



THÈSE

Présentée en vue de l'obtention du diplôme de Doctorat

Spécialité Actuariat

Estimation dans les modèles Autorégressifs Exponentiels Périodiques

Par

BECILA Sabah

Sous la direction de

MERZOUGUI Mouna

MCA. U.B.M. ANNABA

Devant le jury

PRESIDENT: BOUTABIA Hacène

Prof. U.B.M. ANNABA

EXAMINATEUR: ZEGHDOUDI Halim

Prof. U.B.M. ANNABA

EXAMINATEUR: CHADLI Assia

Prof. U.B.M. ANNABA

EXAMINATEUR: EZZEBSA Abdelali

MCA. U. GUELMA

Année 2021

TABLE DES MATIÈRES

Table des figures	iv
Liste des tableaux	vi
Liste des abréviations et notations	vii
Résumé en Arabe	xi
Abstract	xii
Résumé en Français	xiii
Remerciements	xiv
Dédicaces	xvi
Introduction générale	xx
1 Théories des MC non linéaires et du QMV	1
1.1 Introduction	1
1.2 Eléments de base de la théorie asymptotique	3

Table des matières

1.2.1 Fonctions aléatoires	3
1.2.2 Développement de Taylor en probabilité	4
1.2.3 Lois des Grands Nombres et théorème centrale limite pour les processus dépendants	4
1.3 Théorie des moindres carrés non linéaires (MCN)	9
1.3.1 Modèles de régression non linéaire et estimateurs des MCN	10
1.3.2 Application pour des modèles de ST non linéaires	13
1.4 Théorie du quasi-maximum de vraisemblance	17
1.4.1 Critère d'information Kullback-Leibler	17
1.4.2 Propriétés asymptotiques de l'EQMV	21
2 Estimation du modèle PEXPAR(p) restreint par la méthode des MCC linéaires	24
2.1 Introduction	24
2.2 Modèle <i>EXPAR</i> restreint périodique	26
2.3 Estimation des paramètres	35
2.4 Résultats de la simulation	40
3 Estimation par la méthode des MCC non linéaires du modèle <i>PEXPAR</i>(1)	43
3.1 Introduction	43
3.2 Définition de l'estimateur des MCN pour le modèle <i>PEXPAR_S</i> (1)	45
3.3 Comportement asymptotique de l'estimateur des MCN	50
3.3.1 Consistance forte de l'estimateur des MCN	50
3.3.2 Distribution asymptotique de l'estimateur MCN	53
3.4 Résultats de la simulation	56

4 Estimation par la méthode du QMV du modèle $PEXPAR(1)$	62
4.1 Introduction	62
4.2 Estimateur QMV pour le modèle $PEXPAR_S(1)$	63
4.3 Comportement asymptotique de l'estimateur QMV	66
4.3.1 Consistance de l'estimateur QMV	66
4.3.2 Distribution asymptotique de l'EQMV	68
Conclusion et perspectives	72
Annexe : Théorie de mesure de la théorie de probabilité	74
Bibliographie	77

TABLE DES FIGURES

2.1 Réalisation du processus (A) avec nuage de points, histogramme	
et corrélogramme correspondants, n=500, valeur p (Shapiro-Wilk	
normality test) = 3.596e-07.	29
2.2 Réalisation du processus (B) avec nuage de points, histogramme	
et corrélogramme correspondants, n=500, valeur p (Shapiro-Wilk	
normality test) = 0.004525.	30
2.3 Réalisation du processus (C) avec nuage de points, histogramme	
et corrélogramme correspondants, n=500, valeur p (Shapiro-Wilk	
normality test) = 3.596e-07.	31
2.4 Cycle limite du modèle $PEXPAR_2(1)$	32
2.5 Simulation du $PEXPAR_2(2)$ restreint et racines inverses de l'équa-	
tion caractéristique.	34
2.6 Monthplot de la série temporelle.	34
2.7 Nuage de points de la série temporelle.	34
2.8 FAC et FACP périodiques de la série temporelle.	34
2.9 Série simulée et racines inverses de l'équation caractéristique du	
$PEXPAR_2(2)$ restreint et $N = 1000$	41

Table des figures

2.10 Boîtes à moustaches d'erreurs d'estimations de 200 répétitions du	
<i>PEXPAR</i> ₂ (2) restreint et $N = 1000$.	42
2.11 Le graphe Q-Q d'erreurs d'estimations de 200 répétitions du	
<i>PEXPAR</i> ₂ (2) restreint et $N = 1000$.	42
3.1 Série simulée à partir du modèle <i>PEXPAR</i> ₂ (1).	46
3.2 Série simulée avec nuage de points, densité marginale et corrélo-	
gramme correspondants.	47
3.3 Boîtes à moustaches d'erreurs.	58
3.4 Le graphe Q-Q d'erreurs.	58
3.5 Comparaison entre les valeurs réelles et leurs valeurs prévues.	60

LISTE DES TABLEAUX

2.1	Résultats d'estimations pour $PEXPAR_2(2)$ restreint.	41
3.1	Résultats d'estimation du $PEXPAR_2(1)$ non restreint	57
3.2	Estimation des paramètres du $PEXPAR_{12}(1)$ non restreint.	60
3.3	Variances résiduelles mensuelles.	61

└────────────────── Liste des abréviations et notations

AR : Modèle AutoRégressif (AutoRegressive model).

PAR : Modèle AutoRégressif Périodique (Periodic AutoRegressive model).

MPAR : Modèle de Mélange AutoRégressif périodique (Mixture Periodic AutoRegressive model).

PEXP(\mathbf{p}) : Modèle AutoRégressif EXPonentiel Periodique d'ordre \mathbf{p} (Periodic EXPonential AutoRegressive model of order \mathbf{p}).

TAR : Modèle AutoRégressif à seuil (Threshold AutoRegressive model).

ARCH : Modèle d'Hétéroscédasticité Conditionnelle AutoRégressive (AutoRegressive Conditional Heteroskedasticity model).

GARCH : Modèle d'Hétéroscédasticité Conditionnelle AutoRégressive Généralisée (Generalized AutoRegressive Conditional Heteroskedasticity model).

ARMA : Modèle Moyenne Mobile AutoRégressif (AutoRegressive Moving Average model).

PARMA : Modèle Moyenne Mobile AutoRégressif Périodique (Periodic AutoRegressive Moving Average model).

SARIMA : Modèle Moyenne Mobile Intégrée Saisonnière AutoRégressif

sive (Seasonal Autoregressive Integrated Moving Average model).

SMASTAR : Modèle AutoRégressif à seuil Spline Adaptative Semi-Multivariée (Semi-Multivariate Adaptive Spline Threshold AutoRegressive).

MC : Moindres Carrés conditionnels (**LS** : conditional Least Squares).

MCL : Moindres Carrés Linéaires (**LLS** : Linear Least Squares).

MCN : Moindres Carrés Non linéaires (**NLS** : Nonlinear Least Squares).

EMC : Estimateur des Moindres Carrés (**LSE** : Least Squares Estimator).

EMCL : Estimateur des Moindres Carrés Linéaires (**LLSE** : Linear Least Squares Estimator).

EMCN : Estimateur Moindres Carrés Non linéaires (**NLSE** : Nonlinear Least Squares Estimator).

MV : Maximum de Vraisemblance (**ML** : Maximum Likelihood).

QMV : Quasi Maximum de Vraisemblance (**QML** : Quasi Maximum Likelihood).

EMV : Estimateur du Maximum de Vraisemblance (**MLE** : Maximum Likelihood Estimator).

EQMV : Estimateur du Quasi Maximum de Vraisemblance (**QMLE** : Quasi Maximum Likelihood Estimator).

TCL : Théorème Centrale Limite (**CLT** : Central Limit Theorem).

LGNU : Loi des Grands Nombres Uniforme (**ULLN** : Uniform Law of Large Numbers).

V.A : Variable Aléatoire (**R.V** Random Variable).

I.I.D : Indépendants Identiquement Distribués (Independent Identically Distributed).

ST : Série Temporelle (**TS** : Time Series).

DM : Différence de Martingale (**MD** : Martingale Difference).

- DT** : Développement de Taylor (**TD** : Taylor Development).
- N.A** : Normalité Asymptotique (**A.N** : Asymptotic Normality).
- FAC** : Fonction d'AutoCorrélation (**ACF** : AutoCorrelation Function).
- FACP** : Fonction d'AutoCorrélation Partielle (**PACF** : Partial AutoCorrelation Function).
- CV** : ConVergence ou ConVerge (ConVergence or ConVerges).
- p.s** : presque sûrement (**a.s** : almost surely).
- p.p** : presque partout (**a.e** : almost everywhere).
- AIC** : Critère d'Information d'Akaike (Akaike Informarion Criterion).
- CIKL** : Critère d'Information Kullback-Leibler (**KLIC** : Kullback-Leibler Information Criterion).
- PGD** : Processus de Génération de Données (**DGP** : Data Generation Process).
- FKE** : Filtre de Kalman Etendu (**EKF** : Extended Kalman Filter).
- nlm** : minimisation non linéaire (nonlinear minimisation).
- \mathcal{B}^∞ : La σ -algèbre de Borel de \mathbb{R}^∞ (The Borel σ -algebra of \mathbb{R}^∞).
- B** : Les ensembles boréliens $\in \mathcal{B}^\infty$ de \mathbb{R}^∞ (The Borel sets $\in \mathcal{B}^\infty$ of \mathbb{R}^∞).
- $\mathbf{o}_p(1)$: Convergence en Probabilité (convergence in Probability).
- $\mathbf{O}_p(1)$: Borné en Probabilité (bounded in Probability).
- ∇^2 : La matrice des dérivées partielles secondes (The matrix of the second partial derivatives).
- H_n**: La matrice Hesienne limite (The limit Hessian matrix).
- $\mathbb{I}(\mathbf{g} : \mathbf{f})$: (**CIKL**) de **g** par rapport à **f** ((**KLIC**) of **g** with respect to **f**).
- $\underline{\varphi}$: Les paramètres du modèle **PEXP**AR restreint (The parameters of the restricted **PEXP**AR model).
- $L_N(\underline{\varphi})$: La somme conditionnelle des carrés (The conditional sum of squares).

Liste des abréviations et notations

Les titres des références, les noms des savants, les termes techniques ou les mots que nous voulons souligner sont indiqués en gras.

في هذه الأطروحة، نقوم بتوسيع فئة نموذج الانحدار الذاتي الأسّي الدوري المقيد (1) PEXPAR إلى الرتبة p . هدفنا هو بناء مقدرات لنموذج (p) EXPAR المقيد الدوري بواسطة طريقة المربعات الصغرى المشروطة الخطية (MCC) ونموذج (1) EXPAR الدوري غير المقيد بواسطة طريقة المربعات الصغرى المشروطة غير الخطية وطريقة شبه المعقولية العظمى. بعد ذلك، ندرس خصائص المقاربة للمقدرات وكذلك نقوم بعمليات محاكاتها لإثبات أدائها.

الكلمات المفتاحية: نموذج EXPAR الدوري، طريقة MC الخطية وغير الخطية، طريقة QMV ، إرجوديك.

In this thesis, we extend the class of the restricted periodic exponential autoregressive model (**PEXP****AR**(**1**)) to the order **p**. Our objective is the construction of the estimators of the periodically restricted **EXPAR**(**p**) model by the linear least squares (**MC**) method and of the unrestricted periodic **EXPAR**(**1**) model by the nonlinear least squares method and the quasi maximum likelihood. Then, obtain the asymptotic properties of the estimators and eventually make simulations which prove their performance.

Key words : Periodic **EXPAR** model, Linear and nonlinear **LS** method, **QML** method, Ergodicity.

Dans cette thèse, nous étendons la classe du modèle autorégressif exponentiel périodique restreint (**PEXP****AR**(**1**)) à l'ordre **p**. Notre objectif est la construction des estimateurs du modèle **EXPAR**(**p**) restreint périodique par la méthode des moindres carrés (**MC**) linéaires et du modèle **EXPAR**(**1**) périodique non restreint par la méthode des moindres carrés non linéaires et la méthode du quasi maximum de vraisemblance. Ensuite, obtenir les propriétés asymptotiques des estimateurs et éventuellement faire des simulations qui prouvent leur performance.

Mots clés : Modèle **EXPAR** périodique, Méthode des **MC** linéaire et non linéaire, Méthode du **QMV**, Ergodicité.

Remerciements

Je tiens tout d'abord à remercier le bon Dieu tout puissant de m'avoir octroyé cette force avec laquelle j'ai pu terminer ce travail.

J'adresse mes vifs remerciements à ma directrice de thèse Madame **Mouna MERZOUGUI, Maître de conférence A** à l'Université **Badji Mokhtar d'Annaba** d'avoir accepté de m'encadrer et de me diriger avec beaucoup d'attention pendant les années de la thèse. Sa disponibilité, son soutien et sa patience ont été les clés de la réussite de ce travail. Je lui exprime également toute ma gratitude de m'avoir fait bénéficier de ses larges connaissances scientifiques. Sincèrement, j'ai appris avec elle la rigueur, la précision, le travail minutieux et l'exploitation rationnelle du temps. Ses qualités intellectuelles et humaines font d'elle une femme exemplaire.

J'adresse mes sincères remerciements à Monsieur **Hacène BOUTABIA Professeur** à l'Université **Badji Mokhtar d'Annaba** d'avoir accepté de présider le jury de ma soutenance de thèse et je tiens à lui exprimer toute ma profonde reconnaissance.

Je tiens à remercier chaleureusement Monsieur **Halim ZEGHDOUDI Professeur** à l'Université **Badji Mokhtar d'Annaba** et **directeur du Laboratoire de Probabilités et Statistique d'Annaba (LaPS)** d'avoir

Remerciements

accepté d'être membre de jury de ma soutenance de thèse. Aussi, je le remercie pour sa participation scientifique et pour le temps qu'il a consacré à ma recherche.

Egalement, Je tiens à remercier très chaleureusement Madame **Assia CHADLI Professeur** à l'**Université Badji Mokhtar d'Annaba** d'avoir accepté de participer à ce jury ainsi que pour le temps qu'elle a bien voulu consacré pour la lecture de ma thèse.

Je suis très reconnaissante en la respectueuse personne de Monsieur **Abdelali AZZESBA Maître de conférence A** à l'**Université 8 Mai 1945 de Guelma** pour m'avoir fait l'honneur d'accepter d'être examinateur de ce jury et d'avoir apporté un regard éclairé sur mon travail de thèse.

Mes remerciements vont aussi au **chef de Département de Mathématiques** le **Professeur Mohammed Chérif BOURAS**, au personnel administratif du Département pour leur soutien moral et leur sympathie ainsi qu'à mes chers professeurs qui m'ont éclairé le chemin de la science.

Je voudrais remercier également ma famille et à tous ceux qui m'ont inspiré et soutenu au cours de la réalisation de ce travail de thèse.

Enfin, une pensée particulière à mes chers parents pour leurs soutien permanent, leur compréhension et leurs sacrifices.

Toutes les lettres ne sauront trouver les mots qu'il faut...

Tous les mots ne sauraient exprimer la gratitude,

L'amour, le respect, la reconnaissance...

Aussi, c'est tout simplement que :

Je dédie cette thèse ...

A mon très cher père

Tout l'encre du monde ne pourrait suffire pour exprimer mes sentiments envers un être très cher.

Vous avez toujours été mon école de patience, de confiance et surtout d'espoir et d'amour.

Vous êtes et vous resterez pour moi ma référence, la lumière qui illumine mon chemin.

Ce travail est le résultat de l'esprit de sacrifice dont vous avez fait preuve, de l'encouragement et le soutien que vous ne cessez de manifester, j'espère que vous y trouverez les fruits de votre semence et le témoignage de ma grande fierté de vous avoir comme père.

J'implore Dieu, tout puissant, de vous accorder une bonne santé, une longue vie et beaucoup de bonheur.

A ma très chère mère

Aucune dédicace très chère maman, ne pourrait exprimer la profondeur des sentiments que j'éprouve pour vous, vos sacrifices innombrables et votre dévouement firent pour moi un encouragement.

Vous avez guidé mes pas, et m'avez couvé de tendresse, ta prière et ta bénédiction m'ont été d'un grand secours pour mener à bien mes études.

Vous m'avez aidé et soutenu pendant de nombreuses années avec à chaque fois une attention renouvelée.

Puisse Dieu, tout puissant vous combler de santé, de bonheur et vous procurer une longue vie.

**A mon frère Nabil, ma belle soeur Ghizlaine et mon adorable
neveu Nazim**

En signe de l'affection et du grand amour que je vous porte, les mots sont insuffisants pour exprimer ma profonde estime.

Je vous dédie ce travail en témoignage de ma profonde affection et de mon attachement indéfectible, je n'oublierais jamais vos encouragements et votre soutien.

Que Dieu vous accorde santé, succès et félicité pour faire de vous un couple uni et heureux à jamais.

Que Dieu protège Nazim et le bénisse.

**A mon frère Zidane, ma belle soeur Sabrina, mon adorable nièce
Melissa et mon adorable neveu Kenzi**

En témoignage de mon affection fraternelle, de ma profonde tendresse et reconnaissance, je vous souhaite une vie pleine de bonheur et de succès.

Que Dieu, le tout puissant, vous protège et vous garde ainsi que vos enfants.

A mon frère Sofiane

Ces quelques lignes, ne sauraient traduire le profond amour que je te porte. Je trouve en toi le conseil du frère et le soutien de l'ami.

Que ce travail soit l'expression de mon estime pour toi et que Dieu te protège, t'accorde santé, succès et plein de bonheur dans ta vie.

A ma grande mère

Qui m'a accompagné par ses prières, sa douceur, puisse Dieu lui prêter longue vie parmi nous et beaucoup de santé.

A la mémoire de mes grands pères et ma grande mère

J'aurais tant aimé que vous soyez présents.

Que Dieu ait vos âmes dans sa sainte miséricorde et que ce travail soit une prière pour vos âmes.

A mes tantes et oncles paternels et leurs conjoints

A mes tantes et oncles maternels et leurs conjoints

A mes cousins et cousines

En témoignage de mon attachement et de ma grande considération. J'espère que vous trouverez à travers ce travail l'expression de mes sentiments les plus chaleureux.

Que ce travail vous apporte l'estime, et le respect que je porte à votre égard, et soit la preuve du désir que j'aie depuis toujours pour vous honorer.

Tous mes vœux de bonheur et de santé.

A mes chères amies

En souvenir des moments heureux passés ensemble, avec mes vœux sincères de réussite, bonheur, santé et de prospérité.

Une série chronologique est un enregistrement des valeurs de toute quantité fluctuante mesurée à différents moments. Par exemple, nous pouvons avoir un enregistrement des températures quotidiennes, un enregistrement de la tension dans le circuit électrique mesurée à des intervalles de 1sec, un enregistrement de l'indice de prix mensuel d'une marchandise, ou un enregistrement EEG : ElectroEncéphaloGraphe mesurant l'activité électrique entre deux points dans le cerveau d'une personne. Les trois premiers enregistrements sont des exemples de séries chronologiques enregistrées à un ensemble discret de points temporels (chaque jour, chaque seconde, chaque mois), tandis que le quatrième est un exemple de série chronologique enregistrée en continu sur une période de temps. Les deux types de séries se présentent dans des applications pratiques, mais le type "discret" est plus fréquent que le type "continu" et dans ce travail nous nous intéresserons principalement au premier type de série.

Depuis les années **20**, par exemple l'étude de **Yule** sur les taches solaires publiée en **1927**, les modèles linéaires gaussiens des séries chronologiques ont été considérées comme un outil puissant pour l'analyse générale des séries temporelles avec des procédures d'estimation robustes, une théorie statistique

complète ainsi qu'une bonne capacité prédictive. Cependant, les quarante dernières années ont été marquées par des développements majeurs dans ce domaine. Les hypothèses jumelles de linéarité et de stationnarité ont finalement été abandonnées et le sujet a évolué dans une nouvelle direction qui est l'étude des modèles non linéaires qui offrent à la fois une grande excitation et de grands défis. Les limitations des modèles linéaires ont été soulignées et les données réelles ayant souvent des caractéristiques non linéaires ne sont pas prises en compte. De nombreux processus physiques et naturels sont plus précisément modélisés par un modèle non linéaire que par une représentation linéaire. Par exemple, dans les séries financières on observe des périodes très volatiles suivies de périodes plus stables. Ce phénomène appelé "**volatility clustering**" ne sera donc pas expliqué par un modèle à volatilité constante dans le temps. Pour d'autres séries on peut avoir d'autres non linéarités comme l'asymétrie, les domaines de la biologie, l'hydrologie et la météorologie qui présentent des phénomènes proches de la théorie des oscillations non linéaires comme le comportements de cycle limite remarqué dans les séries célèbres de taches solaires ou du lynx qui ont conduit à envisager des modèles autorégressifs non linéaires particuliers.

Les idées mathématiques impliquées dans ces modèles sont beaucoup plus complexes que celles des modèles linéaires, les problèmes statistiques d'identification et d'estimation des paramètres sont également plus complexes. Cependant, une fois ces dernières saisies, la richesse des nouveaux modèles est assez fascinante. L'expérience acquise jusqu'à présent a montré que même les modèles non linéaires simples peuvent capturer des types de comportement de séries chronologiques qu'il serait impossible de décrire avec des modèles linéaires. La plupart des modèles non linéaires ont été introduits pour répondre à des préoccupations pragmatiques telles que :

- (i) Le modèle peut-il donner une meilleure prédiction que les modèles linéaires ?
- (ii) Pouvons-nous estimer le modèle à partir des données ?
- (iii) Pouvons-nous produire une série similaire aux données originales en simulant le modèle estimé ?

Un modèle qui ne peut être estimé à partir des données ou qui ne peut pas reproduire une série similaire par simulation n'est certainement pas intéressant pour les statisticiens bien qu'il puisse encore intéresser certains mathématiciens. Les modèles non linéaires qui ne peuvent pas donner de meilleures performances de prédiction que les modèles linéaires sont également peu attrayants et dénués de sens pour l'analyse des séries chronologiques appliquées. En fait, une première représentation possible pour les séries non linéaires est la série de **Volterra** qui joue le même rôle que la représentation de **Wold** pour les séries linéaires. L'intérêt de cette représentation est plutôt théorique que pratique, pour cette raison et pour résoudre les problèmes évoqués, des modèles plus spécifiques ont été proposés à savoir : les modèles bilinéaire (**Granger et Andersen, 1978**) [24], les modèles autorégressifs à seuils (**TAR**) (**Tong, 1978**) [51], les modèles autorégressifs exponentiels (**EXPAR**) (**Haggan et Ozaki, 1981**) [26], les modèles **ARCH** (**Engle, 1982**) [17],...etc. Ces modèles non linéaires ont été appliqués avec un grand succès dans de nombreux problèmes importants de la vie réelle. Les bases de l'analyse des séries temporelles non linéaires peuvent être trouvées dans **Priestley (1988)** [44], **Tong, (1990)** [52] et **Douc et al. (2013)** [16].

Les modèles autorégressifs exponentiels (**EXPAR**) ont été introduits par **Ozaki (1980)** [41] et **Haggan et Ozaki (1981)** [26] pour expliquer et reproduire les phénomènes stochastiques non linéaires typiques de la théorie des vibrations tels que les décalages de fréquence dépendant de l'amplitude, les

phénomènes de saut et les cycles limites perturbés. Le début a été de prendre un modèle autorégressif Y_t , et ensuite de rendre les coefficients dépendants de manière exponentielle de Y_{t-1}^2 . Parmi les avantages de ce modèle c'est que ces paramètres expliquent clairement quel type de distribution marginale va avoir la série chronologique générée.

Divers propriétés théoriques probabilistes et statistiques tel que la stationnarité, l'ergodicité géométrique, l'estimation, la prévision et les tests ont été l'objet d'étude par plusieurs auteurs. Une méthode d'estimation directe est proposé par **Haggan et Ozaki (1981)** [26] qui consiste à fixer le paramètre non linéaire dans une grille de valeurs, à estimer les autres paramètres par les moindres carrés (**MC**) linéaires et utiliser le critère **AIC** pour sélectionner les paramètres finaux. Cependant, aucun résultat concernant les propriétés asymptotiques de ces estimations n'est donné. **Chan et Tong (1985)** [14] ont montré l'existence d'une solution stationnaire et géométriquement ergodique pour le modèle **EXPAR(1)**, **Tjøstheim (1986)** [51] a traité le problème d'estimation des séries temporelles non linéaires dans un contexte générale par les méthodes des **MC** conditionnel et du maximum de vraisemblance (**MV**) avec application dans les modèles **EXPAR**, **Al- Kassam et Lane (1989)** [1] ont proposé une méthode de prévision pour ce modèle, **Koul et Schick (1997)** [33] ont construit des estimateurs asymptotiquement efficace pour l'**EXPAR(1)** restreint, **Shi et Aoyama (1997)** [46], **Baragona et al. (2002)** [6] ont utilisé l'algorithme génétique pour estimer les paramètres des modèles **EXPAR(p)**. **Allal et El Melhaoui (2006)** [2] ont construit un test paramétrique et non paramétrique pour la détection d'une composante exponentielle dans un modèle **AR(1)**, l'analyse bayésienne de ces modèles a été introduite dans **Ismail (2001)** [30], **Françq et al. (2008)** [18] ont construit des tests de linéarité avec la trilogie rapport de vraisemblance, **Wald et**

multiplicateur de Lagrange pour l'**EXPAR** (1), **Amondela** et **Franco** (2009) [4] ont discuté l'estimation avec la méthode du **quasi maximum de vraisemblance**, **Ghosh et al.** (2015) [22] ont analysé ce modèle avec le **filtre de Kalman étendu**. Étant donné que l'estimation non linéaire prend du temps, **Shi et al.** (2001) [47] ont proposé d'estimer heuristiquement le paramètre non linéaire à partir des données et c'est une remarque très intéressante car lorsque le paramètre non linéaire est connu nous obtenons le modèle **EXPAR** restreint. Cette classe de séries temporelles non linéaires a également trouvé des applications réussies dans l'analyse de données dans un large éventail de domaines notamment l'écologie (**Haggan** et **Ozaki**, 1981 [26], **Priestley**, 1988 [44]), l'hydrologie (**Ozaki**, 1985 [43], **Gurung**, 2013 [25]), le signal vocal (**Ishizuka et al.**, 2005 [29]) et macroéconomique (**Terui et Van Dijk**, 1999 [48], **Amiri**, 2012 [3], **Katsiampa**, 2014 [31]). Récemment, **Azouagh** et **El Melhaoui** (2019) [5] ont montré que «le processus **EXPAR** est le modèle le plus approprié pour la modélisation et la prévision des nombres annuels de taches solaires».

D'un autre côté, la plupart des résultats et des méthodes utilisés dans l'analyse des séries chronologiques sont basés sur l'hypothèse de stationnarité faible du processus qui génère la série temporelle. En pratique, plusieurs séries chronologiques saisonnières rencontrées dans divers domaines tels que l'économie, l'hydrologie, la météorologie ... etc, présentent un caractère non régulier qui ne peut être stationnarisé par une transformation adéquate (différence ordinaire, différence saisonnière, différence mixte, la famille de **Box-Cox** ... etc.). Par conséquent, le traitement de ce genre de séries utilisé par les modèles linéaires classiques **SARIMA** de **Box** et **Jenkins** serait inefficace. Ces derniers ont des autocorrélations qui ne dépendent pas de la saison mais les données désaisonnalisées peuvent encore montrer des variations saison-

nières, la solution était donc de considérer des modèles à coefficients évolutif dans le temps. En effet, la généralisation du théorème de **Wold (1938)** par **Cramer (1961)** aux processus non stationnaires a permis d'élargir la classe des modèles linéaires **ARMA** à coefficient constant à la classe des modèles linéaires **ARMA** à coefficients dépendant du temps. En **1961 Gladyshev [23]** a introduit, pour la première fois dans la littérature, la notion des processus périodiquement corrélés qui sont des processus du second ordre dont la moyenne et la fonction d'autocorrélation sont des fonctions périodiques dans le temps.

Cette notion de périodicité a été exploitée pour analyser les séries hydrologiques et financières à caractère non linéaires et a permis l'émergence de nouvelles classes de modèles de séries chronologiques. C'est ainsi que nous avons l'extension des modèles classiques aux cas périodiques : modèles **GARCH Périodiques (Bollerslev et Ghysels (1996)) [12]**, modèles **AR à seuils Périodiques (Lewis et Ray (2002)) [35]**, modèles **Bilinéaires Périodiques (Bibi et Gautier (2005)) [9]**, modèles de **Mélange Périodiques (Shao (2006)) [45]** **MPAR** et plus récemment le modèle **EXPAR(1)** restreint périodique qui a été introduit dans **Merzougui et al. (2016) [37]** pour modéliser les séries chronologiques saisonnières présentant un comportement non linéaire tel que la fréquence dépendante de l'amplitude, les phénomènes de saut et le cycle limite ayant une structure d'autocovariance périodique.

L'estimation des paramètres est une étape principale dans la modélisation de séries chronologiques qui est compliquée dans un contexte non linéaire, ce problème a été abordé par de nombreux auteurs voir par exemple **Tong (1990) [52]**, **Bierens (1994) [10]**, **Kuan (2000) [34]** et tant parmi d'autres. Les estimateurs des moindres carrés non linéaires (**MCN**) et du quasi maximum de vraisemblance (**QMV**) sont essentiellement utilisés dans ce cas.

L'approche de base sur l'estimation des **MCN** adoptée ici est celle de **Galant** et **White (1988)** [21], une discussion complète peut être trouvée dans **Bierens (1994)** [10] et **Fuller (1996)** [20]. L'estimateur **MCN** du paramètre θ est celui qui minimise la somme des carrés des résidus. L'application nécessite la stricte stationnarité du modèle et nous utiliserons les outils standards comme la compacité et le développement de Taylor. Dans l'estimation traditionnelle du **MV**, on suppose que la densité spécifiée est la vraie fonction et quand ce n'est pas le cas, nous ajoutons le préfixe "**quasi**", bien sûr si la densité est spécifiée correctement alors l'estimateur **QMV** est le même que l'estimateur **MV**. Dans notre travail, la vraisemblance s'écrit comme si la loi de l'innovation était gaussienne. Le **QMV** Gaussien qui en résulte fournit des estimateurs consistants et asymptotiquement normaux et coïncide avec l'estimateur **MCN**.

Dans cette thèse, nous étendons la classe **PEXP**(1) restreint à l'ordre p . Ces modèles satisfont une équation de différence non linéaire semblable à celle des modèles **EXPAR**(p) avec des paramètres et la variance du bruit blanc qui changent périodiquement avec la saison et en fixant le paramètre non linéaire γ dans le terme exponentiel. Notre premier objectif est d'abord le problème d'estimation de ce modèle en utilisant la méthode des moindres carrés linéaires. Comme second objectif, on étudie le problème d'estimation par les **MCN** et le **QMV** du modèle **PEXP**(1) non restreint c'est à dire le paramètre non linéaire γ est supposé inconnu et périodique.

L'intitulé de cette thèse est "**Estimation dans les modèles AutoRégressifs EXponentiels Périodiques**". Pour accomplir nos objectifs, nous avons suivi la méthodologie suivante :

Apport et présentation de la thèse

Chapitre 1 : Théories des moindres carrés non linéaires

et quasi maximum de vraisemblance

Dans ce chapitre, on présente les notions qui ont permis de faire ce travail c'est à dire les éléments de base de la théorie asymptotique de probabilité, les théorie des moindres carrés non linéaires et le quasi maximum de vraisemblance d'un modèle de régression non linéaire selon l'approche de **Gallant** et **White** (1988) [21].

Chapitre 2 : Estimation MC du modèle $\text{EXPAR}(p)$ restreint périodique

Dans le deuxième chapitre, une généralisation du modèle autorégressif exponentiel restreint périodique ($\text{PEXP}(1)$) à l'ordre p est introduite. La méthode des moindres carrés est utilisée pour estimer les paramètres linéaires. Les propriétés asymptotiques des estimateurs pour un PEXP restreint strictement stationnaire sont dérivées. Une petite étude de simulation est réalisée pour vérifier les propriétés asymptotiques.

Chapitre 3 : Estimation MCN du modèle $\text{EXPAR}(1)$ périodique

Dans ce chapitre, nous étudions les propriétés de la consistance forte et de la normalité asymptotique de l'estimateur des MCN du modèle périodique non restreint $\text{EXPAR}(1)$. La littérature statistique générale sur l'estimation des modèles non linéaires de **Gallant** et **White** (1988) [21] est utilisée. Une étude de simulation et un exemple réel sont donnés pour évaluer les performances de cet estimateur.

Chapitre 4 : Estimation QMV du modèle $\text{EXPAR}(1)$ périodique

Ce chapitre est consacré à l'estimation non linéaire du modèle $\text{PEXP}(1)$ par la méthode du quasi maximum de vraisemblance. Sous des hypothèses de régularité, on montre que l'estimateur QMV Gaussien est presque sûrement

et fortement consistant et asymptotiquement normal.

Cette contribution est couronnée par une publication scientifique dans une revue renommée internationale à savoir :

Becila, S and Merzougui, M. (2020). Nonlinear Least Squares estimation of the Periodic EXPAR(1) model. *Communications in Statistics-Theory and Methods*.

CHAPITRE 1

Théories des MC non linéaires et du QMV

Sommaire

1.1 Introduction	1
1.2 Eléments de base de la théorie asymptotique . .	3
1.3 Théorie des moindres carrés non linéaires (MCN)	9
1.4 Théorie du quasi-maximum de vraisemblance .	17

1.1 Introduction

Les développements récents dans la théorie et l'application des modèles de séries chronologiques linéaires nous ont fourni des outils très puissants et pratiques pour la prévision et le contrôle des séries chronologiques générales. Cependant, de tels modèles ne peuvent pas expliquer les phénomènes non linéaires typiques des séries temporelles dans le monde réel. Il est généralement

difficile d'estimer les spécifications non linéaires et d'analyser les propriétés des estimateurs résultants. Les résultats d'estimation de la spécification non linéaire peuvent ne pas être facilement interprétés. Malgré ces difficultés, de plus en plus de preuves empiriques montrent que de nombreuses relations économiques sont en fait non linéaires. Dans ce travail, nous nous concentrons sur l'estimation des spécifications non linéaires. Pour plus de détails sur les régressions non linéaires, on se réfère à **Gallant et White (1988)** [21], **Bierens (1994)** [10] et **Fuller (1996)** [20].

Nous allons essayer de développer une approche plus systématique et discuter d'un cadre général pour l'estimation de séries chronologiques non linéaires. L'approche est basée sur le **développement de Taylor** d'une fonction de pénalité générale qui est ensuite spécialisée à un critère de type des moindres carrés conditionnels et un critère de type de maximum de vraisemblance. **Klimko et Nelson (1978)** [32] ont précédemment considéré de telles extensions de **Taylor** dans le cas des moindres carrés conditionnels dans un contexte général.

Dans ce chapitre, nous considérons les propriétés asymptotiques, i.e. la consistance faible et forte et la normalité asymptotique, des estimateurs des moindres carrés et du quasi maximum de vraisemblance des paramètres d'un modèle de régression non linéaire. Une bonne partie de la présentation est prise de **Bierens (1994)** [10], **Kuan (2000)** [34] et **Amondela et Franq (2009)** [4].

Pour faciliter une discussion plus approfondie de l'estimation des modèles non linéaires, résumons d'abord certains des concepts fondamentaux et bien connus de la théorie asymptotique et quelques concepts et résultats importants en théorie des probabilités. Les théorèmes limites qui sont des outils puissants pour analyser les comportements de convergence des estimateurs

nous intéressent plus particulièrement.

1.2 Eléments de base de la théorie asymptotique

La théorie asymptotique des modèles de régression non linéaires, en particulier les résultats de consistance, dépendent fortement des lois des grands nombres uniforme (LGNU). Comprendre ces lois exige la connaissance de la théorie des probabilités abstraite, une partie de cette théorie est donnée en annexe.

1.2.1 Fonctions aléatoires

Une fonction aléatoire est une fonction qui est une variable aléatoire pour chaque valeur de ses arguments. Généralement les fonctions aléatoires se produisent comme fonctions de variables aléatoires et paramètres.

Définition 1.1

Soit $(\Omega, \mathcal{F}, \mathcal{P})$ un espace de probabilité et soit Θ un sous ensemble de \mathbb{R}^k . La fonction réelle $f(\cdot) = f(\cdot, \omega)$ sur $\Theta \times \Omega$ est une fonction aléatoire sur Θ si : $\forall t \in \mathbb{R}$ et $\forall \theta_0 \in \Theta : \{\omega \in \Omega : f(\theta_0, \omega) \leq t\} \in \mathcal{F}$.

Mais si $f(\cdot)$ est une fonction aléatoire sur un ensemble non dénombrable Θ alors $\sup_{\theta \in \Theta} f(\theta)$ et $\inf_{\theta \in \Theta} f(\theta)$ ne sont pas automatiquement des variables aléatoires. Pour surpasser ce problème on va s'intéresser au fonction aléatoire de type $f(\cdot) = \varphi(\cdot, X)$ où φ est une fonction réelle mesurable sur $\Theta \times \mathbb{R}^m$ avec Θ un ensemble borélien compact sur \mathbb{R}^k et on a le théorème suivant :

Théorème 1.1

Soit Θ un ensemble compact sur \mathbb{R}^k et soit $\varphi(\theta, x)$ une fonction borélienne mesurable sur $\Theta \times \mathbb{R}^m$ qui est continue en $\theta, \forall x \in \mathbb{R}^m$. Il existe une identification (mapping) $\theta(x)$ de \mathbb{R}^k dans Θ avec des composantes mesurables boréliennes tel que

$$\varphi(\theta(x), x) = \sup_{\theta \in \Theta} \varphi(\theta, x),$$

par conséquent le supremum est une fonction mesurable sur \mathbb{R}^m . En plus si $\varphi(\theta, x)$ est continue sur $\Theta \times \mathbb{R}^m$ alors $\sup_{\theta \in \Theta} \varphi(\theta, x)$ est une fonction continue sur \mathbb{R}^m .

Pour la preuve voir **Bierens** [10] p.16.

1.2.2 Développement de Taylor en probabilité

Soit $\{X_n\}$ une suite de variables aléatoires tel que $X_n = a + O_p(r_n)$ où $a \in \mathbb{R}$ et $0 < r_n \rightarrow 0$ quand $n \rightarrow \infty$. Si g est une fonction dérivable jusqu'à l'ordre s , on a alors :

$$g(X_n) = \sum_{j=0}^{s-1} \frac{g^{(j)}(a)}{j!} (X_n - a)^j + O_p(r_n^s),$$

où $g^{(j)}$ est la $j^{\text{ème}}$ dérivée de g et $g^{(0)} = g$.

1.2.3 Lois des Grands Nombres et théorème centrale limite pour les processus dépendants

Plusieurs processus en économétrie et statistique sont corrélés et donc ils ne sont pas pris en compte par les LGN standards. Nous allons présenter la LGN forte pour les processus strictement stationnaires.

Soit $Z_t, t \in \mathbb{N}$ une suite de variables aléatoires avec $E(Z_t) = \mu_t$, la loi des

grands nombres (**LGN**) spécifie les conditions sous lesquelles

$$\frac{1}{n} \sum_{t=1}^n (Z_t - \mu_t),$$

converge vers 0 en probabilité ou presque sûrement. Si la convergence est en probabilité on parle de **LGN** faible et si la convergence est *p.s.* on parle de **LGN** forte. Dans les applications les variables aléatoires Z_t peuvent être des fonctions d'autres variables aléatoires. L'utilité des **LGN** provient du fait que plusieurs estimateurs s'écrivent en fonction de moyennes empiriques de variables aléatoires alors pour établir la limite de l'estimateur on commence par établir les limites de ces moyennes. On commence par donner une série de définitions.

Définition 1.2 (Suites de différences de martingales)

Soit $I_t, t \geq 0$ une suite de σ -algèbre tel que $I_0 \subseteq I_1 \subseteq \dots \subseteq I$. Soit Z_t une suite de V.A. Alors Z_t est une suite de DM si Z_t est I_t -mesurable, $E|Z_t| < \infty$ et

$$E(Z_t | I_{t-1}) = 0, \forall t \geq 1.$$

On note que si Z_t est une suite de DM alors

$$E(Z_t) = E(E(Z_t | I_{t-1})) = 0,$$

par la loi des espérances itérées. En plus puisque

$$E(Z_t Z_{t+k}) = E(Z_t E(Z_{t+k} | I_{t+k-1})) = 0, \forall k \geq 1,$$

la suite de DM est toujours non corrélée et si

$$S_n = \sum_{t=1}^n Z_t,$$

alors S_n est une martingale. Inversement si S_n est une martingale alors $X_n = S_n - S_{n-1}$ est une DM (d'où le nom).

Définition 1.3 (Stationnarité stricte)

La suite de variables aléatoires $Z_t, t \geq 1$ est strictement stationnaire si (Z_1, Z_2, \dots, Z_n) a la même distribution que $(Z_{1+k}, Z_{2+k}, \dots, Z_{n+k}), \forall k \geq 1$ et $n \geq 1$.

Définitions 1.4 (Invariance et ergodicité de suites strictement stationnaires)

(i) Soit l'événement suivant

$$A = \{\omega \in \Omega, (Z_1(\omega), Z_2(\omega), \dots) \in B\},$$

où $B \in B^\infty$ sont les ensembles boréliens de \mathbb{R}^∞ . Alors A est un événement invariant si

$$A = \{\omega \in \Omega, (Z_{1+k}(\omega), Z_{2+k}(\omega), \dots) \in B\}, \forall k \geq 1.$$

(ii) La suite Z_t est ergodique si $\forall A$ événement invariant alors $P(A) = 0$ ou $P(A) = 1$.

(iii) Le processus réel strictement stationnaire Z_t est ergodique si et seulement si \forall l'ensemble borélien B et $\forall k$

$$\frac{1}{n} \sum_{t=1}^n 1_B(Z_t, Z_{t+1}, \dots, Z_{t+k}) \rightarrow P((Z_1, \dots, Z_{1+k}) \in B).$$

On note que toute suite de V.A i.i.d. est strictement stationnaire et ergodique. En plus si Z_t est strictement stationnaire et ergodique et si $g : \mathbb{R}^\infty \rightarrow \mathbb{R}$ une fonction mesurable alors la suite Y_t avec $Y_t = g(Z_t, Z_{t+1}, \dots)$ est aussi strictement stationnaire et ergodique. On donne maintenant la LGN forte suivante appelée théorème ergodique.

Théorème 1.2

Soit Z_t une suite strictement stationnaire et ergodique avec $E|Z_1| < \infty$ et $E(Z_1) = \mu$. Alors :

$$\frac{1}{n} \sum_{t=1}^n Z_t \xrightarrow{p.s} \mu,$$

quand $n \rightarrow \infty$. (**LGN** de **Kolmogorov** cas particulier).

Comme corollaire on a le résultat suivant : Z_t strictement stationnaire et ergodique et f une fonction mesurable satisfaisant $E|f(Z_t, Z_{t+1}, \dots)| < \infty$ alors

$$\frac{1}{n} \sum_{t=1}^n f(Z_t, Z_{t+1}, \dots) \xrightarrow{p.s} E f(Z_t, Z_{t+1}, \dots) \text{ quand } n \rightarrow \infty.$$

Lois des Grands Nombres Uniforme (LGNU)

Les variables aléatoires dans les suites doivent avoir un bon comportement pour assurer des **LGN**. Intuitivement, les V.A sans certains moments bornés peuvent exhiber un comportement aberrant.

Souvent on doit traiter des fonctions de V.A et des paramètres d'un modèle. Par exemple la fonction $q(Y_t(\omega); \theta)$ est une V.A pour un paramètre donné θ et c'est une fonction de θ pour un ω donné. Quand θ est fixé, on peut imposé des conditions sur q et Y_t tel que $\{q(Y_t(\omega); \theta)\}$ obéit aux **LGN**. Quand θ prend des valeurs dans l'espace des paramètres Θ on aura besoin de **LGN** qui ne dépendront pas de θ . Plus précisément, supposons que $\{q(Y_t; \theta)\}$ obéit à une **LGN** forte quelque soit $\theta \in \Theta$:

$$Q_n(\omega; \theta) = \frac{1}{n} \sum_{t=1}^n q(Y_t(\omega); \theta) \xrightarrow{p.s} Q(\theta),$$

où $Q(\theta)$ est une fonction non stochastique de θ . On note Ω_0 l'ensemble de ω tel que $Q_n(\omega; \theta)$ converge vers $Q(\theta)$, comme le comportement de convergence peut dépendre de θ alors $\Omega_0^c(\theta) = \{\omega : Q_n(\omega; \theta) \not\rightarrow Q(\theta)\}$ varie avec θ . Quand Θ est un intervalle de \mathbb{R} , $\bigcup_{\theta \in \Theta} \Omega_0^c(\theta)$ est l'union non dénombrable d'ensembles de non convergence et donc peut ne pas avoir la probabilité 0 bien que chaque $\Omega_0^c(\theta)$ est de probabilité 0. Ainsi l'événement $Q_n(\omega; \theta) \rightarrow Q(\theta), \forall \theta$ c.à.d : $\bigcap_{\theta \in \Theta} \Omega_0(\theta)$ peut se produire avec une probabilité inférieure à 1. En fait $\bigcup_{\theta \in \Theta} \Omega_0^c(\theta)$ peut même ne pas être dans la σ -algèbre \mathcal{F} . Si c'est le cas on

ne peut rien conclure par rapport à la convergence de $Q_n(\omega; \theta)$. Pire encore quand θ dépend de n , comme dans le cas où on remplace θ par son estimateur $\hat{\theta}_n$: on peut ne pas trouver $n^* < \infty$ tel que $Q_n(\omega; \hat{\theta}_n)$ soit proche de $Q(\omega; \hat{\theta}_n)$, $\forall n > n^*$. Toutes ces observations montrent que nous devons étudier la convergence qui est uniforme sur l'espace des paramètres : $Q_n(\omega; \theta)$ converge uniformément vers $Q(\theta)$ presque sûrement si

$$\sup_{\theta \in \Theta} |Q_n(\theta) - Q(\theta)| \xrightarrow{p.s} 0,$$

on suppose toujours que ce supremum est une V.A $\forall n$.

Pour montrer cette **LGNU** on commence par montrer la **LGN** standard pour chaque θ en utilisant le théorème ergodique et on utilise un argument de compacité de Θ .

Théorème 1.3

Soit X_t une suite de V.A i.i.d. dans \mathbb{R}^k avec fonction de distribution G . Soit $f(x, \theta)$ une fonction mesurable sur $\mathbb{R}^k \times \Theta$ où Θ est un ensemble borélien compacte de \mathbb{R}^m qui est continu en $\theta, \forall x$. Si $E[\sup_{\theta \in \Theta} |f(X_t, \theta)|] < \infty$ alors

$$\frac{1}{n} \sum_{t=1}^n f(X_t, \theta) \rightarrow \int f(X, \theta) dG(x), p.s \text{ uniformément sur } \Theta.$$

(Pour la preuve voir : **Bierens**[\[10\]](#)).

Théorème centrale limite

Il y a une grande littérature pour les **TCL** pour les processus dépendants.

Théorème 1.4

Soit Z_t une suite de différence de martingale strictement stationnaire et ergodique avec $Var(Z_t) = \sigma^2 < \infty$. Alors

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n Z_t \xrightarrow{D} N(0, \sigma^2).$$

On termine cette section par deux lemmes utiles

1) Lemme de sommabilité de Cesàro

Si $(S_n)_{n \in \mathbb{N}}$ suite réelle ou complexe avec $S_n \rightarrow r$, alors

$$\frac{1}{n} \sum_{k=1}^n S_k \rightarrow r,$$

c.à.d. lorsque la suite (S_n) a une limite, la moyenne de Cesàro possède la même limite.

2) Lemme de Borel-Cantelli

Soit S_n une suite de variables aléatoires tel que

$$\sum_{n \geq 1} P(|S_n| > \varepsilon) < \infty, \forall \varepsilon > 0,$$

alors

$$S_n \xrightarrow[n \rightarrow \infty]{} 0, \text{ p.s.}$$

1.3 Théorie des moindres carrés non linéaires (MCN)

Dans cette section, nous considérons les propriétés asymptotiques c.à.d. la consistance forte et la normalité asymptotique, des estimateurs des moindres carrés des paramètres d'un modèle de régression non linéaire. Partout, nous supposons que le processus de génération de données est indépendant identiquement distribué. L'approche de régression non linéaire actuelle est celle de **Gallant et White (1988)** [21]. Lorsque l'on cherche à établir les propriétés asymptotiques des estimateurs des MCN toute la difficulté réside dans le fait que l'on ne dispose pas d'expression analytique pour ces estimateurs. En fait, il est nécessaire de recourir à des algorithmes d'optimisation numériques pour approcher la solution. En effet, il n'est pas possible d'exprimer

les estimateurs, solutions de la condition du premier ordre de minimisation, comme des fonctions simples des observations. Il est alors difficile d'étudier la convergence de ces estimateurs comme nous allons le faire dans le chapitre suivant dans un contexte linéaire standard. Il convient ainsi d'adopter une démarche particulière où l'on va chercher à étudier la convergence du critère des **MCN** afin de démontrer la convergence des estimateurs des **MCN**, solutions du programme de minimisation de ce critère.

1.3.1 Modèles de régression non linéaire et estimateurs des MCN

Considérons un processus de génération de données indépendantes $\{(Y_j, X_j)\}$, $j = 1, 2, \dots$. Où $Y_j \in \mathbb{R}$ est la variable dépendante et $X_j \in \mathbb{R}^k$ est un vecteur de régresseurs. On suppose que :

$$Y_j = f(X_j, \theta_0) + U_j, \quad j = 1, 2, \dots$$

où $f(x, \theta)$ est une fonction, non linéaire réelle borélienne mesurable connue, sur $\mathbb{R}^k \times \Theta$ avec $E(Y_j^2) < \infty$ et Θ un ensemble borélien compact dans \mathbb{R}^m contenant θ_0 , appelé espace des paramètres. Pour chaque $x \in \mathbb{R}^k$ la fonction $f(x, \theta)$ est continue sur Θ . Nous rappelons que la compacité d'un ensemble implique que tout recouvrement ouvert contient un sous-recouvrement fini, et qu'un sous-ensemble d'un espace euclidien est compact si et seulement s'il est fermé et borné. On suppose aussi que les erreurs U_j sont i.i.d. et normalement distribués.

En supposant que le processus de génération de données soit observable pour $j = 1, 2, \dots, n$, l'estimateur non linéaire des moindres carrés pour θ_0 est maintenant défini comme une solution mesurable du problème de minimisation suivant :

$$\theta \in \Theta : Q_n(\hat{\theta}) = \inf_{\theta \in \Theta} Q_n(\theta),$$

où

$$Q_n(\theta) = (1/n) \sum_{j=1}^n [Y_j - f(X_j, \theta)]^2.$$

D'après le théorème 1.1 de la section précédente, une telle solution mesurable existe toujours.

La condition du premier ordre du problème de minimisation des **MCN** :

$\frac{\partial Q_n(\theta)}{\partial \theta} = 0$ est un système de n équations non linéaires avec n inconnues, donc en général, il n'a pas de solution exacte. La question qui se pose est de savoir comment montrer que l'estimateur est convergent.

Consistance et normalité asymptotique : théorie générale

Consistance Comme l'estimateur des **MCN** n'a pas de forme analytique en général, une approche différente est donc nécessaire pour établir la consistance. Intuitivement, lorsque la fonction objective $Q_n(\theta)$ soit proche de $E[Q_n(\theta)]$ pour tout θ , il est raisonnable de s'attendre à ce que le minimiseur de $Q_n(\theta)$, soit également proche du minimum de $E[Q_n(\theta)]$. Etant donné que Q_n est non linéaire en θ , une **LGNU** doit être invoquée pour justifier la proximité entre $Q_n(\theta)$ et $E[Q_n(\theta)]$. La preuve de consistance comprend deux étapes principales. Tout d'abord, nous établissons des conditions telles que

$$Q_n(\theta) \rightarrow Q(\theta) \text{ p.s. uniformément sur } \Theta, \quad (1.1)$$

où

$$Q(\theta) = \lim_{n \rightarrow \infty} E[Q_n(\theta)] \text{ est continue sur } \Theta, \quad (1.2)$$

puis nous énonçons des conditions telles que

$$\theta_* \in \Theta : Q(\theta_*) = \inf_{\theta \in \Theta} Q(\theta) \implies \theta_* = \theta_0. \quad (1.3)$$

Ainsi, cette dernière condition dit que $Q(\theta)$ a un infimum unique sur Θ en θ_0 .

Les conditions (1.2) et (1.3) garantissent que $\hat{\theta} \rightarrow \theta_0$ p.s., à cause du théorème fondamental suivant.

Théorème 1.5

Soit $(Q_n(\theta))$ une suite de fonctions aléatoires sur un ensemble compact $\Theta \in \mathbb{R}^m$ tel que pour une fonction réelle continue $Q(\theta)$ sur Θ ,

$$Q_n(\theta) \rightarrow Q(\theta) \text{ p.s. pseudo-uniformément sur } \Theta.$$

Soit θ_n tout vecteur aléatoire dans Θ satisfaisant $Q_n(\theta_n) = \inf_{\theta \in \Theta} Q_n(\theta)$ et soit θ_0 un point unique dans Θ tel que $Q(\theta_0) = \inf_{\theta \in \Theta} Q(\theta)$.

Alors $\theta_n \rightarrow \theta_0$ p.s.

Pour la preuve voir Bierens [10] p.64.

La normalité asymptotique Nous passons ensuite à la normalité asymptotique. Nous donnons le théorème suivant dont la démonstration peut être trouvée dans Bierens [10] p.65.

Théorème 1.6

Supposons que les conditions du théorème 1.5 soient satisfaites (pour la convergence en cas de probabilité) et supposons en plus que :

- (i) Θ est convexe et θ_0 est un point intérieur de Θ ;
- (ii) $\frac{\partial Q_n(\theta)}{\partial \theta}$ et $\frac{\partial^2 Q_n(\theta)}{\partial \theta \partial \theta'}$ sont bien définis comme vecteur et matrice, respectivement, de fonctions aléatoires sur Θ ;
- (iii) $\sqrt{n} \frac{\partial Q_n(\theta_0)}{\partial \theta} \xrightarrow{D} N_m(0, A_1)$, où A_1 est une matrice $m \times m$ semi-définie positive ;
- (iv) pour $i_1, i_2 = 1, 2, \dots, m$,

$$\frac{\partial^2 Q_n(\theta)}{\partial \theta_{i_1} \partial \theta_{i_2}} \xrightarrow{P} \frac{\partial^2 Q(\theta)}{\partial \theta_{i_1} \partial \theta_{i_2}}$$

uniformément sur un voisinage de θ_0 , où la fonction limite impliquée est continue dans θ_0 .

(v) la matrice $A_2 = \frac{\partial^2 Q(\theta_0)}{\partial \theta \partial \theta'}$ est non singulière.

Alors

$$\sqrt{n}(\theta_n - \theta_0) \rightarrow N_m(0, A_2^{-1} A_1 A_2^{-1}).$$

Algorithmes d'optimisation non linéaire

Les algorithmes utilisés dans les principaux logiciels d'économétrie sont généralement construits à partir de la méthode de **Newton Raphson** ou la méthode du **score**. Nous n'évoquerons ici que la première méthode.

Les méthodes d'optimisation numérique sont utilisées pour maximiser une fonction $f(\theta)$ lorsque la condition du premier ordre $\frac{\partial f(\theta)}{\partial \theta} = 0$ n'admet pas de solution analytique; le θ optimal doit être déduit par un algorithme itératif.

Ce dernier utilise trois principaux éléments :

- (i) Des valeurs initiales θ_0 pour amorcer le processus itératif.
- (ii) Une règle de passage d'un vecteur θ au suivant.
- (iii) Une règle d'arrêt s'il y a convergence.

Dans la méthode de **Newton Raphson**, la direction est déterminée par le gradient de la fonction $f(\theta)$, noté $\nabla f(\theta)$, tandis que le pas est déterminé par le hessien, noté $H(\theta)$ et la règle de passage est

$$\theta_{i+1} = \theta_i - H^{-1}(\theta_i) \nabla f(\theta),$$

où $|H(\theta)| \neq 0$ et définie positive.

1.3.2 Application pour des modèles de ST non linéaires

Supposons des modèles de séries chronologiques univariés de la forme suivantes

$$Y_t = m_{\theta_0}(Y_{t-1}, Y_{t-2}, \dots) + \sigma \eta_t,$$

où θ_0 est un paramètre inconnu appartenant à l'ensemble Θ de \mathbb{R}^s , $\eta_t \sim iid(0, 1)$ avec η_t indépendant des Y_{t-i} pour $i > 0$ et σ une constante positive.

Sous ces hypothèses nous avons

$$\begin{aligned} E(Y_t | Y_{t-1}, Y_{t-2}, \dots) &= m_{\theta_0}(Y_{t-1}, Y_{t-2}, \dots) = m_t(\theta_0), \\ V(Y_t | Y_{t-1}, Y_{t-2}, \dots) &= \sigma^2. \end{aligned}$$

Sachant les valeurs initiales Y_0, Y_{-1}, \dots , le critère des **MCN** des observations Y_1, \dots, Y_n évalué en $\theta \in \Theta$ est

$$Q_n(\theta) = (1/n) \sum_{t=1}^n l_t(\theta) \quad \text{où} \quad l_t(\theta) = (Y_t - m_t(\theta))^2.$$

Ce critère théorique n'est pas opérationnel car il dépend des valeurs initiales souvent inconnues. L'estimateur des **MCN** est obtenu en remplaçant la moyenne conditionnelle par l'approximation mesurable

$$\tilde{m}_t(\theta) = m_\theta(Y_{t-1}, \dots, Y_1, 0, \dots)$$

où on suppose que les valeurs initiales sont nulles. L'estimateur des **MCN** de θ_0 est défini comme toute solution mesurable $\hat{\theta}_n$ de

$$\hat{\theta}_n = \arg \inf_{\theta \in \Theta} \tilde{Q}_n(\theta), \tag{1.4}$$

où

$$\tilde{Q}_n(\theta) = (1/n) \sum_{t=1}^n \tilde{l}_t(\theta) \quad \text{où} \quad \tilde{l}_t(\theta) = (Y_t - \tilde{m}_t(\theta))^2.$$

L'existence d'une solution à (1.4) est garantie sous l'hypothèse :

(i) Θ est un ensemble compact et la fonction $\tilde{m}_t(\theta)$ est continue.

Consistance

Supposons que

(ii) $\{Y_t\}$ est un processus strictement stationnaire et ergodique avec $E(Y_t^2) < \infty$.

Alors, les processus $\{m_t(\theta)\}$ et $\{l_t(\theta)\}$ sont aussi des processus stationnaires ergodiques. Par le théorème ergodique, cité dans la section précédente, le critère théorique $Q_n(\theta)$ converge presque sûrement vers le critère asymptotique

$$Q_\infty(\theta) = E_{\theta_0}(l_t(\theta)),$$

à condition que l'espérance existe.

Les valeurs initiales sont souvent uniformément négligeables dans le sens où :

(iii) $\sup_{\theta \in \Theta} |l_t(\theta) - \tilde{l}_t(\theta)| \xrightarrow{p.s} 0$ quand $t \rightarrow \infty$.

Alors nous avons

$$\tilde{Q}_n(\theta) - Q_n(\theta) \rightarrow 0, \text{ p.s uniformément en } \theta.$$

et le critère opérationnel $\tilde{Q}_n(\theta)$ converge aussi vers le critère asymptotique $Q_\infty(\theta)$.

Nous avons besoin aussi d'une hypothèse d'identifiabilité :

(iv) Si $\theta \neq \theta_0$ alors $m_t(\theta) \neq m_t(\theta_0)$, avec probabilité non nulle.

Alors nous avons

$$\begin{aligned} Q_\infty(\theta) - Q_\infty(\theta_0) &= E(m_t(\theta) - m_t(\theta_0))^2 + 2E((Y_t - m_t(\theta_0))(m_t(\theta_0) - m_t(\theta))) \\ &= E(m_t(\theta) - m_t(\theta_0))^2 \geq 0. \end{aligned}$$

La dernière égalité résulte de l'orthogonalité entre η_t et $(m_t(\theta_0) - m_t(\theta))$.

L'inégalité est une égalité si et seulement si $\theta = \theta_0$. Ainsi, nous avons montré que le critère asymptotique est minimisé à la vraie valeur θ_0 . Nous venons de montrer que

$$\theta_0 = \arg \inf_{\theta \in \Theta} \lim_{n \rightarrow \infty} \tilde{Q}_n(\theta), \text{ p.s.}$$

En utilisant la compacité de Θ et l'ergodicité de $l_t(\theta)$, on pourra montrer que

$$\theta_0 = \lim_{n \rightarrow \infty} \arg \inf_{\theta \in \Theta} \tilde{Q}_n(\theta), \text{ p.s.}$$

L'idée de la démonstration est que toute suite définie sur un ensemble compact admet au moins une valeur d'adhérence, il suffit alors de montrer que cette valeur d'adhérence est unique et correspond à la vraie valeur θ_0 . Une démonstration détaillée est donnée au chapitre 3.

Distribution asymptotique de l'estimateur des MCN

En plus des hypothèses précédentes (i)-(iv), on ajoute les conditions suivantes :

(v) θ_0 appartient à l'intérieur $\overset{\circ}{\Theta}$ de Θ .

(vi) $\theta \rightarrow m_t(\theta)$ admet des dérivées continues à l'ordre 3 et

$$E \sup_{\theta \in \Theta} \left| \frac{\partial^3 l_t(\theta_0)}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| < \infty, \forall i, j, k.$$

(vii) $I = E \left(\frac{\partial l_t(\theta_0)}{\partial \theta} \frac{\partial l_t(\theta_0)}{\partial \theta'} \right)$ et $J = E \left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta'} \right)$ existent et sont non singulières.

Nous avons $\frac{\partial \tilde{Q}_n(\hat{\theta}_n)}{\partial \theta} = 0$, parce que $\tilde{Q}_n(\hat{\theta}_n)$ est minimisé en $\hat{\theta}_n$ qui est un point intérieur de Θ , on voit bien la nécessité de l'hypothèse (v). Un développement de Taylor nous donne

$$0 = \sqrt{n} \frac{\partial \tilde{Q}_n(\hat{\theta}_n)}{\partial \theta} = \sqrt{n} \frac{\partial \tilde{Q}_n(\theta_0)}{\partial \theta} + \left(\frac{\partial^2 \tilde{Q}_n(\theta_*)}{\partial \theta \partial \theta'} \right) \sqrt{n} (\hat{\theta}_n - \theta_0),$$

où θ_* est entre θ_0 et $\hat{\theta}_n$.

Les valeurs initiales étant uniformément négligeables, on pourra montrer que :

(viii) $\left\| (1/n) \sum_{t=1}^n \left(\frac{\partial l_t(\theta_0)}{\partial \theta} - \frac{\partial \tilde{l}_t(\theta_0)}{\partial \theta} \right) \right\| \xrightarrow{P} 0$ quand $t \rightarrow \infty$.

(ix) $\sup_{\theta \in V(\theta_0)} \left\| (1/n) \sum_{t=1}^n \left(\frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta'} - \frac{\partial^2 \tilde{l}_t(\theta_0)}{\partial \theta \partial \theta'} \right) \right\| \xrightarrow{P} 0$ quand $t \rightarrow \infty$.

Avec ces conditions on pourra montrer que

$$n^{-1/2} \sum_{t=1}^n \frac{\partial \tilde{l}_t(\theta_0)}{\partial \theta} \xrightarrow{D} N(0, I),$$

et que

$$(1/n) \sum_{t=1}^n \frac{\partial^2 \tilde{l}_t(\theta_*)}{\partial \theta \partial \theta'} \xrightarrow{P} J.$$

d'où le résultat final

$$\sqrt{n} (\hat{\theta}_n - \theta_0) \xrightarrow{D} N(0, J^{-1} I J^{-1}).$$

Dans le chapitre 3, on appliquera cette théorie pour le modèle **EXPAR(1)** périodique, cela revient à vérifier les conditions (i)-(ix).

1.4 Théorie du quasi-maximum de vraisemblance

La méthode du quasi-maximum de vraisemblance (**QMV**) est essentiellement la même que la méthode du **MV** habituellement vue dans les livres de statistiques et d'économétrie. Comment tirer des inférences statistiques sous la possibilité d'une spécification erronée du modèle est la préoccupation majeure de la méthode du **QMV**. La méthode du **MV** suppose que la fonction de densité spécifié est la vraie et qu'il n'y a pas d'erreurs de spécifications. Ainsi les résultats dans la méthode du **MV** ne sont que des cas particuliers de la méthode du **QMV**.

1.4.1 Critère d'information Kullback-Leibler

Soit g la fonction de densité de la variable aléatoire z et f une autre fonction de densité. On définit le **critère d'information Kullback-Leibler (CIKL)**

de g par rapport à f comme

$$\mathbb{I}(g : f) = \int_{\mathbb{R}} \log \left(\frac{g(\zeta)}{f(\zeta)} \right) g(\zeta) d\zeta.$$

Quand f est utilisé pour décrire z , la valeur $\mathbb{I}(g : f)$ est la "surprise" attendue sachant que g est en fait la véritable densité de z . Le résultat suivant montre que le **CIKL** de g par rapport à f est non négatif.

Théorème 1.7

Soit g la fonction de densité de la variable aléatoire z et f une autre fonction de densité. Alors $\mathbb{I}(g : f) \geq 0$, avec l'égalité tenant si, et seulement si, $g = f$ presque partout (c'est-à-dire $g = f$ sauf sur un ensemble de mesure de Lebesgue nulle).

Pour la preuve voir **Kuan**[34] p.231.

On note que le **CIKL** n'est pas une métrique car il n'est pas symétrique en général, i.e., $\mathbb{I}(g : f) \neq \mathbb{I}(f : g)$, et il n'obéit pas à l'inégalité triangulaire. Par conséquent, le **CIKL** n'est qu'une mesure brute de la proximité entre f et g . (**CIKL** est compris comme une distance entre distributions).

Soit $\{z_t\}$ une suite de variables aléatoires de \mathbb{R}^ν définies sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P}_0)$. Soit

$$z^n = (z_1, z_2, \dots, z_n),$$

l'ensemble d'informations générées par z_t . Le vecteur z_t est divisé en deux parties y_t et w_t , où y_t est un scalaire et w_t est $(\nu - 1) \times 1$. Correspondant à \mathbb{P}_0 , il existe une fonction de densité conjointe g^n pour z^n :

$$g^n(z^n) = g(z_1) \prod_{t=2}^n \frac{g^t(z^t)}{g^{t-1}(z^{t-1})} = g(z_1) \prod_{t=2}^n g_t(z_t | z^{t-1}),$$

où g^t désigne la densité conjointe des t variables aléatoires z_1, \dots, z_t , et g_t est la fonction de densité de z_t conditionnelle aux informations passées

z_1, \dots, z_{t-1} . La fonction g^n est le mécanisme aléatoire gouvernant le comportement de z^n et sera appelée processus de génération de données (**PGD**) de z^n .

Comme g^n est inconnue, nous pouvons postuler une fonction de densité conditionnelle $f_t(z_t|z^{t-1}; \theta)$, où $\theta \in \Theta \subset \mathbb{R}^k$, et approximer $g^n(z^n)$ par

$$f^n(z^n; \theta) = f(z_1) \prod_{t=2}^n f_t(z_t|z^{t-1}; \theta).$$

La fonction f^n est appelée **la fonction quasi-vraisemblance**, dans le sens où f^n n'a pas besoin de coïncider avec g^n . Clairement, la densité postulée f^n serait utile si elle est proche du **PGD** g^n . Il est donc naturel d'envisager de minimiser le **CIKL** de g^n par rapport à f^n :

$$\mathbb{I}(g^n : f^n; \theta) = \int_{\mathbb{R}^n} \log \left(\frac{g^n(\zeta^n)}{f^n(\zeta^n; \theta)} \right) g^n(\zeta^n) d(\zeta^n). \quad (1.5)$$

Cela revient à minimiser le niveau de "surprise" résultant de la spécification d'un f^n à la place du **PGD** g^n . Comme g^n ne dépend pas de θ , minimiser le **CIKL** par rapport à θ équivaut à maximiser

$$\int_{\mathbb{R}^n} \log(f^n(\zeta^n; \theta)) g^n(\zeta^n) d(\zeta^n) = E[\log f^n(z^n; \theta)],$$

où E est l'opérateur d'espérance par rapport au **PGD** $g^n(z^n)$. Cela équivaut à son tour à maximiser la moyenne de $\log f_t$:

$$\bar{L}_n(\theta) = \frac{1}{n} E[\log f^n(z^n; \theta)] = \frac{1}{n} E \left(\sum_{t=1}^n \log f_t(z_t|z^{t-1}; \theta) \right). \quad (1.6)$$

Soit θ^* le maximiseur de $\bar{L}_n(\theta)$. Alors, θ^* est aussi le minimiseur du **CIKL** (1.5).

S'il existe un $\theta_0 \in \Theta$ unique tel que $f^n(z^n; \theta_0) = g^n(z^n)$ pour tout n , on dit que $\{f_t\}$ est spécifié correctement dans son intégralité pour $\{z_t\}$. Dans ce cas, le **CIKL** résultant $\mathbb{I}(g^n : f^n; \theta_0) = 0$ et atteint le minimum. Cela montre que $\theta^* = \theta_0$ quand $\{f_t\}$ est correctement spécifié dans son intégralité.

La maximisation de \bar{L}_n n'est cependant pas un problème facilement résoluble car la fonction objective (1.6) implique l'opérateur d'espérance et dépend donc du PGD g^n inconnu. Il est donc naturel d'envisager de maximiser la contrepartie d'échantillon de $\bar{L}_n(\theta)$:

$$L_n(z^n; \theta) = \frac{1}{n} \sum_{t=1}^n \log f_t(z_t | z^{t-1}; \theta), \quad (1.7)$$

qui est connue sous le nom de **fonction quasi-log-vraisemblance**. Le maximiseur de $L_n(z^n; \theta), \tilde{\theta}_n$, est connu sous le nom d'estimateur du quasi-maximum de vraisemblance (**EQMV**) de θ . Le préfixe «quasi» est utilisé pour indiquer que cette solution peut être obtenue à partir d'une **fonction de log-vraisemblance** mal spécifiée. Lorsque $\{f_t\}$ est spécifiée correctement dans son intégralité pour $\{z_t\}$, l'**EQMV** n'est que l'**EMV** standard.

Spécifier un modèle de probabilité complet pour z^n n'est pas possible en pratique car il implique trop de variables aléatoires (n vecteurs aléatoires z_t , chacun avec ν variables aléatoires). Au lieu de cela, les économétriciens sont généralement intéressés par la modélisation d'une variable d'intérêt, disons, y_t , conditionnelle à un ensemble de variables "prédéterminées", par exemple, x_t , où x_t comprend certains éléments de w_t et z^{t-1} . C'est à dire seul le comportement conditionnel de y_t doit être pris en compte. Comme les w_t ne sont pas explicitement modélisés, la densité conditionnelle $g_t(y_t | x_t)$ ne fournit qu'une description partielle de $\{z_t\}$. On trouve alors une **fonction de quasi-vraisemblance** $f_t(y_t | x_t; \theta)$ pour approcher $g_t(y_t | x_t)$. De manière analogue à (1.5), le **CIKL** moyen résultant de g_t par rapport à f_t est

$$\bar{\mathbb{I}}_n(\{g_t : f_t\}; \theta) = \frac{1}{n} \sum_{t=1}^n \mathbb{I}(g_t : f_t; \theta). \quad (1.8)$$

Soit $y^n = (y_1, \dots, y_n)$ et $x^n = (x_1, \dots, x_n)$. Minimiser $\bar{\mathbb{I}}_n(\{g_t : f_t\}; \theta)$ dans (1.8) équivaut donc à maximiser

$$\bar{L}_n(\theta) = \frac{1}{n} \sum_{t=1}^n E[\log f_t(y_t | x_t; \theta)]; \quad (1.9)$$

cf. (1.6). Le paramètre θ^* qui maximise (1.9) doit également minimiser le CIKL moyen (1.8). S'il existe un $\theta_0 \in \Theta$ tel que $f_t(y_t|x_t; \theta_0) = g_t(y_t|x_t)$ pour tout t , on dit que $\{f_t\}$ est correctement spécifiée pour $\{y_t|x_t\}$. Dans ce cas, $\bar{\mathbb{I}}_n(\{g_t : f_t\}; \theta_0)$ est nul, et $\theta^* = \theta_0$.

Comme précédemment, $\bar{L}_T(\theta)$ n'est pas directement observable, nous maximisons plutôt :

$$L_n(y^n, x^n; \theta) = \frac{1}{n} \sum_{t=1}^n \log f_t(y_t|x_t; \theta), \quad (1.10)$$

qui sera également appelée **fonction quasi-log-vraisemblance**. La solution résultante est l'**EQMV** $\tilde{\theta}_n$. Lorsque $\{f_t\}$ est correctement spécifiée pour $\{y_t|x_t\}$, l'**EQMV** $\tilde{\theta}_n$ est également l'**EMV** habituel.

Un exemple majeur est la spécification suivante de la normalité conditionnelle avec $\mu_t(x_t; \theta)$ comme spécification de sa moyenne :

$$y_t|x_t \sim N(\mu_t(x_t; \beta), \sigma^2).$$

En posant $\theta = (\beta' \sigma^2)'$, il est facile de voir que le maximiseur de la fonction quasi-log-vraisemblance $n^{-1} \sum_{t=1}^n \log f_t(y_t|x_t; \theta)$ est aussi la solution de

$$\min_{\beta} \frac{1}{T} \sum_{t=1}^n [y_t - \mu_t(x_t; \beta)]' [y_t - \mu_t(x_t; \beta)].$$

L'**EQMV** résultant de β est donc l'estimateur des **MCN**. Par conséquent, l'estimateur des **MCN** peut être considéré comme un **EQMV** sous l'hypothèse de normalité conditionnelle avec homoscedasticité conditionnelle.

1.4.2 Propriétés asymptotiques de l'EQMV

La fonction quasi-log-vraisemblance est, en général, une fonction non linéaire en θ , de sorte que l'**EQMV** doit être calculé numériquement à l'aide

d'un algorithme d'optimisation non linéaire. Etant donné que maximiser L_n équivaut à minimiser $-L_n$, les algorithmes discutés dans la section précédente, sont facilement appliqués. Pour l'étude des propriétés asymptotiques de l'**EQMV**, nous supposons toujours que la fonction quasi-log-vraisemblance spécifiée est deux fois continument différentiable sur l'espace de paramètres compact Θ avec probabilité 1 et que l'intégration et la différenciation peuvent être interchangées.

Consistance Pour établir le résultat de la consistance c.à.d.

$$\tilde{\theta}_n \xrightarrow[n \rightarrow \infty]{p.s} \theta^*,$$

on a besoin de faire 3 hypothèses :

Hypothèse 1 : $\theta \in \Theta$, $\Theta \subset \mathbb{R}^k$ compact.

Hypothèse 2 : Le critère $L_n(z^n; \theta)$ (ou $L_n(y^n, x^n; \theta)$) converge presque sûrement et uniformément par rapport à θ vers $\bar{L}_n(\theta)$.

Hypothèse 3 : Il existe un θ^* unique qui minimise le **CIKL** (1.5) ou (1.8).

Cela montre que l'**EQMV** est un estimateur consistant pour le minimiseur de **CIKL** $\mathbb{I}(g^n : f^n; \theta)$, que la spécification soit correcte ou non. Lorsque $\{f_t\}$ est correctement spécifiée dans son intégralité, le minimiseur θ^* est également le vrai paramètre θ_0 . Par conséquent, les conditions requises pour assurer la consistance de l'**EQMV** sont également celles d'une **LGNU**.

Normalité asymptotique Considérons la spécification de la fonction quasi-log-vraisemblance pour $z^n : L_n(z^n; \theta)$. Quand θ^* est à l'intérieur de Θ , le développement de Taylor standard de $\frac{\partial L_n(z^n; \tilde{\theta}_n)}{\partial \theta}$ autour de θ^* à l'ordre 1 est

$$\frac{\partial L_n(z^n; \tilde{\theta}_n)}{\partial \theta} = \frac{\partial L_n(z^n; \theta^*)}{\partial \theta} + \frac{\partial^2 L_n(z^n; \theta^*)}{\partial \theta^2} (\tilde{\theta}_n - \theta^*) = 0, \quad (1.11)$$

où θ_n^\dagger est entre $\tilde{\theta}_n$ et θ^* . Alors, tant que $\nabla^2 L_n(z^n; \theta_n^\dagger)$ est inversible avec probabilité 1, (1.11) peut s'écrire

$$\sqrt{n}(\tilde{\theta}_n - \theta^*) = -\nabla^2 L_n(z^n; \theta_n^\dagger)^{-1} \sqrt{n} \nabla L_n(z^n; \theta^*).$$

La distribution asymptotique de $\sqrt{n}(\tilde{\theta}_n - \theta^*)$ est essentiellement déterminée par la distribution asymptotique du score : $\sqrt{n} \nabla L_n(z^n; \theta^*)$. Soit B_n la matrice de variance-covariance du score $\sqrt{n} \nabla L_n(z^n; \theta)$:

$$B_n(\theta) = \text{var}(\sqrt{n} \nabla L_n(z^n; \theta)),$$

Alors à condition que $\log f_t(z_t | z^{t-1})$ vérifie un **TCL**, on a

$$B_n(\theta^*)^{-1/2} \sqrt{n}(\nabla L_n(z^n; \theta^*) - E[\nabla L_n(z^n; \theta^*)]) \xrightarrow{D} N(0, I_k). \quad (1.12)$$

Soit la matrice Hessienne limite

$$E[\nabla^2 L_n(z^n; \theta)] = H_n(\theta).$$

Quand $\nabla^2 L_n(z^n; \theta_n^\dagger)$ vérifie une **LGNU**, on a

$$\nabla^2 L_n(z^n; \theta) - H_n(\theta) \xrightarrow{P} 0,$$

uniformément dans Θ . On pourra montrer finalement que

$$C_n(\theta^*)^{-1/2} \sqrt{n}(L_n(z^n; \theta^*) - E[L_n(z^n; \theta^*)]) \xrightarrow{D} N(0, I_k),$$

où $C_n(\theta^*) = H_n(\theta^*)^{-1} B_n(\theta^*) H_n(\theta^*)^{-1}$, avec $H_n(\theta^*) = E[\nabla^2 L_n(z^n; \theta^*)]$ et $B_n(\theta^*) = \text{var}(\sqrt{n} \nabla L_n(z^n; \theta^*))$.

CHAPITRE 2

Estimation des MC du modèle **PEXP**AR(**p**)

restreint

Sommaire

2.1 Introduction	24
2.2 Modèle <i>EXPAR</i> restreint périodique	26
2.3 Estimation des paramètres	35
2.4 Résultats de la simulation	40

2.1 Introduction

La classe des modèles **EXPAR** introduits par **Ozaki (1980)**[\[41\]](#) et **Haggan et Ozaki (1981)**[\[26\]](#) ont montré leur pertinence pour capturer certaines caractéristiques bien connues de la théorie des vibrations non linéaires telles que la fréquence dépendante de l'amplitude, les phénomènes de saut

et le comportement de cycle limite, ces modèles sont de forme autorégressive avec des coefficients exponentiels dépendant de l'amplitude. De nombreux résultats et méthodes y compris la stationnarité, l'ergodicité géométrique, l'estimation, la prévision et les tests ont été étudiés comme **Ozaki (1982, 1985)** [41] [43], **Chan et Tong (1985)** [13], **Al-Kassam et Lane (1989)** [1], **Koul et Schick (1997)** [33] et **Allal et El Melhaoui (2006)** [2].

Des modèles de séries chronologiques périodiques ont été largement utilisés au cours des dernières décennies pour décrire de nombreuses séries à dynamique périodique. L'incapacité des modèles **SARIMA** à représenter adéquatement de nombreuses séries temporelles saisonnières présentant une structure d'autocovariance périodique a motivé la recherche dans les processus périodiquement corrélés. Cette notion, introduite par **Gladyshev (1960)** [23], a été exploitée dans une variété de nouvelles classes de modèles chronologiques parmi lesquels le **GARCH** périodique (**Bollerslev et Ghysels (1996)**) [12], le bilinéaire périodique (**Bibi et Gautier (2005)**) [9] et le modèle autorégressif périodique de mélange (**Shao (2006)**) [45]. Dans ce chapitre, nous étendons la classe du modèle autorégressif exponentiel périodique restreint (**PEXP(1)**) discutée dans **Merzougui et al. (2016)** [37] à l'ordre p . Les séries de **PEXP** satisfont une équation de différence non linéaire semblable à celle pour des modèles **EXPAR**(p) avec des paramètres et des variances de bruit blanc qui changent périodiquement avec la saison.

Ce chapitre traite l'estimation par la méthode des moindres carrés du modèle périodique **EXPAR**(\mathbf{p}) restreint. **Haggan et Ozaki (1981)** [26] ont proposé la méthode la plus rapide qui consiste à fixer le paramètre non linéaire dans le terme exponentiel au niveau d'une grille de valeurs puis à estimer les autres paramètres par la méthode des moindres carrés linéaires en utilisant le critère **AIC** pour sélectionner les paramètres finaux. **Tjøstheim (1986)** [50] a traité

le problème de l'estimation des séries temporelles non linéaires dans un cadre général, il donne les estimations conditionnelles MC et MV, **Amondela et Francq (2009)** [4] donnent l'estimateur QMV pour le modèle $\text{EXPAR}(1)$, **Shi et Aoyama (1997)** [46] et **Baragona et al. (2002)** [6] ont utilisé un algorithme génétique pour estimer les paramètres, **Shi et al. (2001)** [47] ont remarqué que la fonction objective du coefficient non linéaire n'est pas convexe et que de multiples minima locaux peuvent exister et qu'une telle estimation prend du temps, ils ont proposé, lorsque l'estimation en temps réel est nécessaire, une estimation heuristique des données pour le paramètre non linéaire. Nous pouvons trouver une application pour cette méthode dans **Messaoud et al. (2006)** [40] pour la modélisation des vibrations et des perturbations lors du forage. **Ghosh et al. (2015)** [22] ont utilisé le filtre de Kalman étendu (FKE) pour estimer les modèles EXPAR .

Le reste de ce chapitre est organisé comme ceci. Dans la section 2.2, nous présentons les modèles $\text{PEXP}(\mathbf{p})$ restreints et fournissons quelques notations de base et hypothèses techniques. Dans la section 2.3, nous estimons les paramètres par la méthode des moindres carrés, la consistance et la normalité asymptotique des estimateurs MC y sont établies. Enfin, la performance des estimateurs est présentée via une petite simulation dans la section 2.4.

2.2 Modèle EXPAR restreint périodique

Soit $\{Y_t; t \in \mathbb{Z}\}$ un processus stochastique saisonnier de période S ($S \geq 2$).

Définition 2.1

Le processus $\{Y_t; t \in \mathbb{Z}\}$ est un modèle AutoRégressif EXPonentiel Restreint Périodique d'ordre 1 ($\text{PEXP}AR(1)$) s'il est solution de l'équation aux

différences stochastique non linéaire donnée par

$$Y_t = (\varphi_{t,1} + \varphi_{t,2} \exp(-\gamma Y_{t-1}^2)) Y_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (2.1)$$

où $\{\varepsilon_t; t \in \mathbb{Z}\}$ est un processus iid $(0, \sigma_t^2)$. $\varphi_{t,1}$ et $\varphi_{t,2}$ sont les paramètres autorégressifs et $\gamma > 0$ est le paramètre non linéaire supposé connu. Une détermination heuristique de γ à partir des données est $\hat{\gamma} = -\frac{\log \epsilon}{\max_{1 \leq t \leq n} Y_t^2}$, où ϵ est un nombre petit et n est le nombre des observations. (Shi et al. (2001)) [47]. Les paramètres autorégressifs et la variance sont périodiques de période S , c.à.d :

$$\varphi_{t+kS,1} = \varphi_{t,1}, \quad \varphi_{t+kS,2} = \varphi_{t,2} \quad \text{et} \quad \sigma_{t+kS}^2 = \sigma_t^2, \quad \forall k, t \in \mathbb{Z}.$$

Pour mettre en valeur la périodicité, soit $t = i + S\tau, i = 1, \dots, S$ et $\tau \in \mathbb{Z}$, alors l'équation (2.1) devient

$$Y_{i+S\tau} = (\varphi_{i,1} + \varphi_{i,2} \exp(-\gamma Y_{i+S\tau-1}^2)) Y_{i+S\tau-1} + \varepsilon_{i+S\tau}, \quad i = 1, \dots, S, \quad \tau \in \mathbb{Z} \quad (2.2)$$

Dans l'équation (2.2), $Y_{i+S\tau}$ est la valeur de Y_t durant la $i^{\text{ème}}$ saison du cycle τ et $\varphi_{i,1}, \varphi_{i,2}$ sont les paramètres du modèle à la saison i . C'est clair que les paramètres dépendent de $Y_{i+S\tau-1}$ dans le sens où pour $|Y_{i+S\tau-1}|$ grand nous avons $\varphi_{i,1} + \varphi_{i,2} \exp(-\gamma Y_{i+S\tau-1}^2) \sim \varphi_{i,1}$ et pour $|Y_{i+S\tau-1}|$ petit: $\varphi_{i,1} + \varphi_{i,2} \exp(-\gamma Y_{i+S\tau-1}^2) \sim \varphi_{i,1} + \varphi_{i,2}$. En application, le modèle **PEXP**AR(1) restreint est approprié aux séries temporelles saisonnières affichant des caractéristiques non linéaires tels que les cycles limites.

Ces formes de modèles sont nouvelles dans la littérature des séries temporelles, il est intéressant de faire plusieurs simulations pour voir leurs caractéristiques. Un fait important est leur propriété de non normalité comme on peut le voir dans les histogrammes des Figures [2.1 – 2.3] et confirmé par le test de **Shapiro Wilk** où la valeur p est inférieure à 0.05. Les réalisations

des processus A et B sont données dans les figures [2.1 – 2.2] et à partir de leurs corrélogrammes on peut voir que les processus sont stationnaires dans chaque saison à cause de la décroissance rapide vers 0 quand h augmente. Le modèle (C) est non stationnaire par construction, le corrélogramme montre une décroissance lente vers 0 qui est typiquement une signature des séries non stationnaires. Un autre fait intéressant, que ces type de modèles peuvent exhibés, est le comportement de cycle limite. Un tel phénomène est montré dans la figure [2.4] à partir du modèle (D).

$$\text{Modèle (A) : } \begin{cases} Y_{1+2\tau} = (-0.9 + 1.1 \exp(-Y_{2\tau}^2)) Y_{2\tau} + \varepsilon_{1+2\tau} \\ Y_{2+2\tau} = (0.9 - \exp(-Y_{1+2\tau}^2)) Y_{1+2\tau} + \varepsilon_{2+2\tau} \end{cases}$$

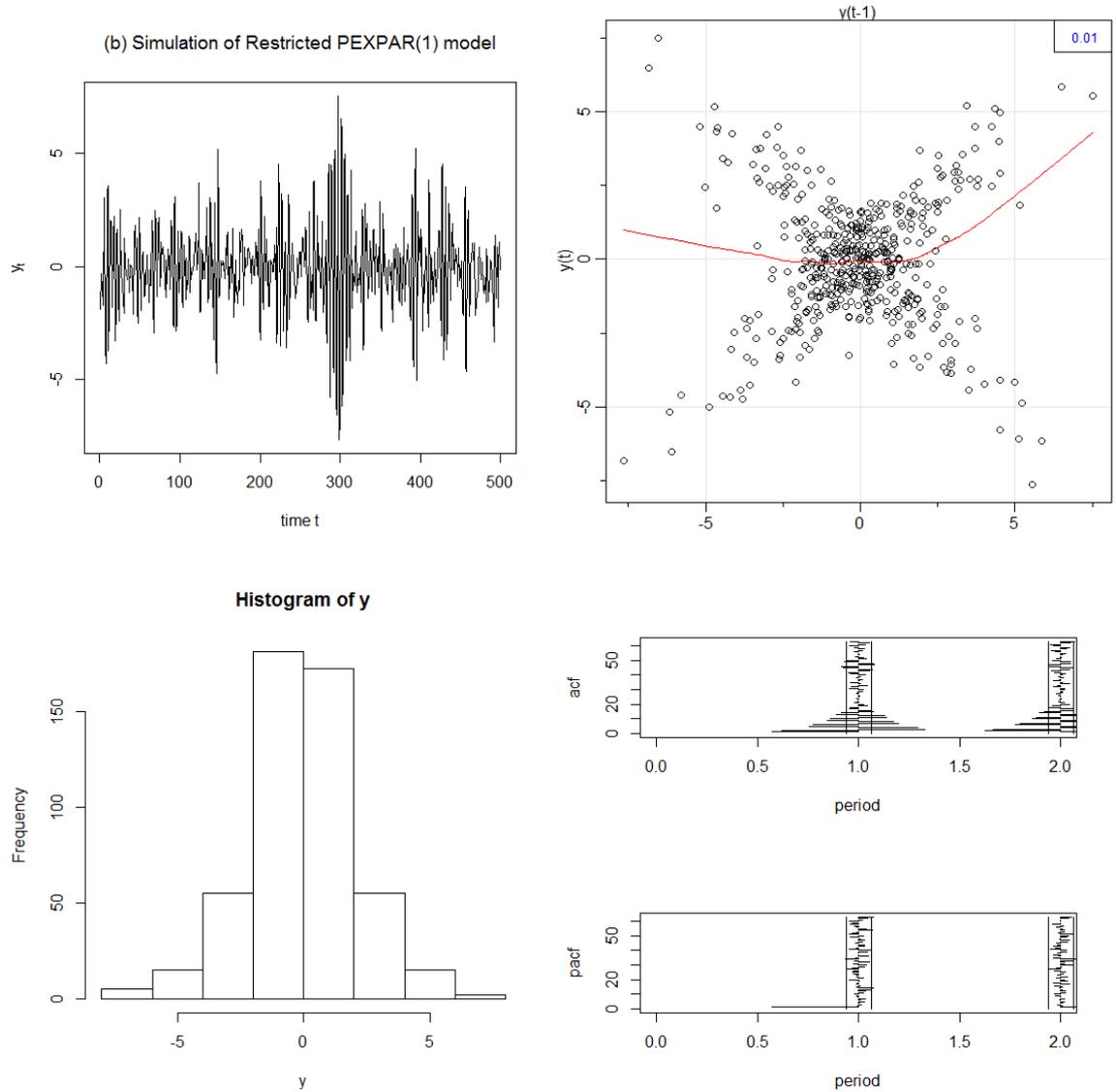


Figure 2.1 : Réalisation du processus (A) avec nuage de points, histogramme et corrélogramme correspondants, $n=500$, valeur p (Shapiro-Wilk normality test) = $3.596e-07$.

$$\text{Modèle (B) : } \begin{cases} Y_{1+2\tau} = (-0.3 + 2 \exp(-Y_{2\tau}^2)) Y_{2\tau} + \varepsilon_{1+2\tau} \\ Y_{2+2\tau} = (-0.8 + \exp(-Y_{1+2\tau}^2)) Y_{1+2\tau} + \varepsilon_{2+2\tau} \end{cases}$$

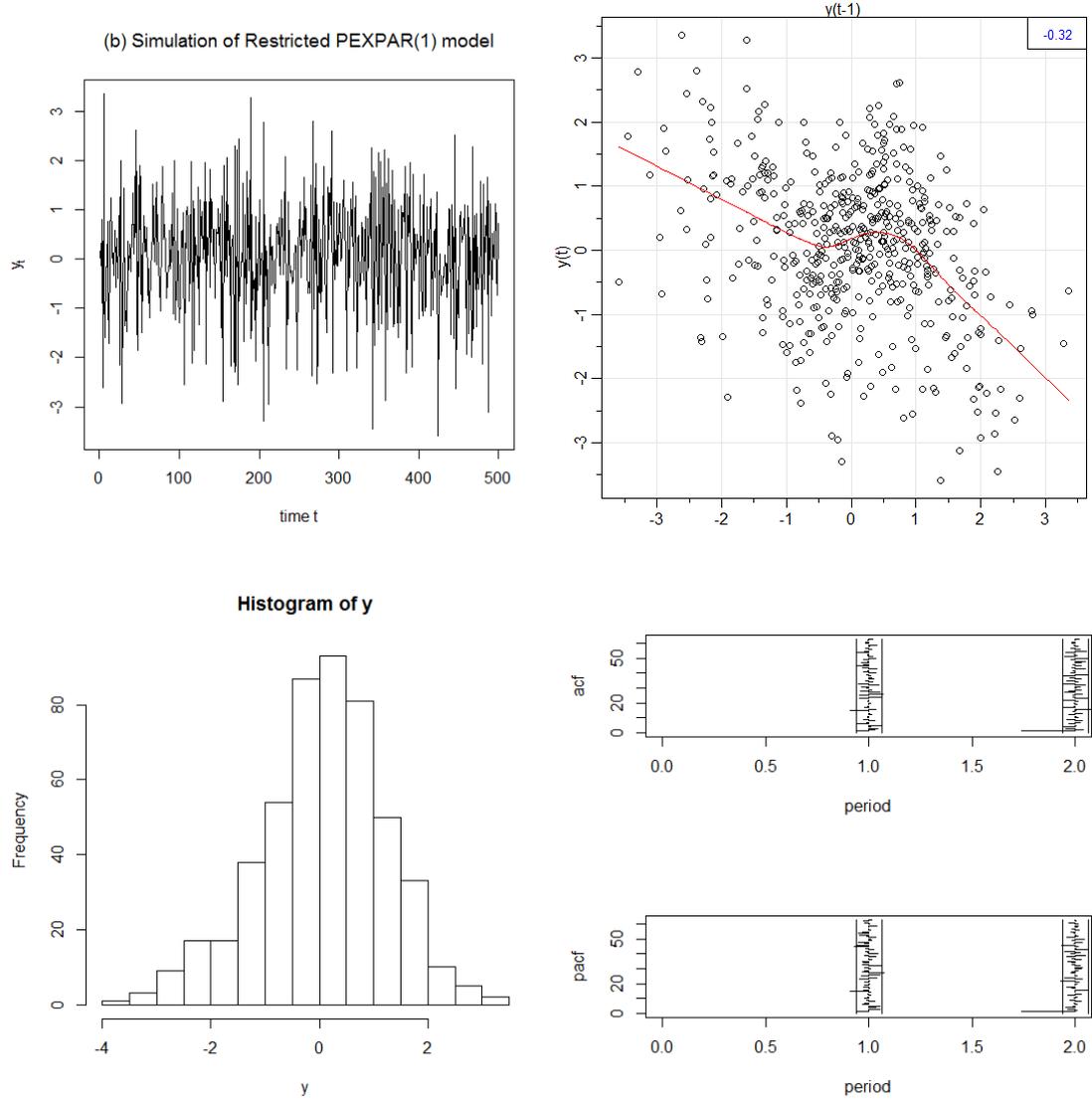


Figure 2.2 : Réalisation du processus (B) avec nuage de points, histogramme et corrélogramme correspondants, $n=500$, valeur p (Shapiro-Wilk normality test) = 0.004525.

$$\text{Modèle (C) : } \begin{cases} Y_{1+2\tau} = (1 - 0.7 \exp(-Y_{2\tau}^2)) Y_{2\tau} + \varepsilon_{1+2\tau} \\ Y_{2+2\tau} = (0.95 - 2 \exp(-Y_{1+2\tau}^2)) Y_{1+2\tau} + \varepsilon_{2+2\tau} \end{cases}$$

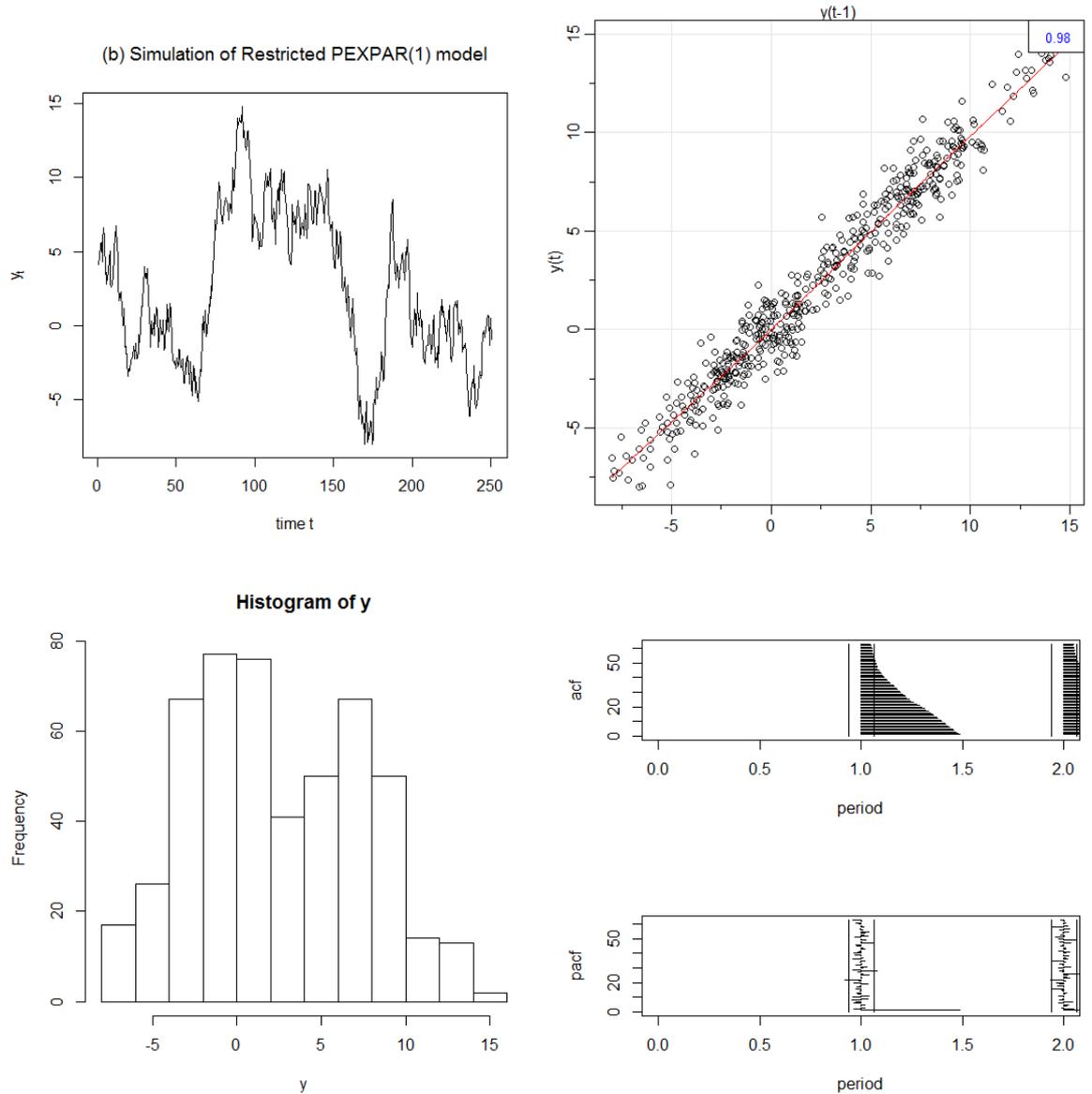


Figure 2.3 : Réalisation du processus (C) avec nuage de points, histogramme et corrélogramme correspondants, $n=500$, valeur p (Shapiro-Wilk normality test) = $3.596e-07$.

$$\text{Modèle (D) : } \begin{cases} Y_{1+2\tau} = (0.2 - 1.5 \exp(-Y_{2\tau}^2)) Y_{2\tau} + \varepsilon_{1+2\tau} \\ Y_{2+2\tau} = (0.8 + 0.3 \exp(-Y_{1+2\tau}^2)) Y_{1+2\tau} + \varepsilon_{2+2\tau} \end{cases}$$

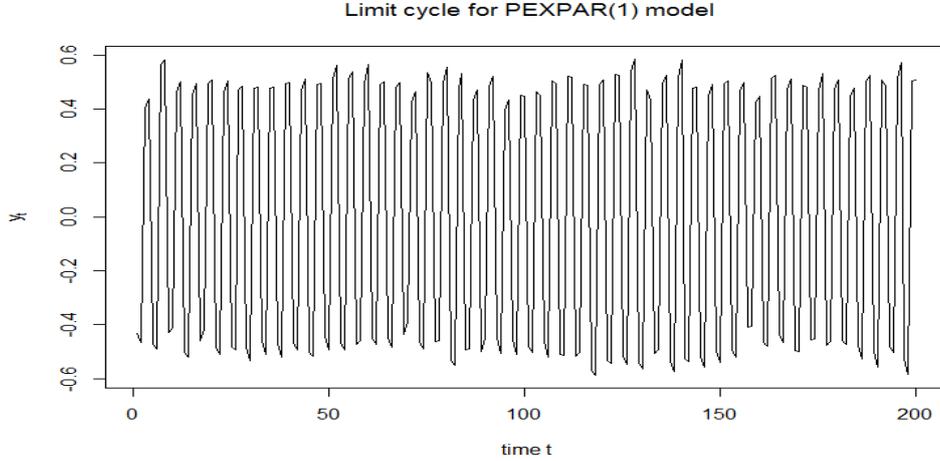


Figure 2.4 : Cycle limite du modèle $PEXP$ AR₂(1).

Augmenter l'ordre du modèle **PEXP**AR(**1**) restreint peut capturer toutes les variations non linéaires dans la série temporelle et comme pour le modèle **PEXP**AR cela permettra aux paramètres de dépendre de Y_{t-1} , d'où la définition suivante.

Définition 2.2

Le processus $\{Y_t; t \in \mathbb{Z}\}$ est dit suivre le modèle AutoRégressif *EX*Ponentiel Restreint Périodique $PEXP$ AR_S(p_t), de période S , s'il est solution de l'équation aux différences non linéaire périodique stochastique de la forme :

$$Y_t = \sum_{j=1}^{p_t} (\varphi_{t,j} + \pi_{t,j} \exp(-\gamma Y_{t-1}^2)) Y_{t-j} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (2.3)$$

où $\{\varepsilon_t; t \in \mathbb{Z}\}$ est un processus périodique *i.i.d.* avec une densité continue $f_{\sigma_t}(\cdot)$, pas nécessairement Gaussienne, de moyenne 0 et de variance finie σ_t^2 .

Les paramètres autorégressifs $\varphi_{t,j}, \pi_{t,j} \forall t \in \mathbb{Z}$ et $j = 1, \dots, p$, l'ordre p_t et la variance des innovations σ_t^2 sont périodiques, dans le temps, de période S , *i.e.*,

$\varphi_{t+kS,j} = \varphi_{t,j}$, $\pi_{t+kS,j} = \pi_{t,j}$, $p_{t+kS} = p_t$ et $\sigma_{t+kS}^2 = \sigma_t^2$, $\forall k, t \in \mathbb{Z}$ et $j = 1, \dots, p_t$.

Le paramètre non linéaire, $\gamma > 0$, est connu.

En mettant $t = i + S\tau$, $i = 1, 2, \dots, S$ et $\tau \in \mathbb{Z}$ et en prenant $p = \max_{i \in \{1, 2, \dots, S\}} p_i$

où $\varphi_{i,j} = 0$, $\pi_{i,j} = 0$, pour chaque $j > p_i$, on peut réécrire l'équation (2.3)

sous la forme équivalente :

$$Y_{i+S\tau} = \sum_{j=1}^p (\varphi_{i,j} + \pi_{i,j} \exp(-\gamma Y_{i+S\tau-1}^2)) Y_{i+S\tau-j} + \varepsilon_{i+S\tau}, \quad i = 1, \dots, S, \quad \tau \in \mathbb{Z} \quad (2.4)$$

Pour Y_{i+rS-1} grand, nous avons $\varphi_{i,j} + \pi_{i,j} \exp(-\gamma Y_{i+rS-1}^2) \rightsquigarrow \varphi_{i,j}$, $j = 1, \dots, p$

et pour X_{i+rS-1} petit: $\varphi_{i,j} + \pi_{i,j} \exp(-\gamma X_{i+rS-1}^2) \rightsquigarrow \varphi_{i,j} + \pi_{i,j}$, $j = 1, \dots, p$.

En pratique, quand nous avons une série temporelle saisonnière avec des caractéristiques de non linéarité comme les cycles limites et en plus la moyenne et la structure de covariance qui change avec la saison, on pourra modéliser avec un modèle **EXPAR** $\mathbf{AR}(\mathbf{p})$ périodique qui représente une fusion entre la présence de non linéarité et la configuration périodique. La figure [2.5] montre une série simulé à partir du modèle **PEXP** $\mathbf{AR}_2(2)$ restreint avec $\underline{\varphi} = (0.5, -2, 0.3, 1; -0.5, 1, -0.4, 1.8)'$, $\gamma = 1$ et $n = 500$ qui vérifie la condition suffisante de la stricte stationnarité (voir l'hypothèse (A_1)). Dans la figure [2.6] nous avons le graphe saisonnier (un monthplot) de la série où on voit que la moyenne dépend de la saison. La figure [2.7] donne les nuages de points qui indiquent clairement un comportement non linéaire et la figure [2.8] donne les corrélogrammes simples et partiels des séries.

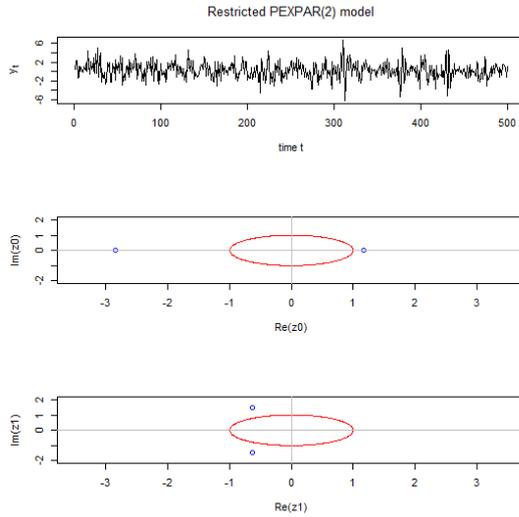


Figure 2.5 : Simulation du $PEXP\text{AR}_2(2)$ restreint et racines inverses de l'équation caractéristique.

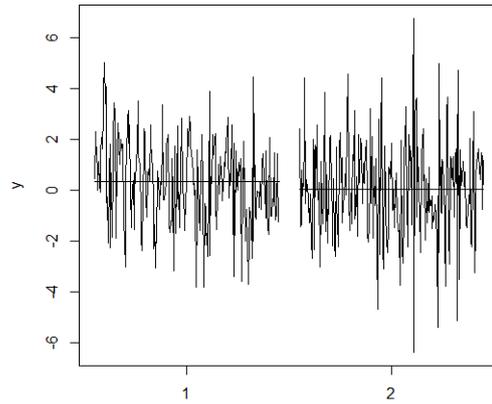


Figure 2.6 : Monthplot de la série temporelle.

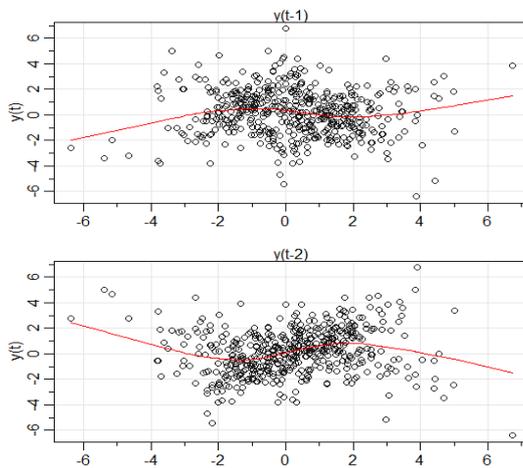


Figure 2.7 : Nuage de points de la série temporelle.

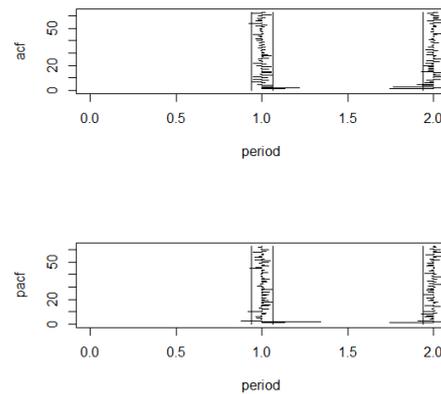


Figure 2.8 : FAC et FACP périodiques de la série temporelle.

Soit

$$\underline{\varphi}_i = (\varphi_{i,1}, \pi_{i,1}, \dots, \varphi_{i,p}, \pi_{i,p})', i = 1, \dots, S, \text{ et } \underline{\varphi} = (\underline{\varphi}'_1, \dots, \underline{\varphi}'_S)' \in \mathbb{R}^{2pS}.$$

Nous considérons les hypothèses suivantes :

A1 : Les paramètres du modèle AutoRégressifs EXPonentiels Périodiques $\underline{\varphi}$ satisfont la condition de stationnarité stricte périodique de (2.2). Une condition suffisante est : Toutes les racines de l'équation caractéristique associée

$$z^p - c_{i,1}z^{p-1} \dots - c_{i,p} = 0,$$

sont à l'intérieur du cercle unité, où $c_{i,j} = \max\{|\varphi_{i,j}|, |\varphi_{i,j} + \pi_{i,j}|\}$, $j = 1, \dots, p$; $i = 1, \dots, S$. Dans le cas non périodique, voir, par exemple, **De Gooijer (2017) [15]**, page 37.

A2 : Le processus périodiquement ergodique $\{Y_t; t \in \mathbb{Z}\}$ est tel que $E(Y_t^4) < \infty$, pour tout $t \in \mathbb{Z}$.

2.3 Estimation des paramètres

Nous considérons le problème de l'estimation des paramètres $\underline{\varphi}$ du modèle (2.4), qui est un problème d'optimisation linéaire, nous pouvons le résoudre en utilisant la procédure des moindres carrés.

Supposons que nous avons des observations $\{Y_1, \dots, Y_N\}$ de (2.4), $N = mS$, et définissons la somme conditionnelle des carrés

$$\begin{aligned}
 L_N(\underline{\varphi}) &= \sum_{i=1}^S L_{i,m}(\underline{\varphi}) \\
 &= \sum_{i=1}^S \left(\sum_{\tau=r+1}^{m-1} \left(Y_{S\tau+i} - E_{\underline{\varphi}}(Y_{S\tau+i} \mid B_{S\tau+i-1}) \right)^2 \right) \\
 &= \sum_{i=1}^S \left(\sum_{\tau=r+1}^{m-1} \left(Y_{S\tau+i} - \sum_{j=1}^p (\varphi_{i,j} + \pi_{i,j} \exp(-\gamma Y_{S\tau+i-1}^2)) Y_{S\tau+i-j} \right)^2 \right).
 \end{aligned}$$

où $r = \lfloor \frac{p}{S} \rfloor$, avec $[x]$ désigne la partie entière de x , $B_{S\tau+i-1}$ est la σ -algèbre générée par le passé du processus jusqu'au temps $S\tau + i - 1$ et $E_{\underline{\varphi}}(\cdot \mid \cdot)$ est l'espérance conditionnelle en supposant que $\underline{\varphi}$ est le vrai paramètre.

L'estimation $\widehat{\underline{\varphi}}_i = (\widehat{\varphi}_{i,1}, \widehat{\pi}_{i,1}, \dots, \widehat{\varphi}_{i,p}, \widehat{\pi}_{i,p})'$, pour une saison fixée i , est une solution aux équations d'estimation

$$\frac{\partial L_{i,m}(\underline{\varphi})}{\partial \varphi_{i,j}} = 0 \text{ et } \frac{\partial L_{i,m}(\underline{\varphi})}{\partial \pi_{i,j}} = 0, \quad j = 1, \dots, p.$$

La solution pour une saison fixée i est

$$\widehat{\underline{\varphi}}_i = \begin{pmatrix} M_{i,1,1} & \cdots & M_{i,1,p} \\ \vdots & \ddots & \vdots \\ M_{i,p,1} & \cdots & M_{i,p,p} \end{pmatrix}^{-1} \times \begin{bmatrix} \sum_{\tau=r+1}^{m-1} Y_{S\tau+i-1} Y_{S\tau+i} \\ \sum_{\tau=r+1}^{m-1} Y_{S\tau+i-1} Y_{S\tau+i} \exp(-\gamma Y_{S\tau+i-1}^2) \\ \vdots \\ \sum_{\tau=r+1}^{m-1} Y_{S\tau+i-p} Y_{S\tau+i} \\ \sum_{\tau=r+1}^{m-1} Y_{S\tau+i-p} Y_{S\tau+i} \exp(-\gamma Y_{S\tau+i-1}^2) \end{bmatrix}, \quad (2.5)$$

$$\widehat{\sigma}_i^2 = \frac{1}{m-r-1} \sum_{\tau=r+1}^{m-1} \left(Y_{S\tau+i} - \sum_{j=1}^p (\widehat{\varphi}_{i,j} + \widehat{\pi}_{i,j} \exp(-\gamma Y_{S\tau+i-1}^2)) Y_{S\tau+i-j} \right)^2, \quad (2.6)$$

où pour $j, k = 1, \dots, p$.

$$M_{i,j,k} = \begin{pmatrix} \sum_{\tau=r+1}^{m-1} Y_{S\tau+i-j} Y_{S\tau+i-k} & \sum_{\tau=r+1}^{m-1} Y_{S\tau+i-j} Y_{S\tau+i-k} \exp(-\gamma Y_{S\tau+i-1}^2) \\ \sum_{\tau=r+1}^{m-1} Y_{S\tau+i-j} Y_{S\tau+i-k} \exp(-\gamma Y_{S\tau+i-1}^2) & \sum_{\tau=r+1}^{m-1} Y_{S\tau+i-j} Y_{S\tau+i-k} \exp(-2\gamma Y_{S\tau+i-1}^2) \end{pmatrix}.$$

Théorème 2.1

Supposons que $\{Y_i\}$, satisfait (2.4), est périodiquement strictement stationnaire, alors les estimateurs des moindres carrés (2.5) et (2.6) sont fortement consistants avec $m \rightarrow \infty$. C'est à dire

$$\widehat{\underline{\varphi}}_i \xrightarrow{p.s} \underline{\varphi}_i \text{ et } \widehat{\sigma}_i^2 \xrightarrow{p.s} \sigma_i^2.$$

et on a

$$\sqrt{m}(\widehat{\underline{\varphi}}_i - \underline{\varphi}_i) \xrightarrow[m \rightarrow \infty]{D} N(0_{2p}, \sigma_i^2 \Gamma_i^{-1}).$$

où

$$\Gamma_i = \begin{pmatrix} \Gamma_{i,1,1} & \cdots & \Gamma_{i,1,p} \\ \vdots & \ddots & \vdots \\ \Gamma_{i,p,1} & \cdots & \Gamma_{i,p,p} \end{pmatrix},$$

et

$$\Gamma_{i,j,k} = \begin{pmatrix} E(Y_{i-j} Y_{i-k}) & E(Y_{i-j} Y_{i-k} \exp(-\gamma Y_{i-1}^2)) \\ E(Y_{i-j} Y_{i-k} \exp(-\gamma Y_{i-1}^2)) & E(Y_{i-j} Y_{i-k} \exp(-2\gamma Y_{i-1}^2)) \end{pmatrix}, \quad j, k = 1, \dots, p.$$

Preuve

En remplaçant par zéro les bornes inférieures des sommes de (2.5) et (2.6) on obtient les estimateurs comparables :

$$\underline{\tilde{\varphi}}_i = \begin{pmatrix} M_{i,1,1} & \cdots & M_{i,1,p} \\ \vdots & \ddots & \vdots \\ M_{i,p,1} & \cdots & M_{i,p,p} \end{pmatrix}^{-1} \times \begin{bmatrix} \sum_{\tau=0}^{m-1} Y_{S\tau+i-1} Y_{S\tau+i} \\ \sum_{\tau=0}^{m-1} Y_{S\tau+i-1} Y_{S\tau+i} \exp(-\gamma Y_{S\tau+i-1}^2) \\ \vdots \\ \sum_{\tau=0}^{m-1} Y_{S\tau+i-p} Y_{S\tau+i} \\ \sum_{\tau=0}^{m-1} Y_{S\tau+i-p} Y_{S\tau+i} \exp(-\gamma Y_{S\tau+i-1}^2) \end{bmatrix}, \quad (2.7)$$

$$\tilde{\sigma}_i^2 = \frac{1}{m} \sum_{\tau=0}^{m-1} \left(Y_{S\tau+i} - \sum_{j=1}^p (\tilde{\varphi}_{i,j} + \tilde{\pi}_{i,j} \exp(-\gamma Y_{S\tau+i-1}^2)) Y_{S\tau+i-j} \right)^2, \quad (2.8)$$

où pour $j, k = 1, \dots, p$.

$$M_{i,j,k} = \begin{pmatrix} \sum_{\tau=0}^{m-1} Y_{S\tau+i-j} Y_{S\tau+i-k} & \sum_{\tau=0}^{m-1} Y_{S\tau+i-j} Y_{S\tau+i-k} \exp(-\gamma Y_{S\tau+i-1}^2) \\ \sum_{\tau=0}^{m-1} Y_{S\tau+i-j} Y_{S\tau+i-k} \exp(-\gamma Y_{S\tau+i-1}^2) & \sum_{\tau=0}^{m-1} Y_{S\tau+i-j} Y_{S\tau+i-k} \exp(-2\gamma Y_{S\tau+i-1}^2) \end{pmatrix},$$

comme une approximation. $\underline{\tilde{\varphi}}_i$ et $\underline{\hat{\varphi}}_i$ ont la même distribution limite, il n'y a pas de différence si nous observons ou non les valeurs initiales en plus des données pour m large. (cf. **Brockwell et Davis (1991)** [13]; chapitre 8 est utilisé dans **Basawa et Lund (2001)** [7] pour le cas **PARMA**). Nous étudions donc les propriétés asymptotiques de $\underline{\tilde{\varphi}}_i$.

En remplaçant $Y_{S\tau+i}$ dans la formule (2.7) on obtient :

$$\underline{\tilde{\varphi}}_i = \underline{\varphi}_i + \begin{pmatrix} M_{i,1,1} & \cdots & M_{i,1,p} \\ \vdots & \ddots & \vdots \\ M_{i,p,1} & \cdots & M_{i,p,p} \end{pmatrix}^{-1} \times \begin{bmatrix} \sum_{\tau=0}^{m-1} Y_{S\tau+i-1} \varepsilon_{S\tau+i} \\ \sum_{\tau=0}^{m-1} Y_{S\tau+i-1} \exp(-\gamma Y_{S\tau+i-1}^2) \varepsilon_{S\tau+i} \\ \vdots \\ \sum_{\tau=0}^{m-1} Y_{S\tau+i-p} \varepsilon_{S\tau+i} \\ \sum_{\tau=0}^{m-1} Y_{S\tau+i-p} \exp(-\gamma Y_{S\tau+i-1}^2) \varepsilon_{S\tau+i} \end{bmatrix}. \quad (2.9)$$

De l'ergodicité périodique de $Y_{S\tau+i}$, on a :

$$\frac{1}{m} \sum_{\tau=0}^{m-1} \varepsilon_{S\tau+i} Y_{S\tau+i-j} \xrightarrow{p.s} E(\varepsilon_i) E(Y_{i-j}) = 0, \quad j = 1, \dots, p$$

$$\frac{1}{m} \sum_{\tau=0}^{m-1} \varepsilon_{S\tau+i} Y_{S\tau+i-j} \exp(-\gamma Y_{S\tau+i-1}^2) \xrightarrow{p.s} E(\varepsilon_i) E(Y_{i-j} \exp(-\gamma Y_{i-1}^2)) = 0, \quad j = 1, \dots, p.$$

Donc $\tilde{\varphi}_i \xrightarrow{p.s} \varphi_i$ quand $m \rightarrow \infty$.

De (2.8), on a

$$\begin{aligned} \tilde{\sigma}_i^2 &= \frac{1}{m} \sum_{\tau=0}^{m-1} \left(\sum_{j=1}^p (\varphi_{i,j} + \pi_{i,j} \exp(-\gamma Y_{S\tau+i-1}^2)) Y_{S\tau+i-j} + \varepsilon_{S\tau+i} \right. \\ &\quad \left. - \sum_{j=1}^p (\tilde{\varphi}_{i,j} + \tilde{\pi}_{i,j} \exp(-\gamma Y_{S\tau+i-1}^2)) Y_{S\tau+i-j} \right)^2 \\ &= \frac{1}{m} \sum_{\tau=0}^{m-1} \left(\varepsilon_{S\tau+i} - \sum_{j=1}^p (\tilde{\varphi}_{i,j} - \varphi_{i,j}) Y_{S\tau+i-j} - \sum_{j=1}^p (\tilde{\pi}_{i,j} - \pi_{i,j}) Y_{S\tau+i-j} \exp(-\gamma Y_{S\tau+i-1}^2) \right)^2. \end{aligned}$$

A partir des résultats précédents, on a : quand $m \rightarrow \infty$, $\tilde{\sigma}_i^2 \xrightarrow{p.s} E(\varepsilon_i^2) = \sigma_i^2$.

De (2.9) :

$$\sqrt{m}(\tilde{\varphi}_i - \varphi_i) = \left(\frac{1}{m} \begin{pmatrix} M_{i,1,1} & \cdots & M_{i,1,p} \\ \vdots & \ddots & \vdots \\ M_{i,p,1} & \cdots & M_{i,p,p} \end{pmatrix} \right)^{-1} \times \left(\frac{1}{\sqrt{m}} \begin{bmatrix} \sum_{\tau=0}^{m-1} Y_{S\tau+i-1} \varepsilon_{S\tau+i} \\ \sum_{\tau=0}^{m-1} Y_{S\tau+i-1} \exp(-\gamma Y_{S\tau+i-1}^2) \varepsilon_{S\tau+i} \\ \vdots \\ \sum_{\tau=0}^{m-1} Y_{S\tau+i-p} \varepsilon_{S\tau+i} \\ \sum_{\tau=0}^{m-1} Y_{S\tau+i-p} \exp(-\gamma Y_{S\tau+i-1}^2) \varepsilon_{S\tau+i} \end{bmatrix} \right).$$

Puisque $\{Y_{S\tau+i-j} \varepsilon_{S\tau+i}\}_\tau$ et $\{Y_{S\tau+i-j} \exp(-\gamma Y_{S\tau+i-1}^2) \varepsilon_{S\tau+i}\}_\tau$, $j = 1, \dots, p$ sont des suites de différences de martingale, alors nous appliquons le théorème

centrale limite pour les différences de martingale (cf. **Ibragimov, 1963**) [28] :

$$\frac{1}{\sqrt{m}} \begin{bmatrix} \sum_{\tau=1}^{m-1} Y_{S\tau+i-1} \varepsilon_{S\tau+i} \\ \sum_{\tau=1}^{m-1} Y_{S\tau+i-1} \exp(-\gamma Y_{S\tau+i-1}^2) \varepsilon_{S\tau+i} \\ \vdots \\ \sum_{\tau=1}^{m-1} Y_{S\tau+i-p} \varepsilon_{S\tau+i} \\ \sum_{\tau=1}^{m-1} Y_{S\tau+i-p} \exp(-\gamma Y_{S\tau+i-1}^2) \varepsilon_{S\tau+i} \end{bmatrix} \xrightarrow[m \rightarrow \infty]{D} N(\mathbf{0}_{2p}, \sigma_i^2 \Gamma_i),$$

ce qui complète la preuve avec le théorème de **Slutsky**. ■

2.4 Résultats de la simulation

La performance de l'estimation est montrée par une petite simulation. Le modèle **PEXP**₂(**2**) restreint est utilisé pour générer des séries temporelles pour les tailles $n = 600, 800, 1000$. Nous considérons $N = 1000$ répliques **Monte Carlo** et rapportons les estimateurs **MC**, leurs biais et leurs écarts-types. Le tableau [2.1] donne l'estimation avec les paramètres $\underline{\varphi} = (0.6, -1, 0.3, -0.5; -0.5, 1, -0.4, 0.8)'$, $\gamma = 1$ et le bruit blanc normal avec $(\sigma_1^2, \sigma_2^2)' = (0.6, 1)'$. Le choix des valeurs des paramètres a été fait de sorte que le modèle remplisse la condition A1, voir la figure [2.9]. Les boîtes à moustaches et les graphes Q-Q des erreurs sont donnés respectivement aux figures [2.10] et [2.11]. Le tableau [2.1] montre que les estimations sont proches des vraies valeurs et que l'écart-type décroît lorsque n devient plus grand et nous remarquons que l'écart-types de $\varphi_{i,j}$ sont plus petits que ceux de $\pi_{i,j}$. Ceci est confirmé par les boîtes à moustaches où nous observons que les erreurs sont plus consistantes pour $\varphi_{i,j}$ et la gamme est plus large pour $\pi_{i,j}$, mais dans tous les cas, les erreurs sont centrées sur 0. D'un autre côté, les graphes Q-Q montrent que les erreurs sont normales.

EMC	$\hat{\varphi}_{1,1}$	$\hat{\pi}_{1,1}$	$\hat{\varphi}_{1,2}$	$\hat{\pi}_{1,2}$	$\hat{\varphi}_{2,1}$	$\hat{\pi}_{2,1}$	$\hat{\varphi}_{2,2}$	$\hat{\pi}_{2,2}$
$n = 600$	0.5995	-0.9951	0.3019	-0.5036	-0.4953	0.9875	-0.3936	0.7876
<i>biais</i>	-0.0004	0.0048	0.0019	-0.0036	0.0046	-0.0124	0.0063	-0.0123
<i>écart – type</i>	0.0539	0.1957	0.0832	0.1409	0.1107	0.2793	0.1094	0.1726
$n = 800$	0.5988	-0.9893	0.3023	-0.5049	-0.5030	1.0045	-0.3950	0.7894
<i>biais</i>	-0.0011	0.0106	0.0023	-0.0049	-0.0030	0.0045	0.0049	-0.0105
<i>écart – type</i>	0.0461	0.1643	0.0723	0.1226	0.0979	0.2463	0.0959	0.1443
$n = 1000$	0.5993	-0.9921	0.3014	-0.5058	-0.5010	0.9907	-0.3949	0.7925
<i>biais</i>	-0.0006	0.0078	0.0014	-0.0058	-0.0010	-0.0092	0.0050	-0.0074
<i>écart – type</i>	0.0406	0.1457	0.0628	0.1067	0.0862	0.2198	0.0869	0.1326

TAB. 2.1 – Résultats d’estimations pour $PEXP\text{AR}_2(2)$ restreint.

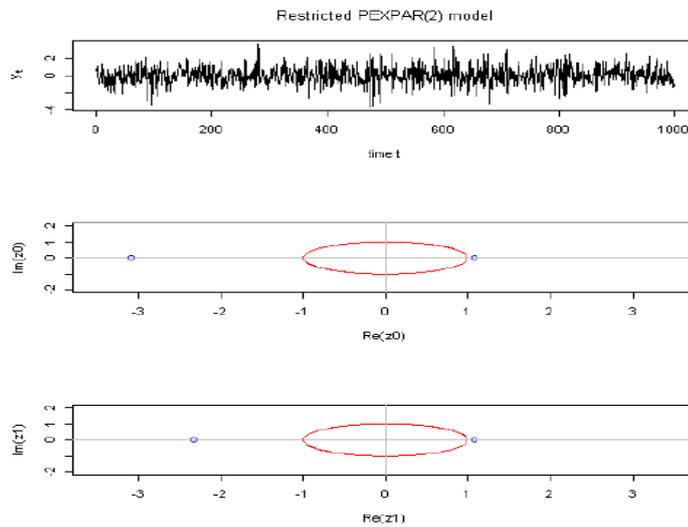


Figure 2.9 : Série simulée et racines inverses de l’équation caractéristique du $PEXP\text{AR}_2(2)$ restreint et $N = 1000$.

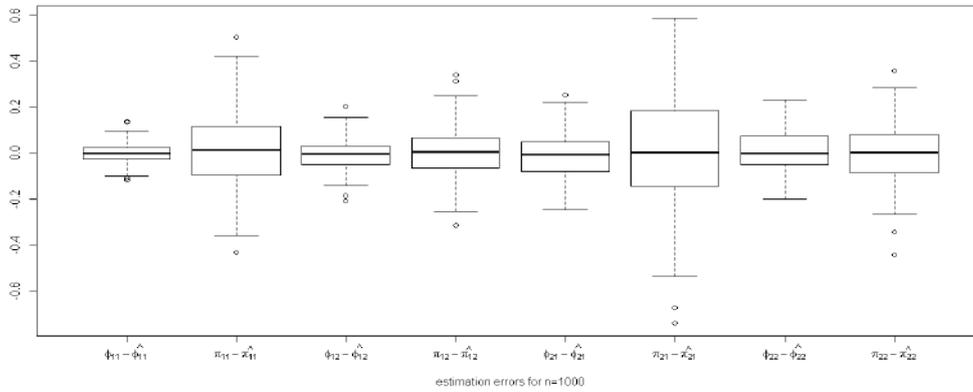


Figure 2.10 : Boîtes à moustaches d’erreurs d’estimations de 200 répétitions du $PEXP\text{AR}_2(2)$ restreint et $N = 1000$.

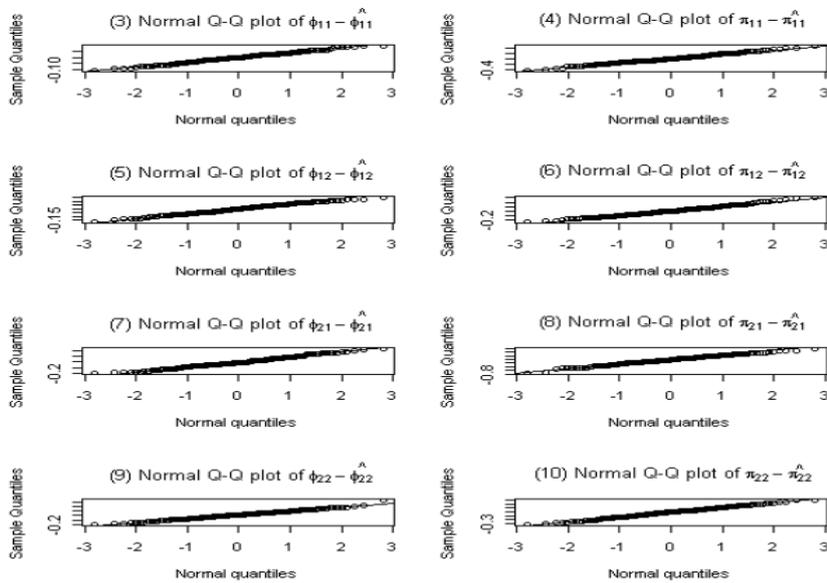


Figure 2.11 : Le graphe Q-Q d’erreurs d’estimations de 200 réplifications du $PEXP\text{AR}_2(2)$ restreint et $N = 1000$.

CHAPITRE 3

Estimation des MC non linéaire du modèle

PEXPAR(1)

Sommaire

3.1 Introduction	43
3.2 Définition de l'estimateur des MCN pour le mo- dèle PEXPAR_s(1)	45
3.3 Comportement asymptotique de l'estimateur des MCN	50
3.4 Résultats de la simulation	56

3.1 Introduction

Pour les modèles de séries chronologiques linéaires, il existe différentes procédures d'estimation mais de nombreux processus physiques et naturels

sont modélisés plus précisément par un modèle non linéaire que par une représentation linéaire. Le problème de l'estimation en contexte non linéaire est plus compliqué, il a été abordé par de nombreux auteurs voir par exemple **Tong (1990)** [52] et **Bierens (1994)** [10]. Les estimateurs des moindres carrés non linéaires (**MCN**) et du quasi maximum de vraisemblance (**QMV**) sont essentiellement utilisés dans ce cas. La première méthode est présentée dans ce chapitre. **Klimko et Nelson (1978)** [32] ont donné des conditions suffisantes pour assurer la consistance et la normalité asymptotique de l'estimateur des **MCN**. On peut trouver une exposition de cette méthode dans **Tong (1990)** [52] ou **Tjøstheim (1986)** [50]. L'approche de base sur l'estimation des **MCN** adoptée ici est celle de **Gallant et White (1988)** [21], une discussion complète peut être trouvée dans **Bierens (1994)** [10] et **Fuller (1996)** [20]. L'estimateur non linéaire des moindres carrés du paramètre θ est celui qui minimise la somme des carrés des résidus ou celui qui résout les conditions du premier ordre du problème de minimisation. L'application nécessite la stricte stationnarité du modèle et nous utiliserons les outils standards comme la compacité et le développement de Taylor. L'idée derrière le résultat de consistance est que si le critère converge presque sûrement vers une fonction non stochastique qui est minimisé au vrai paramètre, alors la limite des estimateurs devrait être la limite de la fonction. Pour le résultat de normalité asymptotique, les estimateurs sont approximés par une combinaison linéaire de moyennes empiriques, puis il suffit d'appliquer le théorème central limite.

Dans ce chapitre, nous prouvons la consistance et la normalité asymptotique de l'estimateur des **MCN** pour le modèle **PEXP**(1). On obtient les mêmes résultats qu'**Amendola et Francq (2009)** [4] pour leur **EQMV** pour l'**EXPAR**(1) non périodique et pour l'**EMCC** de **Tjøstheim (1986)** [50] qui

utilise des conditions légèrement différentes.

Le reste de ce chapitre est organisé comme ceci. Dans la section 3.2, nous définissons l'estimateur des MCN du modèle **PEXP**_S(1). Dans la section 3.3, nous étudions le comportement asymptotique en montrant la forte consistance et la normalité asymptotique. Enfin, une étude de simulation et un exemple réel sont donnés dans la section 3.4.

3.2 Définition de l'estimateur des MCN pour le modèle **PEXP**_S(1)

Le processus $\{Y_t; t \in \mathbb{Z}\}$ est dit suivre le modèle AutoRégressif EXPonentiel Périodique **PEXP**_S(1), de période S ($S \geq 2$), s'il satisfait :

$$Y_t = (\alpha_{t,1}^0 + \alpha_{t,2}^0 \exp(-\gamma_t^0 Y_{t-1}^2)) Y_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (3.1)$$

où $\{\varepsilon_t; t \in \mathbb{Z}\}$ est un processus périodique i.i.d. de moyenne 0 et de variance finie σ_t^2 et $\gamma_t > 0$. Les paramètres autorégressifs $\alpha_{t,1}, \alpha_{t,2}, \forall t \in \mathbb{Z}$, le paramètre non linéaire γ_t et la variance d'innovation σ_t^2 sont périodiques, dans le temps, de période S , i.e.,

$$\alpha_{t+kS,1} = \alpha_{t,1}, \alpha_{t+kS,2} = \alpha_{t,2}, \gamma_{t+kS} = \gamma_t \text{ et } \sigma_{t+kS}^2 = \sigma_t^2, \quad \forall k, t \in \mathbb{Z}.$$

En mettant $t = i + S\tau, i = 1, 2, \dots, S$ et $\tau \in \mathbb{Z}$, on peut réécrire l'équation (3.1) sous la forme équivalente :

$$Y_{i+S\tau} = (\alpha_{i,1}^0 + \alpha_{i,2}^0 \exp(-\gamma_i^0 Y_{i+S\tau-1}^2)) Y_{i+S\tau-1} + \varepsilon_{i+S\tau}, \quad i = 1, \dots, S, \quad \tau \in \mathbb{Z}. \quad (3.2)$$

Dans l'équation (3.2), $Y_{i+S\tau}$ est la valeur de Y_t pendant la $i^{\text{ème}}$ saison du cycle τ . Il est clair que les paramètres dépendent de $Y_{i+S\tau-1}$ dans le sens que pour $|Y_{i+S\tau-1}|$ grand on a :

$$\alpha_{i,1}^0 + \alpha_{i,2}^0 \exp(-\gamma_i^0 Y_{i+S\tau-1}^2) \rightsquigarrow \alpha_{i,1}^0,$$

tandis que pour $|Y_{i+S\tau-1}|$ petit:

$$\alpha_{i,1}^0 + \alpha_{i,2}^0 \exp(-\gamma_i^0 Y_{i+S\tau-1}^2) \rightsquigarrow \alpha_{i,1}^0 + \alpha_{i,2}^0.$$

Pour montrer que le comportement de cycle limite peut être présenté par de tels modèles, nous simulons le modèle **PEXP**_{AR}₂(1) ci-dessous :

$$\begin{cases} Y_{1+2\tau} = (0.3 - 2 \exp(-Y_{2\tau}^2)) Y_{2\tau} + \varepsilon_{1+2\tau} \\ Y_{2+2\tau} = (0.7 + 0.2 \exp(-2Y_{1+2\tau}^2)) Y_{1+2\tau} + \varepsilon_{2+2\tau} \end{cases}$$

Pour ignorer l'effet du bruit blanc, on pose $\varepsilon_t \sim N(0, 0.001)$. Une série simulée est donnée à la figure [3.1] avec une longueur $n = 200$. La figure [3.2] montre une série simulée du même modèle avec $\varepsilon_t \sim N(0, 0.1)$ avec le nuage de points correspondant, la densité marginale de Y_t et le corrélogramme. Nous pouvons voir que le processus est périodiquement stationnaire, non linéaire et non normal.

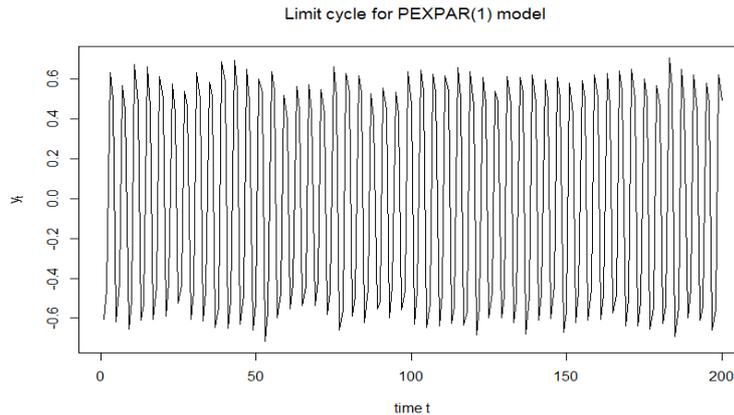


Figure 3.1 : Série simulée à partir du modèle $PEXP$ _{AR}₂(1).

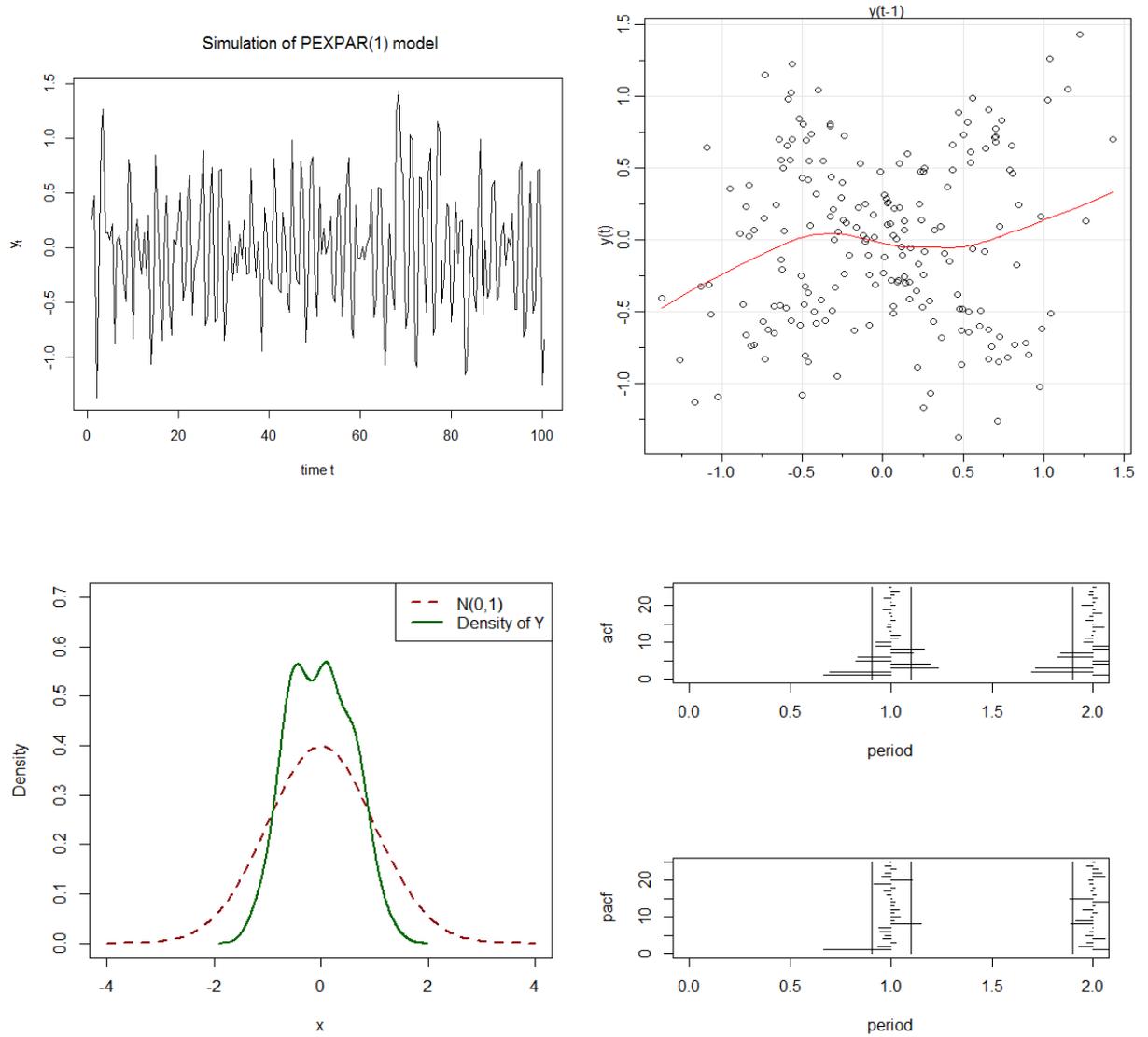


Figure 3.2 : Série simulée avec nuage de points, densité marginale et corrélogramme correspondants.

Le vecteur des paramètres

$$\theta = (\theta'_1, \dots, \theta'_S)' \in \mathbb{R}^{3S} \text{ où } \theta_i = (\alpha_{i,1}, \alpha_{i,2}, \gamma_i)', i = 1, \dots, S,$$

appartient à un espace de paramètres de la forme :

$$\Theta^S \subset (]-1, +1[\times]-1, +1[\times]0, \infty[)^S.$$

La vraie valeur des paramètres est inconnue et désignée par :

$$\theta_{i,0} = (\alpha_{i,1}^0, \alpha_{i,2}^0, \gamma_i^0)', \quad i = 1, \dots, S, \text{ et } \theta_0 = (\theta'_{1,0}, \dots, \theta'_{S,0})' \in \mathbb{R}^{3S}.$$

Supposons que les observations Y_1, \dots, Y_n constituent une réalisation de longueur n à partir de (3.2) et $n = mS$ ce qui signifie que nous avons m cycle complet de données. L'estimateur non linéaire des moindres carrés de θ est celui qui minimise la somme des carrés des résidus. Soit $\hat{\theta}_n$ l'estimateur des MCN qui est défini comme une solution mesurable du problème de minimisation suivant par rapport à θ :

$$\hat{\theta}_n = \arg \inf_{\theta \in \Theta^S} Q_n(\theta).$$

Où

$$Q_n(\theta) = \frac{1}{n} \sum_{t=1}^n \ell_t \text{ et } \ell_t(\theta) = (Y_t - (\alpha_{t,1} + \alpha_{t,2} \exp(-\gamma_t Y_{t-1}^2)) Y_{t-1})^2.$$

La valeur initiale est inconnue mais son choix n'est pas important pour le comportement asymptotique de l'estimateur des MCN donc on met $Y_0 = 0$ qui définit le critère opérationnel $\tilde{Q}_n(\theta)$, alors

$$\hat{\theta}_n = \arg \inf_{\theta \in \Theta^S} \tilde{Q}_n(\theta). \quad (3.3)$$

$$\begin{aligned} \tilde{Q}_n(\theta) &= \frac{1}{S} \sum_{i=1}^S \frac{1}{m} \sum_{\tau=0}^{m-1} (Y_{i+S\tau} - (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1})^2 \\ &= \frac{1}{S} \sum_{i=1}^S \tilde{Q}_{i,m}(\theta_i). \end{aligned}$$

Où

$$\begin{aligned} \tilde{Q}_{i,m}(\theta_i) &= \frac{1}{m} \sum_{\tau=0}^{m-1} \tilde{\ell}_{i,\tau}(\theta_i) \\ &= \frac{1}{m} \sum_{\tau=0}^{m-1} (Y_{i+S\tau} - (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1})^2. \end{aligned} \quad (3.4)$$

L'existence d'une solution à (3.3) est garantie sous l'hypothèse :

A1 : Θ^S est un ensemble de Borel compact $\subset \mathbb{R}^{3S}$.

L'hypothèse de compacité de Θ^S joue un rôle clé en garantissant l'existence de l'infimum sur Θ^S voir, par exemple, le **théorème 1.6.1 Bierens**[10] p.16.

Pour obtenir les estimateurs, nous différencions $\tilde{Q}_{i,m}(\theta_i)$ par rapport à chaque paramètre θ_i , $i = 1, \dots, S$. La condition du premier ordre du problème de minimisation des MCN est un système de $3S$ équations non linéaires avec $3S$ inconnues. Pour une saison i donnée, le vecteur gradient de $\tilde{Q}_{i,m}(\theta_i)$ est

$$\frac{\partial \tilde{Q}_{i,m}(\theta_i)}{\partial \theta_i} = -2 \frac{1}{m} \sum_{\tau=0}^{m-1} (Y_{i+S\tau} - (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1}) \begin{pmatrix} Y_{i+S\tau-1} \\ Y_{i+S\tau-1} \exp(-\gamma_i Y_{i+S\tau-1}^2) \\ -\alpha_{i,2} Y_{i+S\tau-1}^3 \exp(-\gamma_i Y_{i+S\tau-1}^2) \end{pmatrix}. \quad (3.5)$$

Alors les solutions analytiques ne peuvent être trouvées pour $\frac{\partial \tilde{Q}_{i,m}(\theta_i)}{\partial \theta_i} = 0$ et les estimateurs des MCN $\hat{\theta}_{i,m}$ doivent être calculés par optimisation telle que la méthode du **gradient** ou du **quasi Newton**.

Sous l'hypothèse que l'espérance existe, nous avons

$$E \left((Y_{i+S\tau} - (\alpha_{i,1}^0 + \alpha_{i,2}^0 \exp(-\gamma_i^0 Y_{i+S\tau-1}^2)) Y_{i+S\tau-1}) \begin{pmatrix} Y_{i+S\tau-1} \\ Y_{i+S\tau-1} \exp(-\gamma_i^0 Y_{i+S\tau-1}^2) \\ -\alpha_{i,2}^0 Y_{i+S\tau-1}^3 \exp(-\gamma_i^0 Y_{i+S\tau-1}^2) \end{pmatrix} \right) = 0, \quad (3.6)$$

qui est la version asymptotique de $\frac{\partial \tilde{Q}_{i,m}(\theta_i)}{\partial \theta_i} = 0$ en $\theta_{i,0}$ par le théorème ergodique.

3.3 Comportement asymptotique de l'estimateur des MCN

3.3.1 Consistance forte de l'estimateur des MCN

Nous faisons les hypothèses suivantes pour montrer la consistance forte.

A2 : i) Y_t est périodiquement strictement stationnaire, une condition suffisante pour cela est

$$|\alpha_{i,1}| + |\alpha_{i,2}| < 1 \text{ pour } i = 1, \dots, S.$$

ii) $E(\varepsilon_t^2) < \infty$, pour tout $t \in \mathbb{Z}$.

Le premier théorème énonce la forte consistance de l'estimateur MCN $\hat{\theta}_{i,m}$.

Théorème 3.1

Sous les hypothèses A1 – A2, les estimateurs non linéaires des moindres carrés $\hat{\theta}_{i,m}$ sont fortement consistant, i.e. presque sûrement,

$$\hat{\theta}_{i,m} \rightarrow \theta_{i,0} \text{ quand } m \rightarrow \infty.$$

Preuve

D'après A2 i), $\{\ell_{i,\tau}\}$ est un processus ergodique stationnaire et par ergodicité le critère théorique

$$Q_{i,m}(\theta_i) = \frac{1}{m} \sum_{\tau=0}^{m-1} (Y_{i+S\tau} - (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1})^2,$$

converge presque sûrement vers le critère asymptotique

$$Q_{i,\infty}(\theta_i) = E_{\theta_{i,0}}(\ell_{i,\tau}(\theta)) = E(Y_{i+S\tau} - (\alpha_{i,1}^0 + \alpha_{i,2}^0 \exp(-\gamma_i^0 Y_{i+S\tau-1}^2)) Y_{i+S\tau-1})^2.$$

Pour simplifier la notation, Soit

$$\begin{aligned} m_{i,\tau}(\theta_i) &= E(Y_{i+S\tau} | Y_{i+S\tau-1}, Y_{i+S\tau-2}, \dots) \\ &= (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1}. \end{aligned}$$

En itérant cette formule, nous obtenons

$$\begin{aligned}
 m_{i,\tau}(\theta_i) &= (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) \cdots (\alpha_{1,1} + \alpha_{1,2} \exp(-\gamma_1 Y_0^2)) Y_0 \\
 &\quad + (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) \cdots (\alpha_{2,1} + \alpha_{2,2} \exp(-\gamma_2 Y_1^2)) \varepsilon_1 \\
 &\quad + \cdots + (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) \varepsilon_{i+S\tau-1} \\
 &= \prod_{s=1}^i (\alpha_{s,1} + \alpha_{s,2} \exp(-\gamma_s Y_{s+S\tau-1}^2)) \left(\prod_{s=1}^S (\alpha_{s,1} + \alpha_{s,2} \exp(-\gamma_s Y_{s+S\tau-1}^2)) \right)^\tau Y_0 \\
 &\quad + T(\theta_i, \varepsilon_1, \dots, \varepsilon_{i+S\tau-1}).
 \end{aligned}$$

En remplaçant $Y_0 = 0$, on obtient $\tilde{m}_{i,\tau}(\theta_i)$ alors on a

$$m_{i,\tau}(\theta_i) - \tilde{m}_{i,\tau}(\theta_i) = \prod_{s=1}^i (\alpha_{s,1} + \alpha_{s,2} \exp(-\gamma_s Y_{s+S\tau-1}^2)) \left(\prod_{s=1}^S (\alpha_{s,1} + \alpha_{s,2} \exp(-\gamma_s Y_{s+S\tau-1}^2)) \right)^\tau Y_0,$$

par la compacité de Θ , l'hypothèse A2 il s'ensuit

$$\sup_{\theta_i \in \Theta} |m_{i,\tau}(\theta_i) - \tilde{m}_{i,\tau}(\theta_i)| \leq C \rho^\tau,$$

où C et ρ sont des constantes telles que $0 < \rho < 1$.

Puisque $\tilde{\ell}_{i,\tau} = \ell_{i,\tau}$ pour $\tau > 0$, alors

$$\sup_{\theta_i \in \Theta} \left| \ell_{i,\tau} - \tilde{\ell}_{i,\tau} \right| \rightarrow 0 \text{ p.s. quand } \tau \rightarrow \infty, \quad (3.7)$$

pour voir ça, on a :

$$\begin{aligned}
 &\sup_{\theta_i \in \Theta} \left| \ell_{i,\tau} - \tilde{\ell}_{i,\tau} \right| \\
 &= \sup_{\theta_i \in \Theta} \left| (Y_{i+S\tau} - m_{i,\tau}(\theta_i))^2 - (Y_{i+S\tau} - \tilde{m}_{i,\tau}(\theta_i))^2 \right| \\
 &= \sup_{\theta_i \in \Theta} \left| (m_{i,\tau}^2(\theta_i) - \tilde{m}_{i,\tau}^2(\theta_i) - 2Y_{i+S\tau} (m_{i,\tau}(\theta_i) - \tilde{m}_{i,\tau}(\theta_i))) \right| \\
 &= \sup_{\theta_i \in \Theta} \left| ((2m_{i,\tau}(\theta_i) + (\tilde{m}_{i,\tau}(\theta_i) - m_{i,\tau}(\theta_i))) (m_{i,\tau}(\theta_i) - \tilde{m}_{i,\tau}(\theta_i)) - 2Y_{i+S\tau} (m_{i,\tau}(\theta_i) - \tilde{m}_{i,\tau}(\theta_i))) \right| \\
 &\leq \sup_{\theta_i \in \Theta} C \{ (2|m_{i,\tau}(\theta_i)| + C\rho^\tau) + 2|Y_{i+S\tau}| \} \rho^\tau \rightarrow 0, \text{ p.s.}
 \end{aligned}$$

La convergence presque sûre découle du lemme de **Borel-Cantelli** et de A2, en utilisant le lemme de **Cesàro**, nous avons :

$$\tilde{Q}_{i,m}(\theta_i) - Q_{i,m}(\theta_i) \rightarrow 0 \text{ p.s.}, \quad (3.8)$$

uniformément dans θ_i , et le critère opérationnel $\tilde{Q}_{i,m}(\theta_i)$ converge également vers le critère asymptotique $Q_{i,\infty}(\theta_i)$.

On ajoute une hypothèse d'identifiabilité

$$\theta_i \neq \theta_{i,0} \implies \alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2) \neq \alpha_{i,1}^0 + \alpha_{i,2}^0 \exp(-\gamma_i^0 Y_{i+S\tau-1}^2).$$

Ensuite nous avons

$$\begin{aligned} & Q_{i,\infty}(\theta_i) - Q_{i,\infty}(\theta_{i,0}) \\ &= E_{\theta_0}(Y_{i+S\tau} - m_{i,\tau}(\theta_i))^2 - E_{\theta_0}(Y_{i+S\tau} - m_{i,\tau}(\theta_{i,0}))^2 \\ &= E(m_{i,\tau}(\theta_i) - m_{i,\tau}(\theta_{i,0}))^2 + 2E((Y_{i+S\tau} - m_{i,\tau}(\theta_{i,0})) (m_{i,\tau}(\theta_{i,0}) - m_{i,\tau}(\theta_i))) \\ &= E(m_{i,\tau}(\theta_i) - m_{i,\tau}(\theta_{i,0}))^2 \geq 0. \end{aligned} \tag{3.9}$$

La dernière égalité résulte de l'orthogonalité entre $\varepsilon_{i+S\tau}$ et $(m_{i,\tau}(\theta_{i,0}) - m_{i,\tau}(\theta_i))$.

L'inégalité dans (3.9) est une égalité si et seulement si $\theta_i = \theta_{i,0}$. Ainsi, nous avons montré que le critère asymptotique est minimisé à la vraie valeur $\theta_{i,0}$.

Maintenant nous montrons que

$$\theta_{i,0} = \lim_{m \rightarrow \infty} \arg \inf_{\theta_i \in \Theta} \tilde{Q}_{i,m}(\theta_i), \text{ p.s.}$$

Pour tout $\theta_i \in \Theta$ et tout entier positif k , soit $V_k(\theta_i)$ la boule ouverte du centre θ_i et de rayon $\frac{1}{k}$. Soit $\theta_{i,1} \neq \theta_{i,0}$, alors nous avons (propriétés des inf)

$$\inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} Q_{i,m}(\theta_i) \geq \frac{1}{m} \sum_{\tau=0}^{m-1} \inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i).$$

Le processus $\left\{ \inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i) \right\}_\tau$ est stationnaire et ergodique, alors

$$\frac{1}{m} \sum_{\tau=0}^{m-1} \inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i) \xrightarrow{p.s} E \left(\inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i) \right).$$

La suite $\inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i)$ est positive et croît jusqu'à $\ell_{i,\tau}(\theta_{i,1})$ quand $k \rightarrow \infty$. Par le théorème de convergence monotone,

$$\lim_{k \rightarrow \infty} E \left(\inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i) \right) = E \lim_{k \rightarrow \infty} \inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i) = E \ell_{i,\tau}(\theta_{i,1}) > Q_{i,\infty}(\theta_{i,0}).$$

Ainsi, pour tout $\theta_{i,j} \neq \theta_{i,0}$, il existe un voisinage $V(\theta_{i,j})$ tel que

$$\lim_{m \rightarrow \infty} \inf_{\theta_i \in V(\theta_{i,j}) \cap \Theta} Q_{i,m}(\theta_i) > \lim_{m \rightarrow \infty} Q_{i,m}(\theta_{i,0}). \quad (3.10)$$

L'ensemble compact Θ est couvert par un nombre fini d'ensembles ouverts $V(\theta_{i,1}), \dots, V(\theta_{i,M})$ satisfaisant (3.10). Alors avec la probabilité 1, nous avons pour m assez grand

$$\inf_{\theta_i \in \Theta} Q_{i,m}(\theta_i) = \min_{j=1, \dots, M} \inf_{\theta_i \in V(\theta_{i,j})} Q_{i,m}(\theta_i) = \inf_{\theta_i \in V(\theta_{i,0})} Q_{i,m}(\theta_i).$$

La preuve est complète en utilisant (3.8). ■

3.3.2 Distribution asymptotique de l'estimateur des MCN

Les hypothèses suivantes sont ajoutées :

A3 : $\theta_{i,0}$ appartient à l'intérieur $\overset{\circ}{\Theta}$ de Θ .

A4 : $E(\varepsilon_t^6) < \infty$, pour tout $t \in \mathbb{Z}$.

La distribution limite de $\widehat{\theta}_{i,m}$ est donnée par le résultat suivant.

Théorème 3.2

Sous les hypothèses A1 – A4,

$$\sqrt{m}(\widehat{\theta}_{i,m} - \theta_{i,0}) \xrightarrow[m \rightarrow \infty]{D} N(\underline{0}_{3S}, 2\sigma_i^2 J_i^{-1}),$$

où

$$J_i = E \left(\frac{\partial^2 \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i \partial \theta_i'} \right),$$

est une matrice définie positive. De plus, $\widehat{\theta}_{i,m}$ et $\widehat{\theta}_{j,m}$ sont asymptotiquement indépendants, $i \neq j$, $i, j = 1, \dots, S$.

Preuve

La preuve est basée sur un DT standard du critère (3.4) en $\theta_{i,0}$. Puisque $\widehat{\theta}_{i,m}$ converge vers $\theta_{i,0}$, qui se situe à l'intérieur de l'espace des paramètres

par l'hypothèse A3, la dérivée du critère est égale à zéro en $\widehat{\theta}_{i,m}$. Nous avons donc

$$\begin{aligned} 0 &= \sqrt{m} \frac{\partial \widetilde{Q}_{i,m}(\widehat{\theta}_{i,m})}{\partial \theta_i} \\ &= \sqrt{m} \frac{\partial \widetilde{Q}_{i,m}(\theta_{i,0})}{\partial \theta_i} + \left(\frac{\partial^2 \widetilde{Q}_{i,m}(\theta_i^*)}{\partial \theta_i \partial \theta_i'} \right) \sqrt{m} (\widehat{\theta}_i - \theta_{i,0}). \end{aligned}$$

Où θ_i^* est une valeur moyenne satisfaisant $|\theta_i^* - \theta_{i,0}| \leq \left| \widehat{\theta}_{i,m} - \theta_{i,0} \right|$ p.s.

L'hypothèse A4 implique que $E(Y_t^6) < \infty$, alors les dérivées partielles du critère en $\theta_{i,0}$ jusqu'à l'ordre 2 sont intégrables.

Puisque $\widetilde{\ell}_{i,\tau} = \ell_{i,\tau}$ pour $\tau > 0$, alors, à partir de (3.7) et du fait qu'en $\theta_i = \theta_{i,0}$ on a :

$$\begin{aligned} (Y_{i+S\tau} - (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1}) &= \varepsilon_{i+S\tau}, \\ \frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \alpha_{i,1}} &= -2Y_{i+S\tau-1} \varepsilon_{i+S\tau}. \end{aligned}$$

En itérant $Y_{i+S\tau-1}$ jusqu'à Y_0 , on a

$$\left| \frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \alpha_{i,1}} - \frac{\partial \widetilde{\ell}_{i,\tau}(\theta_{i,0})}{\partial \alpha_{i,1}} \right| \leq 2 |\varepsilon_{i+S\tau}| C \rho^\tau.$$

Puis

$$\left| \frac{1}{m} \sum_{\tau=0}^{m-1} \left(\frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \alpha_{i,1}} - \frac{\partial \widetilde{\ell}_{i,\tau}(\theta_{i,0})}{\partial \alpha_{i,1}} \right) \right| \leq \frac{2C}{m} \sum_{\tau=0}^{m-1} |\varepsilon_{i+S\tau}| \rho^\tau.$$

On peut montrer que le dernier terme converge en probabilité vers 0 en utilisant l'**inégalité de Chebychev** d'où le résultat final

$$\left\| \frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} \left(\frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i} - \frac{\partial \widetilde{\ell}_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \right) \right\| \xrightarrow{P} 0 \text{ quand } m \rightarrow \infty,$$

et de la même manière que nous avons

$$\sup_{\theta_i \in V(\theta_{i,0})} \left\| \frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} \left(\frac{\partial^2 \ell_{i,\tau}(\theta_i)}{\partial \theta_i \partial \theta_i'} - \frac{\partial^2 \widetilde{\ell}_{i,\tau}(\theta_i)}{\partial \theta_i \partial \theta_i'} \right) \right\| \xrightarrow{P} 0 \text{ quand } m \rightarrow \infty.$$

On a

$$E\left(\frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i}\right) = E\left(\frac{\partial m_{i,\tau}(\theta_{i,0})}{\partial \theta_i}(Y_i - m_{i,\tau}(\theta_{i,0}))\right) = 0,$$

et

$$\begin{aligned} V\left(\frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i}\right) &= E\left(\frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i'}\right) \\ &= 4\sigma_i^2 E\left(\frac{\partial m_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \frac{\partial m_{i,\tau}(\theta_{i,0})}{\partial \theta_i'}\right) \\ &= 2\sigma_i^2 J_i, \end{aligned}$$

où

$$J_i = 2 \begin{pmatrix} E(Y_{i-1}^2) & E(Y_{i-1}^2 \exp(-\gamma_i^0 Y_{i-1}^2)) & E(-\alpha_{i,2}^0 Y_{i-1}^4 \exp(-\gamma_i^0 Y_{i-1}^2)) \\ E(Y_{i-1}^2 \exp(-\gamma_i^0 Y_{i-1}^2)) & E(Y_{i-1}^2 \exp(-2\gamma_i^0 Y_{i-1}^2)) & E(-\alpha_{i,2}^0 Y_{i-1}^4 \exp(-2\gamma_i^0 Y_{i-1}^2)) \\ E(-\alpha_{i,2}^0 Y_{i-1}^4 \exp(-\gamma_i^0 Y_{i-1}^2)) & E(-\alpha_{i,2}^0 Y_{i-1}^4 \exp(-2\gamma_i^0 Y_{i-1}^2)) & E((\alpha_{i,2}^0)^2 Y_{i-1}^6 \exp(-2\gamma_i^0 Y_{i-1}^2)) \end{pmatrix}$$

Pour montrer que J_i est non singulière, soit $\lambda \in \mathbb{R}^3$ on a $\lambda' J_i \lambda = 0$ si et seulement si

$$\lambda' \frac{\partial m_{i,\tau}(\theta_{i,0})}{\partial \theta_i} = 0, \text{ p.s.},$$

i.e.

$$Y_{i+S\tau-1} (\lambda_1 + \exp(-\gamma_i Y_{i+S\tau-1}^2) (\lambda_2 - \alpha_{i,2}^0 \lambda_3 Y_{i+S\tau-1}^2)) = 0, \text{ p.s.},$$

puisque $\alpha_{i,2}^0 \neq 0$ (quand $\alpha_{i,2}^0 = 0$ le paramètre γ_i^0 n'est pas identifié) et $E(Y_{i+S\tau-1}^2) \geq E(\varepsilon_{i+S\tau-1}^2) > 0$ alors $\lambda_1 = \lambda_2 = \lambda_3 = 0$, ce qui implique que J_i est inversible.

Le vecteur de score est centré et

$$\frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} \frac{\partial \tilde{\ell}_{i,\tau}(\theta_{i,0})}{\partial \theta_i} = \frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} Z_{i+S\tau},$$

où

$$\begin{aligned} Z_{i+S\tau} &= -2\varepsilon_{i+S\tau} \frac{\partial m_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \\ &= -2\varepsilon_{i+S\tau} \begin{pmatrix} Y_{i+S\tau-1} \\ Y_{i+S\tau-1} \exp(-\gamma_i^0 Y_{i+S\tau-1}^2) \\ -\alpha_{i,2}^0 Y_{i+S\tau-1}^3 \exp(-\gamma_i^0 Y_{i+S\tau-1}^2) \end{pmatrix}, \end{aligned}$$

alors $\{Z_{i+S\tau}\}$ est une différence de martingale stationnaire et par le **théorème central limite de Billingsley (1961)** [III] nous avons :

$$\frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} \frac{\partial \tilde{\ell}_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \xrightarrow{D} N(0, 2\sigma_i^2 J_i),$$

et par un **DT** et l'**ergodicité** nous avons

$$\frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} \frac{\partial^2 \tilde{\ell}_{i,\tau}(\theta_i^*)}{\partial \theta_i \partial \theta_i'} \rightarrow J_i \text{ p.s.}$$

Puis

$$\sqrt{m}(\hat{\theta}_i - \theta_{i,0}) = - \left(\frac{\partial^2 \tilde{Q}_{i,m}(\theta_i^*)}{\partial \theta_i \partial \theta_i'} \right)^{-1} \sqrt{m} \frac{\partial \tilde{Q}_{i,m}(\theta_{i,0})}{\partial \theta_i} = -J_i^{-1} \frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} Z_{i+S\tau} + o_p(1).$$

Il suffit d'appliquer le lemme de **Slutsky** pour avoir

$$\sqrt{m}(\hat{\theta}_i - \theta_{i,0}) \xrightarrow[m \rightarrow \infty]{D} N(\underline{0}_{3S}, 2\sigma_i^2 J_i^{-1}).$$

L'indépendance des $\varepsilon_{i+S\tau}$ implique que tous les termes pour $i \neq j$ sont nuls, cela implique que $\sqrt{m}(\hat{\theta}_i - \theta_{i,0})$ et $\sqrt{m}(\hat{\theta}_j - \theta_{j,0})$, $i \neq j$, sont asymptotiquement non corrélés. ■

3.4 Résultats de la simulation

Les propriétés asymptotiques de l'estimateur des **MCN** sont illustrées par une petite étude de simulation. Nous générons des séries temporelles à partir

<i>EMCN</i>	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{1,2}$	$\hat{\gamma}_1$	$\hat{\alpha}_{2,1}$	$\hat{\alpha}_{2,2}$	$\hat{\gamma}_2$
$n = 300$	-0.7005	0.2012	2.0006	0.4975	-0.4013	1.0001
<i>écart - type</i>	0.0309	0.0307	0.0147	0.0597	0.0586	0.0147
$n = 600$	-0.7003	0.2000	1.9997	0.4999	-0.3989	0.9995
<i>écart - type</i>	0.0270	0.0271	0.0142	0.0436	0.0411	0.0144

TAB. 3.1 – Résultats d’estimation du $PEXPAR_2(1)$ non restreint.

des modèles **PEXPAR**₂(1) avec des tailles $n = 300$ et 600 et rapportons la valeur moyenne des estimateurs des **MCN** et leurs écarts types. Le tableau [3.1](#) donne l’estimation des paramètres $\theta = (-0.7, 0.2, 2; 0.5, -0.4, 1)'$. Les boîtes à moustaches et les graphes Q-Q plot d’erreurs d’estimation sont donnés respectivement aux figures [3.3] et [3.4] pour le dernier modèle avec 1000 réplifications. Les programmes sont écrits en **R** en utilisant la fonction **nlm**, il est clair que le programme est sensible aux valeurs initiales donc on fait une perturbation aléatoire sur le vrai paramètre. A partir du tableau [3.1](#), nous pouvons voir que les paramètres sont bien estimés et que l’écart-type diminue lorsque n augmente, ce qui montre que les estimations sont consistantes. Les figures [3.3] et [3.4] montrent que les erreurs sont centrées sur 0 et sont normales.

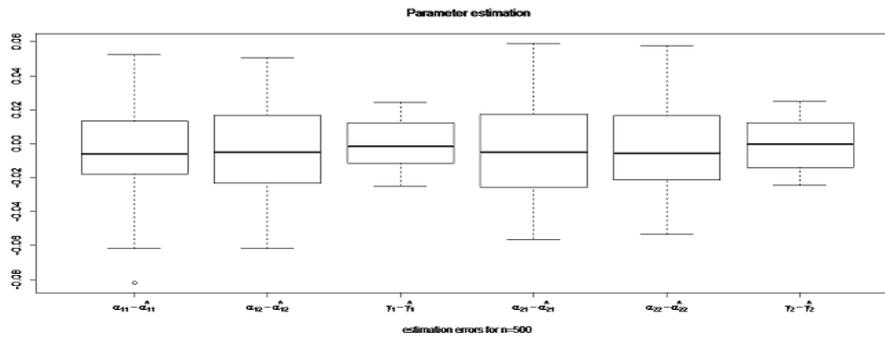


Figure 3.3 : Boîtes à moustaches d'erreurs.

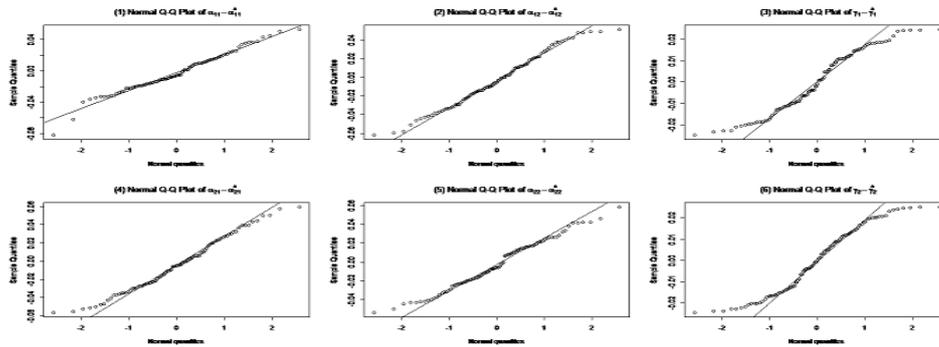


figure 3.4 : Le graphe Q-Q d'erreurs.

A titre d'exemple, nous ajustons un modèle **PEXP**₁₂(1) aux fameuses **données mensuelles de débit du fleuve Fraser**. L'ensemble de données comprend **912** observations des débits mensuels moyens du **fleuve Fraser** à **Hope, Colombie-Britannique, Canada** de **janvier 1915** à **décembre 1990**. Cet ensemble de données est obtenu à partir du lien

<http://lib.stat.cmu.edu/datasets/fraser-river>.

La présence de corrélation périodique a été montrée par **Vecchia** et **Ballerini** (1991) [53] et autres. Cette série a été analysée par plusieurs auteurs : **McLeod** (1994) [36] a utilisé un modèle **PAR** d'ordre (1, 1, 1, 3, 2, 1, 1, 3, 1, 1, 1, 1), **Lewis** et **Ray** (2002) [35] l'ont modélisé avec **SMAS****TAR** périodique, **Tesfaye** et al. (2006) [49] ont fait un ajustement avec un modèle **PARMA**₁₂(1, 1) et **Merzougui** (2017) [36] avec le modèle **PEXP**_{AR}(1) restreint avec $\gamma = 4$. Les données ont été transformées par le log et centrées. L'ajustement a été fait sur les données sauf la dernière année qui a été laissée pour être comparé aux prévisions. Pour les valeurs initiales, nous prenons l'**EMC** linéaire pour les paramètres AR et $\hat{\gamma}_i = -\frac{\log \epsilon}{\max_{0 \leq \tau \leq m-1} Y_{S\tau+i-1}^2}$, où ϵ est un petit nombre. Les résultats de l'estimation des paramètres sont donnés dans le tableau 3.2. Dans la figure [3.5], nous donnons une comparaison entre les valeurs réelles et leurs valeurs prévues pour la dernière année. Les variances résiduelles mensuelles pour le modèle **PEXP**_{AR}₁₂(1) restreint avec $\gamma = 4$ obtenu par la méthode des **MC** et le **PEXP**_{AR}₁₂(1) non restreint avec la méthode des **MCN** sont comparables comme indiqué dans le tableau 3.3. L'explication de ces résultats est que lorsque $\gamma_i \rightarrow \infty$, $\alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2) \rightarrow 0$, en revanche la simulation nous amène à penser que lorsque $\gamma_i > 5$, les résultats deviennent comparables.

i	Janvier	Février	Mars	Avril	Mai	Juin
$\hat{\alpha}_{i,1}$	0.6829	0.8145	0.7652	0.7069	0.1805	0.2368
$\hat{\alpha}_{i,2}$	-0.4936	-1.2685	0.2472	2.5652	0.0450	-4.6453
$\hat{\gamma}_i$	46.6180	242.1705	16.8679	104.7684	27.9312	440.3529
i	Juillet	Août	Septembre	Octobre	Novembre	Décembre
$\hat{\alpha}_{i,1}$	0.6748	0.7616	0.7066	0.7842	0.7406	0.7369
$\hat{\alpha}_{i,2}$	0.9969	-1.4752	1.3841	0.3392	1.5638	1.6418
$\hat{\gamma}_i$	146.5914	248.9337	153.4060	6.5568	97.6726	153.5859

TAB. 3.2 – Estimation des paramètres du $PEXP$ $AR_{12}(1)$ non restreint.

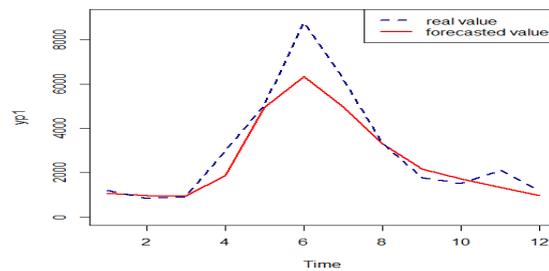


Figure 3.5 : Comparaison entre les valeurs réelles et leurs valeurs prévues.

i	Janvier	Février	Mars	Avril	Mai	Juin
$\hat{\sigma}_{EMCL}^2$	0.0268	0.0238	0.0281	0.0850	0.0451	0.0282
$\hat{\sigma}_{EMCN}^2$	0.0265	0.0241	0.0280	0.0851	0.0455	0.0273
i	Juillet	Août	Septembre	Octobre	Novembre	Décembre
$\hat{\sigma}_{EMCL}^2$	0.0258	0.0131	0.0241	0.0363	0.0480	0.0388
$\hat{\sigma}_{EMCN}^2$	0.0270	0.0129	0.0237	0.0369	0.0473	0.0388

TAB. 3.3 – Variances résiduelles mensuelles.

CHAPITRE 4

Estimation par la méthode du QMV du modèle

PEXPAR(1)

Sommaire

4.1 Introduction	62
4.2 Estimateur du QMV pour le modèle PEXPAR_S(1)	63

4.1 Introduction

L'estimation des paramètres est une étape principale dans la modélisation des séries chronologiques et dans un contexte non linéaire, les méthodes QMV et MCN sont les plus populaires car elles facilitent l'estimation des paramètres. Dans l'estimation traditionnelle du maximum de vraisemblance (EMV) on suppose que la densité spécifiée est la vraie fonction et quand ce n'est pas le cas, nous ajoutons le préfixe **quasi**, bien sûr si la densité est spéci-

fiée correctement alors l'**EQMV** est le même que l'**EMV**. Ce travail est une généralisation de l'estimation **QMV** de l'**EXPAR** classique d'**Amondela** et **Françq** (2009) [4] au cas périodique, la vraisemblance s'écrit comme si la loi de l'innovation était gaussienne. Le **QMV** qui en résulte fournit des estimateurs consistants et asymptotiquement normaux et coïncide avec l'estimateur non linéaire des moindres carrés (**MCN**). Les preuves impliquent la normalité asymptotique de la première dérivée des fonctions de vraisemblance et la convergence de la dérivée du second ordre qui sont évaluée aux vraies valeurs.

Dans la première section, nous donnons l'estimateur **QMV**. Ensuite, on présente les principaux résultats de la forte consistance et la normalité asymptotique de cet estimateur.

4.2 Estimateur du QMV pour le modèle $\text{PEXP}AR_S(1)$

On commence par rappeler le modèle $\text{PEXP}AR_S(1)$. Le processus $\{Y_t; t \in \mathbb{Z}\}$ est un $\text{PEXP}AR_S(1)$ s'il satisfait :

$$Y_{i+S\tau} = (\alpha_{i,1}^0 + \alpha_{i,2}^0 \exp(-\gamma_i^0 Y_{i+S\tau-1}^2)) Y_{i+S\tau-1} + \varepsilon_{i+S\tau}, \quad i = 1, \dots, S, \quad \tau \in \mathbb{Z}. \quad (4.1)$$

où $\{\varepsilon_t; t \in \mathbb{Z}\}$ est un processus *iid* $(0, (\sigma_t^0)^2)$ avec une densité de probabilité $f(\cdot)$.

Le vecteur des paramètres

$$\theta = (\theta'_1, \dots, \theta'_S)' \in \mathbb{R}^{3S} \quad \text{où } \theta_i = (\alpha_{i,1}, \alpha_{i,2}, \gamma_i)', \quad i = 1, \dots, S,$$

appartient à un espace de paramètres de la forme :

$$\Theta^S \subset (]-1, +1[\times]-1, +1[\times]0, \infty[)^S.$$

$(\sigma_t^0)^2$ peut être considéré comme un paramètre de nuisance. La vraie valeur des paramètres est inconnue et notée par :

$$\theta_{i,0} = (\alpha_{i,1}^0, \alpha_{i,2}^0, \gamma_i^0)', i = 1, \dots, S, \text{ et } \theta_0 = (\theta'_{1,0}, \dots, \theta'_{S,0})' \in \mathbb{R}^{3S}.$$

Nous voulons estimer le vrai paramètre θ_0 à partir des observations Y_1, \dots, Y_n où $n = mS$. Etant donné la valeur initiale Y_0 , le log-vraisemblance conditionnelle des observations évaluées en θ dépend de f qui est inconnue et il est donc impossible de la déterminer. L'estimateur **QMV** est obtenu en remplaçant f par $N(0, \sigma_t^2)$:

$$L_n(\theta, Y_1, \dots, Y_n) = -\frac{mS}{2} \log(2\pi) - \frac{m}{2} \sum_{i=1}^S \log(\sigma_i^2) - \sum_{i=1}^S \sum_{\tau=0}^{m-1} \frac{(Y_{i+S\tau} - (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1})^2}{2\sigma_i^2},$$

en supposant $\sigma_i \neq 0$.

Soit $\hat{\theta}_n$ l'estimateur **QMV**, on voit que maximiser L_n équivaut au problème de minimisation suivant par rapport à θ :

$$\hat{\theta}_n = \arg \inf_{\theta \in \Theta^S} Q_n(\theta).$$

Où

$$Q_n(\theta) = \frac{1}{n} \sum_{t=1}^n \ell_t \text{ et } \ell_t(\theta) = \frac{(Y_t - (\alpha_{t,1} + \alpha_{t,2} \exp(-\gamma_t Y_{t-1}^2)) Y_{t-1})^2}{\sigma_t^2}.$$

La valeur initiale est inconnue mais son choix n'est pas important pour le comportement asymptotique de l'estimateur **QMV** donc on met $Y_0 = 0$, alors

$$\hat{\theta}_n = \arg \inf_{\theta \in \Theta^S} \tilde{Q}_n(\theta). \quad (4.2)$$

où

$$\begin{aligned}\tilde{Q}_n(\theta) &= \frac{1}{S} \sum_{i=1}^S \tilde{Q}_{i,m}(\theta_i) \\ &= \frac{1}{S} \sum_{i=1}^S \frac{1}{m} \sum_{\tau=0}^{m-1} \left(\frac{(Y_{i+S\tau} - (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1})^2}{\sigma_i^2} \right).\end{aligned}$$

et

$$\begin{aligned}\tilde{Q}_{i,m}(\theta_i) &= \frac{1}{m} \sum_{\tau=0}^{m-1} \tilde{\ell}_{i,\tau}(\theta_i) \tag{4.3} \\ &= \frac{1}{m} \sum_{\tau=0}^{m-1} \left(\frac{(Y_{i+S\tau} - (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1})^2}{\sigma_i^2} \right).\end{aligned}$$

L'existence d'une solution à (4.2) est garantie sous l'hypothèse :

A1 : Θ^S est un ensemble de Borel compact $\subset \mathbb{R}^{3S}$.

La condition du premier ordre du problème de minimisation du QMV est un système de $3S$ équations non linéaires avec $3S$ inconnues. Le vecteur gradient de $\tilde{Q}_{i,m}(\theta_i)$ est :

$$\frac{\partial \tilde{Q}_{i,m}(\theta_i)}{\partial \theta_i} = -2 \frac{1}{m} \sum_{\tau=0}^{m-1} \left(\frac{Y_{i+S\tau} - (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1}}{\sigma_i^2} \right) \begin{pmatrix} Y_{i+S\tau-1} \\ Y_{i+S\tau-1} \exp(-\gamma_i Y_{i+S\tau-1}^2) \\ -\alpha_{i,2} Y_{i+S\tau-1}^3 \exp(-\gamma_i Y_{i+S\tau-1}^2) \end{pmatrix}$$

Il n'y a pas de solution analytique pour $\frac{\partial \tilde{Q}_{i,m}(\theta_i)}{\partial \theta_i} = 0$, et les estimateurs $\hat{\theta}_{i,m}$ pour la période i doivent être calculés par optimisation numérique comme la méthode du **gradient** ou du **quasi Newton**.

Remarque : L'estimateur QMV est l'estimateur MCN.

4.2.1 Consistance de l'estimateur du QMV

Nous faisons les hypothèses suivantes pour montrer la consistance.

A2 : i) Y_t est périodiquement strictement stationnaire, une condition suffisante pour cela est

$$|\alpha_{i,1}| + |\alpha_{i,2}| < 1 \text{ pour } i = 1, \dots, S.$$

ii) $E(\varepsilon_t^2) < \infty$, pour tout $t \in \mathbb{Z}$.

Le premier théorème énonce la forte consistance de l'EQMV $\widehat{\theta}_{i,m}$.

Théorème 4.1

Sous les hypothèses A1 – A2, les estimateurs du quasi maximum de vraisemblance $\widehat{\theta}_{i,m}$ sont fortement consistants, i.e. presque sûrement,

$$\widehat{\theta}_{i,m} \rightarrow \theta_{i,0} \text{ quand } m \rightarrow \infty.$$

Preuve

Puisque $\ell_{i,\tau}$ est une fonction mesurable de $Y_{i+S\tau}$ et $Y_{i+S\tau-1}$ et de A2 i), $\{\ell_{i,\tau}\}$ est aussi un processus ergodique stationnaire de plus $E(\ell_{i,\tau}(\theta))$ existe au vu de A2 ii). Ainsi, par ergodicité le critère théorique

$$Q_{i,m}(\theta_i) = \frac{1}{m} \sum_{\tau=0}^{m-1} \frac{(Y_{i+S\tau} - (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1})^2}{\sigma_i^2},$$

converge presque sûrement vers le critère asymptotique

$$Q_{i,\infty}(\theta_i) = E_{\theta_{i,0}}(\ell_{i,\tau}(\theta)) = E \left(\frac{(Y_i - (\alpha_{i,1}^0 + \alpha_{i,2}^0 \exp(-\gamma_i^0 Y_{i-1}^2)) Y_{i-1})^2}{\sigma_i^2} \right).$$

Puisque $\ell_{i,\tau} = \widetilde{\ell}_{i,\tau}$ pour $\tau > 0$, alors les valeurs initiales sont uniformément négligeables et on a :

$$\widetilde{Q}_{i,m}(\theta_i) - Q_{i,m}(\theta_i) \rightarrow 0 \text{ p.s,} \tag{4.4}$$

uniformément dans θ_i , et le critère opérationnel $\widetilde{Q}_{i,m}(\theta_i)$ converge également vers le critère asymptotique $Q_{i,\infty}(\theta_i)$.

Pour l'identifiabilité du paramètre, nous pouvons montrer

$$\text{si } \theta_i \neq \theta_{i,0} \implies \alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2) \neq \alpha_{i,1}^0 + \alpha_{i,2}^0 \exp(-\gamma_i^0 Y_{i+S\tau-1}^2).$$

Maintenant, pour simplifier la notation, soit

$$\begin{aligned} m_{i,\tau}(\theta_i) &= E(Y_i | Y_{i-1}, Y_{i-2}, \dots) \\ &= (\alpha_{i,1} + \alpha_{i,2} \exp(-\gamma_i Y_{i+S\tau-1}^2)) Y_{i+S\tau-1}. \end{aligned}$$

Puis

$$\begin{aligned}
 Q_{i,\infty}(\theta_i) - Q_{i,\infty}(\theta_{i,0}) &= E_{\theta_{i,0}} \left(\frac{Y_i - m_{i,\tau}(\theta_i)}{\sigma_i} \right)^2 - 1 \\
 &= E_{\theta_{i,0}} \left(\frac{m_{i,\tau}(\theta_{i,0}) - m_{i,\tau}(\theta_i)}{\sigma_i} \right)^2 + E_{\theta_{i,0}} \left(\frac{\sigma_i^{(0)}}{\sigma_i} \right)^2 - 1 \\
 &= E_{\theta_{i,0}} \left(\frac{m_{i,\tau}(\theta_i) - m_{i,\tau}(\theta_{i,0})}{\sigma_i} \right)^2 + E_{\theta_{i,0}} \left(\frac{\sigma_i - \sigma_i^{(0)}}{\sigma_i} \right)^2 \geq 0.
 \end{aligned} \tag{4.5}$$

On utilise le fait que $Y_i - m_{i,\tau}(\theta_{i,0}) = \varepsilon_{i+S\tau}$ et l'orthogonalité entre $\varepsilon_{i+S\tau}$ et $(m_{i,\tau}(\theta_{i,0}) - m_{i,\tau}(\theta_i))$. L'inégalité dans (4.5) est une égalité si et seulement si $\theta_i = \theta_{i,0}$. Ainsi, nous avons montré que le critère asymptotique est minimisé à la vraie valeur $\theta_{i,0}$. Pour compléter la preuve, il faut montrer que

$$\theta_{i,0} = \lim_{m \rightarrow \infty} \arg \inf_{\theta_i \in \Theta} \tilde{Q}_{i,m}(\theta_i), \text{ p.s.}$$

Pour tout $\theta_i \in \Theta$ et tout entier positif k , soit $V_k(\theta_i)$ la boule ouverte de centre θ_i et de rayon $\frac{1}{k}$. Soit $\theta_{i,1} \neq \theta_{i,0}$, alors nous avons (propriétés des inf)

$$\inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} Q_{i,m}(\theta_i) \geq \frac{1}{m} \sum_{\tau=0}^{m-1} \inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i).$$

Le processus $\left\{ \inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i) \right\}$ est stationnaire et ergodique, alors

$$\frac{1}{m} \sum_{\tau=0}^{m-1} \inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i) \xrightarrow{p.s.} E \left(\inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i) \right).$$

La suite $\inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i)$ est positive et croît jusqu'à $\ell_{i,\tau}(\theta_{i,1})$ quand $k \rightarrow \infty$.

Par le théorème de convergence monotone,

$$\lim_{k \rightarrow \infty} E \left(\inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i) \right) = E \lim_{k \rightarrow \infty} \inf_{\theta_i \in V_k(\theta_{i,1}) \cap \Theta} \ell_{i,\tau}(\theta_i) = E \ell_{i,\tau}(\theta_{i,1}) > Q_{i,\infty}(\theta_{i,0}).$$

Ainsi, pour tout $\theta_{i,j} \neq \theta_{i,0}$, il existe un voisinage $V(\theta_{i,j})$ tel que

$$\liminf_{m \rightarrow \infty} \inf_{\theta_i \in V_k(\theta_{i,j}) \cap \Theta} Q_{i,m}(\theta_i) > \lim_{m \rightarrow \infty} Q_{i,m}(\theta_{i,0}). \tag{4.6}$$

L'ensemble compact Θ est couvert par un nombre fini d'ensembles ouverts $V(\theta_{i,1}), \dots, V(\theta_{i,M})$ satisfaisant (4.6). Alors avec la probabilité 1, nous avons pour m assez grand

$$\inf_{\theta_i \in \Theta} Q_{i,m}(\theta_i) = \min_{j=1, \dots, M} \inf_{\theta_i \in V(\theta_{i,j})} Q_{i,m}(\theta_i) = \inf_{\theta_i \in V(\theta_{i,0})} Q_{i,m}(\theta_i).$$

La preuve est complète en utilisant (4.4). ■

4.2.2 Distribution asymptotique de l'estimateur du QMV

En plus des hypothèses précédentes, les conditions suivantes sont prises en compte.

A3 : $\theta_{i,0}$ appartient à l'intérieur $\overset{\circ}{\Theta}$ de Θ .

A4 : $E(\varepsilon_t^6) < \infty$, pour tout $t \in \mathbb{Z}$.

L'hypothèse A3 est standard pour la normalité asymptotique et A4 est nécessaire pour l'existence de la variance du vecteur de score. La distribution limite de $\hat{\theta}_{i,m}$ est donnée par le résultat suivant.

Théorème 4.2

Sous les hypothèses ci-dessus,

$$\sqrt{m}(\hat{\theta}_{i,m} - \theta_{i,0}) \xrightarrow[m \rightarrow \infty]{D} N(\mathbf{0}_3, 2J_i^{-1}),$$

où

$$J_i = E \left(\frac{\partial^2 \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i \partial \theta_i'} \right),$$

est une matrice définie positive.

Preuve

La preuve est basée sur un **DT** standard du critère (4.3) en $\theta_{i,0}$. Puisque $\hat{\theta}_{i,m}$ converge vers $\theta_{i,0}$, qui se situe à l'intérieur de l'espace des paramètres par l'hypothèse A3, la dérivée du critère est égale à zéro en $\hat{\theta}_{i,m}$. Nous avons

donc

$$\begin{aligned} 0 &= \sqrt{m} \frac{\partial \tilde{Q}_{i,m}(\hat{\theta}_{i,m})}{\partial \theta_i} \\ &= \sqrt{m} \frac{\partial \tilde{Q}_{i,m}(\hat{\theta}_{i,0})}{\partial \theta_i} + \left(\frac{\partial^2 \tilde{Q}_{i,m}(\theta_i^*)}{\partial \theta_i \partial \theta_i'} \right) \sqrt{m} (\hat{\theta}_{i,m} - \theta_{i,0}). \end{aligned}$$

Où θ_i^* est une valeur moyenne satisfaisant

$$|\theta_i^* - \theta_{i,0}| \leq \left| \hat{\theta}_{i,m} - \theta_{i,0} \right| p.s.$$

L'hypothèse A4 implique que $E(Y_t^6) < \infty$, alors les dérivées partielles du critère en $\theta_{i,0}$ jusqu'à l'ordre 3 sont intégrables soit :

$$E \sup_{\theta_i \in \Theta} \left| \frac{\partial^3 \ell_{i,\tau}(\theta_i)}{\partial \theta_{i,j} \partial \theta_{i,k} \partial \theta_{i,l}} \right| < \infty \quad \forall j, k, l. \quad (4.7)$$

Puisque $\tilde{\ell}_{i,\tau} = \ell_{i,\tau}$ pour $\tau > 0$, alors

$$\left\| \frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} \left(\frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i} - \frac{\partial \tilde{\ell}_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \right) \right\| \xrightarrow{P} 0 \text{ quand } m \rightarrow \infty,$$

et

$$\sup_{\theta_i \in V(\theta_{i,0})} \left\| \frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} \left(\frac{\partial^2 \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i \partial \theta_i'} - \frac{\partial^2 \tilde{\ell}_{i,\tau}(\theta_{i,0})}{\partial \theta_i \partial \theta_i'} \right) \right\| \xrightarrow{P} 0 \text{ quand } m \rightarrow \infty.$$

On a

$$E \left(\frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \right) = \sigma_i^{-2} E \left(\frac{\partial m_{i,\tau}(\theta_{i,0})}{\partial \theta_i} (Y_i - m_{i,\tau}(\theta_{i,0})) \right) = 0,$$

et

$$\begin{aligned} V \left(\frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \right) &= E \left(\frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \frac{\partial \ell_{i,\tau}(\theta_{i,0})}{\partial \theta_i'} \right) \\ &= 4 (\sigma_i^0)^{-2} E \left(\frac{\partial m_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \frac{\partial m_{i,\tau}(\theta_{i,0})}{\partial \theta_i'} \right) \\ &= 2J_i, \end{aligned}$$

où

$$J_i = 2(\sigma_i^0)^{-2} \begin{pmatrix} E(Y_{i-1}^2) & E(Y_{i-1}^2 \exp(-\gamma_i^0 Y_{i-1}^2)) & E(-\alpha_{i,2}^0 Y_{i-1}^4 \exp(-\gamma_i^0 Y_{i-1}^2)) \\ E(Y_{i-1}^2 \exp(-\gamma_i^0 Y_{i-1}^2)) & E(Y_{i-1}^2 \exp(-2\gamma_i^0 Y_{i-1}^2)) & E(-\alpha_{i,2}^0 Y_{i-1}^4 \exp(-2\gamma_i^0 Y_{i-1}^2)) \\ E(-\alpha_{i,2}^0 Y_{i-1}^4 \exp(-\gamma_i^0 Y_{i-1}^2)) & E(-\alpha_{i,2}^0 Y_{i-1}^4 \exp(-2\gamma_i^0 Y_{i-1}^2)) & E((\alpha_{i,2}^0)^2 Y_{i-1}^6 \exp(-2\gamma_i^0 Y_{i-1}^2)) \end{pmatrix}$$

Pour montrer que J_i est non singulière, soit $\lambda \in \mathbb{R}^3$ on a $\lambda' J_i \lambda = 0$ si et seulement si

$$\lambda' \frac{\partial m_{i,\tau}(\theta_{i,0})}{\partial \theta_i} = 0, \text{ p.s.},$$

i.e.

$$Y_{i+S\tau-1}(\lambda_1 + \exp(-\gamma_i^0 Y_{i+S\tau-1}^2)(\lambda_2 - \alpha_{i,2}^0 \lambda_3 Y_{i+S\tau-1}^2)) = 0, \text{ p.s.},$$

puisque $\alpha_{i,2}^0 \neq 0$ (quand $\alpha_{i,2}^0 = 0$ le paramètre γ_i^0 n'est pas identifié) et $E(Y_{i+S\tau-1}^2) \geq E(\varepsilon_{i+S\tau-1}^2) > 0$ alors $\lambda_1 = \lambda_2 = \lambda_3 = 0$, ce qui implique que J_i est inversible.

Le vecteur de score est centré et

$$\frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} \frac{\partial \tilde{\ell}_{i,\tau}(\theta_{i,0})}{\partial \theta_i} = \frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} Z_{i+S\tau},$$

où

$$\begin{aligned} Z_{i+S\tau} &= -2(\sigma_i^0)^{-2} \varepsilon_{i+S\tau} \frac{\partial m_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \\ &= -2(\sigma_i^0)^{-2} \varepsilon_{i+S\tau} \begin{pmatrix} Y_{i+S\tau-1} \\ Y_{i+S\tau-1} \exp(-\gamma_i Y_{i+S\tau-1}^2) \\ -\alpha_{i,2} Y_{i+S\tau-1}^3 \exp(-\gamma_i Y_{i+S\tau-1}^2) \end{pmatrix}, \end{aligned}$$

alors $\{Z_{i+S\tau}\}$ est une différence de martingale stationnaire carrée intégrable et par le **théorème central limite de Billingsley (1961)** [11] nous avons :

$$\frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} \frac{\partial \tilde{\ell}_{i,\tau}(\theta_{i,0})}{\partial \theta_i} \xrightarrow{D} N(\underline{0}_3, 2J_i),$$

et par **DT**, l'**ergodicité** et (4.7) on a

$$\frac{1}{m} \sum_{\tau=0}^{m-1} \frac{\partial^2 \tilde{\ell}_{i,\tau}(\theta_i^*)}{\partial \theta_i \partial \theta_i'} \rightarrow J_i \text{ p.s.}$$

Puis

$$\sqrt{m}(\hat{\theta}_{i,m} - \theta_{i,0}) = - \left(\frac{\partial^2 \tilde{Q}_{i,m}(\theta_i^*)}{\partial \theta_i \partial \theta_i'} \right)^{-1} \sqrt{m} \frac{\partial \tilde{Q}_{i,m}(\theta_{i,0})}{\partial \theta_i} = -J_i^{-1} \frac{1}{\sqrt{m}} \sum_{\tau=0}^{m-1} Z_{i+S\tau} + o_p(1).$$

Il suffit d'appliquer le lemme de **Slutsky** pour avoir

$$\sqrt{m}(\hat{\theta}_{i,m} - \theta_{i,0}) \underset{m \rightarrow \infty}{\overset{D}{\rightarrow}} N(\mathbf{0}_3, 2J_i^{-1}).$$

■

Remarque : Pour i fixé, nous obtenons les estimations du **QMV** du modèle classique **EXPAR**(1) dans **Amendola et Francq (2009)** [\[4\]](#).

Conclusion et perspectives

Nous avons étendu, dans cette thèse, les modèles autorégressifs exponentiels restreints périodiques à l'ordre p . Ces modèles peuvent répondre à des préoccupations empiriques pour de nombreuses séries chronologiques saisonnières présentant un comportement non linéaire et ayant une structure d'autocovariance périodique.

Vu l'importance du modèle **PEXP** $\mathbf{R}(\mathbf{p})$, nous nous sommes intéressés à l'estimation de ses paramètres qui est clairement une procédure d'optimisation non linéaire impliquant toutes les difficultés de calculs propres à cette procédure. Cependant, on a surmonté ce problème en fixant le paramètre non linéaire γ et estimer le reste des paramètres par la méthode des **MC** linéaire, qui est connue pour délivrer des estimations avec une variance minimale. Ces estimateurs peuvent être utilisés comme estimateurs initiaux dans d'autres méthodes d'estimation.

Bien que les résultats obtenus soient satisfaisants, on ne s'est pas limité au cas restreint et on s'est attaqué au cas non restreint en utilisant la méthode des **MCN** et du **QMV**.

Quelque soit le cas ou la méthode, nos estimateurs sont consistants et asymptotiquement normaux, les résultats des simulations montrent qu'on a

les propriétés asymptotiques.

Dans l'application nous avons travaillé avec la série des débits du **fleuve Fraser** connue pour sa non linéarité et sa périodicité, le modèle estimé est compétitif.

Nous avons considéré, ici, une condition suffisante de stationnarité stricte mais ce sujet mérite de plus amples recherches.

Ce travail dégage des résultats importants qui permettent d'ouvrir de nombreuses perspectives, tout particulièrement pour l'identification de ces modèles non linéaires. Le modèle étant nouveau, plusieurs questions théoriques et pratiques restent encore ouvertes à savoir :

L'utilisation d'autres méthodes d'estimation comme par exemple M-estimation, Z-estimation, pseudo vraisemblance et faire une étude comparative entre ces méthodes.

Annexe : Théorie de mesure de la théorie de probabilité

Quelques définitions indispensables

Définition A.1 (CV en probabilité)

Soit X_n une suite de variables aléatoires. X_n converge en probabilité vers la variable aléatoire X si :

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0.$$

Souvent, on utilise l'inégalité de Chebychev pour montrer la CV en probabilité.

Définition A.2 (CV presque sur)

Soit X_n une suite de variables aléatoires. X_n converge presque sûrement vers la variable aléatoire X si

$$\exists N \in \mathcal{F}/P(N) = 0 \text{ tel que } \forall \omega \in \Omega/N, \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega).$$

Il suffit d'appliquer le lemme de Borel-Cantelli pour avoir la CV p.s qui est difficile à vérifier à partir de la définition.

Définition A.3 (CV en moyenne)

La suite de variables aléatoires $\{X_n\}$ converge en moyenne d'ordre r vers X , noté $X_n \xrightarrow{r} X$, si

$$E |X_n - X|^r \rightarrow 0 \text{ quand } n \rightarrow \infty.$$

Définition A.4 (CV en distribution)

La suite $\{X_n\}$ de vecteurs aléatoires de dimension k avec des fonctions de distributions $\{F_{X_n}(\cdot)\}$ converge en distribution s'il existe un vecteur aléatoire X de dimension k tel que :

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x), \forall x \in C,$$

où C est l'ensemble des points de continuité de la distribution $F_X(\cdot)$ de X .

Inégalité de Chebychev : Si φ est une fonction mesurable sur \mathbb{R} tel que $\varphi(x)$ est positive et strictement croissante sur $(0, \infty)$ et $\varphi(x) = \varphi(-x)$, alors \forall la V.A X et $\forall \delta > 0$ on a :

$$P(|X| > \delta) \leq E \left[\frac{\varphi(X)}{\varphi(\delta)} \right].$$

Théorèmes centrales limites

Soit $Z_t, t \in \mathbb{N}$ une suite de V.A iid avec $E(Z_t) = \mu$ et $Var(Z_t) = \sigma^2 < \infty$. Soit $\bar{Z}_n = \frac{1}{n} \sum_{t=1}^n Z_t$ la moyenne empirique, par la **LGN** forte de **Kolmogorov** on a : $\bar{Z}_n - E(\bar{Z}_n)$ converge vers 0 *p.s* et par conséquent en probabilité. Ce qui implique que la distribution limite de $\bar{Z}_n - E(\bar{Z}_n)$ est dégénérée en 0 et donc cette distribution limite n'est pas utile pour connaître la forme de la distribution de la moyenne empirique pour n fini. On considère la même quantité avec un autre taux de convergence :

$$\sqrt{n} (\bar{Z}_n - E(\bar{Z}_n)) = \frac{1}{\sqrt{n}} \sum_{t=1}^n (Z_t - \mu), \quad (\text{A.1})$$

alors la variance de cette expression est σ^2 , $\forall n$ indiquant que sa distribution limite ne sera pas dégénérée. Les théorèmes qui donnent des résultats concernant les distributions limites des expressions comme (A.1) sont appelés les théorèmes centrales limites (**TCL**).

Processus Indépendants

Théorème A.1 (TCL de Lindeberg-Lévy)

Soit Z_t une suite de V.A iid avec $E(Z_t) = 0$ et $Var(Z_t) = \sigma^2 < \infty$.

Alors

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n Z_t \xrightarrow{D} N(0, \sigma^2).$$

Dans le cas de variables indépendantes mais hétérogènes on utilise le **TCL** de **Lyapounov**.

On termine par rappeler le théorème de **Bolzano-Weierstrass**

Valeurs d'adhérence :

Soit E un espace métrique, Soit $(u_n) \in E^{\mathbb{N}}$, l est une valeur d'adhérence de $(u_n) \iff$ il existe une sous suite $(u_{\varphi(n)})$ qui converge vers l . \iff

$$\forall \varepsilon > 0, \{u_n, n \in \mathbb{N}\} \cap B(l, \varepsilon) \neq \emptyset.$$

Si u_n converge vers $l \implies l$ est la seule valeur d'adhérence.

Théorème A.2 (Théorème de Bolzano-Weierstrass)

E est compact si et seulement si toute suite de $E^{\mathbb{N}}$ admet au moins une valeur d'adhérence.

Si E est compact et si (u_n) admet une seule valeur d'adhérence l alors (u_n) converge vers l .

BIBLIOGRAPHIE

- [1] Al-Kassam, M. S. and Lane, J. A. (1989). Forecasting Exponential Autoregressive Models Of Order 1. *Journal Of Time Series Analysis*. 10(2), 95-113.
- [2] Allal, J. and El Melhaoui, S.(2006).Optimal Detection Of Exponential Component. *Journal Of Time Series Analysis*. 27(6), 793-810.
- [3] Amiri, E., (2012). Forecasting GDP Growth rate with Nonlinear Models. 1st International Conference on Econometrics Methods and Applications. ICEKU2012, August 25- 27. Sanandaj, Iran, pp. 1-18.
- [4] Amondela, A. and Francq, C. (2009) . *Concepts and tools for nonlinear time series modelling*. In Handbook of Computational Econometrics (eds D. A. Belsley and E. J. Kontoghiorghes), John Wiley & Sons, Ltd, Chichester, UK. pp. 377-427.
- [5] Azouagh, N. and El Melhaoui, S. (2019). An Exponential Autoregressive model for the forecasting of annual sunspots number. *Electronic Journal of Mathematical Analysis and Applications*, Vol. 7(3), pp. 17-23.

- [6] Baragona, R., Battaglia, F. and Cuccina, D. (2002). A note on estimating autoregressive exponential models. *Quaderni di Statistica*, Vol.4, 2002.
- [7] Basawa, I. V. and Lund, R. (2001). Large sample properties of parameters estimates for periodic *ARMA* models. *Journal of Time Series Analysis*, **22**, 1-13.
- [8] Becila, S and Merzougui, M. (2020). Nonlinear Least Squares estimation of the Periodic EXPAR(1) model. *Communications in Statistics-Theory and Methods*. DOI : 10.1080/03610926.2020.1839099.
- [9] Bibi, A. and Gautier, A.(2005). Stationarity and asymptotic inference of some periodic bilinear models, *Comptes Rendus Mathematique* 341(11), 679-682.
- [10] Bierens, H. J. (1994). *Estimation, testing, and specification of cross-section and time series models*. Cambridge University Press.
- [11] Billingsley, P. (1961). The Lindeberg-Levy theorem for martingales, *Proceedings of the American Mathematical Society*, 12 788–792.
- [12] Bollerslev, T. and Ghysels, E. (1996). Periodic autoregressive conditional heteroscedasticity, *Journal of Business and Economic Statistics* 14(2), 139-152.
- [13] Brockwell, P. J. and Davis, R. A. (1991). *Time Series : Theory and methods*. Springer-Verlag. New York.
- [14] Chan, K. S. and Tong, H. (1985). On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations. *Advances in applied probability*, 17(3), 666-678.
- [15] De Gooijer, J. G.(2017). Elements of Nonlinear Time Series Analysis and Forecasting, *Springer International Publishing*.

Bibliographie

- [16] Douc.R, Moulines.E and Stoffer.D.S.(2013). Nonlinear Time Series : *Theory, Methods, and Applications with R Examples*.
- [17] Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with estimates of variance of U.K. Inflation. *Econometrica*, 50, 987-1008.
- [18] Francq, C and Zakoian, J.M. (2010).GARCH Models : Structure, Statistical Inference and Financial Applications .2010 John Wiley & Sons, Ltd.
- [19] Franses, P.H. and Ooms, M., (1997). A Periodic Long Memory Model for Quarterly UK Inflation. *International Journal of Forecasting*, 13, 119–128.
- [20] Fuller, W. A. (1996). *Introduction to Statistical Time Series*. (Second Edition). John Wiley & Sons, New York.
- [21] Gallant, A. R., and White, H. (1988). *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. Basil Blackwell, New York.
- [22] Ghosh, H., Gurung, B. and Gupta, P. (2015). Fitting EXPAR Models Through the Extended Kalman Filter. *Sankhyā : The Indian Journal of Statistics*. 77(1), 27-44.
- [23] Gladyshev, E. G. (1961). Periodically correlated random sequences. *Soviet. Math.*, **2**, 385-88.
- [24] Granger, C. and Andersen, A. (1978). An Introduction to Bilinear Time Series Models. *Vandenhoeck und Ruprecht, Gottingen*.
- [25] Gurung, B. (2013). An Application of Exponential Autoregressive (EXPAR) Nonlinear Time-series Model. *International Journal of Information and Computation Technology*. ISSN 0974-2239, 3(4), 261-266.

- [26] Haggan, V. and Ozaki, T. (1981), Modeling nonlinear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika*. 68(1), 96-189.
- [27] Hamdi, F. and Souam, S. (2013). Mixture periodic GARCH models : Applications to exchange rate modeling. *5th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO).IEEE*.
- [28] Ibragimov, I. A. (1963). A central limit theorem for a class of dependent random variables, *Theory of Probability & Its Applications* 8(1), 83-89.
- [29] Ishizuka, K., Kato, H. and Nakatani, T. (2005). Speech signal analysis with exponential autoregressive model, *Proc. the 30th International Conference of Acoustics, Speech and Signal Processing, Institute of Electrical and Electronics Engineers Inc, Philadelphia, PA, United States*. 1, 225-228.
- [30] Ismail, M. A. (2001). Bayesian Analysis of Exponential AR Models. *The Far East Journal of Statistics*, Vol 5, pp 1-15. (India).
- [31] Katsiampa, P. (2014). *A new approach to modelling nonlinear time series : Introducing the ExpAR-ARCH and ExpAR-GARCH models and applications*. 4th Student Conference on Operational Research, Germany, 34-51.
- [32] Klimko, L.A. and Nelson, P.I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.* 6, 629-642.
- [33] Koul, H. L. and Schick, A.(1997). Efficient estimation in nonlinear autoregressive time series models. *Bernoulli*, 3, 247-277.
- [34] Kuan, C. M. (2000). *Introduction to Econometric Theory*. Institute of Economics, Academia Sinica, Taipei 115, Taiwan.

- [35] Lewis, P. A. W., Ray, B. K. (2002). Nonlinear modelling of periodic threshold autoregressions using TSMARS. *Journal of Time Series Analysis* 23(4) :459–471.
- [36] McLeod, A. I. (1994). Diagnostic Checking Periodic Autoregression Models With Application. *Journal of Time Series Analysis* 15(2), 221-233.
- [37] Merzougui, M., Dridi, H. and Chadli, A. (2016). Test for periodicity in restrictive EXPAR models. *Communication in Statistics- Theory and Methods*, 45 :9, 2770-2783.
- [38] Merzougui, M. (2017). Estimation in periodic restricted EXPAR(1) models. *Communications in Statistics - Simulation and Computation*, D.O.I. 10.1080/03610918.2017.1361975.
- [39] Merzougui, M., Becila, S. (2019). Least squares estimation in periodic restricted EXPAR(P) models. *Communication in Statistics- Advances in Theory and Applications*, <http://jyotiacademicpress.org/jyotic/journalview/35/article/43/75>.
- [40] Messaoud, A., Weihs, C., and Hering, F. (2006) : Nonlinear time series modelling : Monitoring a drilling process. In *Spiliopoulou M., Kruse R., Borgelt C., Nürnberger A., Gaul W. (eds) From Data and Information Analysis to Knowledge Engineering*, pages 302-309. Springer, Heidelberg.
- [41] Ozaki, T. (1980). Non-linear time series models for non-linear random vibrations, *Journal of Applied Probability* 17, 84–93.
- [42] Ozaki, T. (1982). The statistical analysis of Perturbed Limit Cycle Processes Using Nonlinear Time Series Models, *Journal of Time Series Analysis*, 3(1), 29-41.

Bibliographie

- [43] Ozaki, T. (1985). Non linear Time Series Models and Dynamical System. *Handbook of statistics*, 5, 25-83.
- [44] Priestley, M. B. (1988). *Non-linear and Non-stationary Time Series Analysis*, Academic Press, New York.
- [45] Shao, Q. (2006). Mixture periodic autoregressive time series models, *Statistics & Probability Letters* 76(6), 609-618.
- [46] Shi, Z. and Aoyama, H. (1997). Estimation of exponential autoregressive time series model by using genetic algorithm. *Journal of Sound and Vibration*, 205(3), 309-321.
- [47] Shi, Z., Tamura, Y., and Ozaki, T. (2001). Monitoring the stability of BWR oscillation by nonlinear time series modeling. *Annals of Nuclear Energy*, 28, 953-966.
- [48] Terui, N. and Van Dijk, H. K. (1999). Combined Forecasts from Linear and Nonlinear Time Series Models. *Tinbergen Institute Discussion Paper Series. Econometric Institute Report EI-9949/A*.
- [49] Tesfaye, Y.G., Meerschaert, M. M. and Anderson, P. L. (2006). Identification of PARMA Models and Their Application to the Modeling of River Flows. *Water Resources Research*, 42 (1), W01419.
- [50] Tjøstheim, D. (1986). Estimation in Nonlinear Time Series Models. *Elsevier Science Publishers B.V. North-Holland. Stochastic Processes and their Applications* 21 (1986), 251-273.
- [51] Tong, H (1978). On threshold models. In *Pattern Recognition and Signal Processing*, (ed. C. H. Chen). Amsterdam : Sijhoff & Noordhoff.
- [52] Tong, H. (1990). *Nonlinear Time Series : a Dynamical System Approach*. Oxford University Press, Oxford.

Bibliographie

- [53] Vecchia, A.V. and Ballerini, R. (1991). Testing for periodic autocorrelations in seasonal time series data. *Biometrika* 78, 53–63.