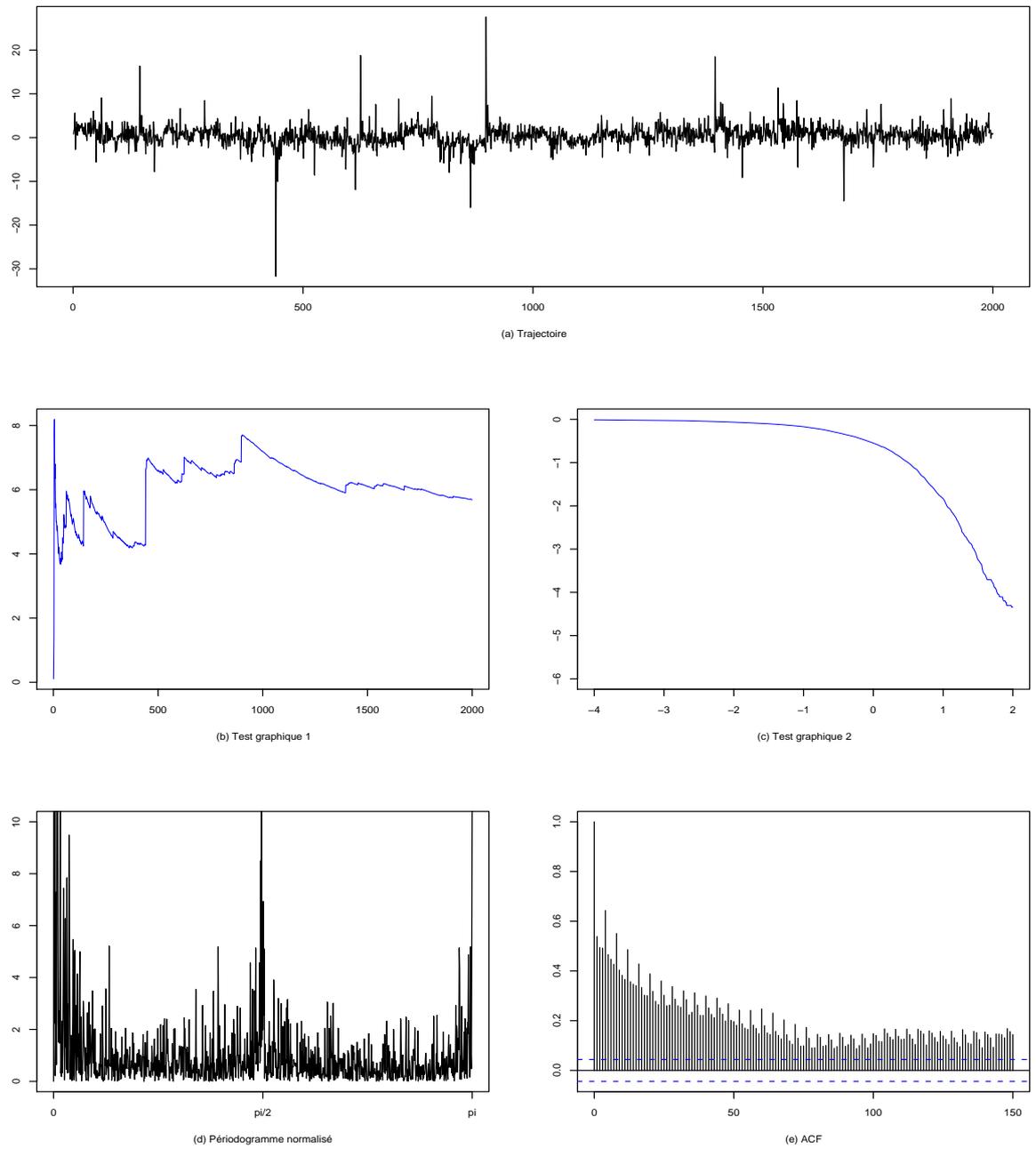
2.5.2 Cas stable

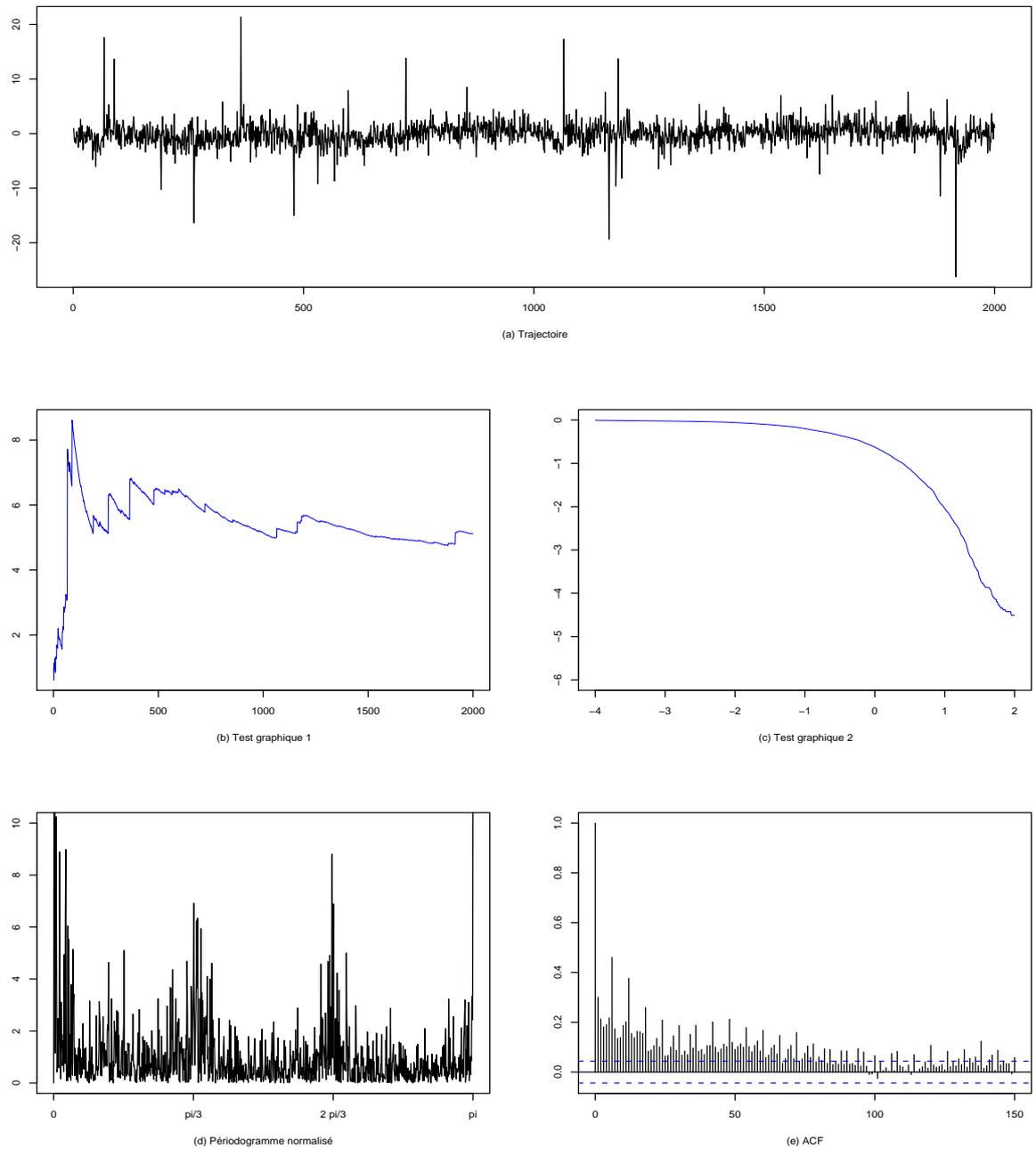
Considérons maintenant que les innovations $(Z_t)_{t \in \mathbb{Z}}$ sont distribuées suivant une loi symétrique α -stable standard d'exposant caractéristique $\alpha = 1.7$. Les graphiques (a), (b) et (c) des Figures 2.4, 2.5 et 2.6 montrent clairement que les données considérées présentent un phénomène de variance infinie. En effet, on observe sur les trajectoires (graphiques (a)) du processus une grande variabilité et les tracés des deux tests graphiques (graphiques (b) et (c)) divergent.

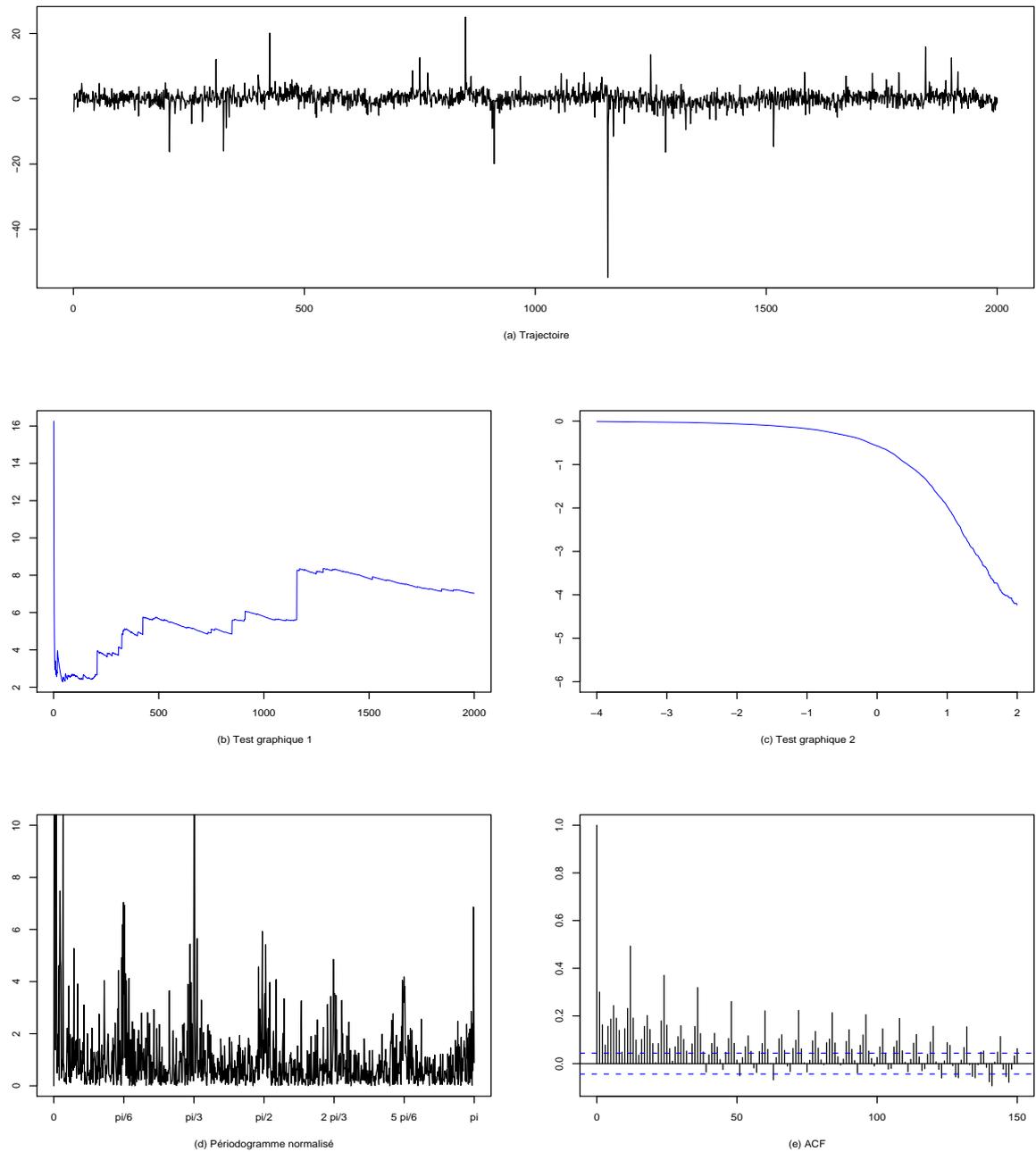
Dans le cas de variance infinie, certaines modifications peuvent être apportées. En effet, les outils de bases utilisés (fonction d'autocorrélation et densité spectrale) n'existent plus. Néanmoins, même si ces quantités n'ont pas d'existence théorique, on peut tout de même calculer leurs valeurs empiriques et les utiliser comme dans le cas classique (voir Mickosch et al.[57], Kokoszka et Taqqu.[46]). Ainsi la fonction d'autocorrélation empirique appropriée dans le cas stable est donnée par :

$$\widehat{\rho}_H(h) = \frac{\sum_{i=1}^{N-|h|} X_i X_{i+h}}{\sum_{i=1}^N X_i^2}. \quad (2.34)$$

On observe sur les graphiques (e) des Figures (2.4, 2.5 et 2.6), que cette fonction d'autocorrélation $\widehat{\rho}_H(h)$ décroît lentement lorsque $h \rightarrow \infty$; montrant par conséquent la présence de mémoire longue. Les graphiques (d) des Figures (2.4, 2.5 et 2.6) donnent le comportement asymptotique du périodogramme normalisé. On observe des pics sur chaque fréquence saisonnière, indiquant ainsi une saisonnalité.

FIG. 2.4 – Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 4$ et $\alpha = 1.7$.

FIG. 2.5 – Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 6$ et $\alpha = 1.7$.

FIG. 2.6 – Trajectoire, Tests graphiques, périodogramme et ACF pour $s = 12$ et $\alpha = 1.7$.

2.6 Conclusions et perspectives

Dans ce chapitre, nous avons introduit les processus ARFISMA symétriques α -stables, dans le but de prendre en compte dans une modélisation la présence éventuelle de comportement de mémoire longue, de saisonnalité et de variance infinie. Nous avons étudié les conditions de stationnarité et d'inversibilité et de causalité de ce processus. Nous avons proposé une méthode de simulation pour ce modèle, et nous avons montré sur quelques exemples le comportement asymptotique de ce processus.

Il reste bien des choses à faire pour cette famille de processus. Il serait d'abord intéressant de s'investir sur l'estimation des paramètres de ce modèle. Le Chapitre suivant fait l'objet de l'estimation des paramètres d'un processus ARFISMA symétrique α -stable.

Estimation des paramètres d'un processus ARFISMA symétrique α -stable

Résumé

Dans ce chapitre, nous présentons différentes méthodes d'estimation des paramètres de mémoire longue d'un processus ARFISMA symétrique α -stable. Nous comparons ces méthodes d'estimation par des simulations de Monte Carlo.

Sommaire

3.1	Introduction	58
3.2	Quelques notations	58
3.3	Estimation semiparamétrique	59
3.4	Estimation par la méthode de Whittle	61
3.4.1	Whittle classique	61
3.4.2	Whittle par la méthode MCMC	61
3.5	Simulations numériques	63
3.5.1	Modèle ARFIMA(0, d, 0)-S α S	64
3.5.2	Modèle ARFISMA(0, D, 0) _s -S α S	65
3.5.3	Modèle ARFISMA(0, d, 0)(0, D, 0) _s -S α S	65
3.6	Conclusions et perspectives	66

3.1 Introduction

Les processus ARFISMA symétriques α -stables introduits dans le Chapitre 2, ont été développés par Diongue, Diop et Ndongo [21], dans le but de modéliser les comportements de mémoire longue, de saisonnalité et de variance infinie. Dans ce chapitre, nous nous proposons d'étudier les différentes méthodes d'estimation des paramètres de mémoire longue d'un processus ARFISMA-S α S et d'évaluer leurs performances par des simulations de Monte Carlo. Ainsi, nous considérons la méthode semiparamétrique proposée par Reisen et al. [67] et deux approches de la méthode de Whittle : la méthode de Whittle classique et une autre alternative basée sur la méthode de Monte Carlo par chaînes de Markov (en abrégé MCMC).

Ce chapitre est organisé comme suit. Dans la Section 3.2, nous introduisons quelques notations et définitions nécessaires pour l'estimation. Dans la Section 3.3, nous présentons la méthode semiparamétrique. La Section 3.4 est consacrée à la présentation des deux approches de la méthode de Whittle. Dans la Section 3.5, nous donnons les résultats de simulation de Monte Carlo pour comparer ces différentes méthodes d'estimation. La Section 3.6 permet de tirer quelques conclusions et perspectives.

3.2 Quelques notations

Dans toute la suite de ce chapitre, nous considérons une trajectoire finie X_1, \dots, X_n issue d'un processus ARFISMA(p, d, q)(P, D, Q)_s-S α S, $(X_t)_{t \in \mathbb{Z}}$ stationnaire, inversible et de moyenne nulle. Nous supposons que les paramètres de la distribution du processus d'innovations sont connus. Nous nous intéressons donc à l'estimation du paramètre vectoriel ψ , incluant les paramètres de mémoire longue (d, D), les paramètres de mémoire courte de la partie non saisonnière $(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ et de la partie saisonnière $(\Phi_1, \dots, \Phi_P, \Theta_1, \dots, \Theta_Q)$. Soit ψ_0 la vraie valeur du paramètre ψ et supposons qu'elle se trouve à l'intérieur de Ψ , où Ψ est un compact de $\mathbb{R}^{p+q+P+Q+2}$ défini par :

$$\Psi = \{\psi \in \mathbb{R}^{p+q+P+Q+2} : \phi_p \neq 0, \theta_q \neq 0, \Phi_P \neq 0, \Theta_Q \neq 0, \text{ les polynômes}$$

$$\Theta(z), \Phi(z), \Phi_s(z^s) \text{ et } \Theta_s(z^s) \text{ n'ont pas de racines communes pour } |z| \leq 1,$$

$$|d + D| < 1 - \frac{1}{\alpha} \text{ et } |D| < 1 - \frac{1}{\alpha}, 1 < \alpha \leq 2\}.$$

La fonction de transfert du processus $(X_t)_{t \in \mathbb{Z}}$ est définie par :

$$g(\lambda, \psi) = \left| \frac{\Theta(e^{-i\lambda}) \Theta_s(e^{-i\lambda s})}{\Phi(e^{-i\lambda}) \Phi_s(e^{-i\lambda s})} \right|^2 \left| 2 \sin\left(\frac{\lambda}{2}\right) \right|^{-2d} \left| 2 \sin\left(\frac{\lambda s}{2}\right) \right|^{-2D}, \quad \forall (\lambda, \psi) \in [-\pi, \pi] \times \Psi. \quad (3.1)$$

3.3 Estimation semiparamétrique

Dans cette section, nous présentons une extension dans le cas saisonnier de la méthode dite du log-périodogramme proposée par Geweke et Porter-Hudak [29] pour estimer le paramètre de mémoire d d'un processus ARFIMA(p, d, q) gaussien. 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. Reisen et al.[67] étendent la méthode GPH dans le cas saisonnier pour estimer les paramètres de mémoire longue d'un processus ARFISMA gaussien et l'idée fondamentale est d'exploiter le comportement asymptotique de la densité spectrale autour de chaque fréquence saisonnière. Notre objectif est d'étendre cette méthode dans le cas stable pour estimer les paramètres de mémoire longue d'un processus ARFISMA symétrique α -stable. Puisque la densité spectrale n'existe pas dans le cas stable, nous considérons la fonction de transfert du processus ARFISMA(0, d, 0)(0, D, 0)_s-SaS pour construire cette régression. Cette fonction de transfert s'obtient en posant $p = q = P = Q = 0$ dans l'équation (3.1) et soit alors :

$$g(\lambda, \psi) = \left| 2 \sin\left(\frac{\lambda}{2}\right) \right|^{-2d} \left| 2 \sin\left(\frac{\lambda s}{2}\right) \right|^{-2D}, \quad \text{avec } \psi = (d, D). \quad (3.2)$$

Pour illustrer cette méthode, considérons d'abord les fréquences suivantes :

$$\lambda_{\nu,j} = \frac{2\pi\nu}{s} + \frac{2\pi j}{n}, \quad \nu = 0, 1, \dots, \left[\frac{s}{2}\right] - 1, \quad j = 1, 2, \dots, m, \quad (3.3)$$

où m représente la largeur de la bande de fréquences utilisées et qui satisfait $\frac{1}{m} + \frac{m}{n} \rightarrow 0$, lorsque $n \rightarrow +\infty$. Comme les fréquences $\lambda_{\nu,j}$ doivent appartenir à l'intervalle $[0, \pi]$, nous posons alors $\lambda_{[\frac{s}{2}],j} = \pi - \frac{2\pi j}{n}$. En évaluant l'équation (3.2) aux fréquences $\lambda_{\nu,j}$, pour $j = 1, \dots, m$ et par passage aux logarithmes, on obtient alors :

$$\log(\tilde{I}_n(\lambda_{\nu,j})) = -d \log \left| 2 \sin\left(\frac{\lambda_{\nu,j}}{2}\right) \right|^2 - D \log \left| 2 \sin\left(\frac{s\lambda_{\nu,j}}{2}\right) \right|^2 + \log \left(\frac{\tilde{I}_n(\lambda_{\nu,j})}{g(\lambda_{\nu,j})} \right), \quad (3.4)$$

où $\tilde{I}_n(\lambda_{\nu,j})$ est le périodogramme normalisé évalué à la fréquence $\lambda_{\nu,j}$.

Différentes méthodes d'estimation de d et D peuvent être obtenues par choix approprié des fréquences de Fourier $\lambda_j = \frac{2\pi j}{n}$, $j = 1, \dots, m$, dans l'équation de régression (3.4). Ainsi, GPH_ν , $\nu = 0, \dots, \lfloor \frac{s}{2} \rfloor - 1$ est l'estimateur obtenu en choisissant $\frac{2\pi j}{n}$ fréquences à droite de chaque fréquence saisonnière $\omega_\nu = \frac{2\pi\nu}{s}$ et $GPH_{\lfloor \frac{s}{2} \rfloor}$ est celui obtenu en utilisant $\frac{2\pi j}{n}$ fréquences à gauche de la fréquence $\omega_{\lfloor \frac{s}{2} \rfloor}$. Par exemple pour $s = 4$, nous obtenons les estimateurs GPH_0 , GPH_1 et GPH_2 en utilisant respectivement les fréquences $\lambda_{0,j} = \frac{2\pi j}{n}$, $\lambda_{1,j} = \frac{\pi}{2} + \frac{2\pi j}{n}$ et $\lambda_{2,j} = \pi - \frac{2\pi j}{n}$, pour $j = 1, \dots, m$. Cependant, il apparaît clairement, par des simulations dans Reisen et al.[67], que les méthodes GPH_ν , $\nu = 0, \dots, \lfloor \frac{s}{2} \rfloor$ ne fournissent pas de bonnes performances. Ceci est surtout dû au nombre de fréquences insuffisant utilisé dans l'équation de régression (3.4). Par conséquent, ces méthodes ne seront pas considérées dans la suite de ce document. Nous nous intéressons ici particulièrement aux méthodes GPH_T et GPH_P qui se distinguent par le choix de la largeur de bande m dans l'équation de régression (3.4).

- La méthode GPH_T , T pour total, calcule les estimateurs de d et D en utilisant toutes les fréquences dans l'équation (3.4). Ainsi, la régression est construite sur les fréquences $\lambda_{\nu,j} = \frac{2\pi\nu}{s} + \frac{2\pi j}{n}$, $\nu = 0, 1, \dots, \lfloor \frac{s}{2} \rfloor - 1$, $j = 1, 2, \dots, m$, avec une largeur de bande $m = \lfloor \frac{s}{2} \rfloor - 1$.
 - La méthode GPH_P , P pour partiel, considère une collection des fréquences de Fourier à droite des fréquences saisonnières (à gauche pour $\nu = \frac{s}{2}$) avec une largeur de bande $m = \lfloor \frac{n}{2s} \rfloor - 1$.
- Pour plus de détails relatifs à ces méthodes, nous renvoyons à la Section 3.1 de Reisen et al.[67].

Remarque 3.3.1 Si nous désirons estimer simultanément les paramètres de mémoire courte et de mémoire longue du processus (2.1) par la méthode semiparamétrique, il suffit de décomposer sa fonction de transfert de la forme suivante :

$$g(\lambda) = |1 - e^{i\lambda}|^{-2d} |1 - e^{is\lambda}|^{-2D} g^*(\lambda), \quad (3.5)$$

où $g^*(\lambda)$ est une fonction positive, continue et périodique de période 2π et contrôle le comportement de mémoire courte. Ainsi, en se servant de l'article de Moulines et Soulier [58], nous pouvons étendre la méthode GPH pour estimer l'ensemble des composantes du paramètre vectoriel ψ du processus $(X_t)_{t \in \mathbb{Z}}$. Dans ce cas, il est commode de considérer

la méthode GPH_T , car elle utilise toutes les fréquences comme le suggèrent Moulines et Soulier [58].

3.4 Estimation par la méthode de Whittle

Dans cette section, nous développons deux méthodes d'estimation des paramètres d'un processus ARFISMA(p, d, q)(P, D, Q) $_s$ -S α S défini par l'équation (2.1), basées sur une évaluation de deux façons différentes de la vraisemblance de Whittle.

Par définition, on appelle estimateur de Whittle de ψ_0 , noté $\widehat{\psi}_n$, la valeur de ψ qui minimise la vraisemblance de Whittle définie par :

$$\sigma_n^2(\psi) = \int_{-\pi}^{\pi} \frac{\widetilde{I}_n(\lambda)}{g(\lambda, \psi)} d\lambda, \quad (3.6)$$

où $g(\lambda, \psi)$ désigne la fonction de transfert du processus $(X_t)_{t \in \mathbb{Z}}$ définie par l'équation (3.1) et $\widetilde{I}_n(\lambda)$ son périodogramme normalisé donné par l'expression (1.17).

Il existe de nombreuses méthodes concurrentes pour approcher l'intégrale (3.6). Nous considérons ici deux approches. La première utilise une méthode de quadrature pour approcher la fonction de vraisemblance de Whittle et aboutit à la méthode de Whittle classique, alors que dans la seconde approche, nous proposons d'évaluer l'intégrale (3.6) en utilisant la méthode de Monte Carlo par chaînes de Markov.

3.4.1 Whittle classique

Afin d'estimer les paramètres d'un modèle ARMA α -stable, plusieurs auteurs ont utilisé la méthode de Whittle classique (voir Klüppelberg et Mikosch ([41], [40]), Mikosch et al. [57], Embrechts et al.[23]). Cette méthode est basée dans le domaine spectral et nécessite l'approximation suivante :

$$\hat{\sigma}_n^2(\psi) = \frac{2\pi}{n} \sum_{j=0}^{n-1} \frac{\widetilde{I}_n(\lambda_j)}{g(\lambda_j, \psi)}, \quad (3.7)$$

où les fréquences λ_j sont données par $\lambda_j = \frac{2\pi j}{n}$, $j = 1, \dots, n-1$.

3.4.2 Whittle par la méthode MCMC

Dans cette section, nous proposons une autre alternative; une approche qui utilise la méthode de Monte Carlo basée sur les chaînes de Markov pour calculer l'intégrale (3.6).

La méthode est la suivante :

Soit C une constante strictement positive telle que $C = \int_{-\pi}^{\pi} \tilde{I}_n(\lambda) d\lambda$, alors $\tilde{I}_n(\lambda)$ est une densité sur $[-\pi, \pi]$, à une constante multiplicative près. Ainsi, la vraisemblance de Whittle (3.6) peut s'écrire sous la forme :

$$\sigma_n^2(\psi) = C \int_{-\pi}^{\pi} \frac{f(\lambda)}{g(\lambda, \psi)} d\lambda = C \mathbb{E}_f \left(\frac{1}{g(\lambda, \psi)} \right), \quad (3.8)$$

où $f(\lambda) = \frac{1}{C} \tilde{I}_n(\lambda)$ est une densité de probabilité sur $[-\pi, \pi]$. En utilisant la loi forte des grands nombres, nous pouvons estimer l'espérance mathématique dans l'équation (3.8) par sa moyenne empirique donnée par :

$$\bar{\sigma}_N^2(\psi) = \frac{1}{N} \sum_{j=1}^N \frac{1}{g(\lambda_j, \psi)}, \quad \text{pour } N \text{ assez grand.} \quad (3.9)$$

L'échantillon $(\lambda_1, \dots, \lambda_N)$ est généré suivant la loi de f qui est connue à une constante multiplicative près. Le principe des méthodes MCMC est de construire une chaîne de Markov ergodique dont la loi stationnaire est f . Un des algorithmes qui permettent de générer $(\lambda_1, \dots, \lambda_N)$ sans simuler directement suivant la loi de f est l'algorithme de Metropolis-Hastings que l'on peut calculer de la manière suivante.

Algorithme :

1. Étant donné λ_t
2. Générer Y_t suivant une loi uniforme sur $\mathcal{U}_{[-\pi, \pi]}$ et notons par y_t la valeur obtenue.
3. Prendre

$$\lambda_{t+1} = \begin{cases} y_t & \text{avec probabilité } \rho(\lambda_t, y_t), \\ \lambda_t & \text{avec probabilité } 1 - \rho(\lambda_t, y_t), \end{cases}$$

$$\text{où } \rho(\lambda_t, y_t) = \min \left\{ \frac{f(y_t)}{f(\lambda_t)}, 1 \right\} = \min \left\{ \frac{\tilde{I}_n(y_t)}{\tilde{I}_n(\lambda_t)}, 1 \right\}.$$

Pour plus de détails relatifs à cet algorithme, nous renvoyons à Robert et Casella [68]. Par suite, une procédure d'estimation de ψ est de minimiser l'équation (3.9) sur le compact Ψ .

3.5 Simulations numériques

Nous présentons maintenant un ensemble de simulations pour valider et comparer les méthodes d'estimation présentées ci-dessous. Pour mener notre étude, nous générons différents processus ARFISMA(0, d , 0)(0, D , 0) _{s} -S α S, pour différentes valeurs de d et D . Les modèles qui seront considérés, sont résumés dans le tableau suivant :

Paramètres	d	D
Modèle 1	0.20	0
Modèle 2	0	0.20
Modèle 3	0.1	0.12

TAB. 3.1 – Processus générateur de données.

Les modèles simulés sont définis par l'équation suivante :

$$X_t = (1 - B)^{-d}(1 - B^s)^{-D}Z_t, \quad (3.10)$$

où le processus $(Z_t)_{t \in \mathbb{Z}}$ est symétrique α -stable. Nous simulons, en utilisant la méthode décrite à la Section 2.4, des échantillons de taille $n = 100$ puis $n = 500$ et $n = 1500$ avec un paramètre de troncature $M = 5000$ ($M > n$ suggéré par Stoev et Taqqu [74]). Pour chaque modèle, nous effectuons 1500 répliques et nous considérons deux cas : $\alpha = 1.3$ et $\alpha = 2$ avec une période de saisonnalité fixée à $s = 4$. Les résultats des simulations donnent la moyenne, la racine de l'erreur quadratique moyenne (RMSE : Root Mean Square Error) et l'erreur moyenne absolue (MAE : Mean Absolute Error) des estimations basées sur 1500 répliques. Pour étudier la dispersion des estimations, nous construisons les box plots correspondants. Tous les calculs ont été menés avec le logiciel R et les résultats de simulations sont donnés par les Tableaux C.1-C.6 et les Figures C.1-C.4 de l'Annexe C. Dans ces tableaux, la première colonne donne les tailles d'échantillon. Dans les colonnes suivantes, nous donnons les paramètres estimés par les différentes méthodes. Les valeurs entre parenthèses représentent les RMSE des estimateurs et celles entre crochets sont les MAE. Les figures montrent la dispersion des 1500 valeurs estimées par les différentes méthodes. L'axe vertical de la figure indique une déviation à la valeur nominale.

3.5.1 Modèle ARFIMA(0, d, 0)-S α S

Dans cette section, nous nous intéressons à l'estimation du paramètre de mémoire longue d'un processus ARFIMA(0,d,0) symétrique α -stable introduit par Kokoswka et Taqqu [45]. Ce modèle est un cas particulier du modèle (2.1), obtenu pour $p = q = P = Q = D = 0$. Pour résoudre le problème d'estimation, Kokoswka et Taqqu [46] utilisent la méthode de Whittle classique (voir Chapitre 1) pour estimer simultanément les paramètres de mémoire longue et de mémoire courte, alors que Geweke et Porter-Hudak [29] emploient la méthode de régression semiparamétrique pour estimer le paramètre de mémoire longue dans le cas gaussien ($\alpha = 2$). A notre connaissance, cette méthode n'est pas encore considérée dans le cas stable. Par conséquent, nous proposons dans cette section d'estimer le paramètre de mémoire de ce processus en utilisant les méthodes GPH_T et GPH_P que nous allons comparer avec les méthodes de Whittle classique et Whittle par MCMC. La méthode GPH_T utilise toutes fréquences de Fourier $\lambda_j = \frac{2\pi j}{n}$, $j = 0, \dots, n-1$. La méthode GPH_P sélectionne partiellement les fréquences de Fourier avec une largeur de bande classique, telle que $m = \lfloor n^\nu \rfloor$ et ν est généralement choisi : $\nu = 0.5, 0.6, 0.7$ (voir Porter-Hudak [63] et Crato et De Lima [20]). Nous considérons le modèle 1 et les résultats de simulation sont présentés dans les Tableaux C.1 et C.2.

Ce qui apparaît en premier est que les RMSE et MAE des paramètres estimés sont tous relativement faibles pour chacune des méthodes d'estimation. De plus, les estimateurs obtenus par les méthodes de Whittle (classique et MCMC) sont meilleurs que ceux obtenus par les méthodes semiparamétriques (GPH_T et GPH_P). Lorsque nous comparons les deux méthodes d'estimation paramétriques, nous notons en général, que l'estimateur de Whittle par MCMC fournit un RMSE faible pour les petites tailles d'échantillon ($n = 100$), alors que les résultats obtenus par la méthode de Whittle classique sont meilleurs pour de grandes tailles d'échantillon ($n = 500, n = 1500$). Concernant, les méthodes de régression, les résultats sont meilleurs pour la méthode GPH_T , car cette méthode utilise toutes les fréquences dans l'équation de régression (3.4). Pour la méthode GPH_P , le cas $\nu = 0.7$ donne une meilleure performance que les autres valeurs de ν , puisqu'il utilise plus de fréquences dans l'équation (3.4). Nous remarquons également, que les RMSE et MAE des estimateurs dans le cas gaussien ($\alpha = 2$) sont, en général, plus larges que ceux obtenus dans le cas stable ($\alpha = 1.3$). Ce phénomène est souvent noté dans la littérature (voir Kokoszka et Taqqu [46]). Les simulations de Monte Carlo montrent également l'impact de la taille de l'échantillon sur les méthodes d'estimation. En effet, quand nous augmentons la taille de l'échantillon

($n = 1500$), les résultats s'améliorent considérablement. La Figure C.1 confirme les résultats précédents. En effet, nous pouvons observer que la dispersion des estimateurs obtenus par les méthodes semiparamétriques est plus large que celle des estimateurs calculés avec les méthodes de Whittle classique et MCMC. Cette dispersion s'améliore lorsque la taille de l'échantillon s'accroît ($n = 1500$).

3.5.2 Modèle ARFISMA(0, D, 0)_s-S α S

Lorsque nous posons $d = p = P = q = Q = 0$ et $\alpha = 2$ dans le modèle (2.1), nous obtenons le processus longue mémoire saisonnier flexible de Hasler [34]. Reisen et al. [67] résolvent le problème d'estimation en considérant les méthodes GPH_T , GPH_P et la méthode du maximum de vraisemblance (MLE). Dans cette section, nous étendons les deux premières méthodes dans le cas stable ($\alpha = 1.3$). Comme la méthode du maximum de vraisemblance n'est pas une possibilité viable dans le cas stable (la densité n'a pas de forme explicite en général), nous allons comparer leurs performances aux méthodes de Whittle classique et MCMC. Pour ce faire, nous considérons le modèle 2 et les résultats sont donnés dans les Tableaux C.3 et C.4.

Dans ces tableaux, nous observons que les méthodes paramétriques sont meilleures que les méthodes semiparamétriques, en terme de RMSE et de MAE. Cependant, la performance de la méthode de Whittle par MCMC est meilleure que la méthode de Whittle classique, pour toutes les tailles d'échantillon. Concernant, les méthodes de régression, la méthode GPH_T est meilleure, comme le note également Reisen et al. [67] dans le cas gaussien. La figure C.2 confirme ces résultats, et montre également l'impact de la taille de l'échantillon sur la dispersion des estimateurs.

3.5.3 Modèle ARFISMA(0, d, 0)(0, D, 0)_s-S α S

Dans ce paragraphe, on s'intéresse au processus ARFISMA symétrique α -stable introduit par Diongue et al. [21] et défini par l'équation (2.1). Nous proposons d'estimer les paramètres de mémoire longue de ce modèle en utilisant les méthodes d'estimation décrites dans les sections précédentes. Nous considérons alors le modèle 3 et les résultats sont résumés par les Tableaux C.5 et C.6.

Nous observons que les estimateurs de Whittle classique et de Whittle par MCMC ont une meilleure performance que les méthodes de régression. Pour la méthode de Whittle classique, on note une grande différence entre la vraie valeur du paramètre et la valeur estimée, pour la taille d'échantillon $n = 100$. cependant, les résultats s'améliorent claire-

ment lorsque la taille de l'échantillon augmente ($n = 500$, $n = 1500$). Pour les méthodes de régression, les meilleurs résultats sont toujours obtenus avec la méthode GPH_T . On note également que le paramètre d a une grande influence sur l'estimation du paramètre D , en particulier pour de petites tailles d'échantillon. Ces résultats peuvent être observés graphiquement à travers les figures C.3 et C.4. En effet, ces figures montrent clairement que les intervalles de confiance obtenus par les méthodes paramétriques sont plus petits que ceux des méthodes de régression et de plus, ils s'améliorent lorsque la taille de l'échantillon s'accroît ($n = 500$, $n = 1500$).

3.6 Conclusions et perspectives

Dans ce chapitre, nous avons considéré différentes méthodes d'estimation des paramètres de mémoire longue d'un processus ARFISMA symétrique α -stable. D'un point de vue empirique, les méthodes de Whittle classique et de Whittle par MCMC donnent une très bonne performance, en terme de RMSE et de MAE. Par conséquent, il nous semble plus adapté d'utiliser ces deux méthodes pour l'estimation des paramètres d'un processus ARFISMA symétrique α -stable. Cependant, les résultats de simulation montrent, en général, que la méthode de Whittle par MCMC est meilleure que la méthode de Whittle classique, en particulier lorsqu'il y a présence de composantes de saisonnalité dans le modèle. On note aussi l'impact du paramètre d sur l'estimation du paramètre D par les différentes méthodes considérées. Cette étude par simulation de Monte Carlo, nous mène également à conclure qu'il y a une similarité entre les méthodes paramétriques et les méthodes semiparamétriques, lorsque que la taille de l'échantillon est grande ($n = 1500$).

D'autres pistes restent à explorer pour ces méthodes d'estimation. Il serait intéressant, par exemple, de faire varier la période de saisonnalité s ($s = 6$, $s = 12, \dots$). En effet, il apparaît clairement dans Reisen et al. [67] que la performance des méthodes semiparamétriques croît avec la valeur de s . Il serait aussi très intéressant de s'investir sur les propriétés asymptotiques de l'estimateur de Whittle basé sur la méthode MCMC. Un sujet qui reste encore flou avec ces types d'estimateur est l'influence de la présence d'une partie autorégressive et d'une partie moyenne-mobile sur l'estimation des paramètres de mémoire longue. Dans le cas d'un processus ARFIMA, plusieurs auteurs (voir par exemple Agiakloglou et al. [3]) ont montré une dégénérescence des performances de l'estimateur du log-périodogramme selon les valeurs d'un paramètres autorégressif. Une étude théorique ou une étude précise par simulations de Monte Carlo semble nécessaire

dans le cas des processus ARFISMA α -stables. Les méthodes d'estimation considérées dans ce chapitre ne permettent pas d'estimer le paramètre de la distribution des innovations, il serait donc intéressant d'étudier les méthodes d'estimation simultanées des paramètres de ce processus.

Estimation simultanée des paramètres d'un processus ARFISMA symétrique α -stable via la méthode de la fonction caractéristique empirique

Résumé

Nous étudions dans ce chapitre une méthode d'estimation basée sur la fonction caractéristique empirique, pour estimer simultanément l'ensemble des paramètres d'un processus ARFISMA symétrique α -stable. Nous établissons les propriétés asymptotiques de cet estimateur et nous évaluons sa performance par des simulations de Monte Carlo.

Sommaire

4.1	Introduction	69
4.2	Procédure d'estimation	69
4.3	Propriétés asymptotiques	75
4.4	Simulations de Monte Carlo	77
4.4.1	Choix empirique de la taille des blocs m	79
4.4.2	Comparaison des estimateurs	85
4.5	Autre alternative	87
4.5.1	Méthode C-GMM	87
4.5.2	Calcul simple de l'estimateur C-GMM	89
4.5.3	Quelques simulations	90
4.6	Conclusions et perspectives	91

4.1 Introduction

La méthode du maximum de vraisemblance est considérée comme l'une des méthodes les plus importantes pour estimer simultanément les paramètres d'un modèle, en série chronologique (voir Brockwell et Davis [12], Box et Jenkins [11] ou Hamilton [32]). Une des raisons est que, sous certaines conditions de régularité appropriées, l'estimateur du maximum de vraisemblance est asymptotiquement efficace. Cependant, pour implémenter la méthode du maximum de vraisemblance, la fonction de vraisemblance doit avoir une forme explicite et est souvent bornée sur l'espace des paramètres à estimer. Il y a plusieurs modèles en économétrie et en finance pour lesquels, il est très difficile d'appliquer la méthode du maximum de vraisemblance, aussi bien dans le cas d'observations dépendantes que dans le cas d'observations indépendantes et identiquement distribuées (i.i.d.). Nous pouvons citer particulièrement les modèles α -stables : modèles ARMA α -stables, ARFIMA α -stables et les processus ARFISMA α -stables, pour lesquels la densité de probabilité n'a pas de forme explicite. Toutefois, la fonction de vraisemblance peut être non bornée mais sa transformée de Fourier est toujours bornée. De plus, si la vraisemblance n'admet pas de forme simple, la transformée de Fourier peut avoir une expression explicite. Puisque, la transformée de Fourier de la densité de probabilité est la fonction caractéristique (CF), on peut alors exploiter la fonction caractéristique empirique (ECF) pour estimer le système de paramètres. Dans ce chapitre, notre objectif est d'estimer simultanément l'ensemble des paramètres d'un processus ARFISMA symétrique α -stable, en utilisant la méthode de la fonction caractéristique empirique (ECF). Nous comparons également la performance de cette approche avec celle de la méthode de Whittle par MCMC proposée dans le Chapitre 3.

Ce chapitre est organisé comme suit. La Section 4.2 est consacrée à la présentation de la méthode de la fonction caractéristique empirique. Dans la Section 4.3, on établit les conditions nécessaires pour obtenir la consistance et la normalité asymptotique de cet estimateur. La Section 4.4 donne les résultats de simulations par Monte Carlo. La Section 4.6 est dédiée aux conclusions et aux perspectives.

4.2 Procédure d'estimation

Nous nous intéressons, dans ce paragraphe à l'estimation simultanée de l'ensemble des paramètres d'un processus ARFISMA symétrique α -stable, par la méthode de la fonction caractéristique empirique (ECF : Empirical Characteristic Function). La motivation de cette méthode ECF apparaît par le fait que, deux distributions sont

égales, si et seulement si, leurs fonctions caractéristiques sont identiques (voir Lukacs [52], page 28). L'idée fondamentale est de minimiser la distance entre la fonction caractéristique empirique et la fonction caractéristique. Il existe plusieurs variétés de méthodes ECF et l'approche dans le cas dépendant n'est pas, exactement, la même dans le cas d'observations i.i.d., car la dépendance doit être prise en compte dans la procédure. Dans la section suivante, nous résumons l'approche dans le cas dépendant et pour plus d'informations, nous renvoyons à Feuerverger [26], Knight et Satchell [42], Knight et Yu [43], Carrasco et Florens [16], Jiang et Knight [39] et Yu [79].

Nous considérons une trajectoire finie X_1, \dots, X_T issue d'un processus ARFISMA symétrique α -stable, centré, stationnaire et inversible, défini par l'équation (2.1). Nous supposons, sans perte de généralité, que le paramètre de dispersion γ du processus d'innovations est connu et égal à l'unité. Nous nous intéressons à l'estimation des paramètres de mémoire longue (d, D) , des paramètres de mémoire courte de la partie non saisonnière $\zeta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$, des paramètres de mémoire courte de la partie saisonnière $\xi = (\Phi_1, \dots, \Phi_P, \Theta_1, \dots, \Theta_Q)$ et de l'exposant caractéristique α du processus d'innovations. Notons alors par $\psi = (\alpha, d, D, \zeta, \xi)'$ le vecteur des paramètres à estimer et par ψ_0 sa vraie valeur. L'espace des paramètres à estimer est alors défini par :

$$\begin{aligned} \Psi = \{ & \psi \in \mathbb{R}^{p+q+P+Q+3} : \phi_p \neq 0, \theta_q \neq 0, \Phi_P \neq 0, \Theta_Q \neq 0, \text{ les polynômes} \\ & \Theta(z), \Phi(z), \Phi_s(z^s) \text{ et } \Theta_s(z^s) \text{ n'ont pas de racines communes} \\ & \text{dans le cercle unité, } |d+D| < 1 - \frac{1}{\alpha} \text{ et } |D| < 1 - \frac{1}{\alpha}, 1 < \alpha \leq 2\}. \end{aligned}$$

Notre objectif est d'estimer le paramètre ψ à partir de la réalisation (X_1, \dots, X_T) , par la méthode ECF. Dans le cas dépendant, la procédure nécessite la construction de blocs adjacents à partir des observations (X_1, \dots, X_T) , de la manière suivante :

$$Y_j = (X_j, \dots, X_{j+m})', \text{ pour } j = 1, \dots, T - m.$$

Ainsi, chaque bloc contient $m + 1$ observations. La fonction caractéristique jointe de chaque bloc est définie par

$$c(r, \psi) = \mathbb{E}(\exp(ir'Y_j)), \text{ où } r = (r_1, \dots, r_{m+1})' \text{ est la variable de transformation.}$$

La fonction caractéristique empirique jointe est donnée par :

$$c_n(r) = \frac{1}{n} \sum_{j=1}^n \exp(ir'Y_j), \text{ où } n = T - m.$$

Par construction, la fonction caractéristique empirique est une contrepartie empirique de la fonction caractéristique et contient toute l'information sur les données, alors que la fonction caractéristique fournit l'information sur les paramètres. Il est alors possible d'avoir une idée sur les paramètres, à partir de la fonction caractéristique empirique $c_n(r)$. Ainsi, on appelle estimateur de ψ par la méthode ECF, noté $\widehat{\psi}_n$, la valeur de ψ qui minimise la fonction objective suivante :

$$I_n(\psi) = \int \dots \int |c_n(r) - c(r; \psi)|^2 g(r) dr, \quad (4.1)$$

où $g(r)$ est une fonction de poids continue. Ou de manière équivalente, on peut minimiser

$$\int \dots \int |c_n(r) - c(r; \psi)|^2 dG(r), \quad (4.2)$$

ou résoudre l'équation suivante :

$$\int \dots \int (c_n(r) - c(r; \psi)) W(r, \psi) dr = 0, \quad (4.3)$$

où $G(r)$ et $W(r, \psi)$ sont des fonctions de poids. Sous certaines conditions de régularité, ces trois procédures sont équivalentes. Cependant, pour faciliter l'application de la méthode ECF, il faut s'assurer d'abord, que la fonction caractéristique jointe admet une expression explicite et ensuite être en mesure de déterminer la valeur de la taille des blocs m et la fonction de poids à utiliser. Pour cela, nous montrons d'abord dans le théorème suivant, que la fonction caractéristique jointe a une forme explicite.

Théorème 4.2.1 Soit $(X_t)_{t \in \mathbb{Z}}$ un processus ARFISMA symétrique α -stable défini par l'équation (2.1). La fonction caractéristique jointe de $X_{t-m}, X_{t-m+1}, \dots, X_t$ est donnée par :

$$c(r_1, r_2, \dots, r_{m+1}; \psi) = \exp \left\{ - \sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^\alpha - \sum_{l=2}^{m+1} \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^\alpha \right\},$$

où les coefficients $(c_j)_{j \geq 0}$ sont déterminés par les équations (2.10), (2.11) et (2.12).

Preuve : Par définition, nous avons :

$$\begin{aligned} c(r_1, \dots, r_{m+1}; \psi) &= \mathbb{E}(\exp(ir'Y_j)) \\ &= \mathbb{E}[\exp\{ir_1 X_{t-m} + ir_2 X_{t-m+1} + \dots + ir_{m+1} X_t\}]. \end{aligned}$$

En utilisant la représentation moyenne mobile infinie donnée par l'équation (2.17), nous pouvons écrire :

$$c(r; \psi) = \mathbb{E} \left[\exp \left\{ ir_1 \sum_{j=0}^{\infty} c_j Z_{t-m-j} + ir_2 \sum_{j=0}^{\infty} c_j Z_{t-m+1-j} + \dots + ir_{m+1} \sum_{j=0}^{\infty} c_j Z_{t-j} \right\} \right].$$

En développant, puis en regroupant les termes en Z_{t-m-j} , pour j allant de 0 à ∞ , nous obtenons :

$$\begin{aligned} c(r; \psi) &= \mathbb{E} \left[\exp \left\{ i \sum_{j=0}^{\infty} \left(\sum_{l=0}^m r_{l+1} c_{j+l} \right) Z_{t-m-j} + i \sum_{h=0}^{m-1} r_{h+2} c_h Z_{t-m+1} + \cdots + i r_{m+1} c_0 Z_t \right\} \right] \\ &= \mathbb{E} \left[\exp \left\{ i \sum_{j=0}^{\infty} \left(\sum_{l=0}^m r_{l+1} c_{j+l} \right) Z_{t-m-j} + i \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} r_{h+l} c_h \right) Z_{t-m-1+l} \right\} \right] \end{aligned}$$

En utilisant l'indépendance des Z_t , nous aurons :

$$\begin{aligned} c(r; \psi) &= \mathbb{E} \left[\exp \left\{ i \sum_{j=0}^{\infty} \left(\sum_{l=0}^m r_{l+1} c_{j+l} \right) Z_{t-m-j} \right\} \right] \\ &\quad \times \mathbb{E} \left[\exp \left\{ i \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} r_{h+l} c_h \right) Z_{t-m-1+l} \right\} \right] \end{aligned}$$

Posons :

$$\begin{aligned} A &= \mathbb{E} \left[\exp \left\{ i \sum_{j=0}^{\infty} \left(\sum_{l=0}^m r_{l+1} c_{j+l} \right) Z_{t-m-j} \right\} \right] \\ B &= \mathbb{E} \left[\exp \left\{ i \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} r_{h+l} c_h \right) Z_{t-m-1+l} \right\} \right]. \end{aligned}$$

Nous avons :

$$\begin{aligned} B &= \mathbb{E} \left[\prod_{l=2}^{m+1} \exp \left\{ i \left(\sum_{h=0}^{m+1-l} r_{h+l} c_h \right) Z_{t-m-1+l} \right\} \right] \\ &= \prod_{l=2}^{m+1} \mathbb{E} \left[\exp \left\{ i \left(\sum_{h=0}^{m+1-l} r_{h+l} c_h \right) Z_{t-m-1+l} \right\} \right], \text{ par indépendance des } Z_t \\ &= \prod_{l=2}^{m+1} \Phi_Z \left(\sum_{h=0}^{m+1-l} r_{h+l} c_h \right), \text{ car les } Z_t \text{ ont même loi que } Z \\ &= \prod_{l=2}^{m+1} \exp \left\{ - \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^\alpha \right\}, \text{ car } Z \sim \text{SaS de dispersion égal à } 1 \\ &= \exp \left\{ - \sum_{l=2}^{m+1} \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^\alpha \right\}. \end{aligned}$$

Pour calculer l'expression de A , posons $Y = \sum_{j=0}^{\infty} \left(\sum_{l=0}^m r_{l+1} c_{j+l} \right) Z_{t-m-j}$ et remarquons que $A = \Phi_Y(1)$, où Φ_Y est la fonction caractéristique de Y .

Nous avons $\sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} < \infty$. En effet

$$\left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} \leq \left(\sum_{l=0}^m |r_{l+1} c_{j+l}| \right)^{\alpha}$$

En appliquant l'inégalité de Hölder, nous avons

$$\begin{aligned} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} &\leq \left(\sum_{l=0}^m |c_{j+l}|^{\alpha} \right) \left(\sum_{l=0}^m |r_{l+1}|^{\frac{\alpha}{\alpha-1}} \right)^{\alpha-1} \\ &\leq K_{\alpha}(r) W_j, \end{aligned}$$

où $K_{\alpha}(r) = \left(\sum_{l=0}^m |r_{l+1}|^{\frac{\alpha}{\alpha-1}} \right)^{\alpha-1}$ est une constante dépendant de r et de α .

$W_j = \sum_{l=0}^m |c_{j+l}|^{\alpha}$ est le terme général d'une série convergente, car

$$\sum_{j=0}^{\infty} W_j = \sum_{j=0}^{\infty} \left(\sum_{l=0}^m |c_{j+l}|^{\alpha} \right) = \sum_{l=0}^m \left(\sum_{j=0}^{\infty} |c_{j+l}|^{\alpha} \right) < \infty,$$

puisque, d'après les conditions de stationnarité du modèle, on a $\sum_{j=0}^{\infty} |c_j|^{\alpha} < 1$. Ce qui

entraîne que, $\sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} < \infty$.

Les variables aléatoires $(Z_t)_t$ sont i.i.d selon une loi symétrique α -stable et de plus

$\sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} < \infty$, donc la variable aléatoire Y est aussi symétrique α -stable et on a

$$Y \stackrel{\mathcal{D}}{=} \left(\sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} \right)^{1/\alpha} Z_1.$$

(Voir Cline[18] ou Cline et Brockwell [19], pour plus de détails.)

Par conséquent nous avons :

$$\begin{aligned}
\Phi_Y(u) &= \mathbb{E} \left[\exp \left\{ iu \sum_{j=0}^{\infty} \left(\sum_{l=0}^m r_{l+1} c_{j+l} \right) Z_{t-m-j} \right\} \right] \\
&= \mathbb{E} \left[\exp \left\{ iu \left(\sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right| \right)^{\alpha} Z_1 \right\} \right] \\
&= \Phi_{Z_1} \left(u \left(\sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right| \right)^{\alpha} \right) \\
&= \exp \left\{ -|u|^{\alpha} \sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} \right\}
\end{aligned}$$

$$\text{D'où } A = \Phi_Y(1) = \exp \left\{ - \sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} \right\}.$$

Finalement, on obtient :

$$\begin{aligned}
c(r; \psi) &= A \times B \\
&= \exp \left\{ - \sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} \right\} \times \exp \left\{ - \sum_{l=2}^{m+1} \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^{\alpha} \right\} \\
&= \exp \left\{ - \sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} - \sum_{l=2}^{m+1} \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^{\alpha} \right\}.
\end{aligned}$$

■

Concernant le choix de la fonction de poids, si nous nous référons à l'article de Knight et Yu [43] ou Yu [79], nous pouvons obtenir une fonction de poids optimale $W^*(r, \psi)$ qui s'écrit sous la forme suivante :

$$W^*(r, \psi) = \int \dots \int \exp(ir'Y_j) \frac{\partial \log f(X_{j+m}|X_j, \dots, X_{j+m-1})}{\partial \psi} dX_j \dots dX_{j+m}, \quad (4.4)$$

où $f(X_{j+m}|X_j, \dots, X_{j+m-1})$ est la fonction de score conditionnelle. Feuerverger [26] montre que, pour m assez grand, cette fonction de poids est optimale, dans le sens où l'estimateur basé sur l'équation (4.3) et $W^*(r, \psi)$ peut atteindre l'efficacité de l'estimateur du maximum de vraisemblance. Par exemple, Knight et Yu [43] ont donné l'expression explicite de $W^*(r, \psi)$, pour un processus ARMA gaussien pour lequel la fonction de score

conditionnelle est connue. Ils ont également montré, par simulation, que l'estimateur basé sur l'équation (4.3) avec la fonction de poids $W^*(r, \psi)$ a de très bonnes performances, en comparaison avec la méthode du maximum de vraisemblance. Cependant, dans le cas stable, cette fonction de score conditionnelle n'admet pas une forme explicite, par conséquent cette procédure ne sera pas considérée dans la suite du document. Ainsi, nous adaptons ici une méthode ECF avec une fonction de poids exponentielle de type : $g(r) = \exp(-r'r)$ qui est très utilisée en pratique. L'avantage majeur de l'utilisation d'une fonction de poids exponentielle est que, l'intégrale (4.1) peut être calculée numériquement, en utilisant une méthode de quadrature ou une méthode d'intégration par Monte Carlo. Cependant, il faut noter que la fonction de poids exponentielle n'est pas, en général, optimale.

En ce qui concerne le choix de la taille des blocs, il est important de noter que l'efficacité de l'estimateur par la méthode ECF dépend fortement du choix de m . En effet, la quantité d'information fournie par les blocs sur l'échantillon croît avec la valeur de m . Il y a, en général, un compromis entre petite et grande valeur de m . Comme le processus (2.1) dont nous disposons, n'admet pas une propriété de Markov, une grande valeur de m marcherait mieux qu'une petite valeur de m , en terme d'efficacité asymptotique, car les blocs retiennent toute l'information sur la série de départ. Cependant, il n'existe pas, dans la littérature un outil théorique qui guide le choix du paramètre m . Ainsi, nous proposons de déterminer m empiriquement, en minimisant l'erreur quadratique moyenne de l'estimateur par la méthode ECF. Plus précisément, soit $\hat{\psi}^{(m)}$ l'estimateur par la méthode ECF, pour m donné. Alors la valeur optimale de m , notée m_{opt} , est telle que

$$m_{opt} = \arg \min_m \frac{1}{R} \sum_{j=1}^R \left(\hat{\psi}_j^{(m)} - \psi_0 \right)^2, \quad (4.5)$$

où R est le nombre de réplcation de l'échantillon, ψ_0 la vraie valeur de ψ et $\hat{\psi}_j^{(m)}$ est l'estimateur pour l'échantillon j .

4.3 Propriétés asymptotiques

Les propriétés asymptotiques de l'estimateur par la méthode ECF, dans le cas i.i.d., ont été établies par Heathcote [35]. Dans le cas dépendant, Knight et Yu [43] ont montré, sous certaines conditions de régularité, la consistance et la normalité asymptotique de cet estimateur. Notre travail consiste, dans ce paragraphe, à établir les conditions nécessaires permettant de vérifier les hypothèses du Théorème 2.1 de Knight et Yu [43]. Les hypothèses suivantes sont alors nécessaires :

Hypothèse 4.3.1 : On suppose que ψ_0 se trouve à l'intérieur de Ψ , où Ψ est un compact de $\mathbb{R}^{p+q+P+Q+3}$.

Hypothèse 4.3.2 : Les coefficients $(c_j)_{j \geq 0}$ de la représentation moyenne mobile infinie (2.17) vérifient les propriétés suivantes :

- 1) $\sum_{j=0}^{\infty} |c_j|^{2(\alpha-1)} < \infty, \quad \forall 1 < \alpha \leq 2.$
- 2) $\sum_{j=0}^{\infty} \left| \frac{\partial c_j}{\partial \delta} \right|^2 < \infty$ et $\sum_{j=0}^{\infty} \left| \frac{\partial^2 c_j}{\partial \delta \partial \delta'} \right|^2 < \infty,$ avec $\delta = (d, D, \zeta, \xi).$
- 3) $\sum_{j=0}^{\infty} \left(\frac{\partial}{\partial \delta} \log \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right| \right)^4 < \infty.$

Hypothèse 4.3.3 : $\int \dots \int |r_j|^6 g(r) dr < \infty, \quad \forall j = 1, \dots, m+1.$

Hypothèse 4.3.4 : Soit $I_0(\psi) = \int \dots \int |c(r; \psi_0) - c(r; \psi)|^2 g(r) dr$ et $I_0(\psi) = 0$ si et seulement si $\psi = \psi_0.$

Hypothèse 4.3.5 : Soit \mathfrak{S}_j la σ -algèbre telle que $\{K_j, \mathfrak{S}_j\}$ est une suite stochastique adaptée, où $K_j = K(Y_j; \psi)$ avec $K(x; \psi)$ est la fonction définie par :

$$K(x; \psi) = \int \dots \int (\cos(r'x) - c(r; \psi)) \frac{\partial}{\partial \psi} c(r; \psi) g(r) dr.$$

On peut considérer que \mathfrak{S}_j est la σ -algèbre engendrée par le passé historique des $K_j.$ Soit $\nu_j = \mathbb{E}[K_0 | K_{-j}, K_{-j-1}, \dots] - \mathbb{E}[K_0 | K_{-j-1}, K_{-j-2}, \dots],$ pour tout $j \geq 0.$ Supposons que $\mathbb{E}[K_0 | \mathfrak{S}_{-l}]$ converge en moyenne quadratique vers 0, lorsque l tend vers ∞ et $\sum_{j=0}^{\infty} \mathbb{E}[\nu'_j \nu_j] < 0.$

Hypothèse 4.3.6 : $g(r)$ est une densité de probabilité sur $\mathbb{R}^{m+1}.$

Remarque 4.3.7

- 1) L'**Hypothèse 4.3.2** et l'**Hypothèse 4.3.3** sont techniques et permettent de vérifier les hypothèses (A2), (A3), (A5) et (A6) du Théorème 2.1 de Knight et Yu [43]. L'**Hypothèse 4.3.2** n'entrave pas les conditions de stationnarité du modèle. En effet, nous avons $2(\alpha - 1) \leq \alpha$ et $|c_j| < 1.$ Ce qui implique que :

$$2(\alpha - 1) \log |c_j| \geq \alpha \log |c_j| \implies |c_j|^{2(\alpha-1)} \geq |c_j|^\alpha$$

Par conséquent, $\sum_{j=0}^{\infty} |c_j|^{2(\alpha-1)} < \infty \implies \sum_{j=0}^{\infty} |c_j|^\alpha < \infty.$

- 2) Les Hypothèses "**Hypothèse 4.3.1**", "**Hypothèse 4.3.4**", "**Hypothèse 4.3.5**" et "**Hypothèse 4.3.6**" sont respectivement analogues aux hypothèses (A1), (A4), (A7) et (A8) de Knight et Yu [43]. L'**Hypothèse 4.3.1** assure la compacité de l'espace des paramètres. L'**Hypothèse 4.3.4** garantit que les paramètres sont identifiables. L'**Hypothèse 4.3.5** fournit les conditions nécessaires à l'application de la loi des grands nombres et du théorème de la limite centrale pour un processus strictement stationnaire et ergodique. Cette hypothèse est également analogue à l'hypothèse "Assumption 3.5" de Hansen [33] qui a établi les propriétés asymptotiques de l'estimateur par la méthode des moments généralisés.

Les propriétés asymptotiques de l'estimateur basé sur la fonction caractéristique jointe sont résumées dans le théorème suivant :

Théorème 4.3.8 Soit $\hat{\psi}_n = \arg \min_{\psi \in \Psi} I_n(\psi)$. Si les Hypothèses "**Hypothèse 4.3.1**" - "**Hypothèse 4.3.4**" et "**Hypothèse 4.3.6**" sont vérifiées, alors $\hat{\psi}_n \xrightarrow{p.s.} \psi_0$. Si de plus, l'**Hypothèse 4.3.5** est vraie, alors

$$\sqrt{n} \left(\hat{\psi}_n - \psi_0 \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, B^{-1}(\psi_0) A(\psi_0) B^{-1}(\psi_0) \right),$$

$$\text{où } A(\psi_0) = \text{Var} (K(Y_1; \psi_0)) + 2 \sum_{j=2}^{\infty} \text{cov} (K(Y_1; \psi_0), K(Y_j; \psi_0))$$

et $B^{-1}(\psi_0)$ est l'inverse de la matrice $B(\psi_0)$ définie par :

$$B(\psi_0) = \int \dots \int \frac{\partial}{\partial \psi} c(r; \psi_0) \frac{\partial}{\partial \psi'} c(r; \psi_0) g(r) dr$$

■

La preuve de ce théorème est présentée dans l'Annexe B.

4.4 Simulations de Monte Carlo

Dans cette section, nous allons essayer de mettre l'accent sur la validité des propriétés de l'estimateur ECF proposé dans ce chapitre, à partir des simulations par la méthode de Monte Carlo. Les expériences de simulation ont aussi pour but de comparer la performance de cet estimateur à celle de l'estimateur de Whittle par MCMC, car les études de simulation effectuées dans le Chapitre 3 justifient ce choix. Pour ce faire, nous effectuons une expérience de 1500 échantillons simulés pour différents processus ARFISMA symétriques α -stables, de taille fixée à $T = 1500$, avec une période de

saisonnalité $s = 4$. Ces différents modèles sont résumés dans le Tableau 4.1 suivant :

Paramètres	α	d	D	ϕ	θ
Modèle 1	1.6	0.15	0.20	0	0
Modèle 2	1.6	0.15	0.20	0.60	0
Modèle 3	1.6	0.15	0.20	0	0.40
Modèle 4	1.6	0.15	0.20	0.60	0.40

TAB. 4.1 – Processus générateur de données.

Pour chaque modèle, nous étudions l'effet de la taille des blocs sur l'estimation, en faisant varier le paramètre m ($m = 1, 2, \dots, 8$). Ceci permet de déterminer la valeur optimale de m suivant le critère défini par l'équation (4.5) et nous comparons l'estimateur obtenu avec l'estimateur de Whittle par MCMC. D'autre part, nous faisons varier les valeurs de ϕ et de θ , afin d'étudier l'influence des paramètres courte mémoire sur les paramètres longue mémoire. Pour étudier l'efficacité des estimateurs, nous déterminons la moyenne (Mean), la racine de l'erreur quadratique moyenne (RMSE) et l'erreur moyenne absolue (MAE).

Nous montrons d'abord, comment on calcule l'estimateur ECF, en pratique pour $m = 1$ et pour les autres valeurs de m , la procédure est la même. Notons pour $m = 1$, on obtient :

$$c(r_1, r_2; \psi) = \exp \left(- \sum_{j=0}^{\infty} |r_1 c_j + r_2 c_{j+1}|^\alpha - |r_2|^\alpha \right) \quad (4.6)$$

et $c_n(r_1, r_2) = \frac{1}{n} \sum_{j=1}^n \exp(ir_1 X_j + ir_2 X_{j+1})$, $n = T - 1$.

Les coefficients $(c_j)_{j \geq 0}$ sont donnés par les équations (2.10), (2.11) et (2.12). En considérant une fonction de poids exponentielle, la procédure consiste alors à minimiser :

$$I_n(\psi) = \int \int |c_n(r_1, r_2) - c(r_1, r_2; \psi)|^2 \exp(-r_1^2 - r_2^2) dr_1 dr_2. \quad (4.7)$$

Par une méthode d'intégration par Monte Carlo, l'intégrale double (4.7) est approchée, pour N assez grand que nous fixons ici à 5000, par la somme suivante :

$$\widehat{I}_n(\psi) = \frac{1}{N} \sum_{j=1}^N |c_n(r_{1,j}, r_{2,j}) - c(r_{1,j}, r_{2,j}; \psi)|^2, \quad (4.8)$$

où $r_{1,j}$ et $r_{2,j}$ sont i.i.d. suivant une loi normale de moyenne nulle et de variance $1/2$. Concernant la troncature de la somme infinie dans l'expression (4.6), il faut noter que Stoev et Taqqu [74] montrent par des simulations, que pour bien représenter le comportement de mémoire longue, un paramètre de troncature supérieur ou égal à la taille de l'échantillon marche mieux, en général. Partant de ce résultat, des moyens de calcul limités dont nous disposons et à la lourdeur des programmes, nous avons tronqué cette somme à la taille de l'échantillon $T = 1500$. Enfin pour obtenir $\hat{\psi}_n$, on minimise sous le compact Ψ , l'équation (4.8). Ceci peut se faire, sous le logiciel R, grâce aux algorithmes d'optimisation disponibles avec les fonctions **optim**, **contrsOptim**,...

4.4.1 Choix empirique de la taille des blocs m

Dans cette section, nous cherchons à déterminer empiriquement la valeur du paramètre m convenable pour chacun des modèles présentés dans le Tableau 4.1. Les Tableaux 4.2-4.5 donnent les résultats de simulations pour ces modèles.

Le Tableau 4.2 donne les paramètres estimés du modèle ARFISMA(0,d,0)(0,D,0) symétrique α -stable, sans partie courte mémoire. Les résultats sont satisfaisants quelle que soit la valeur de m , car les RMSE et les MAE sont tous faibles. Cependant, cette performance varie avec la valeur de m . Ainsi, pour les paramètres d et D , les meilleurs résultats sont obtenus pour $m = 1$, alors que pour le paramètre α , c'est la valeur $m = 4$ qui donne la meilleure performance. De plus, pour le paramètre α , l'écart des RMSE et MAE entre $m = 1$ et $m = 4$ est relativement faible (soit 1.18%). Par conséquent, nous pouvons retenir la valeur $m = 1$, comme valeur optimale de la taille des blocs, pour estimer simultanément l'ensemble des paramètres du Modèle 1.

Le Tableau 4.3 présente les résultats du modèle ARFISMA(1, d, 0)(0, D, 0) symétrique α -stable. Nous constatons que, lorsque nous introduisons une partie autorégressive dans le Modèle 1, les RMSE et les MAE deviennent larges pour toutes les valeurs de m . Ce qui permet de conclure que la présence du paramètre ϕ perturbe l'estimation des autres paramètres et la valeur $m = 1$ ne suffit plus pour estimer correctement les paramètres. En effet, nous constatons que pour les paramètres α , D et ϕ les meilleurs résultats sont obtenus pour $m = 6$, alors que pour le paramètre d , nous avons une meilleure performance pour $m = 5$. De plus, l'écart des RMSE et MAE entre $m = 5$ et $m = 6$, pour le paramètre d , est de 0.2% (RMSE) et 0.1% (MAE). Ainsi, nous choisissons $m = 6$ comme valeur optimale de m pour le Modèle 2.

Nous introduisons maintenant une partie moyenne mobile dans le Modèle 1 et les résultats sont présentés dans le Tableau 4.4. Dans ce tableau, nous notons les mêmes conclusions

que dans l'analyse précédentes. En terme de RMSE et de MAE, nous observons clairement que la valeur optimale de m , pour l'estimation de l'ensemble des paramètres du Modèle 3, est $m = 6$.

Cette dernière expérience consiste à introduire dans le Modèle 1, une partie autorégressive et une partie moyenne mobile, et les résultats sont donnés dans le Tableau 4.5. Nous constatons d'abord que les résultats sont très mauvais pour $m = 1$, surtout pour les paramètres d , D et ϕ . Ce qui montre que, s'il y a présence de paramètres courte mémoire dans le modèle, la quantité d'information contenue dans les blocs pour $m = 1$ est insuffisante pour estimer correctement les paramètres. En comparant, les RMSE et MAE, nous observons bien que la valeur optimale de m , qui donne la meilleure performance, est $m = 6$.

Au vue des résultats de simulations, nous pouvons dire que la performance de la méthode ECF dépend fortement du choix de la taille des blocs m . D'où la nécessité de déterminer la valeur optimale de m . Ainsi, si nous disposons du modèle simple sans partie courte mémoire (Modèle 1), la taille des blocs $m = 1$ suffit pour estimer l'ensemble des paramètres du modèle. Par contre, si nous introduisons dans le Modèle 1 une partie courte mémoire, la taille $m = 6$ semble donner la meilleure performance.

Statistiques	$m = 1$			$m = 2$		
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\alpha}$	\hat{d}	\hat{D}
Mean	1.6158	0.1321	0.1707	1.5942	0.1307	0.1558
RMSE	0.0158	0.0178	0.0292	0.0060	0.0192	0.0441
MAE	0.0158	0.0178	0.0292	0.0057	0.0192	0.0441

	$m = 3$			$m = 4$		
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\alpha}$	\hat{d}	\hat{D}
Mean	1.5737	0.1289	0.1363	1.5960	0.1287	0.1458
RMSE	0.0263	0.0212	0.0637	0.0047	0.0213	0.0542
MAE	0.0262	0.0210	0.0636	0.0040	0.0212	0.0541

	$m = 5$			$m = 6$		
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\alpha}$	\hat{d}	\hat{D}
Mean	1.5915	0.1313	0.1445	1.5934	0.1291	0.1446
RMSE	0.0090	0.0186	0.0554	0.0075	0.0209	0.0554
MAE	0.0084	0.0186	0.0554	0.0066	0.0208	0.0553

	$m = 7$			$m = 8$		
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\alpha}$	\hat{d}	\hat{D}
Mean	1.5963	0.1276	0.1451	1.6016	0.1259	0.1488
RMSE	0.0059	0.0225	0.0549	0.0061	0.0242	0.0512
MAE	0.0047	0.0223	0.0548	0.0048	0.0240	0.0511

TAB. 4.2 – Paramètres estimés du Modèle 1, pour $m = 1, 2, \dots, 8$.

Statistiques	$m = 1$				$m = 2$			
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$
Mean	1.726	0.217	0.233	0.464	1.683	0.125	0.252	0.633
RMSE	0.288	0.208	0.263	0.366	0.171	0.117	0.132	0.192
MAE	0.168	0.174	0.147	0.292	0.122	0.100	0.089	0.162

	$m = 3$				$m = 4$			
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$
Mean	1.663	0.125	0.243	0.641	1.650	0.128	0.235	0.644
RMSE	0.135	0.109	0.101	0.167	0.120	0.106	0.087	0.154
MAE	0.100	0.095	0.073	0.140	0.085	0.093	0.063	0.130

	$m = 5$				$m = 6$			
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$
Mean	1.634	0.160	0.228	0.606	1.624	0.169	0.222	0.591
RMSE	0.098	0.097	0.070	0.119	0.082	0.099	0.057	0.103
MAE	0.068	0.078	0.047	0.089	0.061	0.079	0.040	0.083

	$m = 7$				$m = 8$			
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$
Mean	1.571	0.250	0.187	0.394	1.568	0.248	0.189	0.395
RMSE	0.092	0.135	0.059	0.258	0.093	0.132	0.056	0.247
MAE	0.065	0.114	0.043	0.219	0.067	0.110	0.041	0.210

TAB. 4.3 – Paramètres estimés du Modèle 2, pour $m = 1, 2, \dots, 8$.

Statistiques	$m = 1$				$m = 2$			
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\theta}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\theta}$
Mean	1.694	0.177	0.250	0.339	1.672	0.165	0.240	0.399
RMSE	0.209	0.064	0.130	0.153	0.156	0.056	0.094	0.059
MAE	0.133	0.046	0.077	0.090	0.112	0.041	0.067	0.041

	$m = 3$				$m = 4$			
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\theta}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\theta}$
Mean	1.659	0.169	0.239	0.394	1.637	0.170	0.230	0.398
RMSE	0.135	0.057	0.084	0.052	0.102	0.059	0.070	0.054
MAE	0.097	0.042	0.059	0.039	0.074	0.044	0.047	0.041

	$m = 5$				$m = 6$			
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\theta}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\theta}$
Mean	1.630	0.170	0.225	0.398	1.534	0.126	0.143	0.343
RMSE	0.095	0.064	0.066	0.059	0.066	0.023	0.056	0.057
MAE	0.067	0.046	0.043	0.044	0.065	0.023	0.056	0.056

	$m = 7$				$m = 8$			
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\theta}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\theta}$
Mean	1.622	0.171	0.223	0.396	1.619	0.170	0.222	0.396
RMSE	0.090	0.071	0.070	0.067	0.083	0.069	0.075	0.068
MAE	0.063	0.049	0.043	0.051	0.061	0.050	0.044	0.053

TAB. 4.4 – Paramètres estimés du Modèle 3, pour $m = 1, 2, \dots, 8$.

Stat.	$m = 1$					$m = 2$				
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\theta}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\theta}$
Mean	1.488	0.356	0.0008	0.002	0.331	1.508	0.275	0.026	0.276	0.411
RMSE	0.111	0.206	0.199	0.597	0.069	0.091	0.125	0.173	0.323	0.012
MAE	0.111	0.206	0.199	0.597	0.068	0.091	0.125	0.173	0.323	0.011

Stat.	$m = 3$					$m = 4$				
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\theta}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\theta}$
Mean	1.513	0.262	0.045	0.288	0.402	1.514	0.222	0.089	0.349	0.365
RMSE	0.086	0.112	0.155	0.312	0.005	0.085	0.073	0.110	0.250	0.034
MAE	0.086	0.112	0.154	0.311	0.004	0.085	0.072	0.110	0.250	0.034

Stat.	$m = 5$					$m = 6$				
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\theta}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\theta}$
Mean	1.520	0.178	0.118	0.451	0.351	1.513	0.147	0.133	0.510	0.334
RMSE	0.080	0.031	0.081	0.151	0.049	0.087	0.017	0.066	0.095	0.066
MAE	0.079	0.028	0.081	0.148	0.048	0.086	0.013	0.066	0.089	0.065

Stat.	$m = 7$					$m = 8$				
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\theta}$	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\theta}$
Mean	1.492	0.135	0.139	0.525	0.305	1.456	0.132	0.142	0.503	0.268
RMSE	0.109	0.028	0.062	0.089	0.096	0.148	0.041	0.060	0.124	0.135
MAE	0.107	0.022	0.060	0.077	0.094	0.143	0.032	0.057	0.105	0.131

TAB. 4.5 – Paramètres estimés du Modèle 4, pour $m = 1, 2, \dots, 8$.

4.4.2 Comparaison des estimateurs

Après avoir déterminé la valeur optimale de la taille des blocs, pour chaque modèle du Tableau 4.1, nous comparons la performance de la méthode ECF avec celle de la méthode de Whittle par MCMC définie dans le Chapitre 3 (page 61). Les résultats de simulations sont présentés dans les Tableaux 4.6-4.9. Nous observons d'abord que, pour le modèle simple, la méthode de Whittle par MCMC donne une meilleure performance que la méthode ECF. Nous notons également que ces deux méthodes sont très sensibles à la présence de paramètres courte mémoire, car les RMSE et MAE des Tableaux 4.7-4.9 sont plus larges que ceux du Tableau 4.6. Nous retrouvons un fait souvent noté dans la littérature : l'estimation des paramètres courte mémoire est plus sensible dès que l'on est en présence de paramètres longue mémoire. Malgré la fonction de poids sous optimale utilisée, nous remarquons, en général, que s'il y a présence de paramètre courte mémoire dans le modèle, la méthode ECF donne une meilleure performance. Elle permet également d'avoir une idée sur la distribution du bruit, ce qui n'est pas le cas pour la méthode de Whittle par MCMC. En cela l'approche de la méthode ECF nous semble très intéressante.

Statistiques	ECF (m=1)		\hat{D}	MCMC	
	$\hat{\alpha}$	\hat{d}		\hat{d}	\hat{D}
Mean	1.6158	0.1321	0.1707	0.1565	0.1844
RMSE	0.0158	0.0178	0.0292	0.0132	0.0193
MAE	0.0158	0.0178	0.0292	0.0106	0.0164

TAB. 4.6 – Comparaison entre les méthodes ECF et Whittle-MCMC (Modèle 1)

Statistiques	ECF (m=6)				MCMC		
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	\hat{d}	\hat{D}	$\hat{\phi}$
Mean	1.624	0.169	0.222	0.591	0.164	0.200	0.587
RMSE	0.082	0.099	0.057	0.103	0.164	0.079	0.104
MAE	0.061	0.079	0.040	0.083	0.077	0.026	0.073

TAB. 4.7 – Comparaison entre les méthodes ECF et Whittle-MCMC (Modèle 2)

Statistiques	ECF (m=6)				MCMC		
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\theta}$	\hat{d}	\hat{D}	$\hat{\theta}$
Mean	1.534	0.126	0.143	0.343	0.157	0.204	0.395
RMSE	0.066	0.023	0.056	0.057	0.064	0.068	0.048
MAE	0.065	0.023	0.056	0.056	0.026	0.022	0.029

TAB. 4.8 – Comparaison entre les méthodes ECF et Whittle-MCMC (Modèle 3)

Statistiques	ECF (m=6)					MCMC			
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\theta}$	\hat{d}	\hat{D}	$\hat{\phi}$	$\hat{\theta}$
Mean	1.513	0.147	0.133	0.510	0.334	0.146	0.182	0.610	0.172
RMSE	0.087	0.017	0.066	0.095	0.066	0.050	0.025	0.062	0.233
MAE	0.086	0.013	0.066	0.089	0.065	0.040	0.021	0.049	0.227

TAB. 4.9 – Comparaison entre les méthodes ECF et Whittle-MCMC (Modèle 4)

4.5 Autre alternative

L'estimateur par la méthode ECF converge presque sûrement et est normalement distribué, dans le cas i.i.d. comme dans le cas dépendant. De plus, la vitesse de convergence de cet estimateur est de l'ordre de \sqrt{n} . Ce résultat est remarquablement différent de celui de l'estimateur par la méthode de Whittle, où la vitesse est de l'ordre de $(\frac{n}{\ln n})^{1/\alpha}$. Cependant, pour implémenter la méthode ECF, nous nous opposons à deux grandes difficultés notées par Carrasco et Florens [16]. La première est que la fonction de poids optimale, conduisant à l'efficacité du maximum de vraisemblance, dépend de la fonction score conditionnelle qui est généralement inconnue. La deuxième est que la nécessité d'utiliser un grand ensemble de moments conditionnels aboutit à la singularité de la matrice de covariance. Pour palier à ce manque, Carrasco et Florens [16] propose une alternative qui consiste à utiliser la méthode des moments généralisés avec des moments conditionnels continus que nous notons par méthode C-GMM, pour simplifier. Dans cette section, nous allons décrire brièvement la méthode C-GMM. Vu le temps de calcul coûteux et des moyens de calculs limités dont nous disposons, nous allons présenter une petite simulation pour comparer la performance de cette méthode avec celle de la méthode ECF, sur un modèle simple.

4.5.1 Méthode C-GMM

Pour résoudre le problème lié au choix de la fonction de poids optimale dont souffrent les méthodes d'estimation utilisant la fonction caractéristique, Carrasco et Florens [16] ont suggéré d'utiliser la méthode des moments généralisés basée sur des moments conditionnels continus. Carrasco et al. [15] étendent les travaux de Carrasco et Florens [16], dans le cas où les observations sont dépendantes. Ainsi, dans le cas markovien, Carrasco et al. [15] emploient une méthode C-GMM basée sur la fonction caractéristique conditionnelle et montrent que l'estimateur obtenu atteint l'efficacité du maximum de vraisemblance. Cependant, nous adaptons dans cette partie, une méthode C-GMM basée sur la fonction caractéristique jointe, puisque pour un processus non markovien, la fonction caractéristique conditionnelle est inconnue et est très difficile à estimer. Toutefois, il est important de noter que la méthode C-GMM utilisant la fonction caractéristique jointe n'est pas optimale.

Nous conservons les mêmes notations que dans la Section 4.2. Notre objectif est d'estimer le paramètre $\psi = (\alpha, d, D, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \Phi_1, \dots, \Phi_P, \Theta_1, \dots, \Theta_Q)$ au moyen d'une réalisation (X_1, X_2, \dots, X_T) d'un processus ARFISMA symétrique α -stable dé-

fini par l'équation (2.1), par la méthode C-GMM basée sur la fonction caractéristique jointe de $Y_j = (X_j, \dots, X_{j+m})'$, pour $j = 1, \dots, T - m$. Pour cela, définissons d'abord les moments conditionnels basés sur la fonction caractéristique jointe de la manière suivante :

$$h(r, Y_j; \psi) = \exp(ir'Y_j) - c(r; \psi) := h_j(r; \psi),$$

et posons également

$$h_T(r; \psi) = \frac{1}{T} \sum_{j=1}^T \exp(ir'Y_j) - c(r; \psi) = \frac{1}{T} \sum_{j=1}^T h_j(r; \psi),$$

où $r = (r_1, \dots, r_{m+1})'$ et $c(r; \psi)$ est la fonction caractéristique jointe donnée par le Théorème 4.2.1. Ensuite, notons par K l'opérateur de covariance associé à $\{h_T(r; \psi)\}$, qui peut être estimé en utilisant un estimateur à noyau du type étudié dans Andrews [6], et qui est donné par :

$$\widehat{K}_T(r_1, r_2) = \frac{T}{T-L} \sum_{j=-T+1}^{T-1} \omega\left(\frac{j}{S_T}\right) \widehat{\Gamma}_T(j), \quad \text{avec } L = p + q + P + Q + 3. \quad (4.9)$$

Dans l'équation (4.9), $\omega(\cdot)$ est un noyau, S_T représente la largeur de bande utilisée et $\widehat{\Gamma}_T(j)$ est définie par :

$$\widehat{\Gamma}_T(j) = \begin{cases} \frac{1}{T} \sum_{t=j+1}^T g_t(r_1) \overline{g_{t-j}(r_2)}, & \text{pour } j \geq 0 \\ \frac{1}{T} \sum_{t=-j+1}^T g_{t+j}(r_1) \overline{g_t(r_2)}, & \text{pour } j < 0, \end{cases}$$

$$\text{où } g_t(r) = \exp(ir'Y_t) - \frac{1}{T} \sum_{j=1}^T \exp(ir'Y_j), \quad \text{avec } \bar{g} \text{ le conjugué de } g.$$

Puisque l'inverse de l'opérateur de covariance K_T est non bornée, Carrasco et al.[15] utilisent une version régularisée en introduisant un paramètre de pénalisation α_T et soit $K_T^{\alpha_T}$ la version régularisée de K_T . Ainsi, selon Carrasco et al. [15], l'estimateur C-GMM de ψ associé aux moments conditionnels $\{h_T(r; \psi)\}$ est défini par :

$$\widehat{\psi}_T = \arg \min_{\psi \in \Psi} \|(K_T^{\alpha_T})^{-1/2} h_T(\psi)\|^2, \quad (4.10)$$

où $\|\cdot\|$ est la norme définie sur $\mathbb{L}^2(\pi)$ qui est l'espace de Hilbert des fonctions à valeurs complexes et de carré intégrable :

$$\mathbb{L}^2(\pi) = \left\{ g : \mathbb{R}^d \rightarrow \mathbb{C}; \int |g(\tau)|^2 \pi(\tau) d\tau < \infty \right\},$$

π est une densité de probabilité qui est absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^d et admet tous ses moments, et $\pi(\tau) > 0$ pour tout $\tau \in \mathbb{R}^d$.

Le noyau $\omega(\cdot)$ doit vérifier les conditions de régularité suivantes qui sont basées sur les hypothèses de Andrews [6] :

$$\kappa = \{\omega(\cdot) : \mathbb{R} \rightarrow [-1, 1]; \omega(0) = 1, \omega(x) = \omega(-x), \forall x \in \mathbb{R}, \int_{-\infty}^{+\infty} \omega^2(x) dx < \infty,$$

$\omega(\cdot)$ est continue sauf peut être en un nombre fini de points}.

En effet, Andrews [6] propose une méthode efficace pour le calcul de $\widehat{\Gamma}_T(j)$, en utilisant différents types de noyaux (noyau tronqué, noyau de Bartlett, de Parzen, de Tukey-Hanning et noyau de quadrature spectrale) et il détermine pour chaque type de noyaux, la largeur de bande optimale S_T^* à utiliser. Il montre également que, le noyau de quadrature spectrale offre une meilleure performance, en terme de *MSE*. Il est aussi important de remarquer que, tous ces noyaux peuvent être calculés, sous le logiciel *R*, en utilisant la fonction **kweights** disponible dans le package **sandwich**.

Carrasco et al. [15] ont établi la consistance et la normalité asymptotique de l'estimateur C-GMM. Pour aboutir à ce résultat, le paramètre de pénalisation α_T est choisi tel que : $T^a \alpha_T \rightarrow \infty$ et $\alpha_T \rightarrow 0$, pour tout $a \geq 0$. Cependant, Carrasco et Florens [16], ainsi que Carrasco et al. [15] montrent, par des simulations, que le choix de α_T n'est pas très sensible à l'estimateur $\widehat{\psi}_T$. Toutefois, il est important de trouver un outil pour sélectionner la valeur de α_T , en pratique. Une idée est de choisir α_T suivant le critère défini précédemment par l'équation (4.5).

4.5.2 Calcul simple de l'estimateur C-GMM

Carrasco et al. [15] proposent une méthode simple pour calculer l'estimateur C-GMM, en exprimant la fonction objective en terme de matrices et de vecteurs. Ce résultat est présenté dans la Proposition 4.5.1 suivante. Remarquons d'abord que l'opérateur de covariance (4.9) peut être réécrit, sous la condition : $g_t(r) = 0$, si $t \leq 0$ ou $t \geq T$, comme suit :

$$\widehat{K}_T(r_1, r_2) = \frac{1}{T-L} \sum_{j=1}^T g_t(r_1) U_t(r_2),$$

où $U_t(r) = \omega(0) \overline{g_t(r)} + \sum_{j=1}^T \omega\left(\frac{j}{S_T}\right) \left(\overline{g_{t-j}(r)} + \overline{g_{t+j}(r)}\right)$.

Proposition 4.5.1 Résoudre l'équation (4.10) est équivalent à résoudre le problème suivant :

$$\min_{\psi \in \Psi} \underline{W}'(\psi) [\alpha_T I_T + C^2]^{-1} \underline{V}(\psi),$$

où C est une matrice de dimension $T \times T$ et d'éléments générateurs $c_{tl}/(T - L)$, $t, l = 1, \dots, T$, avec

$$c_{tl} = \int U_t(r) g_l(r) \pi(r) dr.$$

I_T est la matrice identité de dimension $T \times T$, $\underline{W} = [w_1, \dots, w_T]'$ et $\underline{V} = [v_1, \dots, v_T]'$ avec

$$v_t(\psi) = \int U_t(r) h_T(r; \psi) \pi(r) dr, \quad \text{et} \quad w_t(\psi) = \langle g_t(r), h_T(r; \psi) \rangle,$$

où $\langle \dots \rangle$ désigne le produit scalaire sur $\mathbb{L}^2(\pi)$. Ce produit scalaire est défini par :

$$\langle f, g \rangle = \int f(\tau) \overline{g(\tau)} \pi(\tau) d\tau,$$

où $\overline{g(\tau)}$ est le conjugué de $g(\tau)$. ■

La preuve de cette proposition est détaillée dans Carrasco et al. [15].

4.5.3 Quelques simulations

Dans cette section, nous effectuons une petite simulation qui consiste à comparer la performance de la méthode ECF avec celle de la méthode C-GMM pour un modèle ARFISMA simple. Vu le temps de calcul très long nécessaire pour implémenter la méthode C-GMM, nous proposons le plan de simulations suivant : nous générons 1000 échantillons de taille 500 d'un processus ARFISMA(0, d, 0)(0, D, 0) symétrique α -stable, avec $\alpha = 1.3$, $d = 0.10$ et $D = 0.12$. Le paramètre de pénalisation α_T est égal à 0.02 comme dans Carrasco et al. [15]. Les résultats sont présentés dans le Tableau 4.10 suivant :

Statistiques	Méthode ECF			Méthode C-GMM		
	$\hat{\alpha}$	\hat{d}	\hat{D}	$\hat{\alpha}$	\hat{d}	\hat{D}
Mean	1.2679	0.0768	0.0769	1.2636	0.0739	0.0759
RMSE	0.1104	0.0403	0.0526	0.1170	0.0465	0.0567
MAE	0.0843	0.0322	0.0454	0.0901	0.0367	0.0480

TAB. 4.10 – Comparaison entre les méthodes ECF et C-GMM

Malgré le manque de fonction de poids optimale noté, nous observons sur cet exemple que la méthode ECF donne la meilleure performance. Toutefois, force est de reconnaître que ce plan de simulations reste insuffisant pour comparer ces deux méthodes d'estimation. Il serait également intéressant de considérer ces deux méthodes dans des cas, où il y a présence de paramètres courte mémoire dans le modèle.

4.6 Conclusions et perspectives

Dans ce chapitre, nous avons utilisé la méthode ECF pour estimer simultanément l'ensemble des paramètres d'un processus ARFISMA symétrique α -stable. Sous certaines hypothèses, nous avons montré la consistance et la normalité asymptotique de cet estimateur. Nous avons effectué des simulations par la méthode de Monte Carlo pour examiner la performance de la méthode ECF. Les résultats de simulations sont satisfaisants pour les quatre modèles considérés. Ces simulations nous ont également permis de déterminer la valeur optimale de la taille des blocs pour chacun de ces modèles, suivant le critère défini par l'équation (4.5). Ainsi, nous avons retenu que pour le modèle simple sans partie courte mémoire, la valeur $m = 1$ suffit pour estimer les paramètres, alors que s'il y a présence de paramètres courte mémoire dans le modèle, la valeur optimale est $m = 6$. Par comparaison, les résultats de simulations ont montré que la méthode de Whittle par MCMC donne une meilleure performance que la méthode ECF. Cependant, la méthode ECF a l'avantage d'estimer le paramètre de la distribution des innovations, ce qui n'est pas possible avec la méthode de Whittle. Nous avons également considéré dans ce chapitre une autre alternative : la méthode des moments généralisés basée sur des moments conditionnels continus, suggérée par Carrasco et al. [15]. Nous avons

comparé ces deux approches sur un modèle simple sans partie courte mémoire et les résultats de simulations ont montré que la méthode ECF donne la meilleure performance.

Enfin, notons que dans le cadre de futures recherches, il serait intéressant d'améliorer la performance de la méthode ECF, en trouvant une approximation de la fonction de poids optimale donnée par l'équation (4.4). Il serait également intéressant de comparer la performance de la méthode ECF avec celle de la méthode C-GMM, sur des modèles avec partie courte mémoire.

Illustration sur des données réelles

Résumé

Dans ce chapitre, nous présentons une application du modèle ARFISMA α -stable sur des débits mensuels du fleuve Sénégal à Bakel. En particulier, nous comparons ce modèle avec le modèle classique SARIMA de Box et Jenkins [11], dans une optique prévisionnelle.

Sommaire

5.1	Introduction	94
5.2	Méthodologie	94
5.2.1	Tests de validation	94
5.2.2	Méthode de prévision	97
5.3	Application sur des données hydrologiques	99
5.3.1	Première approche	104
5.3.2	Deuxième approche	106
5.3.3	Comparaison des résultats	110
5.4	Conclusions et perspectives	112

5.1 Introduction

A la suite de l'introduction des modèles ARFIMA, de nombreuses applications des processus longue mémoire ont été développées dans la littérature statistique. Cependant, ces modèles sont très limités face à des données présentant une saisonnalité et une grande variabilité, comme en finance, en télécommunication ou en hydrologie. Dans ce chapitre, nous nous intéressons alors à la modélisation du débit du fleuve Sénégal à la station de Bakel, par des processus ARFISMA α -stables.

Ce chapitre est organisé comme suit. Dans la deuxième section, nous révisons les étapes de la modélisation et nous fournissons une méthode de prévision pour un processus ARFISMA α -stable. Dans la troisième section, nous mettons en oeuvre ces étapes dans l'application sur des données hydrologiques. La dernière section est dédiée aux conclusions et perspectives.

5.2 Méthodologie

Dans cette section, nous examinons quelques étapes de la modélisation. En particulier, nous nous focalisons sur les tests de validation et la méthode de prévision pour un processus ARFISMA α -stable.

5.2.1 Tests de validation

1- Test de significativité des paramètres du modèle

Après avoir estimé les paramètres, il est important de tester leur significativité. Dans le cas gaussien, quelle que soit la méthode d'estimation employée, il est possible de calculer la matrice de variance covariance des estimateurs des paramètres du modèle. Partant de cette matrice de variance covariance, il est possible de reconstruire les statistiques de Student associées aux différents paramètres du modèle. Sur le plan appliqué, lorsque le processus est stationnaire, on peut montrer que ces statistiques de Student sont asymptotiquement normales. Donc, on peut appliquer à ces estimateurs les méthodes d'inférence traditionnelles. Cependant, dans le cas stable, la matrice de variance covariance n'existe pas et ce test de Student n'est pas applicable. Néanmoins, on peut construire un test de significativité des paramètres en utilisant le résultat de convergence de l'estimateur de Whittle établi dans le Théorème 1.4.10, que nous étendons ici. En notant par β_n

l'estimateur de Whittle du paramètre β_0 , nous avons :

$$\left(\frac{n}{\ln n}\right)^{1/\alpha} (\beta_n - \beta_0) \xrightarrow{\mathcal{D}} 4\pi W^{-1}(\beta_0) \sum_{k=1}^{\infty} \frac{Y_k}{Y_0} v_k, \quad (5.1)$$

où W^{-1} est l'inverse de la matrice W et v_k est un vecteur qui sont tous deux définis de la même manière que dans le Théorème 1.4.10. Les variables aléatoires Y_0, Y_1, Y_2, \dots sont indépendantes telles que $Y_0 \stackrel{\mathcal{D}}{=} S_{\alpha/2}(C_{\alpha/2}^{-2/\alpha}, 1, 0)$, $Y_k \stackrel{\mathcal{D}}{=} S_{\alpha}S(C_{\alpha}^{-1/\alpha}), \forall k \geq 1$ et la constante C_{α} est donnée par l'équation (1.9).

Nous allons maintenant décrire brièvement la procédure de test. On veut tester l'hypothèse nulle $\mathcal{H}_0 : \beta_n = 0$ contre l'hypothèse alternative $\mathcal{H}_1 : \beta_n \neq 0$, au seuil γ . Pour cela, calculons la p-value :

$$\text{p-value} = \mathbb{P}_{\mathcal{H}_0} (|\beta_n| > \beta_{obs}), \text{ où } \beta_{obs} \text{ est la valeur observée.}$$

$$\Rightarrow \text{p-value} = \mathbb{P}_{\mathcal{H}_0} \left(\left(\frac{n}{\ln n}\right)^{1/\alpha} |\beta_n| > \left(\frac{n}{\ln n}\right)^{1/\alpha} \beta_{obs} \right).$$

En utilisant la propriété asymptotique (5.1), nous pouvons approcher la p-value par :

$$\text{p-value} \approx \mathbb{P}_{\mathcal{H}_0} \left(4\pi W^{-1}(0) \sum_{k=1}^{\infty} \frac{Y_k}{Y_0} v_k > \left(\frac{n}{\ln n}\right)^{1/\alpha} \beta_{obs} \right).$$

En posant $b_k = 4\pi W^{-1}(0)v_k$, on observe alors que la $j^{\text{ième}}$ coordonnée du vecteur aléatoire $\sum_{k=1}^{\infty} b_k \frac{Y_k}{Y_0}$ a la même distribution que $\left(\sum_{k=1}^{\infty} |(b_k)_j|^{\alpha}\right)^{1/\alpha} \frac{Y_1}{Y_0}$ (voir Kokoszka et Taqqu [46]). Ainsi, on peut écrire pour tout $j = 1, \dots, K$ (K est la longueur du vecteur des paramètres à estimer) :

$$\text{p-value} \approx \mathbb{P}_{\mathcal{H}_0} \left(\left(\sum_{k=1}^{\infty} |(b_k)_j|^{\alpha}\right)^{1/\alpha} \frac{Y_1}{Y_0} > \left(\frac{n}{\ln n}\right)^{1/\alpha} \beta_{obs} \right) \quad (5.2)$$

Il s'agit alors essentiellement du test pour la significativité individuelle des paramètres. La relation (5.2) entraîne

$$\begin{aligned} \text{p-value} &\approx \mathbb{P}_{\mathcal{H}_0} \left(\frac{Y_1}{Y_0} > \left(\frac{n}{\ln n}\right)^{1/\alpha} \left(\sum_{k=1}^{\infty} |(b_k)_j|^{\alpha}\right)^{-1/\alpha} \beta_{obs} \right) \\ &\approx 1 - F_{\frac{Y_1}{Y_0}} \left(\left(\frac{n}{\ln n}\right)^{1/\alpha} \left(\sum_{k=1}^{\infty} |(b_k)_j|^{\alpha}\right)^{-1/\alpha} \beta_{obs} \right). \end{aligned}$$

Or d'après Brockwell et Davis [13] (page 539), la fonction de répartition $F_{\frac{Y_1}{Y_0}}$ de la variable aléatoire $\frac{Y_1}{Y_0}$ peut s'écrire de la manière suivante :

$$F_{\frac{Y_1}{Y_0}}(x) = \mathbb{P}\left(\frac{Y_1}{Y_0} \leq x\right) = \int_0^{+\infty} 2^{-1/2}(\pi\omega)^{-3/2} \left[\arctg(x\omega) + \frac{\pi}{2}\right] \exp\left(-\frac{1}{2\omega}\right) d\omega. \quad (5.3)$$

Ainsi nous pouvons calculer la p-value du test et prendre une décision sur la significativité des paramètres. Si la p-value est inférieure à $\gamma\%$, on rejette l'hypothèse nulle \mathcal{H}_0 et si la p-value est supérieure à $\gamma\%$, on accepte \mathcal{H}_0 .

Remarque 5.2.1 La formule (5.3) donne également la possibilité de calculer les quantiles de la loi de $\frac{Y_1}{Y_0}$. Par exemple, on retrouve dans Alder et al. [2] (page 144) le quantile 97,5% pour différentes valeurs de α .

2- Analyse des résidus

Lorsque le processus est bien estimé, les résidus entre les valeurs observées et les valeurs estimées par le modèle doivent vérifier les hypothèses faites sur le processus d'innovations, c'est à dire un bruit blanc symétrique α -stable.

a) Test d'autocorrélation des résidus : Si les résidus obéissent à un bruit blanc, il ne doit pas exister d'autocorrélation dans la série. On étudie souvent la fonction d'autocorrélation (FAC) et la fonction d'autocorrélation partielle (FAP) pour vérifier que dans le cas d'un bruit blanc, il n'existe aucune autocorrélation ou d'autocorrélation partielle significativement non nulle. Cependant, dans le cas stable, ces outils n'ont pas d'existence théorique. Néanmoins, il est montré dans Embrechts et al. [23], qu'on peut calculer leurs quantités empiriques et les utiliser. Cette étude est prolongée par les tests de "portemanteau". En pratique, on utilise deux tests : le test de Box et Pierce et le test de Ljung-Box qui ne sont pas définis dans ce document.

b) Test sur la distribution des résidus : Plusieurs tests peuvent être utilisés pour vérifier la normalité des résidus, mais le plus courant est celui de Jarque et Bera. Ce dernier est fondé sur la notion de skewness (moment d'ordre 3 et asymétrie) et de kurtosis (moment d'ordre 4 et queue de distribution). Ensuite, nous pouvons accompagner ce test par l'estimation des paramètres de la distribution des résidus par la méthode de McCulloch définie dans le chapitre 1 ou effectuer un test sur la convergence de la variance empirique des résidus afin de montrer que le processus résiduel suit une loi stable.

5.2.2 Méthode de prévision

Une fois que l'on a spécifié, estimé et validé le modèle, on désire, en général, l'utiliser pour effectuer des prévisions sur la série. Le but de cette section est d'étudier la prévision d'un processus ARFISMA symétrique α -stable, en minimisant l'erreur de dispersion, puisque l'erreur quadratique moyenne n'existe pas dans le cas stable. Nous allons décrire brièvement cette méthode de prévision et nous renvoyons à Kokoszka [44], Cline et Brockwell [19] ou Brockwell et Davis [13], pour plus de détails. Notons d'abord, par \widehat{X}_{n+k} le prédicteur linéaire à l'horizon k de X_{t+k} basé sur le passé fini X_n, X_{n-1}, \dots, X_1 , pour tout $t \in \mathbb{Z}$ et $k > 0$. Ainsi \widehat{X}_{n+k} peut s'écrire sous la forme suivante :

$$\widehat{X}_{n+k} = a_0 X_n + a_1 X_{n-1} + \dots + a_{n-1} X_1. \quad (5.4)$$

Pour évaluer la performance du prédicteur linéaire (5.4), nous ne pouvons pas considérer l'erreur quadratique moyenne de prévision $\mathbb{E}(\widehat{X}_{n+k} - X_{n+k})^2$, puisque dans le cas stable cette quantité est infinie. Par conséquent, un critère naturel du choix d'un meilleur prédicteur linéaire pour une représentation moyenne mobile infinie est de minimiser la dispersion $\text{disp}(X_{n+k} - \widehat{X}_{n+k})$ (voir Cline et Brockwell [19] ou Kokoszka [44]). Par définition, la dispersion basée sur la représentation moyenne mobile infinie (2.17) est donnée par :

$$\text{disp}(X_n) = C \sum_{j=0}^{\infty} |c_j|^\alpha, \quad (5.5)$$

où les coefficients $(c_j)_{j \geq 0}$ sont déterminés par l'équation (2.9). La constante C représente la dispersion du processus d'innovation (Z_n) et qui est définie par :

$$C = \text{disp}(Z_n) = \gamma C_\alpha,$$

où C_α est donnée par l'équation (1.9) et γ est le paramètre d'échelle de Z_n . En utilisant l'équation (2.17), nous pouvons récrire \widehat{X}_{n+k} et X_{n+k} comme suit :

$$\begin{aligned} \widehat{X}_{n+k} &= \sum_{j=0}^{+\infty} (a_0 c_j + a_1 c_{j-1} + \dots + a_{n-j} c_{j-n+1}) Z_{n-j}, \\ X_{n+k} &= \sum_{j=0}^{+\infty} c_j Z_{n+k-j} = \sum_{j=0}^{k-1} c_j Z_{n+k-j} + \sum_{j=0}^{+\infty} c_{j+k} Z_{n-j}. \end{aligned}$$

En se servant de l'équation (5.5) et de l'indépendance des variables aléatoires Z_t , nous obtenons :

$$\text{disp}(X_{n+k} - \widehat{X}_{n+k}) = C \left[\sum_{j=0}^{k-1} |c_j|^\alpha + \sum_{j=k}^{+\infty} |c_j - (a_0 c_{j-k} + \dots + a_{n-1} c_{j-k-n+1})|^\alpha \right]. \quad (5.6)$$

Ainsi, nous définissons l'erreur minimale de dispersion du prédicteur linéaire de X_{n+k} (basé sur le passé X_n, X_{n-1}, \dots, X_1) comme étant la combinaison linéaire qui minimise l'équation (5.6). Le Théorème 5.2.2 suivant permet de déterminer la dispersion minimale du prédicteur linéaire de X_{n+k} basé sur le passé infini X_n, X_{n-1}, \dots

Théorème 5.2.2 Il existe une unique suite $\{a_0, a_1, \dots\}$ telle que

$$\text{disp} \left(X_{n+k} - \sum_{j=0}^{\infty} a_j X_{n-j} \right) = \min_{u_0, u_1, \dots} \text{disp} \left(X_{n+k} - \sum_{j=0}^{\infty} u_j X_{n-j} \right),$$

où le minimum est pris pour toute suite $\{u_0, u_1, \dots\}$ vérifiant $\sum |u_j|^\alpha < \infty$, c'est à dire sous les conditions $|d + D| < 1 - \frac{1}{\alpha}$ et $|D| < 1 - \frac{1}{\alpha}$ avec $1 < \alpha \leq 2$. La suite $\{a_0, a_1, \dots\}$ est donnée par :

$$a_j = - \sum_{t=0}^{k-1} c_t \tilde{c}_{j+k-t}, \quad (5.7)$$

où les coefficients $(c_j)_{j \geq 0}$ et $(\tilde{c}_j)_{j \geq 0}$ sont déterminés respectivement par les équations (2.9) et (2.19). De plus, on a

$$\text{disp} \left(X_{n+k} - \sum_{j=0}^{\infty} a_j X_{n-j} \right) = C \sum_{j=0}^{k-1} |c_j|^\alpha$$

■

La preuve de ce théorème est détaillée dans Kokoszka [44].

Cependant, en pratique le nombre d'observations disponible est fini et on ne dispose pas également de la valeur des u_n, \dots, u_0 qui minimise $\text{disp} \left(X_{n+k} - \sum_{j=0}^n u_j X_{n-j} \right)$. Ainsi, on utilise en général le prédicteur tronqué :

$$\hat{X}_{n+k} = \sum_{j=0}^{n-1} a_j X_{n-j},$$

où les coefficients $(a_j)_{j \geq 0}$ sont donnée par l'équation (5.7).

5.3 Application sur des données hydrologiques

Dans ce paragraphe, nous nous intéressons à la série du débit du fleuve Sénégal à la station de Bakel. La station de Bakel, située à 14°54' Latitude Nord et 12°27' Longitude Ouest, ouverte depuis 1904, est la station de référence sur le fleuve Sénégal, car étant la première sur le territoire du Sénégal. La Figure 5.1 illustre bien la position de la station de Bakel sur le fleuve Sénégal.

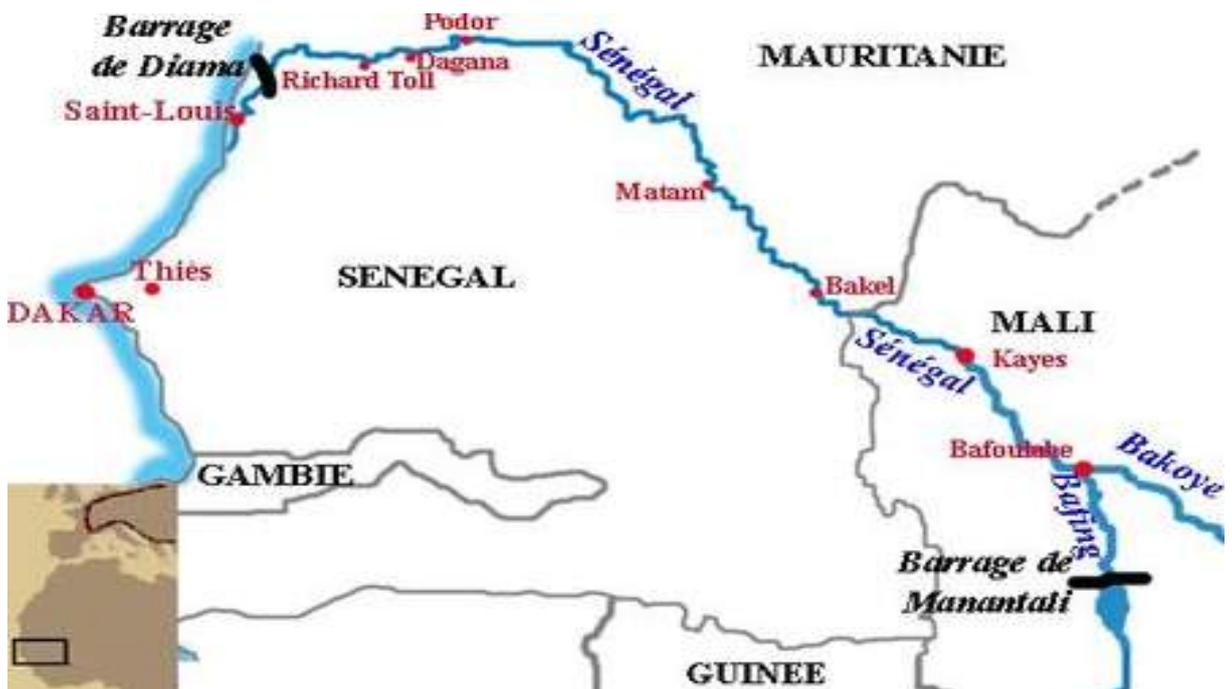


FIG. 5.1 – Position de la station de Bakel sur le fleuve Sénégal.

Les chroniques obtenues sont des données de débits d'eau journaliers mesurés en m^3/s et relevés à la station de Bakel. Cette série a été reconstruite, car les données des hivers avant 1950 n'étaient pas existantes. Les trous multiples ont été comblés à l'aide des données mensuelles du Global Historical Climate Network (disponibles en accès libre) ou par moyenne des deux années antérieures et postérieures (en respectant le cycle hydrologique). Ainsi, nous avons obtenu une série mensuelle qui couvre la période allant de Janvier 1904 à Septembre 2010, représentée sur la Figure 5.2.

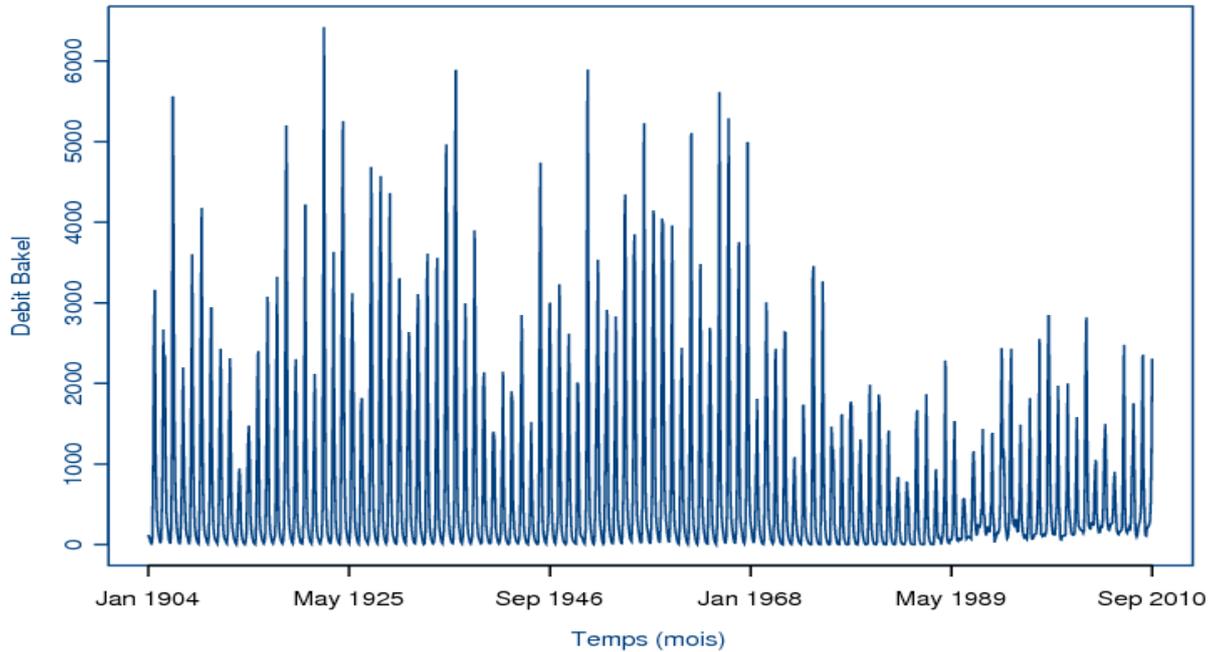


FIG. 5.2 – Evolution de la série du débit du fleuve à Bakel, de Janvier 1904 à Septembre 2010.

Nous allons modéliser cette série en utilisant comme période d'étude les données à partir de Janvier 1904 jusqu'au mois de Septembre 2004. Notons alors par $(X_t)_{t=1,\dots,T}$ cette série, où T est la taille de l'échantillon; soit $T = 1209$ observations. Ensuite, nous effectuerons des prévisions sur la période Octobre 2004 à Septembre 2010, que nous comparerons avec les données réelles observées afin de pouvoir juger de la qualité de ces prévisions. La Figure 5.3 représente l'évolution du débit mensuel moyen sur la période allant de Janvier 1904 à Septembre 2004. Nous constatons une alternative de décroissance et de croissance du débit moyen à Bakel. Nous avons une première décroissance (très faible) du débit jusqu'au mois de Mai, puis une forte croissance de Juin à Septembre et ensuite une décroissance d'Octobre à Décembre. Ces observations sur le débit moyen peuvent s'expliquer par l'effet de la saison des pluies. En effet, la partie croissante du graphe correspond à la période des pluies au Sénégal.

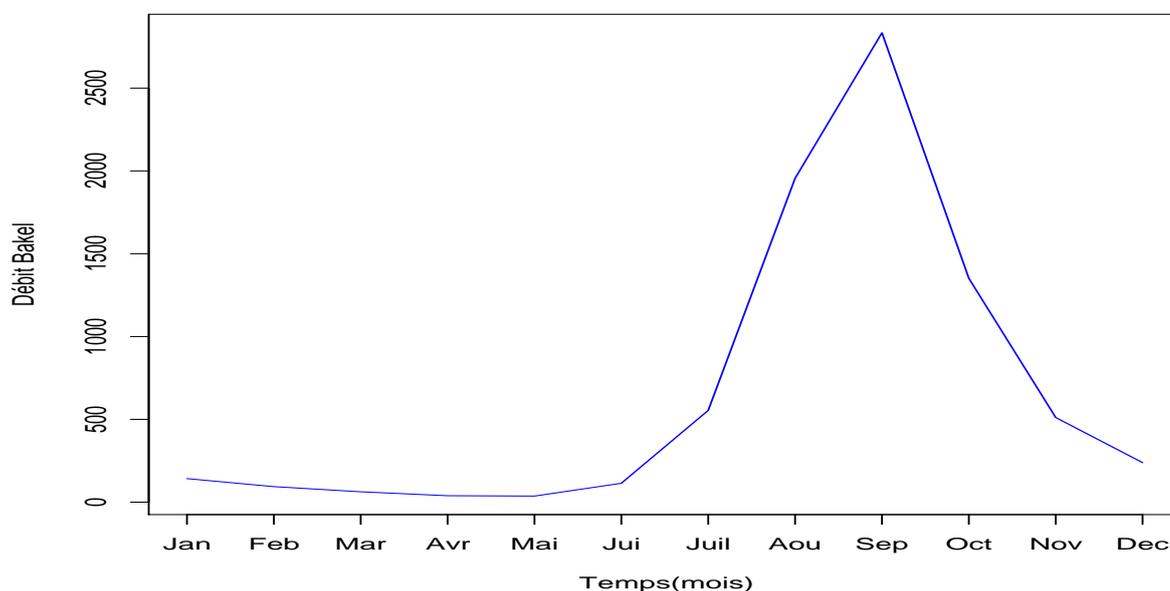


FIG. 5.3 – Débits mensuels moyens du fleuve à Bakel, de Janvier 1904 à Septembre 2004.

Les quatre premiers moments ainsi que la statistique de Jarque et Bera sont résumés dans le Tableau 5.1 suivant :

Moyenne	Variance	Skewness	Kurtosis	Jarque et Bera
662.737	1051507	2.3680	5.9790	2942.5073
				p-value < $2.2 \cdot 10^{-16}$

TAB. 5.1 – Statistiques descriptives de la série du débit du fleuve à Bakel, de Janvier 1904 à Septembre 2004.

L'analyse de ce tableau permet de voir que la distribution des données n'est pas normale. En effet, la statistique de Jarque et Bera est largement supérieure à la valeur critique du Chi-deux (5.99) au niveau de significativité 5%. De plus les valeurs de Skewness et de Kurtosis laissent voir que la série est dissymétrique et présente une queue épaisse. La Figure 5.4 comparant la distribution de la série d'étude avec celle de la loi normale confirme ces résultats.

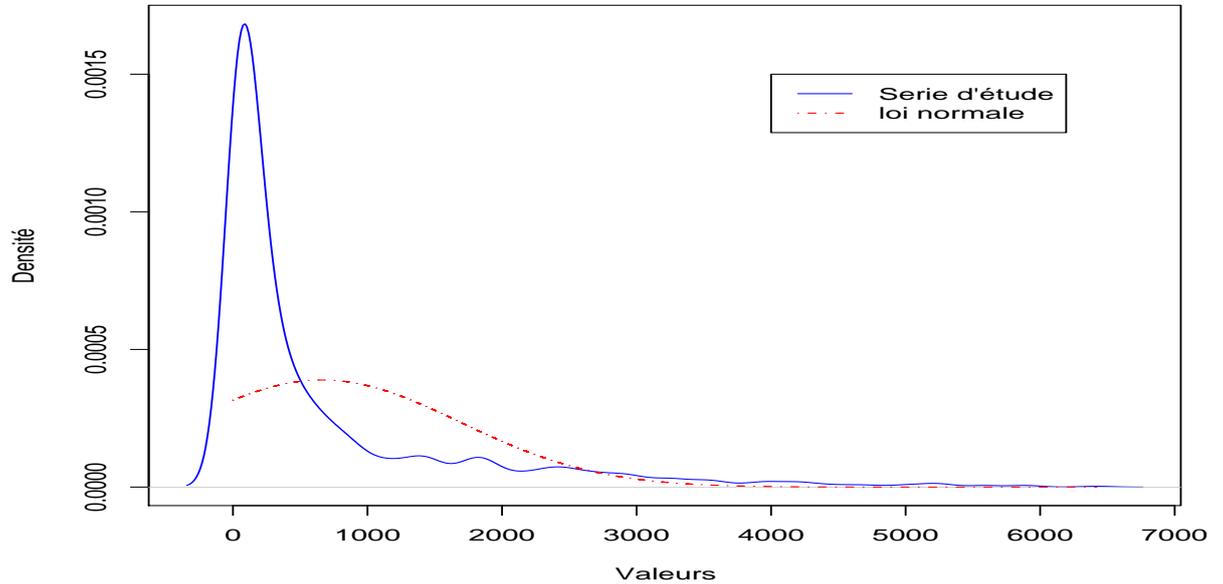


FIG. 5.4 – Densité de la série d'étude, comparée avec celle de la loi normale.

Nous allons examiner d'abord, si nous sommes en présence de données à variance infinie. A cette fin, le test graphique de la convergence de la variance empirique sur la série $(X_t)_{t=1,\dots,T}$, représenté sur la Figure 5.5, indique clairement que les données présentent une variance infinie.

Nous étudions ensuite les propriétés statistiques de la série d'étude, en utilisant l'ACF et le périodogramme normalisé visible sur la Figure 5.6. L'ACF décroît lentement lorsque les retards augmentent, indiquant ainsi un phénomène de mémoire longue sur la série étudiée. Le périodogramme normalisé présente des pics de saisonnalité 12 mois, associés à des fréquences de type $\frac{\pi k}{6}$, pour $k = 0, 1, \dots, 6$. Cette forte saisonnalité de 12 mois est également observable sur le graphique de l'ACF empirique qui montre des pics sur les retards multiples de 12.

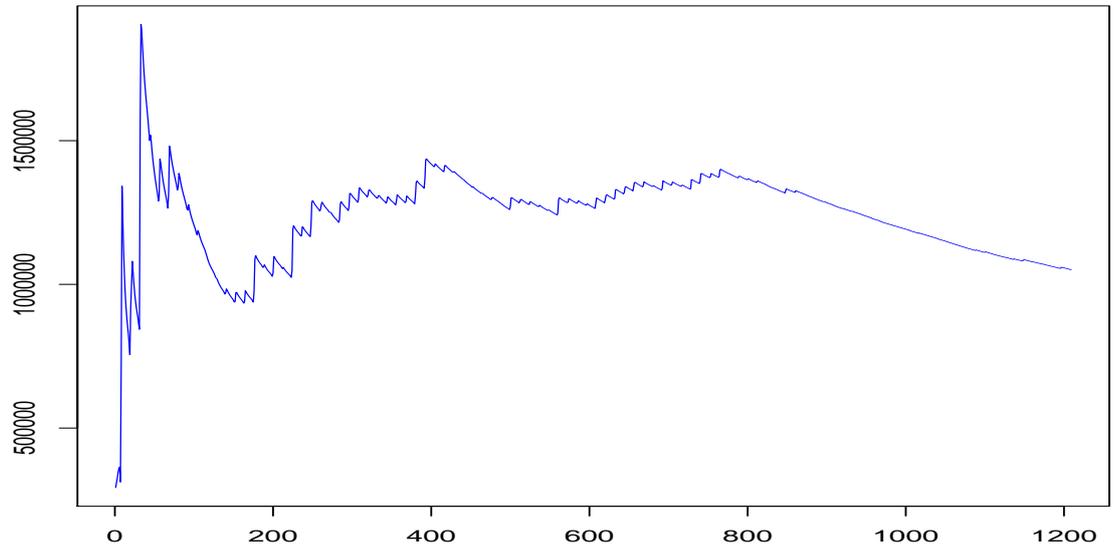


FIG. 5.5 – Test graphique de la variance empirique sur la série d'étude.

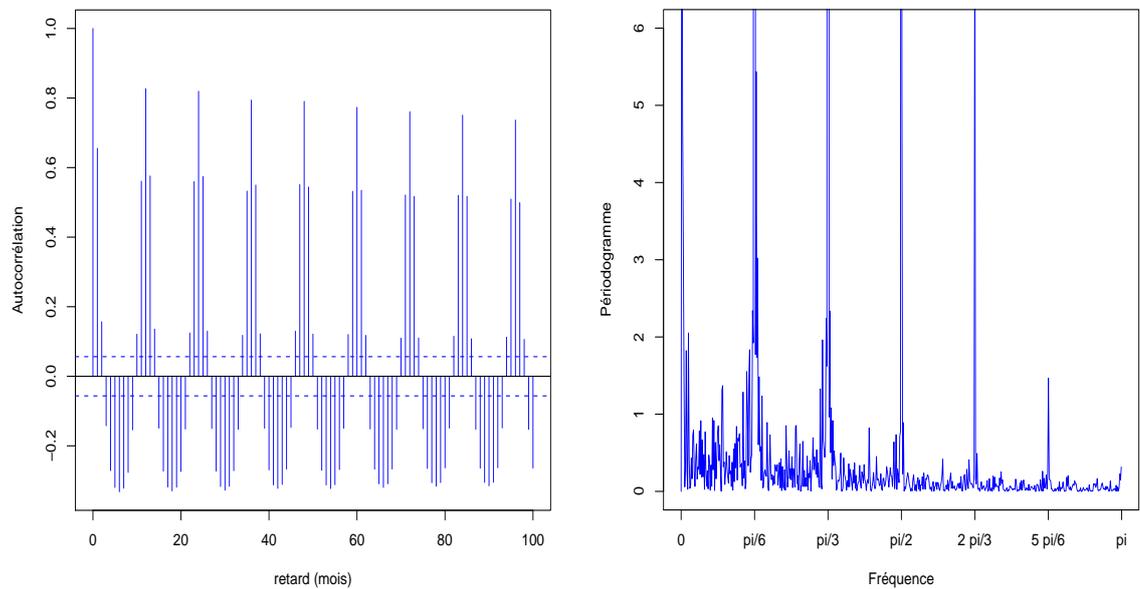


FIG. 5.6 – ACF empirique et périodogramme normalisé de la série d'étude.

L'analyse des données révèle que cette série possède une tendance à long terme ainsi qu'une forte saisonnalité de 12 mois et un phénomène de variance infinie. Toutes ces remarques, nous conduisent à penser qu'un modèle ARFISMA α -stable pourrait bien décrire la série $(X_t)_{t=1,\dots,T}$.

En guise de comparaison, nous allons mettre en compétition deux différentes approches. La première approche est celle basée sur le classique modèle linéaire SARIMA proposé par Box et Jenkins [11]. Dans la seconde approche, nous considérons le modèle ARFISMA α -stable développé par Diongue et al.[21] et détaillé dans le Chapitre 2. Le prédictor utilisé, dans la première approche, est le prédictor des moindres carrés, alors que dans la deuxième approche, nous considérons le prédictor linéaire du minimum de dispersion présenté dans la Section 5.2.2. Nous comparons les capacités prédictives des modèles à l'aide du critère de la racine carrée de l'erreur quadratique moyenne, noté RMSE ("Root Mean Squared Error") et du critère de la moyenne du pourcentage d'erreur absolu, noté MAPE ("Mean Absolute Percentage Error"). Ces deux critères sont respectivement définis comme suit :

$$\text{RMSE}_k = \sqrt{\frac{1}{M} \sum_{j=1}^M \left(X_{t+j+k} - \hat{X}_{t+j+k} \right)^2}, \quad k = 1, \dots, 12$$

$$\text{MAPE}_k = \frac{1}{M} \sum_{j=1}^M \frac{|X_{t+j+k} - \hat{X}_{t+j+k}|}{X_{t+j+k}}, \quad k = 1, \dots, 12,$$

où M est le nombre d'années de l'échantillon de validation des modèles et $\hat{X}_t(k)$ est le prédictor du mois k de X_{t+k} .

Au préalable, nous retranchons à cette série sa moyenne empirique et nous étudions par la suite la série centrée $(X_t - \bar{X})_{t=1,\dots,T}$, où \bar{X} est la moyenne empirique égale à 662.737.

5.3.1 Première approche

Dans ce paragraphe, nous considérons l'approche de Box et Jenkins [11] et nous modélisons la série $(X_t)_{t=1,\dots,T}$ par un processus SARIMA(p, d, q)(P, D, Q)₁₂. Vue l'intense littérature relative à cette théorie, nous résumons ici l'essentiel des étapes de cette modélisation. Ainsi, pour stationnariser la série, nous appliquons sur la série centrée, successivement un filtre de la forme $(1 - B)$ pour enlever la tendance, puis un filtre de la forme

$(1 - B^{12})$ pour faire disparaître la saisonnalité. Ce qui revient donc à choisir les entiers d et D égaux à 1. En ce qui concerne le choix des ordres p , q , P et Q du modèle SARMA obtenu après application des filtres, nous avons testé plusieurs modèles ($p = 1 : 3$, $q = 1 : 3$, $P = 1 : 4$ et $Q = 1 : 4$). L'estimation et la sélection de modèles s'effectuent grâce à la fonction `arima0()` implémentée dans le logiciel R. A l'aide du critère AIC ("Akaike Information Criterion"), le modèle retenu est un modèle SARIMA(2, 1, 1)(3, 1, 4)₁₂ formalisé comme suit :

$$\phi(B)\Phi(B^{12})(1 - B)(1 - B^{12})(X_t - \bar{X}_t) = \theta(B)\Theta(B^{12})\varepsilon_t, \quad (5.8)$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2, \quad \theta(B) = 1 + \theta_1 B,$$

$$\Phi(B^{12}) = 1 - \Phi_1 B^{12} - \Phi_2 B^{24} - \Phi_3 B^{36},$$

$$\Theta(B^{12}) = 1 + \Theta_1 B^{12} + \Theta_2 B^{24} + \Theta_3 B^{36} + \Theta_4 B^{48}.$$

Nous rappelons que l'estimation des paramètres du modèle (5.8) est obtenue par la méthode du maximum de vraisemblance implémentée dans la fonction `arima0()`. Les résultats des paramètres estimés du modèle retenu sont résumés dans le Tableau 5.2.

Paramètres	Paramètres estimés	Ecart-type
ϕ_1	0.5986	0.0297
ϕ_2	-0.1451	0.0260
θ_1	-0.9867	0.1277
Φ_1	-0.2624	0.0132
Φ_2	0.212	0.0756
Φ_3	0.6212	0.04497
Θ_1	-0.6373	0.0364
Θ_2	-0.5246	0.0185
Θ_3	-0.3730	0.0324
Θ_4	0.4966	0.0256

TAB. 5.2 – Paramètre estimés pour le modèle de Box et Jenkins, défini par 5.8 .

Nous effectuons un test de Student pour la significativité des paramètres : si la valeur absolue du paramètre estimé est plus grande que $1.96 \times$ l'écart-type du paramètre, alors on rejette, au rsique de première espèce $\alpha = 5\%$, l'hypothèse de nullité du paramètre.

Ainsi, nous remarquons que tous les paramètres estimés sont significativement non nuls, avec un risque de 5%.

Le modèle étant bien spécifié, estimé et validé, nous l'utilisons pour effectuer des prévisions sur la série. Ainsi, à l'aide de la fonction `predict()` du logiciel R, qui prend comme arguments la sortie de la fonction `arima0()` et le pas de prévision, nous avons effectué des prévisions avec le modèle (5.8). Les résultats relatifs à la qualité de prévision de ce modèle sont commentés à la fin de la Section 5.3 et sont contenus dans le tableau 5.3 suivant :

	Oct.	Nov.	Dec.	Jan.	Fév.	Mars.	Avr.	Mai.	Jui.	Juil.	Aout.	Sept.
RMSE	331.97	126.07	66.16	42.58	46.10	51.56	84.55	90.19	108.85	153.68	368.38	656.26
MAPE	0.438	0.377	0.389	0.271	0.272	0.224	0.313	0.337	0.335	0.169	0.287	0.384

TAB. 5.3 – Résultats en prévision sur la série mensuelle du débit du fleuve Sénégal à Bakel, par le modèle de Box et Jenkins.

5.3.2 Deuxième approche

Comme nous l'avons déjà noté, l'observation de l'ACF empirique, du périodogramme normalisé (Figure 5.6) et du test graphique de la variance empirique (Figure 5.5) suggère de modéliser la série $(X_t)_{t=1,\dots,T}$ par un processus ARFISMA α -stable.

Pour enlever la tendance et la saisonnalité, nous ajustons d'abord sur la série centrée le filtre suivant :

$$Z_t = (1 - B)^d (1 - B^{12})^D (X_t - \bar{X}_t). \quad (5.9)$$

L'estimation des paramètres de mémoire longue du modèle (5.9) par la méthode de Whittle par MCMC donne : $\hat{d} = 0.09220$ et $\hat{D} = 0.4007$. Ensuite, nous calculons les résidus issus de la modélisation (5.9), notés par $(\xi_t)_{t=1,\dots,T}$, en utilisant la représentation autorégressive infinie donnée par l'équation (2.18). La Figure 5.7 décrit l'évolution en fonction du temps du processus résiduel $(\xi_t)_{t=1,\dots,T}$ et le Tableau 5.4 résume les statistiques descriptives de ce processus. L'analyse de ce tableau permet de rejeter l'hypothèse de gaussianité du processus.

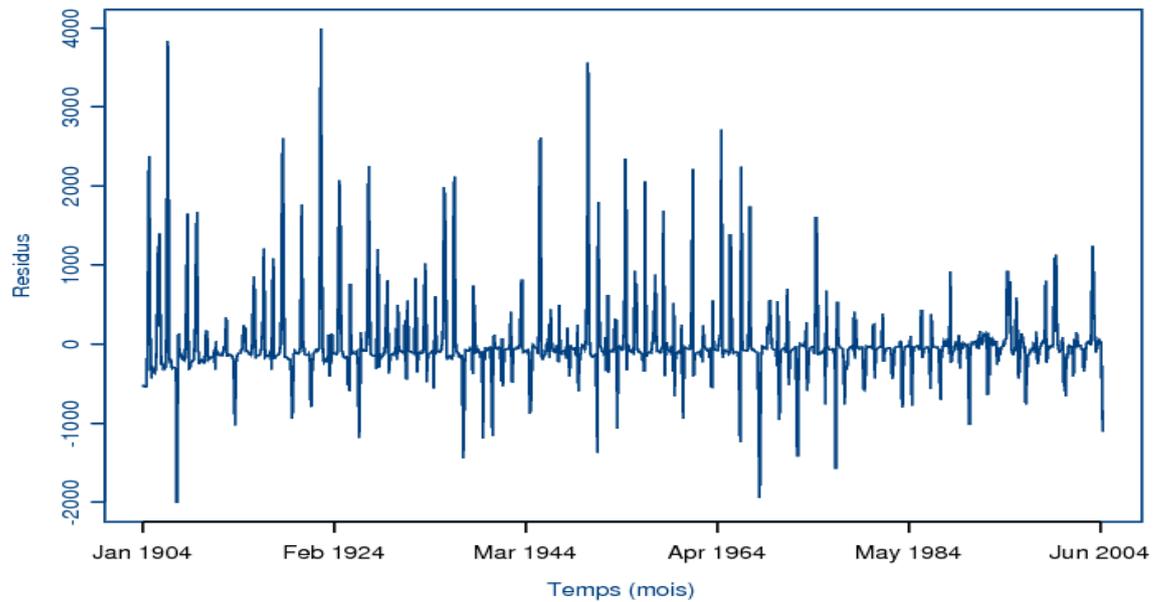


FIG. 5.7 – Evolution du processus résiduel $(\xi_t)_{t=1,\dots,T}$ issu de la modélisation de (5.9)

Moyenne	Variance	Skewness	Kurtosis	Jarque et Bera
-6.1580	238889.3	2.8359	15.9072	14425.8749 p-value < $2.2 \cdot 10^{-16}$

TAB. 5.4 – Statistiques descriptives du processus résiduel $(\xi_t)_{t=1,\dots,T}$.

L'étude de l'ACF et de la PACF empirique représentées sur la Figure 5.8 du processus résiduel, montre clairement qu'un processus à mémoire courte saisonnière est adéquat à la modélisation du processus $(\xi_t)_{t=1,\dots,T}$. Ainsi, nous nous proposons de caler sur ce processus un modèle SARMA(p, q)(P, Q) α -stable. Nous avons testé plusieurs modèles et par le biais du critère d'information d'Akaike corrigé (AICC) ou du critère bayésien d'information (BIC), le modèle optimal sélectionné est un SARMA(1, 1)(1, 2) formalisé de la manière suivante :

$$\phi^*(B)\xi_t = \theta^*(B)\varepsilon_t, \quad (5.10)$$

avec

$$\begin{aligned} \phi^*(B) &= (1 - 0.1971B)(1 - 0.9758B^{12}) \\ \theta^*(B) &= (1 + 0.3387B)(1 - 1.1262B^{12} + 0.1930B^{24}). \end{aligned}$$

L'estimation des paramètres du modèle (5.10) est effectuée par la méthode de Whittle par MCMC.

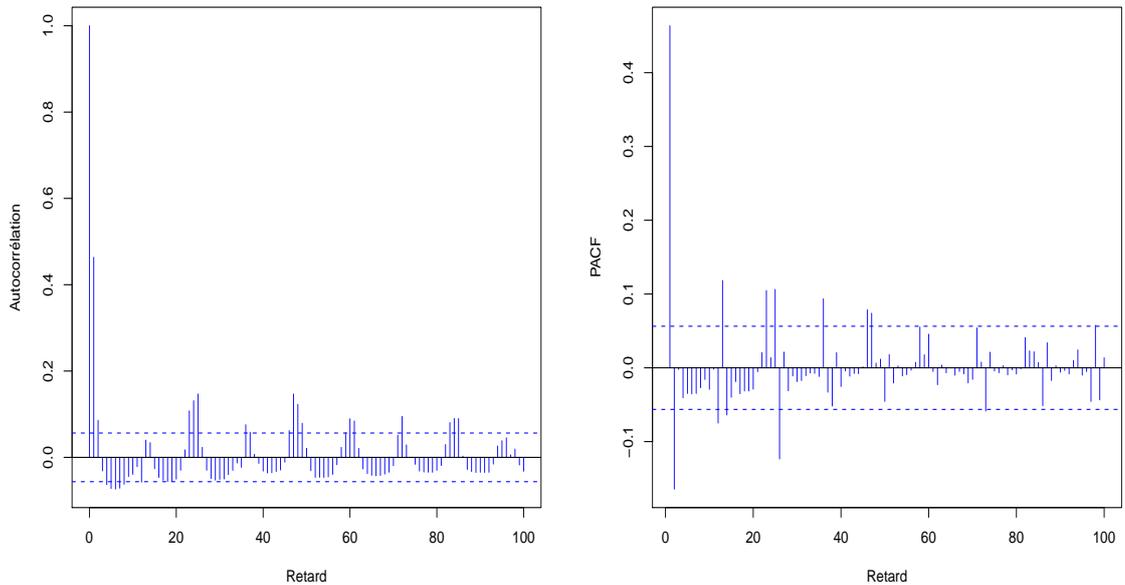


FIG. 5.8 – ACF et PACF empirique du processus résiduel $(\xi_t)_{t=1,\dots,T}$.

Pour tester la blancheur des résidus $(\varepsilon_t)_{t=1,\dots,T}$, issus de la modélisation de (5.10), nous considérons la Figure 5.9, représentant la trajectoire et l'ACF empirique du processus

$(\varepsilon_t)_{t=1,\dots,T}$. Nous observons que la trajectoire du processus résiduel semble être stationnaire autour de zéro, et que l'ACF empirique ne présente pas de pics significativement non nuls, prouvant ainsi que les résidus ne semblent pas être autocorrélés. Ce résultat est confirmé par le test de Box et Pierce visible sur le Tableau 5.5. En effet, la statistique de Box et Pierce $Q_{BP} = 66.6818$ est inférieure au quantile d'ordre 0.95 de la loi de Chi-deux (67.5). En plus, les valeurs du Skewness, du Kurtosis et de la statistique de Jarque et Bera résumées sur le Tableau 5.5, montrent clairement que les résidus ne sont pas normaux. Ainsi, à l'issue de cette analyse, nous pouvons dire que le processus résiduel $(\varepsilon_t)_{t=1,\dots,T}$ est un bruit blanc distribué suivant une loi $S_\alpha(\gamma, \beta, \mu)$.

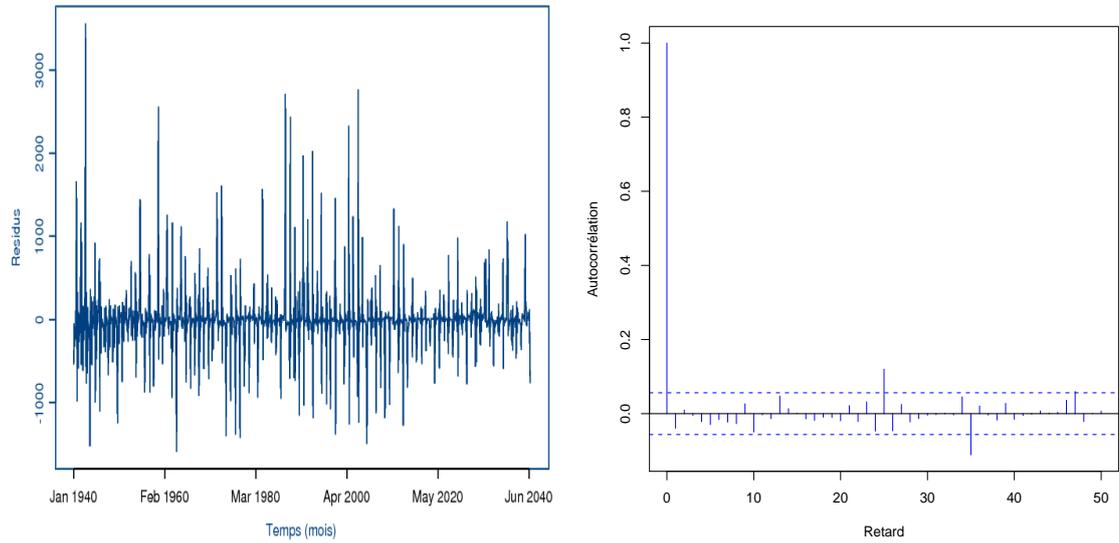


FIG. 5.9 – Trajectoire et ACF empirique du processus résiduel $(\varepsilon_t)_{t=1,\dots,T}$.

Moyenne	Variance	Skewness	Kurtosis	Jarque et Bera	Box et Pierce
-9.5136	170342.7	1.7357	13.7455	10164.93 p-value < $2.2 \cdot 10^{-16}$	66.6818

TAB. 5.5 – Statistiques descriptives du processus $(\varepsilon_t)_{t=1,\dots,T}$.

A l'issue de toutes ces remarques, nous pouvons conclure que le modèle retenu semble prendre en compte les différentes caractéristiques observées sur la série mensuelle du débit du fleuve Sénégal à Bakel. Formellement, le modèle final pour décrire la série d'étude $(X_t)_{t=1,\dots,T}$ s'écrit comme suit :

$$\phi(B) \Phi(B^{12}) (1-B)^{0.09220} (1-B^{12})^{0.4007} (X_t - 662.737) = \theta(B) \Theta(B^{12}) \varepsilon_t, \quad (5.11)$$

$$\text{où } \phi(B) = 1 - 0.1971B \quad \Phi(B^{12}) = 1 - 0.9758B^{12}$$

$$\text{et } \theta(B) = 1 + 0.3387B \quad \Theta(B^{12}) = 1 - 1.1262B^{12} + 0.1930B^{24}$$

et $(\varepsilon_t)_t$ est un processus i.i.d suivant une loi $S_\alpha(\gamma, \beta, \mu)$. L'estimation des paramètres de la distribution des innovations est effectuée par la méthode de McCulloch et les résultats sont résumés dans le Tableaux 5.6 suivant :

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\mu}$
Méthode de McCulloch	1.7903	-0.2092	0.2313	-0.0537

TAB. 5.6 – Estimation des paramètres de la distribution du processus $(\varepsilon_t)_{t=1,\dots,T}$.

Nous allons maintenant utiliser le modèle (5.11) en prévision. Ainsi, les résultats relatifs à la qualité prédictive de ce modèle sont présentés dans le Tableau 5.7 suivant :

	Oct.	Nov.	Dec.	Jan.	Fév.	Mars.	Avr.	Mai.	Jui.	Juil.	Aout.	Sept.
RMSE	341.15	133.67	91.27	57.43	25.90	23.62	45.60	51.67	73.62	144.71	302.48	609.64
MAPE	0.437	0.364	0.512	0.366	0.136	0.101	0.164	0.158	0.219	0.201	0.212	0.319

TAB. 5.7 – Résultats en prévision sur la série mensuelle du débit du fleuve Sénégal à Bakel, par le modèle ARFISMA.

5.3.3 Comparaison des résultats

Dans cette section, nous nous intéressons à la qualité prédictive des modèles présentés dans les Sections 5.3.1 et 5.3.2. Ainsi, nous comparons les performances de prévision

en utilisant les critères de RMSE et de MAPE. Les résultats sont présentés dans les Tableaux 5.3 et 5.7.

Nous observons, d'après ces tableaux, que les RMSE et MAPE sont élevés pour l'ensemble des deux modèles. Cependant, nous notons une diminution relative de ces quantités, surtout pour les mois de Janvier, Février et Mars. Lorsque nous comparons les résultats en prévision obtenus par le modèle de Box et Jenkins (Tableau 5.3) et ceux obtenus par le modèle ARFISMA (Tableau 5.7), nous observons que la performance du modèle ARFISMA se fait sentir à partir du mois de Février. La Figure 5.10 qui représente l'évolution du critère RMSE permet également de voir ce résultat.

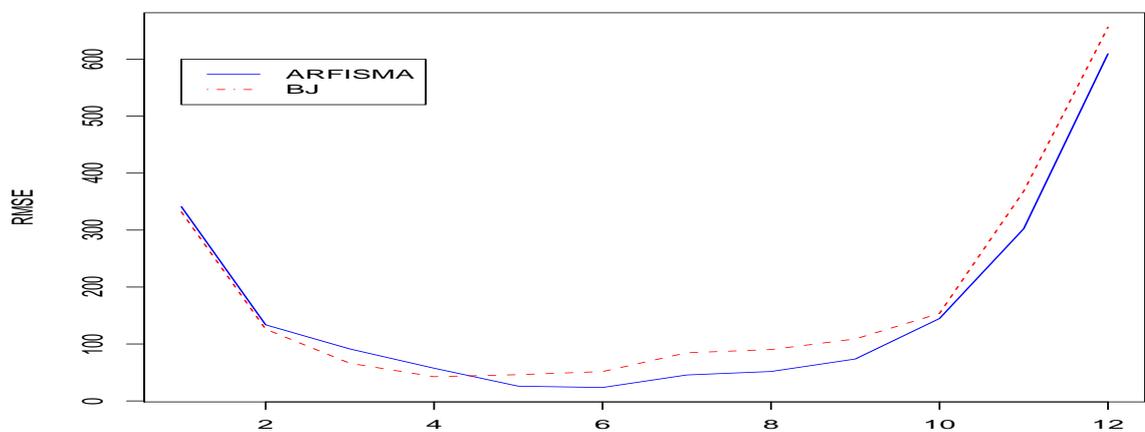


FIG. 5.10 – Comparaison de l'évolution du critère RMSE en fonction des mois, pour les prévisions des deux modèles

5.4 Conclusions et perspectives

Dans ce chapitre, nous avons présenté les résultats théoriques relatifs à la méthodologie. Nous nous sommes également intéressés à l'application sur des données réelles, des processus ARFISMA α -stables développés dans le Chapitre 2. Ainsi, nous avons considéré les débits mensuels du fleuve Sénégal à la station de Bakel. En guise de comparaison, nous étudions le modèle classique SARIMA de Box et Jenkins [11] et le modèle ARFISMA. La performance du modèle ARFISMA semble être la meilleure, en terme de RMSE et MAPE.

Lorsque les paramètres du processus sont estimés, la qualité du prédicteur dépend fortement de la qualité de l'estimation des paramètres. Or la méthode d'estimation utilisée dans ce travail (méthode de Whittle) ne fournit pas une matrice de variance-covariance, comme dans le cas gaussien. Par conséquent, cette méthode ne permet pas de juger la qualité de l'estimation. En pratique, il est donc important d'utiliser une méthode d'estimation qui fournisse un estimateur consistant et qui permette d'obtenir la distribution limite de l'estimateur. Une idée consiste alors à approcher la distribution limite de cet estimateur (équation 5.1) par une méthode de rééchantillonnage de type Bootstrap. Ceci peut faire l'objet d'un futur travail de recherche.

Conclusion générale

Dans ce travail de thèse, nous nous intéressons à une classe de modèles paramétriques de séries chronologiques : *les processus à mémoire longue saisonniers avec variance infinie des innovations*. Ces processus ont la particularité de prendre en compte à la fois les phénomènes de dépendance à long terme, de saisonnalité et de variance infinie que l'on rencontre souvent en finance, en télécommunication ou en hydrologie. Ils constituent une généralisation directe des processus longue mémoire ARFIMA, classiquement utilisés dans la littérature statistique, et possèdent plusieurs extensions.

Après avoir établi les propriétés probabilistes de ce modèle, nous nous sommes intéressés à leurs propriétés statistiques. L'estimation des paramètres d'un processus ARFISMA symétrique α -stable est un problème délicat à régler. La difficulté provient surtout du caractère α -stable du processus. Dans ce travail, nous avons alors proposé plusieurs méthodes d'estimation des paramètres : une première méthode basée sur une extension de la méthode GPH dans le cas saisonnier, la deuxième méthode est basée sur la méthode de Whittle et une troisième basée sur l'évaluation de la vraisemblance de Whittle par la méthode MCMC. Afin de comparer les propriétés asymptotiques de ces estimateurs, nous avons effectué des simulations de Monte Carlo. Les résultats obtenus sont tous satisfaisants, en terme de RMSE et MAE. Cependant, la méthode de Whittle par MCMC semble être privilégiée, surtout lorsque la taille de l'échantillon est faible, car les méthodes semiparamétriques nécessitent un grand nombre de fréquences dans la régression, et donc un grand nombre de données. Malheureusement, le nombre de données disponibles en pratique est souvent limité. Nous nous sommes également intéressés à l'estimation simultanée des paramètres de mémoire longue, des paramètres de la distribution du bruit et des parties autorégressives et moyenne-mobile. Pour cela, nous avons proposé la méthode ECF et la méthode C-GMM qui permettent d'estimer à la fois l'ensemble des paramètres du processus. Pour la méthode ECF, nous avons établi, sous certaines conditions de régularité les propriétés asymptotiques : consistance et normalité asymptotique. Un plan de simulations de Monte Carlo a permis de comparer ces estimateurs et d'étudier l'influence de la présence des parties autorégressives et moyenne-mobile sur l'estimation des paramètres de mémoire longue. Les résultats de simulations ont montré que la méthode de Whittle par MCMC donne une meilleure

performance pour le modèle simple sans partie courte mémoire, alors que pour le modèle avec parties autorégressives et moyenne-mobile, la méthode ECF est préférée. Cependant, une étude précise par simulation de Monte Carlo semble nécessaire pour comparer la méthode ECF et la méthode C-GMM. Nous notons toutefois, que la méthode ECF n'est pas optimale, car la fonction de poids exponentielle utilisée dans la procédure rend les calculs faciles mais elle est sous-optimale. De plus, nous ne disposons pas d'outil théorique permettant de guider le choix de la taille des blocs qui influence fortement la performance de cet estimateur. Une manière d'améliorer les performances de la méthode ECF est d'utiliser une approximation de la fonction de poids optimale (4.4).

En ce qui concerne la prévision avec les processus ARFISMA α -stables, nous avons utilisé la méthode de l'erreur minimale de dispersion développée dans Kokoszka [44], Cline et Brockwell [19] ou Brockwell et Davis [13], pour ces types de processus. En effet, dans le cas stable l'erreur quadratique moyenne n'existe pas. Il est à noter, cependant, que cette méthode ne fournit pas des intervalles de confiance de prévisions. Ceci est également lié au caractère α -stable de ces types de processus. Il serait donc intéressant d'établir les expressions de ces intervalles de confiance et une méthode numérique du type Bootstrap devrait permettre de les obtenir. On se propose de réaliser ce travail dans le cadre de futures recherches.

Enfin, nous avons effectué une application du processus ARFISMA α -stable sur des données de débit du fleuve Sénégal à la station de Bakel. L'étude a également permis de comparer la capacité prédictive de ce modèle avec celle du modèle linéaire classique SARIMA de Box et Jenkins [11]. Les résultats ont montré que la meilleure performance est obtenue avec le modèle ARFISMA α -stable, en terme de RMSE et MAPE. Dans le cadre de futures recherches, nous nous proposons d'améliorer ces résultats, en cherchant plus de précision sur l'estimation. En effet, la méthode de Whittle par MCMC utilisée dans les applications ne fournit pas de moyen pour contrôler l'erreur d'estimation. Nous pourrions également tenter d'adapter dans le cas des processus ARFISMA α -stable, la méthode du maximum de vraisemblance développée récemment par Andrews et al.[5] pour les processus AR(p) α -stables. Il serait également intéressant de tester le modèle sur d'autres données telles que des données de télécommunication ou de température.

En ce qui concerne les extensions possibles de ce travail sur les processus ARFISMA α -stables, plusieurs pistes semblent fructueuses. Dans ce travail, nous avons considéré le modèle avec des cycles fixes. Nous pouvons penser par exemple, à développer

des processus prenant en compte à la fois les phénomènes de cycles plus ou moins fixes, de mémoire longue et de variance infinie. Pour cela, nous avons pensé à développer des processus GARMA à k -facteurs avec un bruit α -stable. Nous nous proposons de réaliser ce travail dans un avenir proche.

Bibliographie

- [1] Adler, R., Feldman, R. E., et Taqqu, M. *A Practical Guide to Heavy Tails : Statistical Techniques and Applications*. Birkhäuser, Boston, 1998.
- [2] Adler R. J., Feldman R. E. et Gallagher C. *Analysing stable time series*. In Adler , Feldman et Taqqu (Eds.), *A Practical Guide to Heavy Tails*, pages 133-158. Birkhäuser, Boston, 1998a. 96
- [3] Agiakloglou, C., Newbold,P. et Wohar, M. Bias in an estimator of the fractional difference parameter. *Journal of Time Series Analysis*, **14**(3) : 235–246, 1993. 66
- [4] Alvarez, A. et Olivares, P. Méthodes d'estimation pour des lois stables avec des applications en finance. *Journal de la Société Française de Statistique*, tome 1, n°4 : 23–54, 2005. 19
- [5] Andrews, B., Calder, M., et Davis, R. A. Maximum likelihood estimation for α -stable autoregressive processes. *Annals of statistics*, **37**(4) : 1946–1982, 2009. 114
- [6] Andrews, D. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, **59** : 817–858, 1991. 88, 89
- [7] Arteche, J. et Robinson, P. M. Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis*, **21**(1) : 1–25, 2000.
- [8] Bassi, F., Embrechts, P. et Kafetzaki, M. *Risk management and quantile estimation*. In Adler , Feldman et Taqqu, editors, *A Practical Guide to Heavy Tails*, pages 111-130. Birkhauser, Boston, 1998. 7
- [9] Bestravos, A., Crovella, M. et Taqqu, M. *Heavy-tailed distributions in the world wide web*. In Adler , Feldman et Taqqu, editors, *A Practical Guide to Heavy Tails*, pages 3-25. Birkhauser, Boston, 1998. 7
- [10] Bitty, A. L. et Hili, O. Estimateurs du minimum de distancede hellinger des processus linéaire à memoire longue. *Comptes Rendus Mathematique*, **348** (7-8) : 445–448, 2010. 2
- [11] Box, G. E. P. et Jenking, G. M. *Time Series Analysis : Forecasting and Control*. San Francisco : Holden Day, 1976. iv, v, 5, 25, 69, 93, 104, 112, 114

- [12] Brockwell, P. J. et Davis, R. A. *Time Series : Theory and Methods*. Springer-Verlag, New York, 1987. 35, 69
- [13] Brockwell, P. J. et Davis, R. A. *Time Series : Theory and Methods*. Springer Series in statistics. Springer-Verlag, second éd, 1991. 7, 25, 26, 27, 36, 37, 96, 97, 114
- [14] Brockwell, P., J et Davis, R., A. *Introduction to Time Series and Forecasting*. Springer, New York, 1996.
- [15] Carrasco, M., Chernov, M., Florens, J. P. et Ghysels, E. Efficient estimation of general dynamic models with a continuum of moment conditions. *Journal of econometrics*, **140** : 529–573, 2007. 87, 88, 89, 90, 91
- [16] Carrasco, M. et Florens, M. Efficient gmm estimation using the empirical characteristic function. *Working paper, CREST, Paris*, 2002. iv, v, 70, 87, 89
- [17] Chambers, J. M., Mallows, C. et Stuck, B. W. A method for simulating stable random variables. *Journal of the American Statistical Association*, **71** (354) : 340–344., 1976. 14
- [18] Cline, D. B. H. *Estimation and linear prediction for regression, autoregression and ARMA with infinite variance data*. Ph. D. Dissertation, Statistics Department, Colorado State University, 1983. 74
- [19] Cline, D. B. H. et Brockwell, P. J. Linear prediction of arma processes with infinite variance. *Stochastic Processes and their Applications*, **19** : 281–296, 1985. 74, 97, 114
- [20] Crato, N. et De Lima, P. J. F. Long-range dependence in the conditional variance of stocks returns. *Economics Letters*, **45** : 281–285, 1994. 64
- [21] Diongue, A. K ., Diop, A. et Ndongo, M. Seasonal fractional ARIMA with stable innovations. *Statistics and Probability Letters*, **78** : 1404–1411., 2008. 3, 58, 65, 104
- [22] DuMouchel, W. H. On the asymptotic normality of the maximum likelihood estimate one sampling from a stable distribution. *Annals of statistics*, **1** : 948–957, 1973.
- [23] Embrechts, P., Klüppelberg, C. et Mikosch, T. *Modelling Extremal Events*. New York : Springer-Verlag, 1997. 7, 61, 96
- [24] Fama, E. Behavior of stock market prices. *J. Bus. U. Chicago*, **38** : 34–105, 1965. 2, 36

-
- [25] Feller, W. *An Introduction to Probability Theory and its Applications*. New York : Wiley (2nd ed.), 1971. 8, 9
- [26] Feuerverger, A. An efficiency result for the empirical characteristic function in stationary time-series models. *Canadian Journal of Statistics*, **18**:155–161, 1990. 70, 74
- [27] Gallagher, C. M. A method for fitting stable autoregressive models using the autocovariation function. *Statistics and Probability Letters*, **53**: 381–390, 2001. 2
- [28] Garcia, R., Renault, E. et Veredas, D. Estimation of stable distributions by indirect inference. *CORE Mimeo*, 2004. 20
- [29] Geweke, J. et Porter-Hudak, S. The estimation and application of long memory time series models. *Journal of Time Series Analysis*, **4**: 221–228, 1983. 2, 3, 59, 64
- [30] Giraitis, L. et Leipus, R. A generalized fractionally differencing approach in long memory modelling. *Lithuanian Mathematical Journal*, **35**:65–81, 1995. 38, 39
- [31] Granger, C.W.J. et Joyeux, R. An introduction to long memory time series and fractional differencing. *Journal of Time Series Analysis*, **1**:15–30, 1980. 1, 29, 35, 36
- [32] Hamilton, J. D. *Time Series Analysis*. Princeton University Press, 1994. 69
- [33] Hansen, L. P. Large sample properties of generalized method of moments estimators. *Econometrica*, **50**(4):1029–1054, 1982. 77
- [34] Hassler, U. (mis)specification of long memory in seasonal time. *Journal of Time Series Analysis*, **15**(1): 19–30, 1994. 65
- [35] Heathcote, C. R. The integrated squared error estimation of parameters. *Biometrika*, **64**:255–264, 1977. 75
- [36] Higuchi, T. Approach to an irregular time series on the basis of the fractal theory. *Physica, D*, **31**: 277–283, 1988. 2
- [37] Hosking, J.R.M. Fractional differencing. *Biometrika*, **68**(1): 165–176, 1981. 1, 29, 35
- [38] Hurst, H. E. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, **116**:770–799, 1951. 2

- [39] Jiang, G. et Knight, J. L. Estimation of continuous time processes via the empirical characteristic function. *Journal of Business Economic Statistics*, **20** : 198–212, 2002. 70
- [40] Klüppelberg, C. et Mikosch, T. . Some limit theory for the self-normalised periodogram of p-stable processes. *Scandinavian Journal of Statistics*, **21** : 485–491, 1994. 27, 61
- [41] Klüppelberg, C. et Mikosch, T. Spectral estimates and stable processes. *Stochastic Processes and their Applications*, **47** : 323–344., 1993. 61
- [42] Knight, J. L. et Satchell, S. E. Estimation of stationary stochastic processes via the empirical characteristic function. *Manuscript, Department of Economics, University of Western Ontario*, 1996. 70
- [43] Knight, J. L. et Yu, J. Empirical characteristic function in time series estimation. *Econometric Theory*, **18**(2) : 691–721, 2002. iv, v, 70, 74, 75, 76, 77, 126, 140
- [44] Kokoszka, P. S. Prediction of infinite variance fractional ARIMA. *Probability and Mathematical Statistics*, **16**(1) : 65–83, 1996. 97, 98, 114
- [45] Kokoszka, P. S. et Taqqu, M. S. Fractional ARIMA with stable innovations. *Stochastic Processes and their Applications*, **60** : 19–47, 1995. iv, v, 1, 3, 7, 29, 30, 35, 37, 41, 42, 44, 64
- [46] Kokoszka, P. S. et Taqqu, M. S. Parameter estimation for infinite variance fractional ARIMA. *The Annals of Statistics*, **24**(5) : 1880–1913, 1996. 2, 31, 32, 33, 52, 64, 95
- [47] Kouamé, E. F. et Hili, O. Minimum distance estimator of k-factors gamma processes. *Statistics and Probability Letters*, **78**(18) : 3254–3261, 2008. 2
- [48] Koutrouvelis, I. A. Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, **75**(372) : 918–928, 1980.
- [49] Lewis, P.A.W, et Ray, B.K. Modeling long-range dependence, nonlinearity and periodic phenomena in sea surface temperatures using tsmars. *Journal of the American Statistical Association*, **92**(439) : 881–893, 1997. 2
- [50] Lo, A. Long-term memory in stock market prices. *Econometrica*, **59** : 1279–1313, 1991. 2

- [51] Lobato, I. et Robinson, P. M. Average periodogram estimation of long memory. *Journal of Econometrics*, **73**:303–324, 1996. 2
- [52] Lukacs, E. *Characteristic Functions*. New York: Griffin, 1970. 70
- [53] Lévy, P. Théorie des erreurs. la loi de gauß et les lois exceptionnelles. *Bulletin de la Société Mathématique de France*, **52**:49–85, 1924. 8
- [54] Magnus, W., Oberhettinger, F. et Soni, R. P. *Formulas and Theorems for the special function of mathematical physics*. 3rd edn. Springer-Verlag, New York, 1966. 37
- [55] Mandelbrot, B. The variation of certain speculative prices. *Journal of business*, **36**:394–419, 1963. 2
- [56] McCulloch, J. H. Simple consistent estimators of stable distribution parameters. *Communications in Statistics*, **15**:1109–1136, 1986. 19
- [57] Mikosch, T, Gadrich, T, Klüppelberg, C et Adler, R. J. Parameter estimation for ARMA models with infinite variance innovations. *Annals of statistics*, **23**(1):305–326, 1995. iv, v, 1, 2, 3, 7, 26, 28, 33, 52, 61
- [58] Moulines, E. et Soulier, P. Broadband log-periodogram regression of time series with long-range dependence. *The Annals of Statistics*, **27**(4):1415–1439, 1999. 60, 61
- [59] Ndongo, M., Diongue, A. K. , Diop, A. et Doussou-Gbété, S. Estimation of long memory parameters for seasonal fractional ARIMA with stable innovations. *Statistical Methodology*, **7**:141–151, 2010. 4
- [60] Nikias, C. L. et Shao, M. *Signal Processing with Alpha-Stable Distributions and Applications*. John Wiley & Sons, New York., 1995. 7
- [61] Nolan, J. P. *Stable Distributions - Models for Heavy Tailed Data*. Birkhauser, Boston., 2004.
- [62] Paulson, A. S., Holcomb, E. W., et Leitch, R.A. The estimation of the parameters of the stable laws. *Biometrika*, **62**:163–170, 1975.
- [63] Porter-Hudak, S. An application to seasonal fractionally differenced model to the monetary aggregates. *Journal of American Statistical Association*, **85**:338–344, 1990. 1, 35, 64

- [64] R. Development core team r: A language and environment for statistical computing. r foundation for statistical computing, vienna, austria. isbn 3-900051-07-0. *URL* <http://www.R-project.org>, 2008. iv, v, 3
- [65] Rainville, E. D. *Special Functions*. Mac Millan, New York, 1960. 37
- [66] Reisen, V. A. Estimation of the fractional difference parameter in the ARIMA(p, d, q) model using the smoothed periodogram. *Journal of Time Series Analysis*, **15**:335–350, 1994. 2
- [67] Reisen, V. A., Rodrigues, A. L. et Palma, W. Estimation os seasonal fractionally integrated processes. *Computational Statistics & Data Analysis*, **50**:568–582, 2006. iv, v, 1, 2, 3, 35, 37, 38, 48, 58, 59, 60, 65, 66
- [68] Robert, C., P. et Casella, G. *Monte Carlo Statistic Methods*. 2nd ed. ,Springer Texts in Statistics, 2004. 62
- [69] Robinson, P. M. Semiparametric analysis of long-memory time series. *The Annals of Statistics*, **22**:515–539, 1994. 2
- [70] Robinson, P. M. Gaussian semiparametric estimation of long range dependence. *Th Annals of Statistics*, **23**:1630–1661, 1995b. 2
- [71] Samorodnitsky, G. et Taqqu, M. *Stable non-gaussian random processes : Stochastic models with infinite variance*. Chapman & Hall, 1994. 9, 13, 14, 36
- [72] Sowell, F. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, **53**:165–188, 1992. 2
- [73] Stanislavsky, A.,A., Burnecki, K., Magdziarz, M., Weron, A. et Weron, K. FARIMA modeling of solar flare activity from empirical time series of soft x-ray solar emission. *The Astrophysical Journal*, **693**:1877–1882, 2009. 2
- [74] Stoev, S. et Taqqu, M. Simulation methods for linear fractional stable motion and FARIMA using the fast fourier transform. *Fractals*, **12**(1):95–121, 2004. 3, 45, 46, 63, 79
- [75] Stuck, B. W. et Kleiner, B. A statistical analysis of telephone noise. *The Bell System Technical Journal*, **53**:1263–1320, 1974. 2, 36

-
- [76] Szegő, G. *Othogonal polynomials*. AMS Colloquium Publications, vol. 23, Amer. Math. Soc. : New York, 1959. 37
- [77] Whittle, P. *Hypothesis Testing in Time Series Analysis*. New York, 1951. 2
- [78] Woodward, W.A., Cheng, Q.C. et Gray, H.L. A k-factor gamma long-memory model. *Journal of Time Series Analysis*, **19**(4) :485–504, 1998. 1, 39
- [79] Yu, J. Empirical characteristic function estimation and its applications. *Econometric Reviews*, **23**(2) :93–123, 2004. 70, 74
- [80] Zolotarev, V. M. *One-dimentional stable distribution*. In *Translations of Mathematical Monographs*, volume 65. American Mathematical Society, 1986. 12

Calcul des coefficients $(c_j)_{j \geq 0}$

Dans cette partie, nous présentons le calcul des coefficients $(c_j)_{j \geq 0}$ de la représentation moyenne mobile infinie d'un processus ARFISMA(p, d, q)(P, D, Q)_s symétrique α -stable. Ces coefficients sont obtenus à partir du développement de l'équation suivante :

$$\Theta_s(z^s)\Theta(z) \sum_{j=0}^{+\infty} \psi_j(d, \nu)z^j = \Phi_s(z^s)\Phi(z) \sum_{j=0}^{+\infty} c_j z^j, \quad \forall |z| < 1.$$

$$\text{Posons : } A(z) = \Phi_s(z^s)\Phi(z) \sum_{j=0}^{+\infty} c_j z^j, \quad B(z) = \Theta_s(z^s)\Theta(z) \sum_{j=0}^{+\infty} \psi_j(d, \nu)z^j$$

• Développons le polynôme $A(z)$:

$$\begin{aligned} A(z) &= \left(1 - \sum_{i=1}^P \Phi_i z^{is}\right) \left(1 - \sum_{j=1}^p \phi_j z^j\right) \sum_{k=0}^{+\infty} c_k z^k \\ &= \sum_{k=0}^{+\infty} c_k z^k - \left(\sum_{j=1}^p \phi_j z^j\right) \left(\sum_{k=0}^{+\infty} c_k z^k\right) - \left(\sum_{i=1}^P \Phi_i z^{is}\right) \left(\sum_{k=0}^{+\infty} c_k z^k\right) \\ &\quad + \left(\sum_{i=1}^P \Phi_i z^{is}\right) \left(\sum_{j=1}^p \phi_j z^j\right) \left(\sum_{k=0}^{+\infty} c_k z^k\right) \end{aligned} \quad (\text{A.1})$$

Développons ensuite chaque terme de l'équation (A.1) :

$$\rightarrow \left(\sum_{i=1}^p \phi_i z^i\right) \left(\sum_{j=0}^{+\infty} c_j z^j\right) = \sum_{j=0}^{\infty} \left(\sum_{i=1}^{\min(j,p)} \phi_i c_{j-i}\right) z^j$$

$$\rightarrow \left(\sum_{i=1}^P \Phi_i z^{is}\right) \left(\sum_{k=0}^{+\infty} c_k z^k\right) = \sum_{\lambda=1}^P \sum_{i=1}^{s-1} \tilde{\beta}_{\lambda s+i} z^{\lambda s+i},$$

$$\text{avec } \tilde{\beta}_{\lambda s+i} = \sum_{k=1}^{\lambda} \Phi_k c_{(\lambda-k)s+i}$$

$$\begin{aligned} \rightarrow \left(\sum_{i=1}^P \Phi_i z^{is} \right) \left(\sum_{j=1}^p \phi_j z^j \right) \left(\sum_{k=0}^{+\infty} c_k z^k \right) &= \left(\sum_{i=1}^P \sum_{j=1}^p \phi_j \Phi_i z^{is+j} \right) \left(\sum_{k=0}^{+\infty} c_k z^k \right) \\ &= \sum_{\lambda=1}^P \sum_{k=1}^p \tilde{H}_{\lambda s+k} z^{\lambda s+k}, \\ \text{avec } \tilde{H}_{\lambda s+k} &= \Phi_\lambda \sum_{i=1}^k \phi_i c_{k-i} + \sum_{l=1}^{\lambda-1} \Phi_l \sum_{i=1}^p \phi_i c_{(\lambda-l)s+k-i} \end{aligned}$$

Donc $A(z)$ s'écrit :

$$\begin{aligned} A(z) &= \sum_{j=0}^{s-1} \left(c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} \right) z^j + \sum_{\lambda=1}^P \sum_{k=0}^{s-1} \left(\tilde{H}_{\lambda s+k} - \tilde{\beta}_{\lambda s+k} \right) z^{\lambda s+k} \\ &+ \sum_{\lambda=1}^{\infty} \sum_{k=0}^{s-1} \left(c_{\lambda s+k} - \sum_{i=1}^{\min(p,\lambda s+k)} \phi_i c_{\lambda s+k-i} \right) z^{\lambda s+k} \end{aligned}$$

Le développement du polynôme $B(z)$ s'effectue de la même manière et on obtient :

$$\begin{aligned} B(z) &= \sum_{j=0}^{s-1} \left(\psi_j(d, \nu) + \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu) \right) z^j + \sum_{\lambda=1}^Q \sum_{k=0}^{s-1} \left(\tilde{\alpha}_{\lambda s+k} + \tilde{K}_{\lambda s+k} \right) z^{\lambda s+k} \\ &+ \sum_{\lambda=1}^{\infty} \sum_{k=0}^{s-1} \left(\psi_{\lambda s+k}(d, \nu) + \sum_{i=1}^{\min(q,\lambda s+k)} \theta_i \psi_{\lambda s+k-i}(d, \nu) \right) z^{\lambda s+k} \\ \text{avec } \tilde{K}_{\lambda s+k} &= \Theta_\lambda \sum_{i=1}^k \theta_i \psi_{k-i}(d, \nu) + \sum_{l=1}^{\lambda-1} \Theta_l \sum_{i=1}^q \theta_i \psi_{(\lambda-l)s+k-i}(d, \nu) \\ \tilde{\alpha}_{\lambda s+k} &= \sum_{i=1}^{\lambda} \Theta_i \psi_{(\lambda-i)s+k}(d, \nu) \end{aligned}$$

On suppose sans perte de généralité que $P \leq Q$. Par identification de polynômes ($A(z) = B(z)$), on obtient alors :

- Pour $j \in [0, s-1] \cup [(Q+1)s, +\infty[$

$$c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} = \psi_j(d, \nu) + \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu).$$

- Pour $j \in [s, (P+1)s-1]$

$$c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} + \tilde{H}_j - \tilde{\beta}_j = \psi_j(d, \nu) + \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu) + \tilde{K}_j + \tilde{\alpha}_j.$$

-
- Pour $j \in [(P + 1)s, (Q + 1)s - 1]$

$$c_j - \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} = \psi_j(d, \nu) + \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu) + \tilde{K}_j + \tilde{\alpha}_j.$$

■

Démonstration du Théorème 4.3.8

Dans cette partie, nous présentons les résultats relatifs à la consistance et à la normalité asymptotique de l'estimateur de la fonction caractéristique empirique pour un processus ARFISMA symétrique α -stable établis dans le Chapitre 4.

La preuve de ce théorème consiste à vérifier les hypothèses du Théorème 2.1 de Knight et Yu [43]. Plus précisément, nous allons montrer, sous les hypothèses "**Hypothèse 4.3.2**" et "**Hypothèse 4.3.3**", que les hypothèses (A2), (A3), (A5) et (A6) de Knight et Yu [43] sont vraies.

(A2) : $I_n(\psi)$ est deux fois continûment dérivable sous le signe intégral, pour tout $\psi \in \Psi$. Pour cela, posons $f(r; \psi) = |c_n(r) - c(r; \psi)|^2 g(r)$. Nous avons alors

$$\begin{aligned} \left| \frac{\partial}{\partial \psi} f(r; \psi) \right| &= 2 g(r) |c_n(r) - c(r; \psi)| \left| \frac{\partial c(r; \psi)}{\partial \psi} \right| \\ &\leq 2 g(r) (|c_n(r)| + |c(r; \psi)|) \left| \frac{\partial c(r; \psi)}{\partial \psi} \right| \\ &\leq 4 g(r) \left| \frac{\partial c(r; \psi)}{\partial \psi} \right|. \end{aligned}$$

En posant $U(r; \psi) = \sum_{j=0}^{\infty} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha} + \sum_{l=2}^{m+1} \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^{\alpha}$, nous obtenons :

$$\begin{aligned} \left| \frac{\partial}{\partial \psi} f(r; \psi) \right| &\leq 4 g(r) \exp(-U(r; \psi)) \left| \frac{\partial U(r; \psi)}{\partial \psi} \right| \\ &\leq 4 g(r) \left| \frac{\partial U(r; \psi)}{\partial \psi} \right|. \end{aligned}$$

▷ **Montrons que $I_n(\psi)$ est de classe \mathcal{C}^1 par rapport à $\delta = (d, D, \zeta, \xi)$**

D'après ce qui précède, nous avons :

$$\left| \frac{\partial}{\partial \delta} f(r; \psi) \right| \leq 4 g(r) \left| \frac{\partial U(r; \psi)}{\partial \delta} \right|.$$

Posons $S(r; \psi) = \sum_{j=0}^{\infty} u_j(r; \psi)$, avec $u_j(r; \psi) = \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha}$. Pour montrer que $U(r; \psi)$ est dérivable par rapport à δ , il suffit de montrer que $S(r; \psi)$ est dérivable. Ceci revient à montrer que $\left| \frac{\partial u_j(r; \psi)}{\partial \delta} \right| \leq w_j$, où w_j est le terme général d'une série convergente.

Nous pouvons écrire :

$$\left| \frac{\partial u_j(r; \psi)}{\partial \delta} \right| = \alpha \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha-1} \left| \sum_{l=0}^m r_{l+1} \frac{\partial c_{j+l}}{\partial \delta} \right|.$$

Posons $a_j = \frac{\partial c_j}{\partial \delta}$, nous avons alors :

$$\left| \frac{\partial u_j(r; \psi)}{\partial \delta} \right| \leq 2 \left(\sum_{l=0}^m |r_{l+1} c_{j+l}| \right)^{\alpha-1} \left(\sum_{l=0}^m |r_{l+1} a_{j+l}| \right).$$

En appliquant l'inégalité de Hölder, nous obtenons :

$$\left| \frac{\partial u_j(r; \psi)}{\partial \delta} \right| \leq 2 \left(\sum_{l=0}^m |r_{l+1}|^{\frac{\alpha}{\alpha-1}} \right)^{\alpha-1} \left(\sum_{l=0}^m |c_{j+l}|^{\alpha} \right)^{1-\frac{1}{\alpha}} \left(\sum_{l=0}^m |a_{j+l}|^{\alpha} \right)^{\frac{1}{\alpha}},$$

or $\sum_{l=0}^m |r_{l+1}|^{\frac{\alpha}{\alpha-1}} < \left(\sum_{l=0}^m |r_{l+1}| \right)^{\frac{\alpha}{\alpha-1}}$, puisque $\frac{\alpha}{\alpha-1} > 1$.

$$\Rightarrow \left| \frac{\partial u_j(r; \psi)}{\partial \delta} \right| \leq 2 \left(\sum_{l=0}^m |r_{l+1}| \right)^{\alpha} \left(\sum_{l=0}^m |c_{j+l}|^{\alpha} \right)^{1-\frac{1}{\alpha}} \left(\sum_{l=0}^m |a_{j+l}|^{\alpha} \right)^{\frac{1}{\alpha}}. \quad (\text{B.1})$$

- Il existe l_0 tel que $|c_{j+l}|^{\alpha} \leq |c_{j+l_0}|^{\alpha}$, pour tout $l = 0, \dots, m$.

$$\Rightarrow \left(\sum_{l=0}^m |c_{j+l}|^{\alpha} \right)^{1-\frac{1}{\alpha}} \leq (m+1)^{1-\frac{1}{\alpha}} |c_{j+l_0}|^{\alpha-1}.$$

- Il existe l_1 tel que $|a_{j+l}|^{\alpha} \leq |a_{j+l_1}|^{\alpha}$, pour tout $l = 0, \dots, m$.

$$\Rightarrow \left(\sum_{l=0}^m |a_{j+l}|^{\alpha} \right)^{\frac{1}{\alpha}} \leq (m+1)^{\frac{1}{\alpha}} |a_{j+l_1}|.$$

L'inégalité (B.1) devient alors :

$$\begin{aligned}
\left| \frac{\partial u_j(r; \psi)}{\partial \delta} \right| &\leq 2(m+1) \left(\sum_{l=0}^m |r_{l+1}| \right)^\alpha |c_{j+l_0}|^{\alpha-1} |a_{j+l_1}| \\
&\leq 2 \frac{(m+1)}{2} \left(\sum_{l=0}^m |r_{l+1}| \right)^\alpha [|c_{j+l_0}|^{2(\alpha-1)} + |a_{j+l_1}|^2] \\
&\leq (m+1) \left(\sum_{l=0}^m |r_{l+1}| \right)^\alpha w_j, \quad \text{avec } w_j = |c_{j+l_0}|^{2(\alpha-1)} + |a_{j+l_1}|^2.
\end{aligned}$$

D'après l'**Hypothèse 4.3.2**, w_j est le terme général d'une série convergente. Ce qui implique que $U(r; \psi)$ est dérivable par rapport à δ et nous avons :

$$\begin{aligned}
\left| \frac{\partial U(r; \psi)}{\partial \delta} \right| &= \frac{\partial S(r; \psi)}{\partial \delta} + \alpha \sum_{l=2}^{m+1} \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^{\alpha-1} \left| \sum_{h=0}^{m+1-l} r_{h+l} \frac{\partial c_h}{\partial \delta} \right| \\
&\leq \sum_{j=0}^{\infty} \left| \frac{\partial u_j(r; \psi)}{\partial \delta} \right| + \alpha \sum_{l=2}^{m+1} \left[\left(\sum_{h=0}^{m+1-l} |r_{h+l} c_h| \right)^{\alpha-1} \left(\sum_{h=0}^{m+1-l} |r_{h+l} a_h| \right) \right] \\
&\leq (m+1) \left(\sum_{l=0}^m |r_{l+1}| \right)^\alpha \sum_{j=0}^{\infty} w_j \\
&\quad + \alpha \sum_{l=2}^{m+1} \left[\left(\sum_{h=0}^{m+1-l} |r_{h+l} c_h| \right)^{\alpha-1} \left(\sum_{h=0}^{m+1-l} |r_{h+l} a_h| \right) \right].
\end{aligned}$$

Or $\sum_{j=0}^{\infty} w_j < \infty \implies \exists M > 0$ tel que $\sum w_j < M$. Ainsi, nous pouvons écrire :

$$\left| \frac{\partial U(r; \psi)}{\partial \delta} \right| \leq M (m+1) \left(\sum_{l=0}^m |r_{l+1}| \right)^\alpha + \alpha \sum_{l=2}^{m+1} \left[\left(\sum_{h=0}^{m+1-l} |r_{h+l} c_h| \right)^{\alpha-1} \left(\sum_{h=0}^{m+1-l} |r_{h+l} a_h| \right) \right].$$

$$\text{Posons } A = \alpha \sum_{l=2}^{m+1} \left[\left(\sum_{h=0}^{m+1-l} |r_{h+l} c_h| \right)^{\alpha-1} \left(\sum_{h=0}^{m+1-l} |r_{h+l} a_h| \right) \right]. \quad (\text{B.2})$$

$$\text{Or } |c_h| < 1 \implies \left(\sum_{h=0}^{m+1-l} |r_{h+l} c_h| \right)^{\alpha-1} < \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^{\alpha-1}.$$

En appliquant l'inégalité de Hölder, nous avons :

$$\sum_{h=0}^{m+1-l} |r_{h+l} a_h| \leq \left(\sum_{h=0}^{m+1-l} |r_{h+l}|^2 \right)^{1/2} \left(\sum_{h=0}^{m+1-l} |a_h|^2 \right)^{1/2}.$$

Ainsi, l'expression A peut être majorée comme suit :

$$\begin{aligned}
A &\leq 2 \sum_{l=2}^{m+1} \left[\left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^{\alpha-1} \left(\sum_{h=0}^{m+1-l} |r_{h+l}|^2 \right)^{1/2} \left(\sum_{h=0}^{m+1-l} |a_h|^2 \right)^{1/2} \right] \\
&\leq 2 \sum_{l=2}^{m+1} \left[\left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^{\alpha-1} \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right) \left(\sum_{h=0}^{m+1-l} |a_h|^2 \right)^{1/2} \right] \\
&\leq 2 \sum_{l=2}^{m+1} \left[\left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^{\alpha} \left(\sum_{h=0}^{m+1-l} |a_h|^2 \right)^{1/2} \right] \\
&\leq 2 \sum_{l=2}^{m+1} \left[\left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^{\alpha} \left(\sum_{h=0}^{+\infty} |a_h|^2 \right)^{1/2} \right] \\
&\leq 2 M' \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^{\alpha}, \quad \text{car } \sum_{h=0}^{+\infty} |a_h|^2 < \infty, \quad \text{d'après l'Hypothèse 4.3.3.}
\end{aligned}$$

Nous obtenons alors :

$$\left| \frac{\partial U(r; \psi)}{\partial \delta} \right| \leq M (m+1) \left(\sum_{l=0}^m |r_{l+1}| \right)^{\alpha} + 2 M' \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^{\alpha}. \quad (\text{B.3})$$

• Si $\sum_{l=0}^m |r_{l+1}| > 1$, alors $\left(\sum_{l=0}^m |r_{l+1}| \right)^{\alpha} < \left(\sum_{l=0}^m |r_{l+1}| \right)^2$. En effet, on a $\alpha \leq 2$ et $\log \left(\sum_{l=0}^m |r_{l+1}| \right) > 0$.

$$\Rightarrow \alpha \log \left(\sum_{l=0}^m |r_{l+1}| \right) \leq 2 \log \left(\sum_{l=0}^m |r_{l+1}| \right)$$

$$\Rightarrow \log \left(\sum_{l=0}^m |r_{l+1}| \right)^{\alpha} \leq \log \left(\sum_{l=0}^m |r_{l+1}| \right)^2$$

$$\text{d'où } \left(\sum_{l=0}^m |r_{l+1}| \right)^{\alpha} \leq \left(\sum_{l=0}^m |r_{l+1}| \right)^2.$$

• De même, si $\sum_{l=0}^m |r_{l+1}| < 1$, alors $\left(\sum_{l=0}^m |r_{l+1}|\right)^\alpha < \sum_{l=0}^m |r_{l+1}|$.

Nous posons alors :

$$\left(\sum_{l=0}^m |r_{l+1}|\right)^\alpha \leq H_1(r) = \begin{cases} \left(\sum_{l=0}^m |r_{l+1}|\right)^2, & \text{si } \sum_{l=0}^m |r_{l+1}| > 1 \\ \sum_{l=0}^m |r_{l+1}|, & \text{si } \sum_{l=0}^m |r_{l+1}| \leq 1. \end{cases}$$

De manière analogue, nous obtenons :

$$\left(\sum_{h=0}^{m+1-l} |r_{h+l}|\right)^\alpha \leq H_1(r; l) = \begin{cases} \left(\sum_{h=0}^{m+1-l} |r_{h+l}|\right)^2, & \text{si } \sum_{h=0}^{m+1-l} |r_{h+l}| > 1 \\ \sum_{h=0}^{m+1-l} |r_{h+l}|, & \text{si } \sum_{h=0}^{m+1-l} |r_{h+l}| \leq 1. \end{cases}$$

Ainsi, l'inégalité (B.3) devient alors :

$$\left|\frac{\partial U(r; \psi)}{\partial \delta}\right| \leq M (m+1)H_1(r) + 2 M' \sum_{l=2}^{m+1} H_1(r; l).$$

En posant $\Gamma_1(r) = M (m+1)H_1(r) + 2 M' \sum_{l=2}^{m+1} H_1(r; l)$, nous obtenons finalement :

$$\left|\frac{\partial}{\partial \delta} f(r; \psi)\right| \leq 4 \Gamma_1(r) g(r).$$

En utilisant l'**Hypothèse 4.3.3**, nous montrons facilement que :

$$\int \dots \int \Gamma_1(r) g(r) dr < \infty.$$

D'où $I_n(\psi)$ est de classe \mathcal{C}^1 par rapport à $\delta = (d, D, \zeta, \xi)$.

▷ **Montrons que $I_n(\psi)$ est de classe \mathcal{C}^1 par rapport à α .**

Pour cela, nous remarquons d'abord que :

$$\begin{aligned} \left| \frac{\partial u_j(r; \psi)}{\partial \alpha} \right| &= \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^\alpha \ln \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right| \\ &\leq \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^\alpha \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right| \\ &\leq \left(\sum_{l=0}^m |r_{l+1} c_{j+l}| \right)^{\alpha+1}. \end{aligned}$$

En appliquant ensuite l'inégalité de Hölder et en utilisant les calculs précédents, nous obtenons :

$$\begin{aligned} \left| \frac{\partial u_j(r; \psi)}{\partial \alpha} \right| &\leq \left(\sum_{l=0}^m |c_{j+l}|^\alpha \right)^{1+1/\alpha} \left(\sum_{l=0}^m |r_{l+1}|^{\frac{\alpha}{\alpha-1}} \right)^{\frac{\alpha^2-1}{\alpha}} \\ &\leq (m+1)^{1+1/\alpha} |c_{j+l_0}|^{\alpha+1} \left(\sum_{l=0}^m |r_{l+1}|^{\frac{\alpha}{\alpha-1}} \right)^{\frac{\alpha^2-1}{\alpha}} \\ &\leq (m+1)^{1+1/\alpha} |c_{j+l_0}|^{\alpha+1} \left(\sum_{l=0}^m |r_{l+1}| \right)^{\alpha+1}. \end{aligned} \quad (\text{B.4})$$

Or pour tout $1 < \alpha \leq 2$, on a $(m+1)^{1+1/\alpha} < (m+1)^2$ et $|c_{j+l_0}|^{\alpha+1} < |c_{j+l_0}|^2$.

En remarquant que :

$$\left(\sum_{l=0}^m |r_{l+1}| \right)^{\alpha+1} \leq H_2(r) = \begin{cases} \left(\sum_{l=0}^m |r_{l+1}| \right)^3, & \text{si } \sum_{l=0}^m |r_{l+1}| > 1 \\ \left(\sum_{l=0}^m |r_{l+1}| \right)^2, & \text{si } \sum_{l=0}^m |r_{l+1}| \leq 1, \end{cases}$$

l'inégalité (B.4) devient alors :

$$\left| \frac{\partial u_j(r; \psi)}{\partial \alpha} \right| \leq (m+1)^2 H_2(r) |c_{j+l_0}|^2.$$

Or $|c_{j+l_0}|^2$ est le terme général d'une série convergente, car d'après les conditions de stationnarité du modèle, nous avons $\sum_{j=0}^{\infty} |c_j|^\alpha < \infty$, pour tout $1 < \alpha \leq 2$. D'où $U(r; \psi)$

est bien dérivable par rapport à α . Donc, nous pouvons écrire :

$$\begin{aligned}
\left| \frac{\partial U(r; \psi)}{\partial \alpha} \right| &= \left| \frac{\partial}{\partial \alpha} S(r; \psi) + \frac{\partial}{\partial \alpha} \sum_{l=2}^{m+1} \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^\alpha \right| \\
&= \left| \frac{\partial}{\partial \alpha} S(r; \psi) + \sum_{l=2}^{m+1} \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^\alpha \ln \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right| \right| \\
&\leq (m+1)^2 H_2(r) \sum_{j=0}^{+\infty} |c_{j+l_0}|^2 + \sum_{l=2}^{m+1} \left[\left| \sum_{h=0}^{m+1-l} r_{l+h} c_h \right|^{\alpha+1} \right] \\
&\leq M (m+1)^2 H_2(r) + \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} |r_{l+h} c_h| \right)^{\alpha+1} \\
&\leq M (m+1)^2 H_2(r) + \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} |r_{l+h}| \right)^{\alpha+1}, \quad \text{car } |c_h| < 1.
\end{aligned}$$

$$\text{Or } \left(\sum_{h=0}^{m+1-l} |r_{l+h}| \right)^{\alpha+1} \leq H_2(r; l) = \begin{cases} \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^3, & \text{si } \sum_{h=0}^{m+1-l} |r_{h+l}| > 1 \\ \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^2, & \text{si } \sum_{h=0}^{m+1-l} |r_{h+l}| \leq 1. \end{cases}$$

Ce qui implique que $\left| \frac{\partial U(r; \psi)}{\partial \alpha} \right| \leq M (m+1)^2 H_2(r) + \sum_{l=2}^{m+1} H_2(r; l)$.

En posant $\Gamma_2(r) = M (m+1)^2 H_2(r) + \sum_{l=2}^{m+1} H_2(r; l)$, nous obtenons alors :

$$\left| \frac{\partial f(r; \psi)}{\partial \alpha} \right| \leq 4 \Gamma_2(r) g(r).$$

Sous l'**Hypothèse 4.3.3**, nous avons $\int \dots \int \Gamma_2(r) g(r) dr < \infty$ et par suite $I_n(r; \psi)$ est de classe \mathcal{C}^1 par rapport à α .

En résumé, nous venons de montrer que $I_n(r; \psi)$ est de classe \mathcal{C}^1 par rapport à ψ . Il reste

à montrer que $I_n(r; \psi)$ est de classe \mathcal{C}^2 par rapport à ψ . Pour cela, nous avons d'abord :

$$\begin{aligned}
\left| \frac{\partial^2}{\partial \psi \partial \psi'} f(r; \psi) \right| &= 2 g(r) \left| \frac{\partial c(r; \psi)}{\partial \psi} \right| \left| \frac{\partial c(r; \psi)}{\partial \psi'} \right| \\
&+ 2 g(r) |c_n(r) - c(r; \psi)| \left| \frac{\partial^2 c(r; \psi)}{\partial \psi \partial \psi'} \right| \\
&\leq 2 g(r) \left| \frac{\partial c(r; \psi)}{\partial \psi} \right| \left| \frac{\partial c(r; \psi)}{\partial \psi'} \right| + 4g(r) \left| \frac{\partial^2 c(r; \psi)}{\partial \psi \partial \psi'} \right| \\
&\leq 2 g(r) \left| \frac{\partial U(r; \psi)}{\partial \psi} \right| \left| \frac{\partial U(r; \psi)}{\partial \psi'} \right| \\
&+ 4 g(r) \left[\left| \frac{\partial U(r; \psi)}{\partial \psi} \right| \left| \frac{\partial U(r; \psi)}{\partial \psi'} \right| + \left| \frac{\partial^2 U(r; \psi)}{\partial \psi \partial \psi'} \right| \right] \\
&\leq 6 g(r) \left| \frac{\partial U(r; \psi)}{\partial \psi} \right| \left| \frac{\partial U(r; \psi)}{\partial \psi'} \right| + 4 g(r) \left| \frac{\partial^2 U(r; \psi)}{\partial \psi \partial \psi'} \right|.
\end{aligned}$$

▷ Montrons que $I_n(\psi)$ est de classe \mathcal{C}^2 par rapport à α .

$$\left| \frac{\partial^2}{\partial \alpha \partial \alpha'} f(r; \psi) \right| \leq 6 g(r) \left| \frac{\partial U(r; \psi)}{\partial \alpha} \right| \left| \frac{\partial U(r; \psi)}{\partial \alpha'} \right| + 4 g(r) \left| \frac{\partial^2 U(r; \psi)}{\partial \alpha \partial \alpha'} \right|.$$

D'après ce qui précède, nous avons :

$$\left| \frac{\partial U(r; \psi)}{\partial \alpha} \right| \leq \Gamma_2(r), \quad \text{pour tout } 1 < \alpha \leq 2.$$

$$\Rightarrow \left| \frac{\partial^2}{\partial \alpha \partial \alpha'} f(r; \psi) \right| \leq 6 g(r) [\Gamma(r)]^2 + 4 g(r) \left| \frac{\partial^2 U(r; \psi)}{\partial \alpha \partial \alpha'} \right|.$$

Or nous pouvons majorer le terme $\left| \frac{\partial^2 U(r; \psi)}{\partial \alpha \partial \alpha'} \right|$ comme suit :

$$\begin{aligned}
\left| \frac{\partial^2 u_j(r; \psi)}{\partial \alpha \partial \alpha'} \right| &= \ln \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right| \left| \frac{\partial}{\partial \alpha'} \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right| \right|^{\alpha'} \\
&= \ln \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right| \ln \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right| \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha'} \\
&\leq \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^2 \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha'} \\
&\leq \left(\sum_{l=0}^m |r_{l+1} c_{j+l}| \right)^{\alpha'+2}.
\end{aligned}$$

En appliquant l'inégalité de Hölder, nous avons :

$$\begin{aligned}
\left| \frac{\partial^2 u_j(r; \psi)}{\partial \alpha \partial \alpha'} \right| &\leq \left(\sum_{l=0}^m |c_{j+l}|^{\alpha'} \right)^{1+2/\alpha'} \left(\sum_{l=0}^m |r_{l+1}|^{\frac{\alpha'}{\alpha'-1}} \right)^{\frac{(\alpha'-1)(\alpha'+2)}{\alpha'}} \\
&\leq \left(\sum_{l=0}^m |c_{j+l}|^{\alpha'} \right)^{1+2/\alpha'} \left(\sum_{l=0}^m |r_{l+1}| \right)^{\alpha'+2} \\
&\leq (m+1)^{1+2/\alpha'} \left(\sum_{l=0}^m |r_{l+1}| \right)^{\alpha'+2} |c_{j+l_0}|^{\alpha'+2}.
\end{aligned}$$

Or $(m+1)^{1+2/\alpha'} \leq (m+1)^3$, $|c_{j+l_0}|^{\alpha'+2} \leq |c_{j+l_0}|^3$ et

$$\left(\sum_{l=0}^m |r_{l+1}| \right)^{\alpha'+2} \leq H_3(r) = \begin{cases} \left(\sum_{l=0}^m |r_{l+1}| \right)^4, & \text{si } \sum_{l=0}^m |r_{l+1}| > 1 \\ \left(\sum_{l=0}^m |r_{l+1}| \right)^3, & \text{si } \sum_{l=0}^m |r_{l+1}| \leq 1. \end{cases}$$

Ce qui entraîne que :

$$\left| \frac{\partial^2 u_j(r; \psi)}{\partial \alpha \partial \alpha'} \right| \leq (m+1)^3 H_3(r) |c_{j+l_0}|^3.$$

Or $|c_{j+l_0}|^3$ est le terme général d'une série convergente. En effet :

$2(\alpha-1) \leq \alpha \leq 2 < 3$ et $|c_j| < 1 \implies |c_j|^3 < |c_j|^{2(\alpha-1)}$ et d'après l'**Hypothèse 4.3.2**

nous avons $\sum_{j=0}^{\infty} |c_j|^{2(\alpha-1)} < \infty$; ce qui entraîne que $\sum_{j=0}^{\infty} |c_j|^3 < \infty$.

Par suite $U(r; \psi)$ est deux fois dérivable par rapport à α et nous avons :

$$\begin{aligned}
\left| \frac{\partial^2 U(r; \psi)}{\partial \alpha \partial \alpha'} \right| &= \left| \frac{\partial^2}{\partial \alpha \partial \alpha'} S(r; \psi) + \frac{\partial}{\partial \alpha'} \sum_{l=2}^{m+1} \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^{\alpha'} \right| \\
&\leq M (m+1)^3 H_3(r) + \sum_{l=2}^{m+1} \left[\left| \sum_{h=0}^{m+1-l} r_{l+h} c_h \right|^{\alpha'+2} \right] \\
&\leq M (m+1)^3 H_3(r) + \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} |r_{l+h} c_h| \right)^{\alpha'+2} \\
&\leq M (m+1)^3 H_3(r) + \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} |r_{l+h}| \right)^{\alpha'+2}, \quad \text{car } |c_h| < 1 \\
&\leq M (m+1)^3 H_3(r) + \sum_{l=2}^{m+1} H_3(r; l),
\end{aligned}$$

où $H_3(r; l)$ est donné par l'expression suivante :

$$H_3(r; l) = \begin{cases} \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^4, & \text{si } \sum_{h=0}^{m+1-l} |r_{h+l}| > 1 \\ \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^3, & \text{si } \sum_{h=0}^{m+1-l} |r_{h+l}| \leq 1. \end{cases}$$

En posant $\Gamma_3(r) = M (m+1)^3 H_3(r) + \sum_{l=2}^{m+1} H_3(r; l)$, nous obtenons :

$$\left| \frac{\partial^2}{\partial \alpha \partial \alpha'} f(r; \psi) \right| \leq [6 \Gamma_2^2(r) + 4 \Gamma_3(r)] g(r)$$

et d'après l'**Hypothèse 4.3.3**, nous avons :

$$\int \dots \int [6 \Gamma_2^2(r) + 4 \Gamma_3(r)] g(r) dr < \infty.$$

Ce qui implique que $I_n(\psi)$ est de classe \mathcal{C}^2 par rapport à α .

▷ **Montrons que $I_n(\psi)$ est de classe \mathcal{C}^2 par rapport à $\delta = (d, D, \zeta, \xi)$.**

$$\left| \frac{\partial^2}{\partial \delta \partial \delta'} f(r; \psi) \right| \leq 6 g(r) \left| \frac{\partial U(r; \psi)}{\partial \delta} \right| \left| \frac{\partial U(r; \psi)}{\partial \delta'} \right| + 4 g(r) \left| \frac{\partial^2 U(r; \psi)}{\partial \delta \partial \delta'} \right|$$

D'après ce qui précède, nous avons : $\left| \frac{\partial U(r; \psi)}{\partial \delta} \right| \leq \Gamma_1(r)$, pour tout δ .

$$\implies \left| \frac{\partial^2}{\partial \delta \partial \delta'} f(r; \psi) \right| \leq 6 g(r) \Gamma_1^2(r) + 4 g(r) \left| \frac{\partial^2 U(r; \psi)}{\partial \delta \partial \delta'} \right|$$

Or nous avons :

$$\begin{aligned} \left| \frac{\partial^2 u_j(r; \psi)}{\partial \delta \partial \delta'} \right| &= \alpha(\alpha - 1) \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha-2} \left| \sum_{l=0}^m r_{l+1} \frac{\partial c_{j+l}}{\partial \delta} \right| \left| \sum_{l=0}^m r_{l+1} \frac{\partial c_{j+l}}{\partial \delta'} \right| \\ &+ \alpha \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha-1} \left| \sum_{l=0}^m r_{l+1} \frac{\partial^2 c_{j+l}}{\partial \delta \partial \delta'} \right|. \end{aligned}$$

Posons :

$$\begin{aligned} A_1 &= \alpha(\alpha - 1) \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha-2} \left| \sum_{l=0}^m r_{l+1} \frac{\partial c_{j+l}}{\partial \delta} \right| \left| \sum_{l=0}^m r_{l+1} \frac{\partial c_{j+l}}{\partial \delta'} \right|, \\ A_2 &= \alpha \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha-1} \left| \sum_{l=0}^m r_{l+1} \frac{\partial^2 c_{j+l}}{\partial \delta \partial \delta'} \right|. \end{aligned}$$

$$\text{Posons } a_j = \frac{\partial c_j}{\partial \delta}, \quad b_j = \frac{\partial^2 c_j}{\partial \delta \partial \delta'}, \quad \Omega_j = \frac{\partial}{\partial \delta'} \log \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|.$$

En utilisant les calculs précédents, nous avons :

$$\begin{aligned} A_1 &= \alpha(\alpha - 1) \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha-1} \left| \sum_{l=0}^m r_{l+1} \frac{\partial c_{j+l}}{\partial \delta} \right| \times \frac{\partial}{\partial \delta'} \log \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right| \\ &= \alpha(\alpha - 1) \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha-1} \left| \sum_{l=0}^m r_{l+1} a_{j+l} \right| \times \Omega_j \\ &\leq 2(m+1) \left(\sum_{l=0}^m r_{l+1} \right)^\alpha |c_{j+l_0}|^{\alpha-1} |a_{j+l_1}| \Omega_j \\ &\leq (m+1) \left(\sum_{l=0}^m r_{l+1} \right)^\alpha \left[|c_{j+l_0}|^{2(\alpha-1)} + \Omega_j^2 \right] |a_{j+l_1}| \\ &\leq \frac{(m+1)}{2} \left(\sum_{l=0}^m r_{l+1} \right)^\alpha \left[|c_{j+l_0}|^{4(\alpha-1)} + \Omega_j^4 + |a_{j+l_1}|^2 \right]. \end{aligned}$$

De manière analogue, nous obtenons :

$$\begin{aligned} A_2 &= \alpha \left| \sum_{l=0}^m r_{l+1} c_{j+l} \right|^{\alpha-1} \left| \sum_{l=0}^m r_{l+1} b_{j+l} \right| \\ &\leq (m+1) \left(\sum_{l=0}^m r_{l+1} \right)^\alpha \left[|c_{j+l_2}|^{2(\alpha-1)} + |b_{j+l_3}|^2 \right]. \end{aligned}$$

Par suite, nous avons :

$$\left| \frac{\partial^2 u_j(r; \psi)}{\partial \delta \partial \delta'} \right| \leq \frac{(m+1)}{2} \left(\sum_{l=0}^m r_{l+1} \right)^\alpha W_j,$$

avec $W_j = |c_{j+l_0}|^{4(\alpha-1)} + \Omega_j^4 + |a_{j+l_1}|^2 + 2|c_{j+l_2}|^{2(\alpha-1)} + 2|b_{j+l_3}|^2$. En remarquant que $|c_{j+l_0}|^{4(\alpha-1)} < |c_{j+l_0}|^{2(\alpha-1)}$ et en utilisant l'**Hypothèse 4.3.2**, nous pouvons conclure que W_j est le terme général d'une série convergente. Ce qui entraîne que $U(r, \psi)$ est deux fois dérivables par rapport à δ et nous avons :

$$\begin{aligned} \left| \frac{\partial^2 U(r; \psi)}{\partial \delta \partial \delta'} \right| &= \frac{\partial^2 S(r; \psi)}{\partial \delta \partial \delta'} + \alpha \frac{\partial}{\partial \delta'} \sum_{l=2}^{m+1} \left| \sum_{h=0}^{m+1-l} r_{h+l} c_h \right|^{\alpha-1} \left| \sum_{h=0}^{m+1-l} r_{h+l} \frac{\partial c_h}{\partial \delta} \right| \\ &\leq \sum_{j=0}^{\infty} \left| \frac{\partial^2 u_j(r; \psi)}{\partial \delta \partial \delta'} \right| + \alpha \sum_{l=2}^{m+1} \left[\left| \sum_{h=0}^{m+1-l} r_{l+h} c_h \right|^{\alpha-1} \left| \sum_{h=0}^{m+1-l} r_{l+h} \frac{\partial^2 c_h}{\partial \delta \partial \delta'} \right| \right] \\ &+ \alpha(\alpha-1) \sum_{l=2}^{m+1} \left[\left| \sum_{h=0}^{m+1-l} r_{l+h} c_h \right|^{\alpha-2} \left| \sum_{h=0}^{m+1-l} r_{l+h} \frac{\partial c_h}{\partial \delta} \right| \left| \sum_{h=0}^{m+1-l} r_{l+h} \frac{\partial c_h}{\partial \delta'} \right| \right] \\ &\leq \frac{M(m+1)}{2} \left(\sum_{l=0}^m r_{l+1} \right)^\alpha + B_1 + B_2, \end{aligned}$$

$$\text{où } B_1 = \alpha \sum_{l=2}^{m+1} \left[\left| \sum_{h=0}^{m+1-l} r_{l+h} c_h \right|^{\alpha-1} \left| \sum_{h=0}^{m+1-l} r_{l+h} \frac{\partial^2 c_h}{\partial \delta \partial \delta'} \right| \right]$$

$$B_2 = \alpha(\alpha-1) \sum_{l=2}^{m+1} \left[\left| \sum_{h=0}^{m+1-l} r_{l+h} c_h \right|^{\alpha-2} \left| \sum_{h=0}^{m+1-l} r_{l+h} \frac{\partial c_h}{\partial \delta} \right| \left| \sum_{h=0}^{m+1-l} r_{l+h} \frac{\partial c_h}{\partial \delta'} \right| \right].$$

En s'inspirant du calcul de l'expression (B.2), nous obtenons :

$$\begin{aligned} B_1 &= \alpha \sum_{l=2}^{m+1} \left[\left| \sum_{h=0}^{m+1-l} r_{l+h} c_h \right|^{\alpha-1} \left| \sum_{h=0}^{m+1-l} r_{l+h} b_h \right| \right] \\ &\leq M_1 \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^\alpha, \quad \text{car } \sum_{h=0}^{+\infty} |b_h|^2 < \infty. \end{aligned}$$

Pour la majoration de B_2 , nous avons :

$$B_2 = \alpha(\alpha - 1) \sum_{l=2}^{m+1} \left[\left| \sum_{h=0}^{m+1-l} r_{l+h} c_h \right|^{\alpha-1} \left| \sum_{h=0}^{m+1-l} r_{l+h} a_h \right| \frac{\partial}{\partial \delta'} \log \left| \sum_{h=0}^{m+1-l} r_{l+h} c_h \right| \right],$$

or $\frac{\partial}{\partial \delta'} \log \left| \sum_{h=0}^{m+1-l} r_{l+h} c_h \right|$ est borné, car étant le terme général d'une série convergente. En utilisant le même calcul qu'en (B.2), nous obtenons :

$$B_2 \leq M_2 \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^\alpha.$$

Ainsi, nous pouvons écrire finalement que :

$$\begin{aligned} \Rightarrow \left| \frac{\partial^2 U(r; \psi)}{\partial \delta \partial \delta'} \right| &\leq \frac{M(m+1)}{2} \left(\sum_{l=0}^m r_{l+1} \right)^\alpha + (M_1 + M_2) \sum_{l=2}^{m+1} \left(\sum_{h=0}^{m+1-l} |r_{h+l}| \right)^\alpha \\ &\leq \frac{M(m+1)}{2} H_1(r) + (M_1 + M_2) \sum_{l=2}^{m+1} H_1(r; l) = \Gamma_1(r). \end{aligned}$$

D'où $\left| \frac{\partial^2}{\partial \delta \partial \delta'} f(r; \psi) \right| \leq 6 g(r) \Gamma_1^2(r) + 4 g(r) \Gamma_1(r)$.

Sous l' **Hypothèse 4.3.3**, nous avons $\int \dots \int (6\Gamma_1^2(r) + 4\Gamma_1(r)) g(r) dr < \infty$. Ainsi, $I_n(\psi)$ est de classe \mathcal{C}^2 par rapport à δ .

D'où $I_n(\psi)$ est deux fois continûment dérivable sous le signe intégral, par rapport à $\psi \in \Psi$.

(A3) : La suite $(X_t)_{t \in \mathbb{Z}}$ est strictement stationnaire et ergodique.

D'après le Théorème 2.3.1, $(X_t)_{t \in \mathbb{Z}}$ est strictement stationnaire et admet une unique représentation moyenne mobile infinie donnée par :

$$X_t = \sum_{j=0}^{\infty} c_j Z_{t-j}.$$

Or les variables aléatoires $(Z_t)_t$ étant indépendantes et identiquement distribuées et $\sum_{j=0}^{\infty} |c_j|^\alpha < \infty$ pour tout $1 < \alpha \leq 2$, donc $(X_t)_{t \in \mathbb{Z}}$ est un processus ergodique. D'où $(X_t)_{t \in \mathbb{Z}}$ est un processus strictement stationnaire et ergodique.

(A5) : $K(x; \psi)$ est une fonction mesurable et bornée de x ,

où $K(x; \psi) = \int \dots \int (\cos(r'x) - c(r; \psi)) \frac{\partial}{\partial \psi} c(r; \psi) g(r) dr$.

Nous avons :
$$\left| (\cos(r'x) - c(r; \psi)) \frac{\partial}{\partial \psi} c(r; \psi) g(r) \right| \leq 2g(r) \left| \frac{\partial}{\partial \psi} c(r; \psi) \right|$$

$$\leq 2g(r) \left| \frac{\partial}{\partial \psi} U(r; \psi) \right|.$$

Or
$$\left| \frac{\partial}{\partial \psi} U(r; \psi) \right| \leq \Gamma(r) = \begin{cases} \Gamma_1(r), & \text{si } \psi = (d, D, \zeta, \xi) \\ \Gamma_2(r), & \text{si } \psi = \alpha. \end{cases}$$

Ainsi, nous avons :

$$\left| (\cos(r'x) - c(r; \psi)) \frac{\partial}{\partial \psi} c(r; \psi) g(r) \right| \leq 2\Gamma(r)g(r), \quad \forall \psi \in \Psi.$$

En utilisant l'**Hypothèse 4.3.3**, nous obtenons $\int \dots \int \Gamma(r)g(r) dr < \infty$. D'où $K(x; \psi)$ est continue et par suite $K(x; \psi)$ est une fonction mesurable de x et bornée.

(A6) : $B(\psi_0) = \int \dots \int \frac{\partial}{\partial \psi} c(r; \psi_0) \frac{\partial}{\partial \psi'} c(r; \psi_0) g(r) dr$ est non singulière et $\frac{\partial^2}{\partial \psi \partial \psi'} c(r; \psi)$ est uniformément bornée par une fonction intégrable.

Nous avons :

$$\left| \frac{\partial}{\partial \psi} c(r; \psi_0) \frac{\partial}{\partial \psi'} c(r; \psi_0) \right| \leq \left| \frac{\partial}{\partial \psi} U(r; \psi_0) \frac{\partial}{\partial \psi'} U(r; \psi_0) \right|$$

$$\leq \Gamma^2(r)$$

et d'après l'**Hypothèse 4.3.3**, nous avons $\int \dots \int \Gamma^2(r)g(r) dr < \infty$.

$$\text{D'où } \int \dots \int \left| \frac{\partial}{\partial \psi} c(r; \psi_0) \frac{\partial}{\partial \psi'} c(r; \psi_0) \right| (r)g(r) dr < \infty.$$

Par conséquent $B(\psi_0)$ est non singulière. Pour montrer que $\frac{\partial^2}{\partial \psi \partial \psi'} c(r; \psi)$ est uniformément bornée, nous remarquons d'abord que :

$$\left| \frac{\partial^2 c(r; \psi)}{\partial \psi \partial \psi'} \right| \leq \left| \frac{\partial U(r; \psi)}{\partial \psi} \right| \left| \frac{\partial U(r; \psi)}{\partial \psi'} \right| + \left| \frac{\partial^2 U(r; \psi)}{\partial \psi \partial \psi'} \right|$$

$$\leq \Gamma^2(r) + \left| \frac{\partial^2 U(r; \psi)}{\partial \psi \partial \psi'} \right|.$$

Or d'après ce qui précède, nous avons :

$$\left| \frac{\partial^2 U(r; \psi)}{\partial \psi \partial \psi'} \right| < \Gamma'(r) = \begin{cases} \Gamma_1(r), & \text{si } \psi = (d, D, \zeta, \xi) \\ \Gamma_3(r), & \text{si } \psi = \alpha. \end{cases}$$

Ce qui implique que :

$$\left| \frac{\partial^2 c(r; \psi)}{\partial \psi \partial \psi'} \right| \leq \Gamma^2(r) + \Gamma'(r)$$

et sous l'**Hypothèse 4.3.3**, nous avons :

$$\int \dots \int (\Gamma^2(r) + \Gamma'(r)) g(r) dr < \infty.$$

D'où $\left| \frac{\partial^2 c(r; \psi)}{\partial \psi \partial \psi'} \right|$ est uniformément bornée par une fonction intégrable.

Ainsi, toutes les hypothèses sont vérifiées, nous pouvons appliquer le Théorème 2.1 de Knight et Yu [43]. Ce qui prouve le théorème. ■

Comparaison des méthodes d'estimation

Dans cette partie, nous présentons les résultats de simulation de Monte Carlo permettant de comparer les méthodes d'estimation des paramètres d'un processus ARFISMA symétrique α -stable décrites dans le Chapitre 3.

n	GPH_T	GPH_P			Whittle	MCMC
	\hat{d}	\hat{d} $\nu = 0.5$	\hat{d} $\nu = 0.6$	\hat{d} $\nu = 0.7$	\hat{d}	\hat{d}
100	0.2080 (0.1142) [0.0791]	0.2175 (0.2575) [0.1901]	0.2184 (0.1980) [0.1454]	0.2142 (0.1544) [0.1106]	0.1709 (0.0903) [0.0664]	0.2260 (0.1086) [0.0700]
500	0.2021 (0.0469) [0.0305]	0.2154 (0.1601) [0.1164]	0.2103 (0.1112) [0.0803]	0.2049 (0.0790) [0.0546]	0.1949 (0.0389) [0.0241]	0.2099 (0.0519) [0.0298]
1500	0.2012 (0.0274) [0.0157]	0.2047 (0.1169) [0.0836]	0.2040 (0.0756) [0.0527]	0.2029 (0.0507) [0.0336]	0.1991 (0.0234) [0.0121]	0.2041 (0.0315) [0.0169]

TAB. C.1 – Paramètres estimés du modèle 1, pour $\alpha = 1.3$.

n	GPH_T	GPH_P			Whittle	MCMC
	\hat{d}	\hat{d} $\nu = 0.5$	\hat{d} $\nu = 0.6$	\hat{d} $\nu = 0.7$	\hat{d}	\hat{d}
100	0.2015 (0.1170) [0.0904]	0.2042 (0.2816) [0.2207]	0.2034 (0.2239) [0.1762]	0.2018 (0.1601) [0.1250]	0.1603 (0.0980) [0.0794]	0.1902 (0.0821) [0.0658]
500	0.1989 (0.0471) [0.0376]	0.2082 (0.1727) [0.1370]	0.2027 (0.1180) [0.0940]	0.1986 (0.0824) [0.0648]	0.1907 (0.0377) [0.0298]	0.1985 (0.0369) [0.0293]
1500	0.2003 (0.0264) [0.0211]	0.1978 (0.1243) [0.0976]	0.1977 (0.0817) [0.0646]	0.1996 (0.0549) [0.0441]	0.1971 (0.0209) [0.0167]	0.1996 (0.0230) [0.0181]

TAB. C.2 – Paramètres estimés du modèle 1, pour $\alpha = 2$.

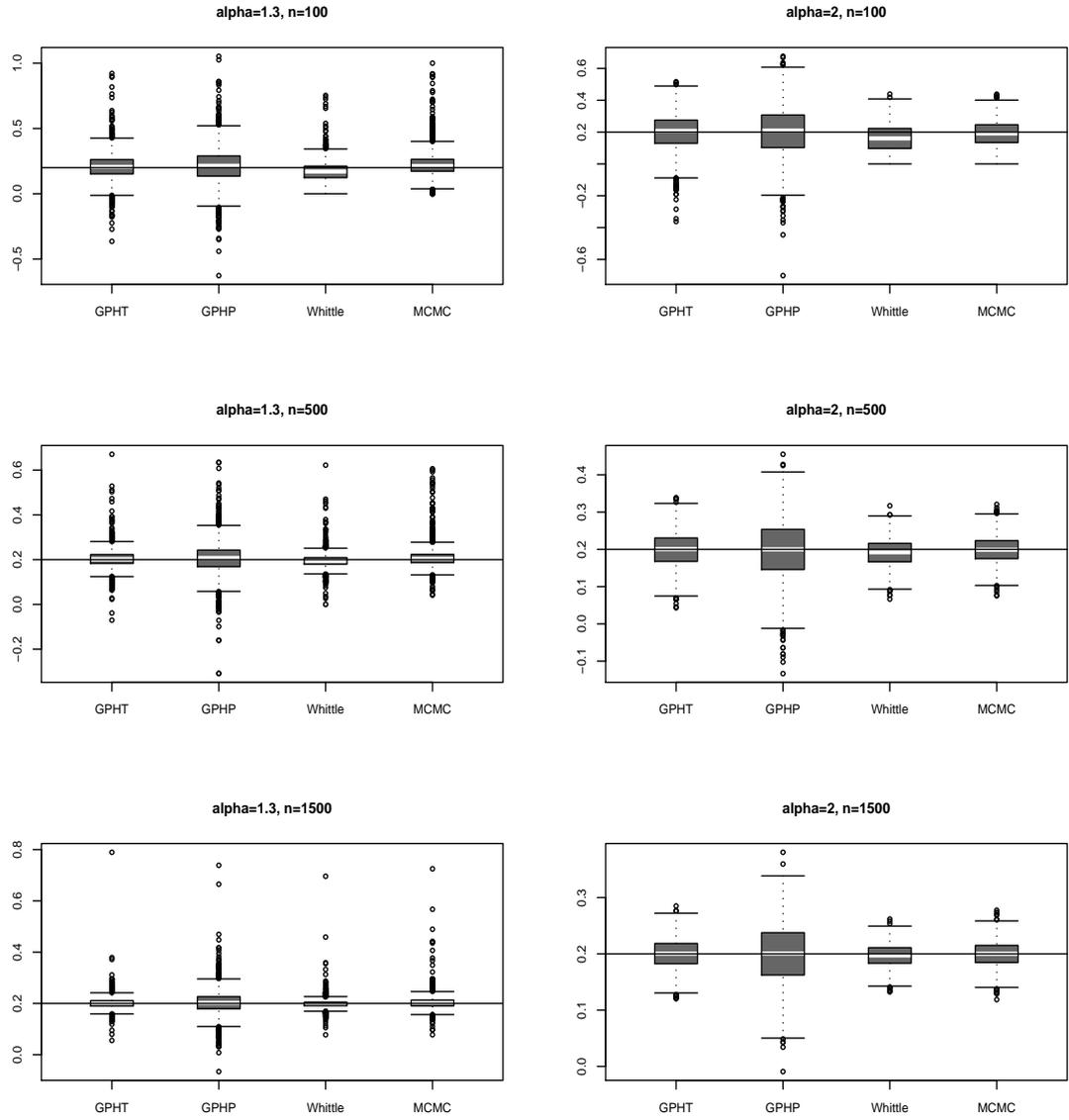


FIG. C.1 – Boxplot des paramètres estimés d du modèle 1, avec $\alpha = 1.3$ et $\alpha = 2$ par les méthodes $GPHT$, GPH , Whittle classique et Whittle par MCMC, pour $n = 100$, $n = 500$ et $n = 1500$, basé sur 1500 répliques.

n	GPH_T \hat{D}	GPH_P \hat{D}	Whittle \hat{D}	MCMC \hat{D}
100	0.1999 (0.1303) [0.0948]	0.1987 (0.1553) [0.1139]	0.0800 (0.1311) [0.1243]	0.1761 (0.0743) [0.0537]
500	0.2011 (0.0431) [0.0305]	0.2006 (0.0500) [0.0370]	0.1710 (0.0434) [0.0351]	0.1967 (0.0325) [0.0229]
1500	0.2004 (0.0251) [0.01564]	0.20001 (0.0281) [0.0187]	0.1895 (0.0210) [0.0147]	0.1988 (0.0213) [0.0140]

TAB. C.3 – Paramètres estimés du modèle 2, pour $\alpha = 1.3$.

n	GPH_T \hat{D}	GPH_P \hat{D}	Whittle \hat{D}	MCMC \hat{D}
100	0.1967 (0.1551) [0.1236]	0.1928 (0.1887) [0.1471]	0.0864 (0.1312) [0.1208]	0.1751 (0.0872) [0.0692]
500	0.1996 (0.0519) [0.0415]	0.1986 (0.0594) [0.0476]	0.1679 (0.0518) [0.0412]	0.1946 (0.0386) [0.0307]
1500	0.1999 (0.0268) [0.0214]	0.2003 (0.0311) [0.0248]	0.1885 (0.0243) [0.0194]	0.1979 (0.0235) [0.0190]

TAB. C.4 – Paramètres estimés du modèle 2, pour $\alpha = 2$.

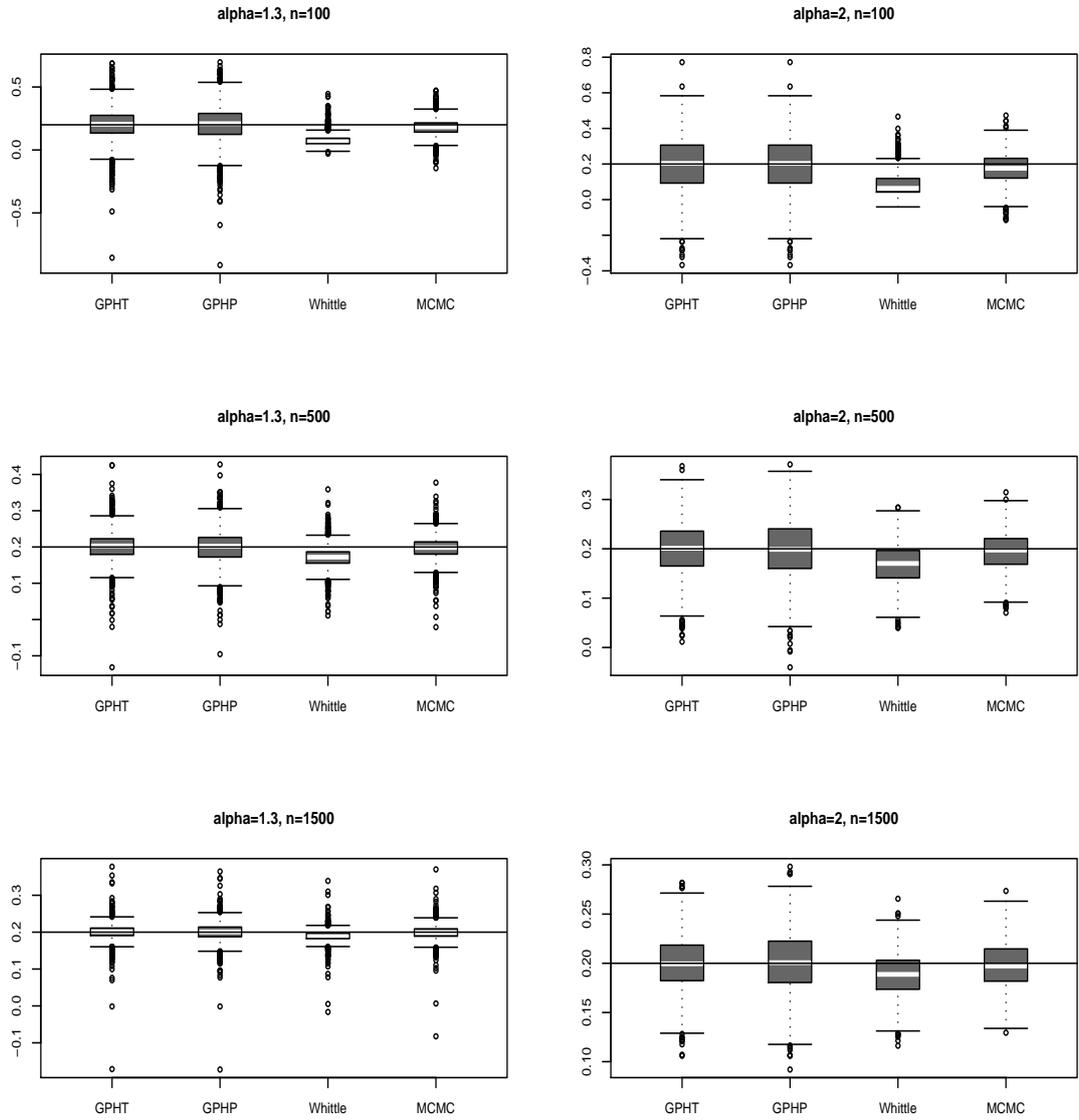


FIG. C.2 – Boxplot des paramètres estimés de D du modèle 2, avec $\alpha = 1.3$ et $\alpha = 2$ par les méthodes $GPHT$, $GPHP$, Whittle classique et Whittle par MCMC, pour $n = 100$, $n = 500$ et $n = 1500$, basé sur 1500 répliques.

n	GPH_T		GPH_P		Whittle		MCMC	
	\hat{d}							
100	0.1004 (0.1030) [0.0739]	0.1190 (0.1309) [0.0948]	0.1002 (0.1066) [0.0766]	0.1181 (0.1589) [0.1157]	0.0766 (0.0835) [0.0604]	0.0461 (0.0812) [0.0779]	0.0926 (0.0723) [0.0518]	0.0969 (0.0738) [0.0529]
500	0.1002 (0.0414) [0.0282]	0.1208 (0.0432) [0.0305]	0.0965 (0.0738) [0.0590]	0.1197 (0.0930) [0.0744]	0.0979 (0.0324) [0.0218]	0.0918 (0.0410) [0.0345]	0.0983 (0.0319) [0.0221]	0.1165 (0.0321) [0.0231]
1500	0.1002 (0.0247) [0.0151]	0.1203 (0.0262) [0.0162]	0.1003 (0.0256) [0.0157]	0.1199 (0.0303) [0.0204]	0.0994 (0.0198) [0.0113]	0.1093 (0.0211) [0.0153]	0.0993 (0.0216) [0.0140]	0.1187 (0.0214) [0.0141]

TAB. C.5 – Paramètres estimés du modèle 3, pour $\alpha = 1.3$.

n	GPH_T		GPH_P		Whittle		MCMC	
	\hat{d}							
100	0.0990 (0.1238) [0.0944]	0.1168 (0.1569) [0.1204]	0.0999 (0.1275) [0.1011]	0.1124 (0.1950) [0.1520]	0.0688 (0.1067) [0.0834]	0.0497 (0.0824) [0.0774]	0.0875 (0.0923) [0.0730]	0.0949 (0.0882) [0.0701]
500	0.0980 (0.0498) [0.0399]	0.1197 (0.0542) [0.0433]	0.0979 (0.0513) [0.0410]	0.1191 (0.0635) [0.0508]	0.0957 (0.0401) [0.0320]	0.0903 (0.0473) [0.0392]	0.0966 (0.0396) [0.0319]	0.1139 (0.0396) [0.0313]
1500	0.0995 (0.0272) [0.0217]	0.1197 (0.0275) [0.0220]	0.0993 (0.0281) [0.0223]	0.1203 (0.0327) [0.0261]	0.0988 (0.0214) [0.0171]	0.1082 (0.0246) [0.0197]	0.0988 (0.0240) [0.0190]	0.1175 (0.0236) [0.0190]

TAB. C.6 – Paramètres estimés du modèle 3, pour $\alpha = 2$.

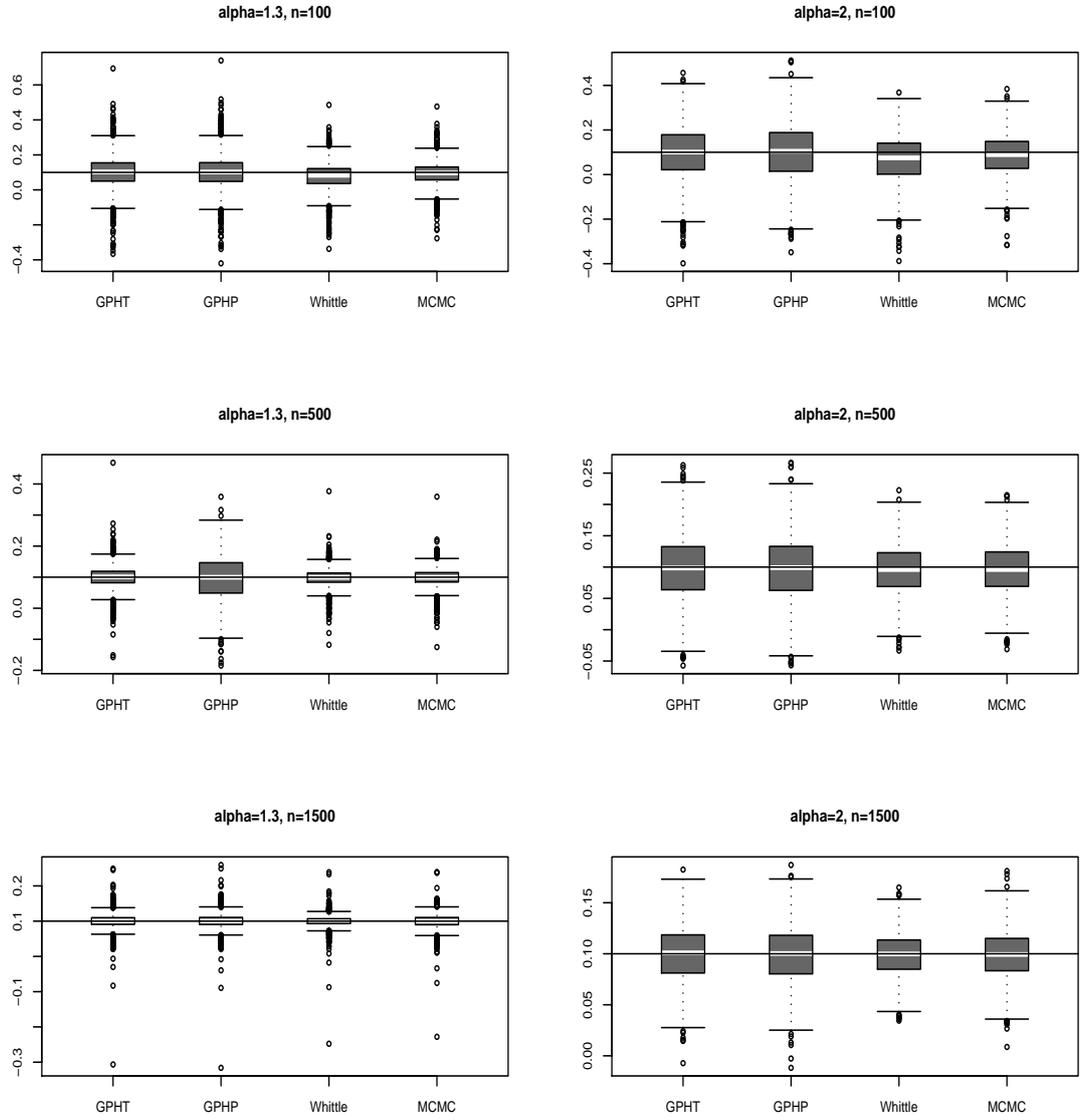


FIG. C.3 – Boxplot des paramètres estimés de d du modèle 3, avec $\alpha = 1.3$ et $\alpha = 2$ par les méthodes $GPHT$, GPH , Whittle classique et Whittle par MCMC, pour $n = 100$, $n = 500$ et $n = 1500$, basé sur 1500 répliques.

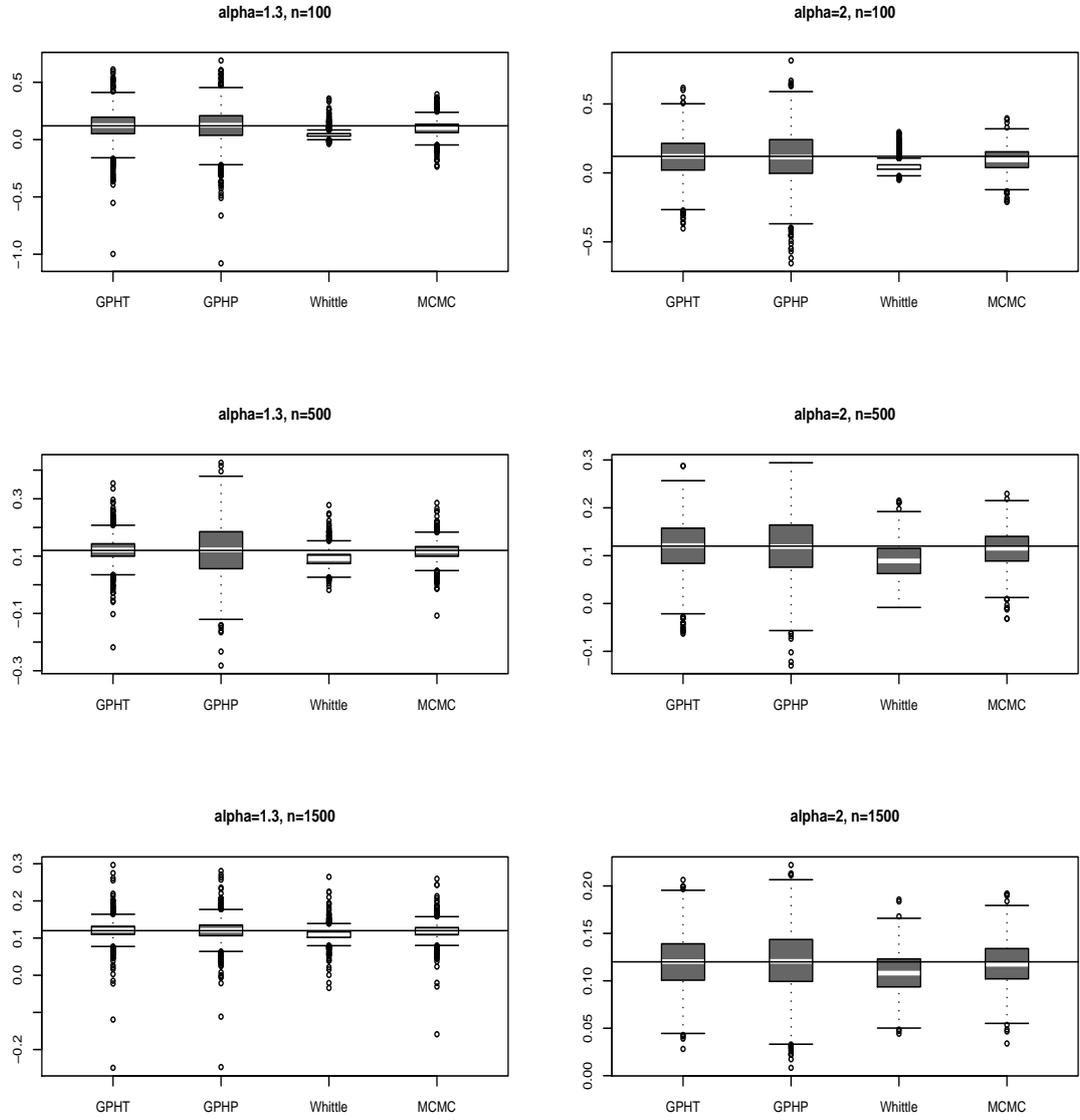


FIG. C.4 – Boxplot des paramètres estimés de D du modèle 3, avec $\alpha = 1.3$ et $\alpha = 2$ par les méthodes $GPHT$, GPH , Whittle classique et Whittle par MCMC, pour $n = 100$, $n = 500$ et $n = 1500$, basé sur 1500 répliques.