



HAL
open science

Modélisation longue mémoire multivariée : applications aux problématiques du producteur d'EDF dans le cadre de la libéralisation du marché européen de l'électricité

Abdou Ka Diongue

► **To cite this version:**

Abdou Ka Diongue. Modélisation longue mémoire multivariée : applications aux problématiques du producteur d'EDF dans le cadre de la libéralisation du marché européen de l'électricité. Mathématiques [math]. École normale supérieure de Cachan - ENS Cachan, 2005. Français. NNT : . tel-00132732

HAL Id: tel-00132732

<https://theses.hal.science/tel-00132732>

Submitted on 22 Feb 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N^o ENSC-2005-54



**THESE DE DOCTORAT
DE L'ECOLE NORMALE SUPERIEURE DE CACHAN**

Présentée par
Monsieur Abdou Kâ DIONGUE

**pour obtenir le grade de
DOCTEUR DE L'ECOLE NORMALE SUPERIEURE DE
CACHAN**

Domaine :
Mathématiques Appliquées

Sujet de la thèse :

**Modélisation longue mémoire multivariée : applications
aux problématiques du producteur d'EDF dans le cadre
de la libéralisation du marché européen de l'électricité**

Thèse présentée et soutenue à Cachan le 3 octobre 2005 devant
le jury composé de :

M. Yves Meyer	Professeur des Universités	Président
M. Rodney Wolff	Professeur des Universités	Rapporteur
M. Vêlayoudom Marimoutou	Professeur des Universités	Rapporteur
M. Aliou Diop	Université Gaston Berger (HDR)	Examineur
M. Bertrand Vignal	Electricite de France	Examineur
Mme Esther Ruiz	Professeure des Université	Examinatrice
Mme Dominique Guégan	Professeure des Universités	Directrice de thèse

Laboratoire IDHE (MORA)
ENS CACHAN/CNRS/UMR 8533
61, avenue du Président Wilson, 94235 CACHAN Cedex (France)

Remerciements

Je tiens à remercier, Dominique Guégan, directrice du département d'économie et gestion de l'Ecole Normale Supérieure (ENS) de Cachan pour m'avoir accueillie au sein de l'équipe Modélisation, Risque et Applications (MORA) et pour avoir dirigé ma thèse. L'enthousiasme, l'intuition scientifique et la tenacité dont elle a fait preuve, ainsi que la confiance qu'elle m'a accordée, ont fortement contribué à l'aboutissement de ce travail.

Toute ma reconnaissance au département OSIRIS d'EDF R&D qui a financé ma thèse. Merci à Bertrand Vignal, ingénieur chercheur au département OSIRIS d'EDF, pour sa disponibilité, ses conseils ainsi que la liberté qu'il a su me donner.

Je remercie Monsieur Vélayoudom Marimoutou, professeur à l'université d'Aix-en-provence et Monsieur Rodney Wolff, professeur à l'université de Queensland Brisbane (Australie), pour avoir été rapporteur de cette thèse.

Je voudrais exprimer ma gratitude à Monsieur Yves Meyer, professeur à l'ENS de Cachan, pour l'intérêt qu'il a bien voulu porter à ce travail et pour avoir accepté de présider mon jury. Tous mes remerciements à Aliou Diop, maître de conférence à l'université Gaston Berger de Saint-Louis du Sénégal, et Esther Ruiz, professeure à l'Université Carlos III de Madrid (Espagne), pour avoir bien voulu être membre du jury.

Je tiens à saluer tous les agents du département OSIRIS en particulier ceux des groupes R39 et R31 ainsi que tous les doctorants et jeunes docteurs de l'équipe MORA qui m'ont soutenu durant ces trois années.

Un grand merci à l'ensemble de mes amis et plus particulièrement à Omar, Mayatta, Jaya, Nadir, Rachid, Bosse, Babacar, Ndome, Thiaré, Moussa Salame qui m'ont encouragé pendant mon travail de recherche.

Enfin, j'adresse mes remerciements à Adjì Seynabou Sow pour son aide, sa présence et sa patience tout au long de mes recherches ainsi que toute ma famille qui a été d'un grand soutien.

Tout travail scientifique est le fruit de multiples collaborations. Une personne

apporte ses connaissances du sujet, une autre ses aptitudes techniques et une dernière son soutien moral dans les moments difficiles. Malheureusement, les publications scientifiques ne retiennent que certains de ces participants.

A

*mon grand-père feu El hadji Alhassane Salame,
mon amour Adjil Seynabou,
ma mère, mon père et mes frères et soeurs.*

Table des matières

Introduction	10
0.1 La libéralisation des marchés de l'électricité	11
0.2 Une revue de la littérature	11
0.3 Notre contribution	14
1 Généralités	19
1.1 Introduction	20
1.2 Introduction aux processus Longue Mémoire Généralisés . . .	23
1.2.1 Les polynômes de Gegenbauer	23
1.2.2 Propriétés probabilistiques	24
1.2.3 Méthodes de simulations numériques	27
1.2.4 Estimation des paramètres	28
1.2.5 Prévision	30
1.3 Processus Longue Mémoire combinés à de l'hétéroscédasticité .	31
1.3.1 Présentation du modèle	31
1.3.2 Propriétés probabilistiques	32
1.3.3 Estimation des Paramètres	34
1.4 Introduction aux processus Longue Mémoire Multivariés . . .	35
1.4.1 Présentation du modèle	35
1.4.2 Propriétés probabilistiques	36
1.4.3 Simulations numériques	37
1.4.4 Estimation des paramètres	38
2 Les Processus GIGARCH à k facteurs	41
2.1 Introduction	42
2.2 Présentation du Modèle	43
2.3 Estimation des paramètres du modèle	46
2.3.1 La méthode des moindres carrés conditionnels	46

2.3.2	Estimation des paramètres par la méthode de Whittle	55
2.4	Simulations numériques	60
2.5	Prévision avec un processus GIGARCH à k facteurs	61
2.5.1	Prévision en moyenne	62
2.5.2	Prévision en variance	62
2.6	Conclusions et perspectives	65
3	Les Processus Longue Mémoire Généralisés Multivariés	66
3.1	Introduction	67
3.2	Les processus GARMA à k -facteurs multivariés	68
3.2.1	Présentation du modèle	68
3.2.2	Propriétés probabilistiques	69
3.3	Procédure d'estimation des paramètres	73
3.4	Prévision	73
3.5	Conclusions et perspectives	78
4	Application aux prix spot de marchés de l'électricité	79
4.1	Introduction	80
4.2	Présentation des données	81
4.3	Approche univariée	86
4.4	Approche multivariée	91
4.5	Comparaison des résultats	93
4.6	Conclusions et perspectives	97
	Conclusion générale	99
	Bibliographie	100
	Annexes	113
	A Comparaison des méthodes d'estimation	114
	B RMSE heure par heure	121

Liste des tableaux

4.1	Statistiques descriptives des prix spot des marchés allemands et français sur la période du 01/01/2002 au 31/12/2003.	85
4.2	Paramètres longue mémoire estimés pour le modèle (4.1) pour les séries log prix $(X_{it})_t$ $i = 1, 2$	87
4.3	Statistiques descriptives des résidus $(\varepsilon_{it})_t$, $i = 1, 2$, issus de la modélisation (4.1).	88
4.4	Estimation des paramètres pour les modèles 4.2-4.3 et 4.4-4.5.	90
4.5	Paramètres estimés pour le modèle VAR défini par 4.13	94
A.1	Paramètres estimés pour un modèle GIGARCH Gaussien centré réduit défini par (2.40)-(2.42) ($n = 300$).	115
A.2	Paramètres estimés pour un modèle GIGARCH Student à 5 degrés de liberté défini par (2.40)-(2.42) ($n = 300$).	116
A.3	Paramètres estimés pour un modèle GIGARCH Gaussien centré réduit défini par (2.40)-(2.42) ($n = 500$).	117
A.4	Paramètres estimés pour un modèle GIGARCH Student à 5 degrés de liberté défini par (2.40)-(2.42) ($n = 500$).	118
A.5	Paramètres estimés pour un modèle GIGARCH Gaussien centré réduit défini par (2.40)-(2.42) ($n = 1000$).	119
A.6	Paramètres estimés pour un modèle GIGARCH Student à 5 degrés de liberté défini par (2.40)-(2.42) ($n = 1000$).	120

Table des figures

4.1	Prix spot horaires moyens des marchés français et allemand pour la période allant du 01/01/2002 au 31/12/2004	82
4.2	Evolution des prix spot horaires des marchés EEX (en haut à gauche), Powernext (en haut à droite) pour la période du 01/01/2002 au 31/12/2004.	83
4.3	Zoom sur un mois du 1/03/2002 au 30/03/2002 (en bas).	83
4.4	Evolution à l'échelle logarithmique des prix spot horaires des marchés EEX (en haut à gauche), Powernext (en haut à droite) pour la période du 01/01/2002 au 31/12/2003.	84
4.5	Histogramme des prix spot horaires pour la période du 01/01/2002 au 31/12/2004 des marchés EEX (en haut à gauche), Powernext (en haut à droite).	84
4.6	FAC et Périodogramme des prix spot d'électricité à l'échelle logarithmique du marché allemand EEX.	85
4.7	FAC et Périodogramme des prix spot d'électricité à l'échelle logarithmique du marché français Powernext.	86
4.8	Évolution des processus résiduels $(\varepsilon_{it})_t$, $i = 1, 2$, issus de la modélisation (4.1).	87
4.9	Histogramme des processus résiduels $(\varepsilon_{it})_t$, $i = 1, 2$, issus de la modélisation (4.1).	88
4.10	FAC et PACF des processus $(\varepsilon_{1t})_t$, issus de la modélisation (4.1).	89
4.11	FAC et PACF du processus $(\varepsilon_{2t})_t$, issus de la modélisation (4.1).	89
4.12	FAC des processus $(\xi_{it})_t$, $i = 1, 2$	90
4.13	FAC croisées des processus $(\varepsilon_{it})_t$, $i = 1, 2$	92
4.14	FAC des processus $(\xi_{it})_t$, $i = 1, 2$	93
4.15	Résultats de prévision en $J + 30$ de prix spot d'électricité du marché allemand pour la période du 01/12/2004 au 31/12/2004 et du 01/07/2004 au 31/07/2004.	96

4.16	Résultats de prévision en $J + 30$ de prix spot du marché français pour la période du 01/12/2004 au 31/12/2004 et du 01/07/2004 au 31/07/2004. .	96
4.17	RMSE en moyenne (EEX à gauche et Powernext à droite) pour une prévision en $J + 30$	97
B.1	RMSE heure par heure des prix spot français pour un horizon de prévision d'un mois.	122
B.2	RMSE heure par heure des prix spot français pour une prévision en $J + 30$.	123
B.3	RMSE heure par heure des prix spot français pour une prévision en $J + 30$.	124
B.4	RMSE heure par heure des prix spot français pour une prévision en $J + 30$.	125
B.5	RMSE heure par heure des prix spot allemand pour une prévision en $J + 30$.	126
B.6	RMSE heure par heure des prix spot allemand pour une prévision en $J + 30$.	127
B.7	RMSE heure par heure des prix spot allemand pour une prévision en $J + 30$.	128
B.8	RMSE heure par heure des prix spot allemand pour une prévision en $J + 30$.	129

Introduction

0.1 La libéralisation des marchés de l'électricité

Après la deuxième guerre mondiale, plusieurs pays dont la France et le Royaume Uni ont décidé de nationaliser les compagnies d'électricité pour les transformer en monopole national d'Etat. Dès 1980, la Communauté Economique Européenne a formulé un projet de démantèlement des entraves aux échanges d'énergies entre états membres. Une directive européenne a imposé dès 1999 aux états membres une ouverture progressive de leur marché de l'électricité à la concurrence. La dérégulation, comme prévue par cette directive européenne, envisage de dissocier une partie de l'activité physique de l'activité commerciale. La partie de l'activité physique qui consiste à transporter-distribuer de l'électricité peut demeurer un monopole naturel et l'activité commerciale qui consiste à vendre ou acheter des kWh peut être soumise à la concurrence. Par exemple en France, les consommateurs éligibles qui consomment plus de 40 GWh par an ont pu dès le 1er janvier 1999 choisir leur fournisseur. Ce seuil fut ensuite abaissé à 20 GWh au 1er janvier 2000 et à 9 GWh au 1er janvier 2003 ce qui a provoqué une ouverture de marché de l'ordre 33%. En 1990, une partie de l'Union Européenne (Suède, Danemark, Finlande et la Norvège) a mis au point une variante de ce modèle sur le nom de Nordpool de l'électricité. En Scandinavie, les entreprises ne sont pas systématiquement privatisées et deux grandes compagnies d'électricité y ont subsisté : la société d'Etat suédoise VATTENFALL et le conglomérat finlandais FORTUM. En Espagne, ce même modèle a conduit à la constitution de deux pôles : le groupe ENDESA et le groupe IBERDROLA. En France, Italie, Grèce et Irlande, les opérateurs historiques ont conservé un poids important sur leur ancienne zone de monopole. En Allemagne 4 grands opérateurs sont présents.

0.2 Une revue de la littérature

Dans le contexte actuel de dérégulation et de libéralisation des marchés de l'électricité, la modélisation et la prévision de prix de l'électricité sont de plus en plus étudiées dans la littérature statistique. Le plus souvent, les travaux sont orientés vers la compréhension de la dynamique de formation des prix et l'obtention de prévisions de prix satisfaisantes. En effet, les prévisions de

prix d'électricité peuvent aider les producteurs à mieux se positionner face au marché et à s'engager efficacement sur leurs contrats bilatéraux. Par exemple, à court terme, un producteur "price-maker" utilise la prévision "day-ahead" pour établir son planning optimal de production et pour construire sa stratégie d'offre. A moyen terme, un producteur "price-taker", au contraire, a besoin de prévoir ce prix d'équilibre sur plusieurs mois afin de rendre ses contrats bilatéraux fructueux. Pour les mêmes raisons que les producteurs, les détaillants et les grands consommateurs ont besoin des prévisions à court et moyen terme du prix d'équilibre de l'électricité. Cependant, l'électricité n'étant pas une comodité stockable, plusieurs travaux ont montré que l'utilisation des approches considérées pour modéliser les comportements observés sur les prix de comodités énergétiques stockables n'était pas réaliste (pour plus de détails, voir Tipping *et al.* (2004), Bunn *et al.* (2004) ou Burger *et al.* (2004)).

Depuis longtemps, la prévision à moyen et court terme de prix de comodités énergétiques comme le gaz, le fuel, le charbon, le pétrole etc... a fait l'objet de plusieurs travaux. On peut citer Fei *et al.* (2000) ou Breipohl (2002) et Koreneff *et al.* (1998) qui ont utilisé les modèles ARIMA (*Auto Regressive Integrated Moving Average*) pour la prévision de prix de pétrole et du gaz naturel. Wang *et al.* (1997) et Huang *et al.* (2000) ont utilisé ces mêmes modèles pour prévoir la demande en électricité. Récemment, les chercheurs ont commencé à s'intéresser à la prévision des prix de l'électricité. Deux grandes classes de modèles ont été retenues : les modèles de fondamentaux et les modèles économétriques ou statistiques. Dans les modèles de fondamentaux, les prix de l'électricité sont obtenus à partir de modèles liant le coût moyen de production et la consommation moyenne d'électricité. Plusieurs chercheurs ont utilisé ces modèles de fondamentaux pour analyser la formation des prix de l'électricité et fournir des résultats de prévision assez satisfaisants. Parmi eux, on peut citer Fleten et Wallace (1998) qui ont proposé une approche basée sur le calcul du prix d'équilibre théorique du marché. Eydeland et German (1999) et Skanttze *et al.* (2000) ont considéré une approche basée sur une modélisation directe de la fonction d'offre. Vehviläinen et Pyykkönen (2005) ont utilisé une approche combinant les modèles stochastiques et les modèles de fondamentaux et l'ont appliqué aux prix d'électricité du marché du Nordpool. Remarquons que l'approche par les modèles de fondamentaux nécessite une connaissance pertinente des données à utiliser. La collecte et la

mise à jour des données étant très complexes, on en déduit alors que l'utilisation des modèles de fondamentaux pour la génération de scénarios reste très difficile. D'autres auteurs ont utilisé les modèles statistiques pour l'analyse et la prévision de prix spot de marchés de l'électricité. On peut citer, Fosso *et al.* (1999) qui ont utilisé le modèle AR (*Auto Regressive*) pour fournir des prévisions à court terme de prix spot du marché norvégien de l'électricité. Szkuta *et al.* (1999), Gao *et al.* (2000), Ni *et al.* (2001) ou Ying *et al.* (2001) ont utilisé les réseaux de neurones pour répondre à la question de la prévision des prix de l'électricité. Ramsey *et al.* (1998) ont proposé une approche hybride combinant les réseaux de neurones et les modèles de logique floue (fuzzy-logic) et ont appliqué leur résultat sur le marché anglais. Pour l'évaluation des prévisions des prix à horizon d'un jour, Zhang *et al.* (2000) ont proposé un modèle stochastique alors que Espínola *et al.* (2002) ont proposé trois types de modèles pour les marchés espagnol et californien (regression dynamique, fonction de transfert et ARIMA) afin de comparer leurs capacités prédictives. Ces derniers ont conclu que les prévisions des prix de l'électricité dépendent très fortement de la demande en électricité. Les ondelettes ont été utilisées par Kim *et al.* (2002) pour des prévisions à court terme des prix de l'électricité. En 2002, Atkins et Chen ont appliqué le test de racine unitaire de Dickey Fuller sur les données de prix de l'électricité du marché de l'Alberta (Canada) pour étudier la stationnarité des données. Ce test a rejeté les hypothèses de stationnarité et de processus intégré d'ordre un. Ils ont alors identifié la présence d'effet longue mémoire sur leurs données et les ont décrites par un modèle FARIMA. Tous ces auteurs ont proposé des modèles permettant de prendre en compte certaines caractéristiques observées sur les prix spot de l'électricité comme par exemple la dépendance long terme (au sens de la décroissance lente de la fonction d'autocorrélation), les cycles plus ou moins fixes et la volatilité non constante. A ce jour, aucun travail n'a proposé un modèle permettant de modéliser à la fois l'ensemble de ces composantes observées sur les prix spot. D'autre part, les marchés européens de l'électricité étant interconnectés, à notre connaissance, il n'y a aucun chercheur qui s'est intéressé à l'apport de la modélisation jointe de prix spot pour la prévision. Nous allons tenter de répondre à ces questions dans ce travail de thèse CIFRE effectué au sein de la division Recherche et Développement d'Electricité de France (EDF).

0.3 Notre contribution

Le chapitre 1 consiste en une suite de rappels permettant de définir l'ensemble des notions que nous utilisons dans la suite de ce travail. Nous rappelons ainsi trois types de processus et donnons leurs propriétés : les processus longue mémoire généralisés (section 1.2), les processus FARIMA(p, d, q) – GARCH(r, s) (section 1.3) et les processus longue mémoire généralisés multivariés (section 1.4). D'une part, nous nous intéressons aux processus longue mémoire généralisés. Nous donnons les expressions analytiques de la densité spectrale et de la fonction d'autocorrélation (voir Giraitis et Leipus (1995) ou Woodward *et al.* (1998)). Nous rappelons deux méthodes de simulations numériques : la première basée sur l'expression moyenne mobile tronquée (Gray *et al.* (1989)) alors que la seconde est basée sur la décomposition de Choleski de la matrice de variance-covariance du processus (Geweke et Porter-Hudak (1983)). Nous rappelons la méthode de pseudo-maximum de vraisemblance de Whittle dans le cas où les G-fréquences sont supposées connues ainsi que les propriétés asymptotiques de l'estimateur, l'expression analytique du prédictor de moindres carrés à l'horizon h et les bornes de confiance de l'erreur de prévision (voir Ferrara (2000) ou Ferrara et Guégan (2001b)). D'autre part, nous rappelons quelques propriétés des processus longue mémoire à variances hétéroskédastiques (FARIMA(p, d, q)-GARCH(r, s)). Nous donnons les conditions nécessaires de stationnarité, d'inversibilité et l'expression analytique de la fonction d'autocorrélation. Nous rappelons aussi la méthode de la somme des carrés conditionnels (*CSS*) ainsi que les propriétés asymptotiques de l'estimateur CSS (voir Ling et Li (1997)). Enfin, nous rappelons les processus longue mémoire multivariés (VARFIMA). Nous donnons les hypothèses nécessaires pour assurer la stationnarité, la causalité et l'inversibilité de ces processus (voir Sowell (1987) et (1989)). Une méthode spectrale de simulation numérique est présentée (voir Chambers (1995)). Nous nous intéressons à la méthode de pseudo-maximum de vraisemblance pour l'estimation des paramètres ainsi qu'aux propriétés asymptotiques de l'estimateur (voir Hosoya (1996) et (1997)).

Le chapitre 2 est consacré aux processus GIGARCH à k facteurs. Dans ce chapitre, la partie estimation et prévision est nouvelle. Nous définissons ce processus et précisons les hypothèses permettant d'assurer les conditions de stationnarité et d'inversibilité (voir Guégan (2000) et (2003)). Nous nous inté-

ressons particulièrement à l'estimation des paramètres du processus et nous proposons deux méthodes d'estimation des paramètres d'un processus GIGARCH à k facteurs dans le cas où les G-fréquences sont supposées connues. La première méthode est basée sur une extension de la méthode de la somme des carrés conditionnels (CSS). Nous étudions les propriétés asymptotiques de l'estimateur CSS dans le cas de bruit gaussien et student (Théorèmes 2.3.1 et 2.3.2). Une seconde méthode basée sur une extension de la méthode de Whittle est étudiée ainsi que ses propriétés asymptotiques (Théorèmes 2.3.3-2.3.5). Nous proposons ensuite un plan de simulations numériques afin de comparer les propriétés asymptotiques de ces méthodes d'estimation. Nous faisons varier la taille de l'échantillon ainsi que les paramètres du modèle. La méthode de pseudo-maximum de Whittle est alors privilégiée par rapport à la méthode CSS. Nous étudions enfin les méthodes de prévision en moyenne et en variance pour les processus GIGARCH à k facteurs. Nous donnons les expressions analytiques du prédicteur des moindres carrés du processus GIGARCH à k facteurs (Théorèmes 2.5.1 et 2.5.2). Ces résultats ont fait l'objet de publications et de présentations à des congrès :

Diongue, A.K., Guégan D., Vignal B. (2003) Processus GIGARCH à k facteurs : estimation et application aux prix spot de marché de l'électricité, Note de Recherche MORA, 2003-14, Ecole Normale Supérieure de Cachan, France, 1-40.

Diongue A.K. (2003) Application des processus GIGARCH à k facteurs aux marchés spot de l'électricité, Note HR-39/03/026/A, Département OSIRIS, EDF R&D, Clamart.

Diongue, A.K. et Guégan D. (2004) Estimating parameters of a k -factor GIGARCH process, C.R.Acad.Sci. Paris I 339(6) : 435-440.

Diongue A.K., Guégan D., Vignal B. (2005a) Forecasting market spot prices electricity with a k -factor GIGARCH process, Note de Recherche MORA, 2004-09, Ecole Normale Supérieure de Cachan, France, 1-31, (*submitted Journal of Forecasting*).

Diongue A.K., Guégan, D. Vignal, B. (2005b) Prévision de prix spot de marchés de l'électricité à l'aide des processus GIGARCH à k facteurs, Note HR-31/05/004/A, Département OSIRIS, EDF R&D, Clamart, 1-21.

Diongue A.K., Guégan D., Vignal B. (2004) A k -factor GIGARCH process : estimation and application on electricity market spot prices, Proceedings of 8th International Conference on Probabilistic Methods Applied to Power Systems, Iowa State University, Ames, 1-7 USA.

Diongue, A.K. Estimation des paramètres d'un processus GIGARCH à k facteurs, Séminaire MORA, mai 2003, Cachan, France.

Diongue, A.K. Processus GIGARCH à k facteurs : Intérêt pour les marchés spot de l'électricité ?, Séminaire AFP, Juin 2003, EDF R&D Clamart, France.

Diongue, A.K. Processus GIGARCH à k facteurs et modélisation des prix spot de l'électricité, Séminaire AFP, novembre 2003, EDF R&D Clamart, France.

Diongue, A.K. Processus GIGARCH à k facteurs : estimation et application aux prix spot de marché de l'électricité, Conférence AFFI, décembre 2003, Paris, France (avec D. Guégan et B. Vignal).

Le chapitre 3 est consacré aux processus longue mémoire généralisés multivariés (MVGARMA) qui sont une généralisation naturelle des processus VARFIMA. Les résultats nouveaux de ce chapitre concernent l'étude probabilistique ainsi que les méthodes d'estimation et de prévision. Nous définissons ces processus. Nous étudions les conditions de stationnarité, d'inversibilité et de causalité de processus MVGARMA à k facteurs. Nous introduisons un certain nombre d'hypothèses essentielles pour assurer ces propriétés probabilistiques. Nous examinons les expressions moyenne mobile infinie et autoregressive infinie. Puis, nous nous intéressons à une procédure empirique d'estimation des paramètres à deux temps. D'abord, nous nous intéressons à l'estimation des paramètres longue mémoire de chaque série en supposant que les G -fréquences sont connues. Puis, en se servant de l'expression autoregressive tronquée, nous récupérons les résidus que nous modélisons ensuite avec un modèle ARMA multivarié (VARMA). Enfin, nous étudions les méthodes de prévisions pour ce type de processus. Nous donnons l'expression analytique du prédicteur de moindres carrés ainsi que l'intervalle de confiance du prédicteur des moindres carrés dans le cas où le processus d'innovations

est Gaussien. Ces résultats ont fait l'objet de prépublications et ont été présentés dans des conférences internationales :

Diongue A.K., Guégan D. Vignal, B. (2005c) Forecasting market spot prices using a k -factor Multivariate GARMA process (en cours rédaction).

Diongue A.K. Préviation de prix spot de marchés de l'électricité avec les processus longue mémoire généralisés multivariés, Séminaire Projet AFP, juin 2005, Clamart France.

Diongue A.K. Forecasting market spot prices using a k -factor Multivariate GARMA process, 25th International Symposium on Forecasting, San Antonio, June 12-15, 2005, USA (with Guégan D. and Vignal, B.)

Le chapitre 4 est consacré à l'application des processus proposés aux chapitres 2 et 3 sur les prix spot "day-ahead" des marchés de l'électricité allemand et français pour la période allant du 1 janvier 2002 au 31 décembre 2005. Les marchés de l'électricité constituent un nouveau champ d'application des processus longue mémoire. Nous commençons par une présentation détaillée des données en rappelant la formation des prix spot de l'électricité. Nous fournissons les propriétés statistiques des données et nous montrons, en utilisant une étude temporelle et spectrale, la présence de phénomène de dépendance long terme sur les prix spot. Cette étude permet de mettre en évidence la présence de cycles plus ou moins fixes, de pics de prix ainsi que de volatilité non constante sur les données de prix de l'électricité allemand et français. Nous considérons dans cette étude la période allant du 1 janvier 2002 au 31 décembre 2003 comme période de calage du modèle. La période allant du 1 janvier 2004 au 31 décembre 2004 sert pour la validation des modèles proposés. Pour chacun des marchés, nous ajustons un modèle GIGARCH à k facteurs. Une modélisation jointe des deux marchés avec un processus longue mémoire généralisé multivarié est également proposée. Pour valider les résultats, nous donnons des prévisions allant jusqu'à un horizon d'un mois. Nous calculons le RMSE (Roots Mean Square Error) dans chaque cas, puis nous donnons les résultats de prévision pour deux mois. Nous comparons ces résultats avec un modèle de Benchmark suggérant que la meilleure prévision d'une heure d'un jour donné est la moyenne de la même heure de l'année 2004. Nous donnons les RMSE pour chaque heure de la journée et ceci pour tous les jours

de l'horizon de prévision. Afin d'étudier le phénomène de causalité entre ces marchés d'électricité, nous comparons pour chacun des marchés les capacités prédictives pour des modèles GIGARCH à k facteurs et MVGARMA à k facteurs.

Chapitre 1

Généralités

1.1 Introduction

Depuis les travaux de Granger et Joyeux (1980) et Hosking (1981), les processus longue mémoire (FARIMA) ont été très abondamment étudiés dans la littérature. Ces processus permettent de prendre en compte dans une modélisation la présence de composante longue mémoire (au sens de la décroissance lente de la fonction d'autocorrélation) dans la série d'étude. Pour plus d'informations sur cette classe de processus, on renvoie aux articles de Adenstedt (1974), Granger et Joyeux (1980) et de Hosking (1981) ainsi qu'aux ouvrages de Brockwell et Davis (1987), Béran (1994) ou Guégan (1994). Les processus longue mémoire ont fait l'objet de plusieurs extensions. On rappelle dans ce chapitre 1 trois grandes classes de processus longue mémoire : les processus longue mémoire généralisés, les processus longue mémoire conditionnellement hétéroskédastiques et les processus longue mémoire multivariés.

La présence à la fois de comportements cycliques (périodiques ou non), de dépendance long terme ou de volatilité apparaît clairement dans de nombreuses séries chronologiques issues de différents champs d'application. Les séries horaires de prix spot de l'électricité des marchés allemand et français en sont un exemple.

Gray *et al.* (1989) ont introduit les processus GARMA (*Gegenbauer Autoregressive Moving Average*) afin de pouvoir modéliser à la fois les phénomènes de cycles plus ou moins fixes et de mémoire longue souvent observés dans une série chronologique. Ces processus ont été généralisés par Woodward *et al.* (1998) aux processus GARMA à k facteurs. Plusieurs auteurs se sont par la suite intéressés aux propriétés statistiques et probabilistiques de ces processus. Par exemple, Giraitis et Leipus (1995) et Woodward *et al.* (1998) ont étudié les propriétés statistiques de ces processus. Giraitis et Leipus (1995), Chung (1996a, 1996b), Yajima (1996), Hosoya (1997), Ferrara et Guégan (1999), Arteche et Robinson (2000), Velasco et Robinson (2000) ou Giraitis *et al.* (2001) ont proposé des méthodes d'estimation des paramètres des processus longue mémoire généralisés de type maximum de vraisemblance ou quasi-maximum de vraisemblance. Une méthode semiparamétrique (méthode locale) a été proposée par Ferrara et Guégan (2000).

D'autres études ont montré la présence à la fois de dépendance à long terme

et d'effet hétéroscédastique sur des données financières ou hydrologiques. Les processus FARIMA(p, d, q)–GARCH(r, s) permettant de prendre en compte simultanément ces deux caractéristiques ont été introduits. Baillie, Chung et Tieslau (1995) ont utilisé de tels processus pour analyser les séries de taux d'inflation de prix mensuels après la seconde guerre mondiale pour dix différents pays. Ling et Li (1997) ont proposé une méthode d'estimation des paramètres basée sur la somme des carrés conditionnels pour estimer les paramètres de ces processus. Ils ont appliqué leur résultat sur les données de l'indice Hang Seng (HSI) journalier sur la période allant de 1983 à 1984. En 2000, les processus GIGARCH ont été introduits par Guégan puis étendus par la même auteure aux processus GIGARCH à k facteurs en 2003. Elle étudie les propriétés probabilistes de ces processus. Chen et Yu (2005) ont introduit le processus ARFIMA-TGARCH. Ils ont proposé une méthode d'estimation des paramètres de ce type de processus basée sur une approche bayésienne et ont appliqué leur résultat sur les rendements journaliers d'actifs de six marchés.

Les processus longue multivariés (VARFIMA : *Vector Autoregressive Fractionally Integrated Moving Average*) permettant de modéliser conjointement plusieurs séries de données possédant un comportement longue mémoire ont été introduits. L'estimation des paramètres d'un processus VARFIMA a attiré l'attention de plusieurs chercheurs. Robinson (1995) a généralisé la méthode de Geweke Porter-Hudak (1983). Heyde et Gray (1993) et Hosoya (1996) ont proposé une méthode de quasi-maximum de vraisemblance (QMLE) généralisant la méthode de Whittle (1951). Hosoya (1996) a montré, sous des hypothèses très générales, n'incluant ni l'hypothèse de gaussianité ni celle de différence martingale des innovations, la consistance et la normalité asymptotique de l'estimateur QMLE. Sethuraman et Basawa (1997) ont proposé une méthode de maximum de vraisemblance (MLE) pour estimer les paramètres d'un processus VARFIMA. Ils ont considéré dans leur étude le cas particulier de processus VARFIMA où le vecteur de paramètre de mémoire longue est supposé constant. Ils ont ainsi établi les propriétés asymptotiques de cet estimateur. Lobato (1999) s'est intéressé à la méthode d'estimation semiparamétrique. Il a proposé une méthode locale à deux étapes et a étudié les propriétés de consistance et de normalité asymptotique de l'estimateur. Enfin, il a appliqué ses résultats sur les taux de change de quatre pays différents.

La question de la prévision avec les processus longue mémoire a connu des développements théoriques et pratiques très importants. Par exemple, Palma et Chan (1997) ont proposé une méthode d'estimation et de prévision de séries ayant des données manquantes. Ferrara et Guégan (2000c) ont étudié la prévision avec un processus longue mémoire généralisé. Ils ont appliqué leur résultat sur les données trimestrielles de trafic passagers de la RATP, sur les données climatiques mensuelles de température de l'air au Château de Nottingham et sur la série de taux de change mensuel du Franc français/Dollar US. Collet et Guégan (2003) ont étudié le prédicteur de moindres carrés de processus Gegenbauer non gaussien à k facteurs et ont appliqué leur résultat à différentes données (nombre de tâches solaires, debits du Rhin et relevés pluviométriques). On peut aussi citer Bhansali et kokoszka (2003) ou Bondon et Palma (2004) et Bhardwaj et Swanson (2005) qui ont montré l'utilité des processus FARIMA pour la prévision de séries macroéconomiques et financières. Ravishanker et Ray (2002) ont suggéré une méthode de prévision de processus VARFIMA basée sur une approche bayésienne. Ils l'ont appliqué aux données journalières de température relevées à la surface de la mer à l'échelle logarithmique mesurée sur la côte Californienne à Granite Canyon, Pacific Grove et Farallon Island.

Dans ce chapitre, on se propose de faire un bilan des connaissances, sous forme de présentation formelle de ces trois types de processus. Le premier paragraphe est consacré aux processus longue mémoire généralisés. On introduit d'abord les polynômes de Gegenbauer. Ensuite, on présente une définition des processus longue mémoire généralisés ainsi que quelques propriétés probabilistiques et statistiques. Des méthodes de simulations numériques sont aussi présentées. Pour finir ce paragraphe, on rappelle une méthode de prédicteur des moindres carrés des processus longue mémoire généralisés. Dans le second paragraphe, on s'intéresse aux processus longue mémoire à variance conditionnelle hétéroskédastique. On définit ici les processus FARIMA-GARCH et on étudie leurs propriétés probabilistiques. On présente enfin la méthode de la somme des carrés conditionnels pour l'estimation des paramètres de ce processus. Dans le dernier paragraphe, on étudie les processus longue mémoire multivariés (VARFIMA). On commence par définir ces processus. On s'intéresse très rapidement aux conditions de stationnarité et d'inversibilité du processus VARFIMA. Pour finir, on présente des méthodes de simulations numériques et on rappelle la méthode de quasi-maximum de vraisemblance

pour l'estimation des paramètres de ce type de processus.

1.2 Introduction aux processus Longue Mémoire Généralisés

Dans cette partie, on rappelle très rapidement les polynômes de Gegenbauer. On donne ensuite une définition du processus GARMA à k facteurs. Enfin, on présente les principales propriétés probabilistiques des processus GARMA à k facteurs.

1.2.1 Les polynômes de Gegenbauer

On va introduire dans cette section les polynômes de Gegenbauer. Les polynômes de Gegenbauer, notés $(C_j^{(d,\nu)})_{j \in \mathbb{N}}$, sont donnés par :

$$(1 - 2\nu z + z^2)^{-d} = \sum_{j \geq 0} C_j^{(d,\nu)} z^j, \quad (1.1)$$

où $|z| \leq 1$ et $|\nu| \leq 1$. Rainville (1960) montre que les coefficients $(C_j^{(d,\nu)})_{j \in \mathbb{N}}$ peuvent être calculés très simplement de la manière suivante :

$$\forall j \in \mathbb{N}, \quad C_j^{(d,\nu)} = \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} \frac{(-1)^k \Gamma(d+j-k) (2\nu)^{j-2k}}{\Gamma(d) \Gamma(k+1) \Gamma(j-2k+1)}, \quad (1.2)$$

où $\lfloor \frac{j}{2} \rfloor$ est la partie entière de $\frac{j}{2}$. On rappelle que la fonction Gamma, notée $\Gamma(\cdot)$, est définie, pour tout réel a positif ou nul, par :

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

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

Il est souvent plus simple et plus rapide de calculer ces polynômes de Gegenbauer $(C_j^{(d,\nu)})_{j \in \mathbb{N}}$ à partir de l'algorithme récursif suivant :

$$\begin{cases} C_0^{(d,\nu)} = 1 \\ C_1^{(d,\nu)} = 2d\nu \\ \forall j > 1, C_j^{(d,\nu)} = 2\nu \left(\frac{d-1}{j} + 1\right) C_{j-1}^{(d,\nu)} - \left(2\frac{d-1}{j} + 1\right) C_{j-2}^{(d,\nu)}. \end{cases} \quad (1.3)$$

Pour plus de détails sur les polynômes de Gegenbauer, on se réfère à Szëgo (1959), Rainville (1960) ou Magnus et *al.* (1966). Dans la suite, on donne une définition des processus GARMA à k facteurs (voir les articles de Gray et *al.* (1989), Robinson (1994c), Giraitis et Leipus (1995) et Woodward et *al.* (1998), Ferrara et Guégan (2000)). On donne dans la suite la définition des processus longue mémoire généralisés.

Définition 1.2.1 Soit $(\varepsilon_t)_{t \in \mathbb{Z}}$ un processus suivant un bruit blanc de variance constante finie σ_ε^2 . On dit que le processus $(X_t)_{t \in \mathbb{Z}}$ est un processus GARMA à k -facteurs s'il vérifie l'équation suivante :

$$\phi(B) \prod_{i=1}^k (I - 2\nu_i B + B^2)^{d_i} (X_t - \mu) = \theta(B) \varepsilon_t, \quad (1.4)$$

où k est un nombre entier fini, $|\nu_i| \leq 1$ pour $i = 1, \dots, k$, les d_i sont des nombres fractionnaires pour $i = 1, \dots, k$, et μ est la moyenne du processus. Les polynômes $\phi(B)$ et $\theta(B)$ sont respectivement les polynômes autorégressif d'ordre p et moyenne mobile d'ordre q . Ici, B est l'opérateur retard et I est l'opérateur identité. Pour $i = 1, \dots, k$, les fréquences $\omega_i = \cos^{-1}(\nu_i)$ sont appelées les fréquences de Gegenbauer (*G-fréquences*).

Remarque 1.2.1 Dans le cas où $k = 1$ et $\nu = 1$ le processus GARMA à k -facteurs devient un processus FARIMA($p, 2d_1, q$). Ainsi, les processus GARMA à k -facteurs sont une généralisation des processus FARIMA.

On rappelle ici quelques hypothèses permettant d'assurer la stationnarité et la causalité du processus.

- (H_0) : $\phi(B) = I - \phi_1 B - \dots - \phi_p B^p$ et $\theta(B) = I - \theta_1 B - \dots - \theta_q B^q$ sont deux polynômes n'ayant pas de racines communes et ayant toutes leurs racines en dehors du cercle unité.
- (H_1) : $|d_i| < \frac{1}{2}$ si $|\nu_i| < 1$ ou $|d_i| < \frac{1}{4}$ si $|\nu_i| = 1$ pour $i = 1, \dots, k$.

1.2.2 Propriétés probabilistiques

Dans cette section, on rappelle les résultats relatifs à la stationnarité et à l'inversibilité des processus longue mémoire généralisés. Ces propriétés, établies par Giraitis et Leipus (1995) et Woodward *al.* (1998), sont résumées dans la Proposition 1.2.1.

Proposition 1.2.1 *Sous les hypothèses (H_0) et (H_1) , le processus GARMA à k facteurs $(X_t)_{t \in \mathbb{Z}}$ défini en (1.4) est stationnaire, inversible et causal. Sa représentation autoregressive infinie est donnée par*

$$\varepsilon_t = \sum_{j=0}^{\infty} \alpha_j(d, \nu, \phi, \theta) X_{t-j}, \quad (1.5)$$

où le vecteur $d = (d_1, \dots, d_k)'$, $\nu = (\nu_1, \dots, \nu_k)'$ et où les poids $(\alpha_j(d, \nu, \phi, \theta))_{j \in \mathbb{N}}$ de cette représentation autorégressive infinie sont tels que $\alpha_0(d, \nu, \phi, \theta) = 1$ et pour tout $j \geq 1$, on a :

$$\alpha_j(d, \nu, \phi, \theta) = \pi_j(d, \nu) - \sum_{i=1}^{\min(j,p)} \phi_i \pi_{j-i}(d, \nu) + \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu), \quad (1.6)$$

et la représentation moyenne mobile infinie du processus $(X_t)_{t \in \mathbb{Z}}$ est donnée par l'équation suivante :

$$X_t = \sum_{j=0}^{\infty} \beta_j(d, \nu, \phi, \theta) \varepsilon_{t-j}, \quad (1.7)$$

où les poids $\beta_j(d, \nu, \phi, \theta)$ de cette représentation moyenne mobile infinie sont tels que $\beta_0(d, \nu, \phi, \theta) = 1$ et pour tout $j \geq 1$ on obtient :

$$\beta_j(d, \nu, \phi, \theta) = \psi_j(d, \nu) + \sum_{i=1}^{\min(j,p)} \phi_i \beta_{j-i}(d, \nu, \phi, \theta) - \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu), \quad (1.8)$$

avec

$$\pi_j(d, \nu) = \sum_{\substack{0 \leq l_1, \dots, l_k \leq j \\ l_1 + \dots + l_k = j}} C_{l_1}^{(-d_1, \nu_1)} \dots C_{l_k}^{(-d_k, \nu_k)}, \quad (1.9)$$

et

$$\psi_j(d, \nu) = \pi_j(-d, \nu). \quad (1.10)$$

Pour $k = 1$ et $\nu = 1$, on obtient le résultat dans le cas d'un processus FARIMA(p,2d,q).

Pour analyser une série chronologique par les processus longue mémoire généralisés, les statisticiens disposent de deux outils : la fonction d'autocorrélation empirique (FAC) et la densité spectrale empirique. Sans perte de généralités, on présente dans la suite des résultats relatifs aux comportements théoriques de ces deux fonctions dans le cas d'un processus Gegenbauer à k -facteurs.

Proposition 1.2.2 Soit $(X_t)_{t \in \mathbb{Z}}$ un processus Gegenbauer à k -facteurs stationnaire défini en (1.4). Alors :

– La densité spectrale de $(X_t)_{t \in \mathbb{Z}}$ est donnée par :

$$f_X(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \prod_{i=1}^k |2(\cos(\omega) - \nu_i)|^{-2d_i}, \quad (1.11)$$

où $0 \leq \omega \leq \pi$ et pour $i = 1, \dots, k$, les $\omega_i = \cos^{-1}(\nu_i)$ sont appelés les fréquences de Gegenbauer (G -fréquences).

– Soient $|\nu_i| < 1$ et $0 < d_i < \frac{1}{2}$ pour $i = 1, \dots, k$, la fonction d'autocovariance, notée $\rho(\cdot)$, de $(X_t)_{t \in \mathbb{Z}}$ a le comportement asymptotique suivant

$$\rho(h) \sim \sum_{i=1}^k h^{2d_i-1} \sin(\pi d_i - h\omega_i), \quad (1.12)$$

pour $h \rightarrow \infty$.

– Soient $\nu_i = 1$ et $0 < d_i < \frac{1}{4}$ pour $i = 1, \dots, k$, alors

$$\rho(h) \sim \sum_{i=1}^k h^{4d_i-1}, \quad (1.13)$$

pour $h \rightarrow \infty$.

– Soient $\nu_i = -1$ et $0 < d_i < \frac{1}{4}$ pour $i = 1, \dots, k$, alors

$$\rho(h) \sim (-1)^h \sum_{i=1}^k h^{4d_i-1}, \quad (1.14)$$

pour $h \rightarrow \infty$.

Ce résultat asymptotique de la fonction d'autocovariance est dû à Giraitis et Leipus (1995). On remarque que l'expression analytique (1.12) de la fonction d'autocovariance se décompose en k termes qui eux-même se décomposent en deux composantes : une composante assurant la décroissance lente de la fonction d'autocovariance et une autre faisant osciller celle-ci. Alors que les expressions (1.2.2) et (1.14) se décomposent en k termes, chacune d'entre elle assurant la décroissance lente de la fonction d'autocorrélation.

1.2.3 Méthodes de simulations numériques

On développe dans ce paragraphe deux méthodes de simulations numériques de processus GARMA à k -facteurs de taille finie n . Dans le chapitre 2, on utilisera des simulations de Monte-Carlo afin de valider et de comparer les méthodes d'estimation de paramètres proposées.

La première méthode, proposée par Gray *et al.* (1989), est une méthode très intuitive de simulation de processus GARMA à k facteurs. Cette méthode consiste à utiliser l'expression moyenne mobile infinie d'un processus Gegenbauer à k facteurs défini par l'équation (1.7) et à la tronquer à partir d'un rang M . Ainsi, les réalisations de ce processus, $(Y_t)_{t \in \mathbb{Z}}$, sont données par :

$$Y_t = \mu + \sum_{j=0}^M \psi_j(d, \nu) \varepsilon_{t-j}, \quad (1.15)$$

où μ est la moyenne du processus, $(\varepsilon_t)_{t \in \mathbb{Z}}$ est la réalisation d'un bruit blanc centré et les coefficients $(\psi_j(d, \nu))_{j \in \mathbb{N}}$ sont les poids de la forme moyenne infinie définis dans la Proposition 1.2.1. Les réalisations du processus GARMA à k facteurs, noté $(X_t)_{t \in \mathbb{Z}}$, sont alors obtenues à partir de la formule récursive suivante :

$$X_t = \sum_{j=1}^p \phi_j X_{t-j} + Y_t - \sum_{j=1}^q \theta_j Y_{t-j}, \quad (1.16)$$

Le choix de M est laissé à l'expérimentateur. Il dépend entièrement de la précision désirée. Gray *et al.* (1989) ont choisi $M = 290000$ dans leurs simulations de processus GARMA. Le principal défaut de cette méthode est qu'elle est très gourmande en temps calcul.

La deuxième méthode est une généralisation de la méthode proposée par Geweke et Porter-Hudak (1983) dans le cas des processus FARIMA. Cette méthode est basée sur la décomposition de Choleski de la matrice de variance-covariance de l'échantillon notée Σ . On a vu précédemment qu'il n'existait pas une expression directe de la fonction d'autocovariance d'un processus GARMA à k facteurs. On utilisera ainsi la densité spectrale pour calculer les éléments de la matrice variance-covariance du processus. En 2002, Bertelli et Caporin proposent une méthode, appelée "splitting method" afin de calculer les éléments de la fonction d'autocovariance de processus longue mémoire.

La représentation de Choleski de Σ est alors :

$$\Sigma = LL', \quad (1.17)$$

où L est une matrice triangulaire inférieure et L' sa transposée. Les n réalisations Y_1, \dots, Y_n du processus GARMA à k -facteurs, $(X_t)_{t \in \mathbb{Z}}$, sont obtenues de la manière suivante :

$$X = (X_1, \dots, X_n)' = L\varepsilon, \quad (1.18)$$

où $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ est la réalisation d'un processus bruit blanc centré. Cette méthode est avantageuse car elle est très simple à mettre en oeuvre et très rapide d'exécution. On note également qu'elle permet d'éviter les troncatures. Pour des informations plus détaillées sur ces méthodes de simulations numériques, on renvoie à Geweke et Porter-Hudak (1983), Beran (1994, p. 215), Ferrara et Guégan (2000) ou Bertelli et Caporin (2002) ou Caporin (2002).

La simulation de processus longue mémoire avec les ondelettes a récemment attiré l'attention de beaucoup de chercheurs. McCoy et Walden (1996) ont introduit une méthode de simulation numérique de processus Gaussiens fractionnairement intégrés en utilisant la transformée d'ondelettes discrètes. Utilisant les vaguelettes ("*Wavelet packets*") introduites par Coifman, Meyer et Wickerhauser (1992), Whitcher (2001) a proposé un algorithme pour la simulation numérique de processus longue mémoire généralisés. Sa méthode dépend à la fois des localisations des fréquences de Gegenbauer et de la taille de l'ondelette utilisée dans la transformée de vaguelettes discrètes. Collet et Fadili (2005) ont amélioré cette méthode en ne faisant dépendre leur algorithme que des localisations des fréquences de Gegenbauer.

1.2.4 Estimation des paramètres

Dans ce paragraphe, on s'intéresse à l'estimation des paramètres d'un processus GARMA à k facteurs dans le cas où les G-fréquences sont connues. On ne présente ici que la méthode de Whittle. Soit X_1, \dots, X_n une trajectoire de taille finie issue d'un processus GARMA à k facteurs, noté $(X_t)_{t \in \mathbb{Z}}$, centré, stationnaire, inversible, introduit dans la Définition (1.2.1), avec k fixé, et dont les G-fréquences ω_i , pour $i = 1, \dots, k$, sont connues. On s'intéresse à l'estimation du paramètre vectoriel $\psi = (\alpha, d)'$, où $\alpha = (\sigma_\varepsilon^2, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$ et $d = (d_1, \dots, d_k)'$ par la méthode du pseudo-maximum de vraisemblance

de Whittle. On note ψ_0 la vraie valeur de ψ . L'estimateur de Whittle est alors obtenu en minimisant la vraisemblance de Whittle définie par :

$$L_W(X, \psi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \log(f_X(\omega, \psi)) + \frac{I_n(X, \omega)}{f_X(\omega, \psi)} \right\} d\omega, \quad (1.19)$$

où $I_n(X, \omega_j)$ représente le périodogramme du processus $(X_t)_{t \in \mathbb{Z}}$ et est défini par :

$$I_n(X, \omega_j) = \frac{1}{2\pi n} \left| \sum_{t=1}^n X_t^2 e^{i\omega_j t} \right|^2. \quad (1.20)$$

Empiriquement, le périodogramme est le plus souvent évalué aux fréquences de Fourier définies par :

$$\omega_j = \frac{2\pi j}{n}, \quad 0 \leq j \leq \left[\frac{n}{2} \right].$$

La fonction f_X est la densité spectrale du processus GARMA à k facteurs définie par :

$$f_X(\omega) = \frac{\sigma_\varepsilon^2 |\theta(e^{i\omega})|^2}{2\pi |\phi(e^{i\omega})|^2} \prod_{j=1}^k |2 \cos(\omega_j) - \nu_j|^{-d_j}. \quad (1.21)$$

Cette méthode d'estimation a été étudiée par Hosoya (1997). Dans son article, il s'intéresse à des processus possédant une densité spectrale à singularités multiples, non restreintes à l'origine. Ferrara (2000) établit les propriétés asymptotiques de l'estimateur de Whittle en supposant l'hypothèse de Gaussianité des innovations. On remarque que ce résultat est un corollaire du Théorème 2.3 de Hosoya (1997) où l'auteur suppose certaines conditions de mélangeance sur le processus d'innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$, qui sont moins fortes que l'hypothèse de gaussianité généralement utilisée.

En supposant le processus Gaussien, on a le résultat suivant.

Proposition 1.2.3 *Soit $(X_t)_{t \in \mathbb{Z}}$ un processus GARMA à k facteurs centré stationnaire, inversible et Gaussien introduit dans la Définition 1.2.1, avec k fixé et dont les G -fréquences ω_i , pour $i = 1, \dots, k$, sont connues. Si $n \rightarrow \infty$, alors les estimateurs $\hat{\alpha}_n$ et \hat{d}_n de α et d sont asymptotiquement indépendants, et*

$$n^{\frac{1}{2}} (\hat{\alpha}_n - \alpha_0) \rightarrow \mathbb{N}(0, 4\pi V(\alpha)^{-1}), \quad (1.22)$$

où $V(\alpha)$ est la matrice de dimension $(p+q+1) \times (p+q+1)$ de terme général :

$$V(\alpha)_{ij} = \int_{-\pi}^{\pi} g^2(\omega) \frac{\partial g^{-1}(\omega)}{\partial \alpha_i} \frac{\partial g^{-1}(\omega)}{\partial \alpha_j} d\omega, \quad (1.23)$$

et

$$n^{\frac{1}{2}} (\hat{d}_n - d_0) \rightarrow \mathbb{N}(0, 4\pi V(d)^{-1}), \quad (1.24)$$

où $V(d)$ est la matrice de dimension $k \times k$ de terme général :

$$V(d)_{ij} = \int_{-\pi}^{\pi} \log \left| 4 \sin \left(\frac{\omega - \omega_i}{2} \right) \sin \left(\frac{\omega + \omega_i}{2} \right) \right| \log \left| 4 \sin \left(\frac{\omega - \omega_j}{2} \right) \sin \left(\frac{\omega + \omega_j}{2} \right) \right| d\omega. \quad (1.25)$$

1.2.5 Prévision

Dans cette section, on rappelle les méthodes de prévision d'un processus GARMA à k -facteurs Gaussien dont les paramètres sont connus. On rappelle aussi les propriétés statistiques de ces prévisions. Pour plus de détails sur ces résultats, on renvoie à Ferrara (2000) ou Ferrara et Guégan (2001b).

Proposition 1.2.4 *Soit $(X_t)_{t \in \mathbb{Z}}$ un processus GARMA à k -facteurs centré stationnaire, inversible et causal défini en (1.4). Le prédicteur des moindres carrés à l'horizon h , pour $h \geq 1$, noté $\hat{X}_t(h)$, est donné par l'expression suivante :*

$$\hat{X}_t(h) = \xi_h(B) X_t, \quad (1.26)$$

où

$$\xi_h(B) = \sum_{j \geq 0} \xi_j(d, \nu, \phi, \theta, h) B^j. \quad (1.27)$$

Les poids $(\xi_j(d, \nu, \phi, \theta, h))_{j \in \mathbb{N}}$ sont donnés par :

$$\xi_j(d, \nu, \phi, \theta, h) = - \sum_{i=0}^{h-1} \beta_i(d, \nu, \phi, \theta) \alpha_{h+j-i}(d, \nu, \phi, \theta) \quad \text{pour } j \geq 0, \quad (1.28)$$

où les coefficients $(\beta_j(d, \nu, \phi, \theta))_{j \in \mathbb{N}}$ et $(\alpha_j(d, \nu, \phi, \theta))_{j \in \mathbb{N}}$ sont donnés dans la Proposition 1.2.1.

Un prédicteur autorégressif naturel à l'horizon h , pour $h \geq 1$, peut être obtenu à partir de la Proposition 1.2.4. En effet, on obtient à partir de la représentation AR(∞) du processus GARMA à k facteurs l'équation suivante :

$$\hat{X}_t^{AR}(h) = - \sum_{j \geq 0} \alpha_{j+h}(d, \nu, \phi, \theta) X_{t-j}. \quad (1.29)$$

Il en découle très facilement, l'expression de $\hat{X}_t(h)$:

$$\hat{X}_t = \sum_{j=0}^{h-1} \beta_j(d, \nu, \phi, \theta) \hat{X}_t^{AR}(h-j). \quad (1.30)$$

On s'intéresse maintenant à l'erreur de prévision du prédicteur des moindres carrés, ainsi qu'à son intervalle de confiance.

Proposition 1.2.5 *Soit $(X_t)_{t \in \mathbb{Z}}$ un processus GARMA à k facteurs centré stationnaire, inversible et causal, dont le processus d'innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc Gaussien de variance σ_ε^2 finie. Alors, on a :*

$$\mathbb{P} \left\{ \hat{X}_t(h) - t_{1-\frac{\alpha}{2}} MSE(h)^{\frac{1}{2}} \leq X_{t+h} \leq \hat{X}_t(h) + t_{1-\frac{\alpha}{2}} MSE(h)^{\frac{1}{2}} \right\} = 1 - \alpha, \quad (1.31)$$

où $t_{1-\frac{\alpha}{2}}$ est le fractile de la loi gaussienne au risque α et $MSE(h)$ est l'erreur quadratique moyenne définie par :

$$MSE(h) = \sigma_\varepsilon^2 \sum_{j=0}^{h-1} \beta_j^2(d, \nu, \phi, \theta).$$

Enfin, notons que Collet et Guégan (2003) se sont intéressés au cas de la prévision de processus longue mémoire généralisés (Gegenbauer à k facteurs) non gaussiens.

1.3 Processus Longue Mémoire combinés à de l'hétéroscédasticité

1.3.1 Présentation du modèle

Dans ce paragraphe, on définit les processus FARIMA(p, d, q)-GARCH(r, s) introduit par Baillie *et al.* (1995). On s'intéresse ensuite aux propriétés probabilistiques et statistiques de ce processus.

Définition 1.3.1 *Soit $(\varepsilon_t)_{t \in \mathbb{Z}}$ un bruit blanc centré de variance σ_ε^2 . Un processus $(X_t)_{t \in \mathbb{Z}}$ est défini comme étant un processus FARIMA(p, d, q)-GARCH(r, s) stationnaire et inversible s'il vérifie l'équation suivante :*

$$\phi(B)(I-B)^d(X_t - \mu) = \theta(B)\varepsilon_t, \quad (1.32)$$

où

$$\varepsilon_t | \mathcal{F}_{t-1} \sim \mathbb{N}(0, \sigma_t^2), \quad \sigma_t^2 = a_0 + \sum_{i=1}^r a_i \varepsilon_{t-i}^2 + \sum_{i=1}^s b_i \sigma_{t-i}^2, \quad (1.33)$$

où d est un nombre fractionnaire, μ est la moyenne du processus, $\phi(B)$ et $\theta(B)$ sont deux polynômes définis dans la Définition 1.3.1 et vérifiant l'hypothèse (H_0) . Ici, B est l'opérateur retard, $a_0 > 0$, $a_1, \dots, a_r, b_1, \dots, b_s \geq 0$, r et s sont des entiers positifs et \mathcal{F}_{t-1} est le σ -algèbre engendré par toute l'information passée $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$.

On remarque que si d est un entier positif ou nul, alors les équations (1.32)-(1.33) définissent un processus ARIMA(p, d, q)-GARCH(r, s). On peut aussi remarquer que si la variance conditionnelle σ_t^2 , défini par (1.33), est constante alors le processus $(X_t)_{t \in \mathbb{Z}}$ suit un processus FARIMA(p, d, q) tel qu'il a été introduit par Granger et Joyeux (1980) et Hosking (1981).

1.3.2 Propriétés probabilistiques

Le processus GARCH(r, s) défini par l'équation (1.33) et introduit par Bollerslev (1986) peut bien sûr être remplacé par un autre modèle à variance conditionnelle hétéroscédastique tel que les modèles ARCH à seuil, introduit par Li et Li (1996). Dans le théorème suivant, on rappelle quelques propriétés de base du processus défini par l'équation (1.32)-(1.33). Donnons d'abord l'hypothèse (H_2) suivante :

- (H_2) : Supposons que $a_0 > 0$ et les coefficients $a_i, i = 1, \dots, r$ et $b_i, i = 1, \dots, s$ tels que $\sum_{i=1}^r a_i + \sum_{i=1}^s b_i < 1$, et $d < \frac{1}{2}$.

Théorème 1.3.1 *Supposons que l'hypothèse (H_2) soit vérifiée. Si $\phi(B) = \theta(B) = I$, alors il existe une solution unique (ε_t, X_t) \mathcal{F}_t -mesurable stationnaire au second ordre pour le modèle défini par les équations (1.32)-(1.33). La solution (ε_t, X_t) est donnée par :*

$$\varepsilon_t = Z_t \left[a_0 + \sum_{j=1}^{\infty} \delta' \left(\prod_{i=1}^j A_{t-i} \right) \xi_{t-j} \right]^{\frac{1}{2}} \quad a.s., \quad (1.34)$$

et

$$X_t = \sum_{k=0}^{\infty} \frac{(k+d-1)!}{k!(d-1)!} \varepsilon_{t-k} \quad a.s., \quad (1.35)$$

où $\xi_t = (a_0 Z_t^2, 0, \dots, 0, a_0, 0 \dots, 0)'_{(r+s) \times 1}$ (c'est à dire la première composante du vecteur vaut $a_0 Z_t^2$ et la $(r+1)$ ième composante est a_0), Z_t est une suite de variable aléatoire indépendante et identiquement distribuée de loi normale centrée réduite, $\delta = (a_1, \dots, a_r, b_1, \dots, b_s)'$ et la matrice A_t est définie par :

$$A_t = \left(\begin{array}{ccc|ccc} \alpha_1 Z_t^2 & \dots & \alpha_r Z_t^2 & \beta_1 Z_t^2 & \dots & \beta_s Z_t^2 \\ \mathbf{I}_{(r-1) \times (r-1)} & & O_{(r-1) \times 1} & & & O_{(r-1) \times s} \\ \hline \alpha_1 & \dots & \alpha_r & \beta_1 & \dots & \beta_s \\ & & O_{(s-1) \times r} & \mathbf{I}_{(s-1) \times (s-1)} & & O_{(s-1) \times 1} \end{array} \right),$$

avec $\mathbf{I}_{r \times r}$ est la matrice identité. Les processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ et $(X_t)_{t \in \mathbb{Z}}$ sont alors strictement stationnaires et ergodiques.

On s'intéresse maintenant aux propriétés statistiques de ce processus FARIMA(p, d, q)-GARCH(r, s). Le théorème suivant résume quelques résultats présentés par Ling et Li (1997).

Théorème 1.3.2 Soit $(X_t)_{t \in \mathbb{Z}}$ un processus FARIMA(p, d, q)-GARCH(r, s), correspondant à la Définition 1.3.1.

1. Si $d < \frac{1}{2}$ alors le processus $(X_t)_{t \in \mathbb{Z}}$ est stationnaire au second ordre et peut se réécrire sous sa forme moyenne mobile suivante :

$$X_t - \mu = \phi^{-1}(B) \theta(B) \sum_{k=0}^{\infty} \frac{(k+d-1)}{k! (d-1)!} \varepsilon_{t-k}, \quad (1.36)$$

où ε_t est défini par l'équation (1.34). Ainsi, le processus $(X_t)_{t \in \mathbb{Z}}$ est ergodique et strictement stationnaire.

2. Si $d > -\frac{1}{2}$ alors le processus $(X_t)_{t \in \mathbb{Z}}$ est inversible. Alors le processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ s'écrit sous sa forme autorégressive infinie suivante :

$$\varepsilon_t = \phi(B) \theta^{-1}(B) \sum_{k=0}^{\infty} \frac{(k-d-1)!}{k! (-d-1)!} (X_{t-k} - \mu). \quad (1.37)$$

On note par $\gamma(k) = \frac{\text{cov}(X_t, X_{t-k})}{\text{var}(X_t)}$ la fonction d'autocorrélation d'un processus FARIMA(p, d, q)-GARCH(r, s). Dans la Proposition 1.3.1, on donne, comme dans le théorème 2(4) de Hosking (1981), l'expression asymptotique de cette fonction d'autocorrélation.

Proposition 1.3.1 *Si $(X_t)_{t \in \mathbb{Z}}$ est un processus FARIMA(p, d, q)–GARCH(r, s) stationnaire et inversible défini par les équations (1.32)-(1.33), alors sa fonction d'autocorrélation vérifie :*

$$\gamma(k) \sim c |k|^{2d-1} \quad \text{quand } k \rightarrow \infty, \quad (1.38)$$

où c est une constante positive.

1.3.3 Estimation des Paramètres

Dans ce paragraphe, on présente la méthode de maximum de vraisemblance proposée par Ling et Li (1997) pour estimer les paramètres d'un processus FARIMA(p, d, q)–GARCH(r, s). Supposons que X_1, \dots, X_n sont générés par un processus FARIMA(p, d, q)–GARCH(r, s), défini par les équations (1.32)-(1.33), avec une moyenne supposée connue égale à zéro. On note par $\gamma = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, d)'$, $\delta = (a_0, a_1, \dots, a_r, b_1, \dots, b_s)'$ et $\lambda = (\gamma', \delta)'$. Supposons que $\lambda_0 = (\gamma'_0, \delta'_0)'$ est la vraie valeur du paramètre λ et se trouve à l'intérieur du compact Θ . L'estimateur par la méthode du maximum de vraisemblance (MLE) $\hat{\lambda}_n$ de λ sur Θ maximise la fonction objective donnée par :

$$L(\lambda) = \frac{1}{n} \sum_{t=1}^n \ell_t, \quad \ell_t = -\frac{1}{2} \ln \sigma_t^2 - \frac{\varepsilon_t^2}{2\sigma_t^2}. \quad (1.39)$$

On résume dans le théorème suivant les propriétés asymptotiques de l'estimateur MLE $\hat{\lambda}$ de λ et de la matrice d'information. On énonce d'abord l'hypothèse (H_3) suivante :

– (H_3) : Supposons que le moment d'ordre 4 du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ existe.

Théorème 1.3.3 *Soient les processus $(X_t)_{t \in \mathbb{Z}}$ et $(\varepsilon_t)_{t \in \mathbb{Z}}$ définis par les équations (1.32)-(1.33) et satisfaisants les hypothèses (H_0)-(H_3). Alors*

$$-\frac{1}{n} \sum_{t=1}^n \begin{bmatrix} \frac{\partial^2 \ell_t}{\partial \gamma \partial \gamma'} & \frac{\partial^2 \ell_t}{\partial \delta \partial \gamma'} \\ \frac{\partial^2 \ell_t}{\partial \gamma \partial \delta'} & \frac{\partial^2 \ell_t}{\partial \delta \partial \delta'} \end{bmatrix} \xrightarrow{p.s} \begin{bmatrix} \Omega_\gamma & 0 \\ 0 & \Omega_\delta \end{bmatrix}, \quad (1.40)$$

lorsque $n \rightarrow \infty$ et Ω_γ et Ω_δ sont des matrices positives définies par :

$$\Omega_\gamma = E \left[\frac{1}{\sigma_t^2} \frac{\partial \varepsilon_t}{\partial \gamma} \frac{\partial \varepsilon_t}{\partial \gamma'} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma'} \right], \quad (1.41)$$

et

$$\Omega_\delta = E \left[\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta} \frac{\partial \sigma_t^2}{\partial \delta'} \right]. \quad (1.42)$$

Théorème 1.3.4 *Sous les hypothèses du Théorème 1.3.3, on a les résultats suivants :*

1. *Il existe un estimateur MLE $\hat{\lambda}_n$ de λ satisfaisant l'équation suivante :*

$$\frac{\partial L(\lambda)}{\partial \lambda} = 0 \text{ et } \hat{\lambda}_n \xrightarrow{P} \lambda_0 \text{ quand } n \rightarrow \infty.$$
2. $\sqrt{n}(\hat{\lambda}_n - \lambda_0) \xrightarrow{D} \mathbb{N}(0, \Omega_0^{-1})$ *quand* $n \rightarrow \infty$, *avec* $\Omega_0 = \text{diag}(\Omega_{\gamma_0}, \Omega_{\delta_0})$ *où* Ω_{γ_0} *et* Ω_{δ_0} *sont les valeurs de* Ω_γ *et de* Ω_δ *en* $\lambda = \lambda_0$.

Remarque 1.3.1 *Si la variance conditionnelle σ_t^2 est une constante positive alors la procédure d'estimation des paramètres se réduit à celle donnée par Béran (1995) pour les processus FARIMA(p, d, q) avec variance constante.*

1.4 Introduction aux processus Longue Mémoire Multivariés

1.4.1 Présentation du modèle

On s'intéresse aux processus FARIMA multivariés. Les modèles FARIMA(p, d, q) ont été introduits dans la littérature par Granger et Joyeux (1980) et Hosking (1981). Dans ce paragraphe, on présente les modèles FARIMA multivariés (VARFIMA) introduits par Sowell (1987 et 1989). Comme leur nom l'indique, ces modèles sont une généralisation des modèles FARIMA(p, d, q). Pour des soucis de clarté, on se limite au cas des processus FARIMA bivariés.

Définition 1.4.1 *Soit $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ un bruit blanc de moyenne nulle et tel que $\text{cov}(\varepsilon_{it}, \varepsilon_{js}) = \delta(t, s) K_{ij}$, pour $i, j = 1, 2$. En supposant que la matrice de variance-covariance $K = (K_{ij})$ est définie positive, alors le processus vectoriel $(X_t)_{t \in \mathbb{Z}}$ de dimension $l = 2$ suit un modèle FARIMA bivarié s'il vérifie l'équation suivante :*

$$\Phi(B) \Delta(B) (X_t - \mu) = \Theta(B) \varepsilon_t, \quad (1.43)$$

où $\Phi(B)$ et $\Theta(B)$ sont des polynômes matriciels d'ordre p et q respectivement. Ici $\Delta(B) = \text{diag} \left[(I - B)^{d_1}, (I - B)^{d_2} \right]$ et $\mu = (\mu_1, \mu_2)'$.

Remarque 1.4.1 Si $d_j = 0$, pour $j = 1, 2$, le processus défini par l'équation (1.43) se réduit à un processus VARMA (Multivariate ARMA). Si $l = 1$ alors le processus devient un processus FARIMA.

1.4.2 Propriétés probabilistiques

On s'intéresse dans cette partie aux conditions de stationnarité et d'inversibilité des processus FARIMA multivariés. On donne tout d'abord les hypothèses nécessaires pour assurer la stationnarité et l'inversibilité du processus VARFIMA.

- (H_4) : On suppose que les paramètres de mémoire longue sont tels que $-\frac{1}{2} < d_i < \frac{1}{2}$ pour $i = 1, 2$.
- (H_5) : Les polynômes $\det(\Phi(B))$ et $\det(\Theta(B))$ n'ont aucune racine commune et ont toutes leurs racines en dehors du cercle unité. Ici $\det(\Phi(B))$ désigne le déterminant de la matrice $\Phi(B)$.

On résume, dans la Proposition 1.4.1 suivante, les propriétés probabilistiques du processus VARFIMA que l'on doit à Sowell (1989).

Proposition 1.4.1 *Sous les hypothèses (H_4) - (H_5) , le processus $(X_t)_{t \in \mathbb{Z}}$, introduit dans la Définition 1.3.1, est stationnaire, inversible et causal.*

On sait que dans le cas des processus à dépendance de long terme, l'approche temporelle est peu utilisée. En effet, seule une approximation de la fonction d'autocorrélation est connue. Dans la suite, on rappelle la densité spectrale d'un processus longue mémoire multivarié.

Soit $f(\omega, \psi)$ la densité spectrale du processus longue mémoire multivarié $(X_t)_{t \in \mathbb{Z}}$ donné dans la Définition 1.4.1. Elle est ainsi définie par :

$$f(\omega, \psi) = \frac{1}{2\pi} \left(\sum_{j=0}^{\infty} e^{i\omega j} \right) K(\mu) \left(\sum_{j=0}^{\infty} e^{i\omega j} \right)^*. \quad (1.44)$$

Ici, $\psi = (\theta, \mu)$ est le vecteur de paramètres tel que θ est constitué des paramètres de mémoire $(d_1, d_2)'$, des paramètres de la partie VARMA Φ_{ij} pour $i = 1, 2$ et $j = 1, \dots, p$ et Θ_{ij} pour $i = 1, 2$ et $j = 1, \dots, q$. Le paramètre vectoriel $\mu = (\mu_1, \mu_2)'$ est la moyenne du processus $(X_t)_{t \in \mathbb{Z}}$. Enfin, A^* est le conjugué de la matrice transposée de A .

1.4.3 Simulations numériques

Dans cette section, on s'intéresse aux méthodes de simulations numériques des processus longue mémoire multivariés de dimension l . Plus particulièrement, on étudie une méthode basée sur l'approche spectrale proposée par Chambers (1995). Ce type d'approche est, comme on l'a remarqué dans la section 1.2.3, souvent utilisé pour les processus longue mémoire. L'objectif est de générer un échantillon X_1, \dots, X_n de dimension l correspondant à la densité spectrale $f(\omega, \psi)$ définie par l'équation (1.44). Sans perte de généralité, on suppose que le bruit $(\varepsilon_t)_{t \in \mathbb{Z}}$ est Gaussien centré et réduit c'est à dire

$$\begin{pmatrix} \operatorname{Re}(\varepsilon_t) \\ \operatorname{Im}(\varepsilon_t) \end{pmatrix} \sim \mathbb{N} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \right].$$

De plus, on suppose que $\varepsilon_{n-k} = \bar{\varepsilon}_k$ pour $1 \leq k \leq \frac{n}{2} - 1$ où $\bar{\varepsilon}_k = \operatorname{Re}(\varepsilon_k) - i\operatorname{Im}(\varepsilon_k)$ est le conjugué de la variable ε_k . Remarquons que pour $k \neq \frac{n}{2}$ et n on a $\mathbb{E}(\varepsilon_k \varepsilon_k') = 0$ et $\mathbb{E}(\varepsilon_k \varepsilon_k^*) = 2I$.

Si $(X_t)_{t \in \mathbb{Z}}$ est un processus réel alors sa densité spectrale est une fonction Hermitienne c'est à dire $f(\omega, \psi) = f^*(\omega, \psi)$. La densité spectrale peut alors se décomposer sous la manière suivante :

$$f(\omega, \psi) = U(\omega) M(\omega) U^*(\omega), \quad (1.45)$$

où $M(\omega)$ est une matrice diagonale d'ordre $l \times l$ dont les composantes sont les vecteurs propres réels associés à la fonction $f(\omega, \psi)$ et $U(\omega)$ est la matrice des vecteurs propres complexes d'ordre $l \times l$ telle que $U^*(\omega) U(\omega) = U(\omega) U^*(\omega) = I$. Cette décomposition de la densité spectrale entraîne que $f(\omega, \psi)^{\frac{1}{2}}$ est égale à :

$$f(\omega, \psi)^{\frac{1}{2}} = U(\omega) M(\omega)^{\frac{1}{2}} U(\omega)^*, \quad (1.46)$$

où $M^{\frac{1}{2}}$ est la matrice diagonale d'ordre $l \times l$ dont les composantes sont les racines carrées des valeurs propres associées à la densité spectrale $f(\omega, \psi)$. Il est alors très facile de montrer que $f(\omega, \psi)^{\frac{1}{2}} f(\omega, \psi)^{\frac{1}{2}} = f(\omega, \psi)$. De plus, la matrice $f(\omega, \psi)$ est symétrique et on a $f(-\omega, \psi) = f(\omega, \psi)$. La série $\{X_t\}_{t=1}^n$ est générée en utilisant la transformation de Fourier discrète sur des combinaisons linéaires des éléments $\{\varepsilon_t\}_{t=1}^{\frac{n}{2}}$ et (ε_n) . Alors, si les fréquences de Fourier ω_k sont telles que $\omega_k = \frac{2\pi k}{n}$ pour $k = 1, \dots, n$ et les variables

aléatoires définies par :

$$V_k = U_k M_k^{\frac{1}{2}} \varepsilon_k, \quad \text{pour } k = 1, \dots, \frac{n}{2}, n, \quad (1.47)$$

et

$$V_{n-k} = \overline{U}_k M_k^{\frac{1}{2}} \overline{\varepsilon}_k = \overline{V}_k \quad \text{pour } 1 \leq k \leq \frac{n}{2} - 1, \quad (1.48)$$

où $U_k = U(\omega_k)$ et $M_k^{\frac{1}{2}} = M^{\frac{1}{2}}(\omega_k)$. Ainsi, l'échantillon X_1, \dots, X_n satisfaisant la densité spectrale $f(\omega, \psi)$ est générée à partir de l'équation suivante :

$$X_t = \left(\frac{\pi}{n}\right)^{\frac{1}{2}} \sum_{k=1}^n V_k e^{it\omega_k}, \quad \text{pour } t = 1, \dots, n. \quad (1.49)$$

1.4.4 Estimation des paramètres

Soit $X_t = (X_{1t}, X_{2t})'$ un processus VARFIMA introduit dans la Définition 1.3.1. Si les hypothèses (H_4) et (H_5) sont vérifiées alors le processus $(X_t)_{t \in \mathbb{Z}}$ admet une représentation moyenne mobile infinie donnée dans l'équation suivante :

$$\begin{aligned} X_t &= D^{-1}(B) \Phi(B) \Theta(B) \varepsilon_t \\ &= G(B) \varepsilon_t \end{aligned} \quad (1.50)$$

où les polynômes d'ordres infinis $G_{ij}(B)$ pour $i, j = 1, 2$ sont définis par :

$$G_{ij}(B) = \sum_{k=0}^{\infty} G_{ij}(k, \theta) B^k, \quad (1.51)$$

et où les coefficients $(G_{ij}(k, \theta))_{k \in \mathbb{N}}$ sont déterminés en utilisant les coefficients des parties AR et MA du processus et de la relation suivante :

$$(I - B)^{-d_i} = I + \sum_{k=1}^{\infty} \frac{\Gamma(d_i + k)}{k! \Gamma(d_i)} B^k. \quad (1.52)$$

Il s'ensuit que :

$$\sum_{k=0}^{\infty} \text{tr}(G(k, \theta)) K G^*(k, \theta) < \infty. \quad (1.53)$$

Dans la suite, on s'intéresse à l'estimation du paramètre ψ basé sur l'échantillon X_1, \dots, X_n issu du processus VARFIMA donné dans la Définition 1.4.1. Pour répondre à cette question, plusieurs travaux ont montré que l'approche

spectrale était plus avantageuse. En effet, la fonction de vraisemblance est très difficile à déterminer dans le domaine temporel. On ne présente alors dans cette section que la méthode de pseudo-maximum de vraisemblance de Whittle (QMLE) proposée par Hosoya (1996, 1997).

Soit $w(\omega_j)$ la transformée de Fourier finie basée sur l'échantillon X_1, \dots, X_n et définie par :

$$w(\omega_j) = \frac{1}{\sqrt{2n\pi}} \sum_{t=1}^n X_t e^{it\omega_j}. \quad (1.54)$$

Les transformées de Fourier finies $w(\omega_j)$, $j = 1, \dots, n$ admettent une distribution Normale multivariée pour n pris très grand. Elles sont indépendantes et admettent chacune la densité de distribution définie par :

$$\frac{1}{\pi^2} \left[\det(f(\omega_j, \psi))^{-\frac{1}{2}} \right] \exp \left[-\frac{1}{2} \text{tr} \left(f^{-1}(\omega_j, \psi) w(\omega_j) w^*(\omega_j) \right) \right], \quad j = 1, \dots, n, \quad (1.55)$$

où $\text{tr}(A)$ est la trace de la matrice A et \exp est la fonction exponentielle. Pour plus de détails sur cette densité de probabilité, on renvoie à l'ouvrage de Hannan (1970), pages 224-225.

Une approximation de la log-vraisemblance $L_n(\psi)$ basée sur l'échantillon X_1, \dots, X_n est alors donnée par l'expression suivante :

$$L_n(\psi) = - \sum_{j=1}^n \left[\log(\det(f(\omega_j, \psi))) + \text{tr}(f^{-1}(\omega_j, \psi) I_n(X, \omega_j)) \right], \quad (1.56)$$

où $I_n(X, \omega_j)$ est le périodogramme donnée par l'équation (1.57) suivante :

$$I_n(X, \omega_j) = w(\omega_j) w^*(\omega_j). \quad (1.57)$$

La forme intégrale, notée $\hat{L}_n(\psi)$, de l'expression $L_n(\psi)$ définie dans l'équation (1.56) est alors donnée par la formule suivante :

$$\hat{L}_n(\psi) = -n \left[\int_{-\pi}^{\pi} \log(\det(f(\omega, \psi))) d\omega + \int_{-\pi}^{\pi} \text{tr} \{ f^{-1}(\omega, \psi) I_n(X, \omega) \} d\omega \right]. \quad (1.58)$$

Cette fonction $\hat{L}_n(\psi)$ est appelée quasi-pseudo-vraisemblance. Supposons dans la suite que $\int_{-\pi}^{\pi} \log \det f(\omega, \psi) d\omega$ est différentiable par rapport à ψ . De plus pour presque tout point ω , on suppose que $f^{-1}(\omega, \psi)$ est aussi différentiable par rapport à ψ . On note par

$$H_j(\psi) = \partial \int_{-\pi}^{\pi} \log \det f(\omega, \psi) d\omega / \partial \psi_j \quad (1.59)$$

et

$$h_j(\omega, \psi) = \frac{\partial f^{-1}(\omega, \psi)}{\partial \psi_j}. \quad (1.60)$$

Ainsi un estimateur QMLE $\hat{\psi}_n$ de ψ est défini tel que $S_n(\hat{\psi}_n) = 0$ avec $S_n(\psi)$ est le gradient de $\hat{L}_n(\psi)$ constitué des composantes $S_{nj}(\psi)$ définies par :

$$S_{nj}(\psi) = H_j(\psi) + \int_{-\pi}^{\pi} \text{tr}[h_j(\omega, \psi) I_n(X, \omega)] d\omega. \quad j = 1, \dots, p + q \quad (1.61)$$

Sous des conditions très générales, Hosoya (1996, 1997) a démontré les propriétés asymptotiques de cet estimateur QMLE.

Chapitre 2

Les Processus GIGARCH à k facteurs

Ce chapitre est essentiellement la version française de deux articles :

- "Estimating parameters for a k -factor GIGARCH process" (2004), *C.R.Acad.Sci. Paris I*, 339(6), 435-440.
- "Forecasting Electricity Spot Market Prices With a k -factor GIGARCH process" (2004), *submitted in Journal of Forecasting*.

2.1 Introduction

L'estimation de processus longue mémoire a fait l'objet de plusieurs travaux dans la littérature. L'estimation des paramètres des modèles de type Gegenbauer a été largement étudiée dans la littérature, mais elle ne concerne que le cas des modèles de Gegenbauer à un ou plusieurs facteurs générés par des bruits blancs. Dans ce cadre on renvoie aux travaux de Giraitis et Leipus (1995), Chung (1996a, 1996b), Yajima (1996), Hosoya (1997) ou Ferrara et Guégan (2001 a et b). Dans le cas des modèles FARMA, l'estimation des paramètres a été étudiée en présence de bruits de type ARCH par Baillie, Chung et Tieslau, Ling et Li (1997). Ils ont appliqué leurs résultats à l'analyse de l'inflation des prix mensuels de différents pays.

Bien que les processus longue mémoire semblent être bien adaptés à la prévision de long terme, peu d'applications sont disponibles à l'heure actuelle sur la capacité prédictive des processus longue mémoire. Pour la prévision avec les processus FARIMA, on renvoie aux articles de Geweke et Porter-Hudak (1983), Barkoulas et Baum (1997) et Franses et Ooms (1997), qui considèrent des séries économiques, et à l'article de Noakes et *al.* (1988) traitant des données environnementales. Une autre approche basée sur des simulations de Monte Carlo a été considérée par plusieurs auteurs afin de comparer des prévisions obtenues à partir de modèle FARIMA et de modèle ARMA, voir Ray (1993a), Smith et Yadav (1994), Crato et Ray (1996) ou Brodsky et Hurvich (1999). Ces derniers montrent l'efficacité des modèles longue mémoire pour la prévision à moyen et long terme. La capacité prédictive des processus longue mémoire généralisés a été étudiée dans les articles de Porter-Hudak (1990), Ray (1993b), Sutcliffe (1994), Woodward et *al.* (1998) et Ferrara et Guégan (2001a, 2000a) ou Collet et Guégan (2003).

En 2000, Guégan a proposé un modèle appelé GIGARCH, permettant de prendre en compte les différents comportements notés précédemment. Ces modèles ont été étendus en 2003, par la même auteure, aux processus GIGARCH à k facteurs. Si l'étude probabiliste de ces modèles est en partie traitée dans ces deux articles, l'étude statistique reste à mettre en oeuvre, en particulier l'estimation des paramètres et la prévision avec ces modèles. Dans ce chapitre, nous allons étudier ces deux questions. Pour l'estimation des paramètres, nous privilégions deux méthodes. La première méthode est

basée sur la méthode des moindres carrés conditionnels. Elle permet d'avoir des résultats consistants si la loi du bruit est conditionnellement symétrique et de moment d'ordre 4 fini. Nous généralisons alors les travaux de Ling et Li (1997) établis pour des processus FARMA avec bruit GARCH. La seconde méthode correspond à la méthode du maximum de vraisemblance de Whittle. Nous généralisons alors les travaux de Ferrara et Guégan (2001a et b), voir aussi Ferrara (2000b), dans le cas où le bruit est dépendant. Nous démontrons la consistance des estimateurs. Une fois l'estimation des paramètres établie, nous nous intéressons à la prévision à partir des modèles GIGARCH à k facteurs. Ici, nous donnons l'expression analytique du prédicteur de moindres carrés pour un processus GARMA à k facteurs avec un bruit GARCH. Nous nous intéressons aussi à la prévision en variance à partir d'un tel processus.

Ce chapitre est organisé de la manière suivante. Dans le paragraphe 2, nous introduisons le modèle GIGARCH à k facteurs. Dans le paragraphe 3, nous discutons les différentes méthodes d'estimation des paramètres : méthode des moindres carrés conditionnels (CSS) et méthode de Whittle. Nous donnons les propriétés asymptotiques des estimateurs et les hypothèses sont discutées. Dans le paragraphe 4, nous présentons les résultats des simulations numériques de Monte Carlo pour ces différentes méthodes. Le paragraphe 5 fait l'objet de la prévision par un modèle GIGARCH à k facteurs. Le paragraphe 6 est dédié aux conclusions et aux perspectives.

2.2 Présentation du Modèle

Ici nous introduisons le modèle avec lequel nous allons travailler. Nous précisons quelques notations ainsi que les hypothèses de base.

Un processus $(X_t)_{t \in \mathbb{Z}}$ est un processus GIGARCH à k facteurs s'il satisfait les équations suivantes :

$$\phi(B) \prod_{i=1}^k (1 - 2\eta_i B + B^2)^{d_i} (X_t - \mu) = \theta(B) \varepsilon_t \quad (2.1)$$

et

$$\varepsilon_t = \xi_t \cdot \sigma_t, \quad (2.2)$$

où μ est la moyenne du processus $(X_t)_{t \in \mathbb{Z}}$, $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un processus supposé centré dont la variance conditionnelle σ_t^2 par rapport à la σ -algèbre $\mathcal{F}_t = \sigma(\varepsilon_s, s < t)$ vérifie l'équation suivante :

$$\forall t, \quad \sigma_t^2 = a_0 + \sum_{i=1}^r a_i \varepsilon_{t-i}^2 + \sum_{j=1}^s b_j \sigma_{t-j}^2. \quad (2.3)$$

$\phi(B)$ est un polynôme autorégressif d'ordre p et $\theta(B)$ est un polynôme moyenne mobile d'ordre q introduits dans la Définition 1.2.1. Les paramètres de ces deux polynômes sont des paramètres de mémoire courte : ils correspondent au comportement en moyenne du processus $(X_t)_t$. B est l'opérateur retard défini par $BX_t = X_{t-1}$. Les paramètres $0 < d_i < \frac{1}{2}, i = 1, \dots, k$ sont les paramètres de mémoire longue, les paramètres $|\eta_i| < 1, i = 1, \dots, k$ sont les paramètres de localisation des fréquences de la longue mémoire, k est un nombre entier fini. Notons que les fréquences $\omega_i = \arccos(\eta_i)$ pour $i = 1, \dots, k$ sont appelées G-fréquences. Le processus $(\xi_t)_{t \in \mathbb{Z}}$ est un processus bruit blanc fort dont la loi sera précisée dans la suite. Ce processus est indépendant de ε_t et du passé de ce processus. Les paramètres qui apparaissent dans l'expression (2.3) sont aussi des paramètres de mémoire courte, mais ils caractérisent le comportement en variance du processus $(X_t)_t$. Il est donc important de noter que le modèle défini par les équations (2.1) et (2.2) contient à la fois des paramètres de mémoire longue et des paramètres de mémoire courte.

Les hypothèses suivantes sont nécessaires pour assurer la stationnarité du processus $(X_t)_{t \in \mathbb{Z}}$ défini par les équations (2.1)-(2.2).

(H_6) : $|d_i| < \frac{1}{2}$ si $|\eta_i| < 1$ ou $|d_i| < \frac{1}{4}$ si $|\eta_i| = 1$ pour $i = 1, \dots, k$.

(H_7) : Les paramètres de localisation $\eta_i, i = 1, \dots, k$, sont supposés connus.

Par la suite, si ces paramètres de localisation sont inconnus, nous utiliserons l'approche développée par Yajima (1995) pour les estimer.

Remarquons enfin que \mathcal{F}_t , la σ -algèbre engendrée par le passé du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ coïncide, sous les conditions de causalité et d'inversibilité (qui découlent des hypothèses (H_0) , (H_2) et (H_6)), avec celle engendrée par le passé

du processus $(X_t)_{t \in \mathbb{Z}}$.

Dans la suite, nous rappelons les expressions des formes autorégressive et moyenne mobile infinies ainsi que les propriétés asymptotiques de la fonction d'autocorrélation et de la densité spectrale des processus GIGARCH à k facteurs définis par les équation (2.1)-(2.3) que nous devons à Guégan (2000) et (2003).

Proposition 2.2.1 *Soit $(X_t)_{t \in \mathbb{Z}}$ un processus généré par les équations (2.1)-(2.3). Si les hypothèses (H_0) , (H_2) et (H_6) sont vérifiées alors le processus $(X_t)_{t \in \mathbb{Z}}$ est stationnaire, causal et inversible.*

1. *Sa forme moyenne mobile infinie est donnée, pour tout t , par l'équation suivante :*

$$X_t = \sum_{j=0}^{\infty} \beta_j(d, \eta, \phi, \theta) \varepsilon_{t-j}, \quad (2.4)$$

où les poids $(\beta_j(d, \eta, \phi, \theta))_{j \in \mathbb{N}}$ sont donnés dans l'équation (1.8).

2. *Sa forme autorégressive infinie s'écrit, pour tout t , de la manière suivante :*

$$\varepsilon_t = \sum_{j=0}^{\infty} \alpha_j(d, \eta, \phi, \theta) X_{t-j}, \quad (2.5)$$

où les poids $(\alpha_j(d, \eta, \phi, \theta))_{j \in \mathbb{N}}$ sont donnés dans l'équation (1.6).

Pour la preuve de la Proposition 2.2.1, on renvoie aux articles de Guégan (2000) et (2003).

Proposition 2.2.2 *Supposons que les hypothèses de la Proposition 2.2.1 soient vérifiées, alors la fonction d'autocorrélation $\rho(h) = \text{cov}(X_t, X_{t-h})$ du processus défini par les équations (2.1)-(2.3), dans le cas où $\phi(B) = \theta(B) = I$ et $\max_{j=1, \dots, k} d_j > 0$ s'écrit sous la forme suivante :*

$$\rho(h) = \sum_{j=1, \dots, k, d_j > 0} V_j^{d, \eta} h^{2d_j^* - 1} (\cos(h\omega_j) + o(1)) \quad \text{si } h \rightarrow \infty, \quad (2.6)$$

avec

$$\begin{cases} V_j^{(d, \nu)} = C'_j \text{ si } \omega_j = 0 \text{ ou } \pi, \\ V_j^{(d, \nu)} = 2C'_j \text{ si } 0 < \omega_j < \pi. \end{cases} \quad (2.7)$$

Ici,

$$C'_j = f_\varepsilon(\omega) \Gamma(1 - 2d_j^*) \sin(d_j^* \pi) D^2(j), \quad (2.8)$$

où $f_\varepsilon(\omega)$ est la densité spectrale du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$. De plus,

$$\begin{cases} d_j^* = 2d_j \text{ si } \omega_j = 0 \text{ ou } \pi, \\ d_j^* = d_j \text{ si } 0 < \omega_j < \pi. \end{cases} \quad (2.9)$$

Enfin,

$$D(j) = \begin{cases} \prod_{j \neq l} |2(\cos(\omega_j) - \cos(\omega_l))|^{-d_j} \text{ si } \omega_j = 0 \text{ ou } \pi, \\ |2 \sin(\omega_j)| \prod_{j \neq l} |2(\cos(\omega_j) - \cos(\omega_l))|^{-d_j} \text{ si } 0 < \omega_j < \pi. \end{cases} \quad (2.10)$$

La densité spectrale $f_X(\omega)$ du processus $(X_t)_{t \in \mathbb{Z}}$ est donnée par l'expression suivante :

$$f_X(\omega) = \left| \frac{\phi(e^{-i\omega})}{\theta(e^{-i\omega})} \right| \prod_{j=1}^k |2(\cos(\omega) - \cos(\omega_j))|^{-2d_j} f_\varepsilon(\omega). \quad (2.11)$$

Son expression asymptotique s'écrit, si $\omega \rightarrow \omega_j$, $j = 1, \dots, k$, de la manière suivante :

$$f_X(\omega) \sim \left| \frac{\phi(e^{-i\omega})}{\theta(e^{-i\omega})} \right| D^2(j) |\omega - \omega_j|^{-2d_j^*} f_\varepsilon(\omega). \quad (2.12)$$

On renvoie à Guégan (2000) ou (2003) pour la preuve de la Proposition 2.2.2.

2.3 Estimation des paramètres du modèle

Dans ce paragraphe, nous présentons deux types de méthodes d'estimation des paramètres : la méthode des moindres carrés conditionnels et la méthode du pseudo-maximum de vraisemblance de Whittle.

2.3.1 La méthode des moindres carrés conditionnels

Nous allons estimer les paramètres du modèle GIGARCH défini par les équations (2.1) - (2.3), incluant les paramètres de mémoire longue (d_1, \dots, d_k) , les paramètres de mémoire courte homoscédastique $(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ et les paramètres de mémoire courte hétéroscédastique $(a_0, a_1, \dots, a_r, b_1, \dots, b_s)$. La méthode qu'on considère est basée sur une maximisation de la somme

des moindres carrés conditionnels. Nous supposons, sans perte de généralité, que la moyenne est connue et nous posons $\mu = 0$. Le modèle de base (2.1) se réécrit :

$$\phi(B) \prod_{i=1}^k (1 - 2\eta_i B + B^2)^{d_i} X_t = \theta(B) \varepsilon_t. \quad (2.13)$$

L'intérêt de la méthode des moindres carrés conditionnels est qu'elle peut être décrite pour n'importe quel type de distribution sur le bruit, moyennant certaines conditions de régularité sur la distribution du bruit. Ceci nous amènera à différencier les résultats obtenus suivant les propriétés de cette loi. Ici les résultats de convergence des estimateurs sont établis pour des bruits à distribution Gaussienne et Student. Nous donnons la démonstration en détail pour le cas Gaussien et nous en déduisons les résultats pour le cas Student. D'autre part nous pouvons remarquer que tous ces résultats peuvent être étendus à une plus grande classe de lois vérifiant la condition suivante :

(C_0) : La distribution conditionnelle du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ est symétrique et ses moments d'ordre 4 sont finis.

En effet, les développements de la théorie asymptotique de l'estimation CSS ne dépendent pas de l'expression explicite de la loi conditionnelle du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$, mais reposent sur la condition (C_0) .

Soit une suite de variables aléatoires X_1, \dots, X_n générée par le modèle défini par les équations (2.2)-(2.13). Notons :

$$\begin{aligned} \gamma &= (d_1, \dots, d_k, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q), \\ \delta &= (a_0, a_1, \dots, a_r, b_1, \dots, b_s), \end{aligned}$$

et $\omega = (\gamma, \delta)$ les paramètres à étudier. Soit $\omega_0 = (\gamma_0, \delta_0)$ la vraie valeur du paramètre ω et supposons qu'elle se trouve à l'intérieur de Θ , où Θ est un compact de $\mathbb{R}^{r+s+p+q+k+1}$.

On rappelle que, si les observations initiales $X_0, X_{-1}, X_{-2}, \dots$ sont supposées fixes ou nulles ce qui est équivalent à supposer que les variables aléatoires $\varepsilon_0, \varepsilon_{-1}, \varepsilon_{-2}, \dots$ sont fixes ou nulles, alors l'estimateur de la somme des carrés conditionnels étudié sera asymptotiquement équivalent à celui du maximum de vraisemblance (MLE), voir Chung (1994). Cette remarque justifie l'approche que nous utilisons ici.

2.3.1.1 Estimation des paramètres par CSS pour un bruit Gaussien

Nous présentons en détail la méthode d'estimation des paramètres du modèle (2.2)-(2.13) quand le bruit $(\varepsilon_t)_{t \in \mathbb{Z}}$ suit une loi conditionnelle Gaussienne. La loi du bruit $(\varepsilon_t)_{t \in \mathbb{Z}}$ conditionnellement à la tribu engendrée par le passé de $(X_t)_{t \in \mathbb{Z}}$, notée \mathcal{F}_{t-1} , est une loi $\mathcal{N}(0, \sigma_t^2)$ où la variance σ_t^2 est donnée par l'expression (2.3). L'estimateur ω des moindres carrés conditionnels (CSS) que nous allons étudier ici, noté $\hat{\omega}$, dans l'espace Θ décrit précédemment, est obtenue en maximisant la log-vraisemblance conditionnelle du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ sur sa σ -algèbre \mathcal{F}_0 . Cette vraisemblance s'écrit :

$$L_G(\omega) = \frac{1}{n} \sum_{t=1}^n \ell_t, \quad \ell_t = -\frac{1}{2} \log(\sigma_t^2) - \frac{\varepsilon_t^2}{2\sigma_t^2}. \quad (2.14)$$

Pour obtenir $\hat{\omega}$, on calcule les dérivées premières de cette vraisemblance et la matrice de l'information associée. Pour tout t , les dérivées premières par rapport à γ et δ sont données par :

$$\frac{\partial \ell_t}{\partial \gamma} = \frac{1}{2\sigma_t^2} \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right) \frac{\partial \sigma_t^2}{\partial \gamma} - \frac{\varepsilon_t}{\sigma_t^2} \frac{\partial \varepsilon_t}{\partial \gamma}, \quad (2.15)$$

et

$$\frac{\partial \ell_t}{\partial \delta} = \frac{1}{2\sigma_t^2} \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right) \frac{\partial \sigma_t^2}{\partial \delta}. \quad (2.16)$$

Ces dérivées font intervenir les dérivées de ε_t et de σ_t^2 , que nous précisons maintenant :

$$\frac{\partial \varepsilon_t}{\partial \phi_j} = -\phi^{-1}(B) \varepsilon_{t-j}, \quad (2.17)$$

$$\frac{\partial \varepsilon_t}{\partial \theta_j} = -\theta^{-1}(B) \varepsilon_{t-j}, \quad (2.18)$$

$$\begin{aligned} \frac{\partial \varepsilon_t}{\partial d} &= \log(1 - 2\eta B + B^2) \varepsilon_t \\ &= \sum_{j=1}^{\infty} \frac{\cos(\omega j)}{j} \varepsilon_{t-j}. \end{aligned} \quad (2.19)$$

De même nous avons :

$$\frac{\partial \sigma_t^2}{\partial \gamma} = 2 \sum_{i=1}^r a_i \varepsilon_{t-i} \frac{\partial \varepsilon_{t-i}}{\partial \gamma} + \sum_{j=1}^s b_j \frac{\partial \sigma_{t-j}^2}{\partial \gamma}, \quad (2.20)$$

et

$$\frac{\partial \sigma_t^2}{\partial \delta} = \tilde{\varepsilon}_t, \quad (2.21)$$

où $\tilde{\varepsilon}_t = (1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-r}^2, \sigma_{t-1}^2, \dots, \sigma_{t-s}^2)'$.

Pour calculer la matrice d'information de Fisher (dont nous montrons dans le Théorème 2.3.1 qu'elle est bien définie pour ces modèles), nous calculons les dérivées d'ordre deux suivantes :

$$\frac{\partial^2 \ell_t}{\partial \delta \partial \delta'} = -\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta} \frac{\partial \sigma_t^2}{\partial \delta'} \left(\frac{\varepsilon_t^2}{\sigma_t^2} \right) + \frac{1}{2} \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right) \frac{\partial}{\partial \delta} \left(\frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \delta'} \right), \quad (2.22)$$

et

$$\begin{aligned} \frac{\partial^2 \ell_t}{\partial \gamma \partial \gamma'} &= -\frac{1}{\sigma_t^2} \frac{\partial \varepsilon_t}{\partial \gamma} \frac{\partial \varepsilon_t}{\partial \gamma'} - \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma'} \frac{\varepsilon_t^2}{\sigma_t^2} \\ &\quad + \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right) \frac{\partial}{\partial \gamma} \left[\frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \gamma'} \right] \\ &\quad - \frac{2\varepsilon_t}{\sigma_t^2} \frac{\partial \varepsilon_t}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma'} + \frac{\varepsilon_t}{\sigma_t^2} \frac{\partial^2 \varepsilon_t}{\partial \gamma \partial \gamma'}. \end{aligned} \quad (2.23)$$

Les estimateurs CSS de γ et de δ peuvent alors être calculés par maximisation de la vraisemblance $L_G(\omega)$ introduite en (2.14). Le résultat est donné dans le théorème suivant :

Théorème 2.3.1 *Soit un processus $(X_t)_{t \in \mathbb{Z}}$ défini par les équations (2.2)-(2.13), avec k un entier quelconque. Supposons que le processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ admette une loi conditionnelle Gaussienne. Etant donné un échantillon de taille T et les hypothèses (H_0) , (H_2) - (H_3) et (H_6) - (H_7) alors l'estimateur des moindres carrés conditionnels de ω , noté $\hat{\omega}_n$, obtenu en maximisant la log vraisemblance (2.14) a les propriétés suivantes :*

1. $\hat{\omega}_n$ existe et vérifie l'équation suivante :

$$\frac{\partial L_G(\omega)}{\partial \omega} = 0,$$

et $\hat{\omega}_n \xrightarrow{P} \omega_0$ quand $n \rightarrow \infty$.

2. $\sqrt{n}(\hat{\omega}_n - \omega_0) \xrightarrow{D} \mathbb{N}(0, \Omega_0^{-1})$ quand $n \rightarrow \infty$, où $\Omega_0 = \text{diag}(\Omega_{\gamma_0}, \Omega_{\delta_0})$,

avec

$$\Omega_{\gamma_0} = E \left(\left[\frac{1}{\sigma_t^2} \frac{\partial \varepsilon_t}{\partial \gamma} \frac{\partial \varepsilon_t}{\partial \gamma'} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma'} \right] \right),$$

et

$$\Omega_{\delta_0} = E \left(\left[\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta} \frac{\partial \sigma_t^2}{\partial \delta'} \right] \right).$$

3. De plus, des estimateurs consistants des matrices d'information Ω_{γ} et Ω_{δ} sont donnés par les expressions suivantes :

$$\hat{\Omega}_{\gamma} = \frac{1}{n} \sum_{t=1}^n \left[\frac{1}{\sigma_t^2} \frac{\partial \varepsilon_t}{\partial \gamma} \frac{\partial \varepsilon_t}{\partial \gamma'} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma'} \right].$$

Et

$$\hat{\Omega}_{\delta} = \frac{1}{n} \sum_{t=1}^n \left[\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta} \frac{\partial \sigma_t^2}{\partial \delta'} \right].$$

Preuve :

1. Pour montrer le théorème 2.3.1, nous énonçons d'abord le lemme 2.3.1 suivant :

Lemme 2.3.1 *Supposons que les hypothèses du théorème 2.3.1 soient vérifiées. Alors :*

$$-\frac{1}{n} \sum_{t=1}^n \begin{pmatrix} \frac{\partial^2 \ell_t}{\partial \gamma \partial \gamma'} & \frac{\partial^2 \ell_t}{\partial \delta \partial \gamma'} \\ \frac{\partial^2 \ell_t}{\partial \gamma \partial \delta'} & \frac{\partial^2 \ell_t}{\partial \delta \partial \delta'} \end{pmatrix} \xrightarrow{p.s} \begin{pmatrix} \Omega_{\gamma} & 0 \\ 0 & \Omega_{\delta} \end{pmatrix} \quad (2.24)$$

et $\Omega_{\gamma}, \Omega_{\delta}$ sont des matrices définies positives avec :

$$\Omega_{\gamma} = E \left[\frac{1}{\sigma_t^2} \frac{\partial \varepsilon_t}{\partial \gamma} \frac{\partial \varepsilon_t}{\partial \gamma'} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma'} \right]$$

et

$$\Omega_{\delta} = E \left[\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta} \frac{\partial \sigma_t^2}{\partial \delta'} \right]$$

Preuve du lemme 2.3.1. On a $\forall i = 1, \dots, k$,

$$\begin{aligned} E \left(\frac{\partial \varepsilon_t}{\partial d_i} \right)^2 &= 4 \sum_{j=1}^{\infty} \frac{\cos^2(j\omega_i)}{k^2} E(\varepsilon_t^2) \\ &= 2 \sum_{j=1}^{\infty} \frac{1 + \cos(2j\omega_i)}{j^2} E(\varepsilon_t^2) \end{aligned} \quad (2.25)$$

$$\begin{aligned} &= 2 \left[\sum_{j=1}^{\infty} \frac{1}{j^2} + \sum_{j=1}^{\infty} \frac{\cos(2j\omega_i)}{j^2} \right] E(\varepsilon_t^2) \\ &= 2 \left(\frac{\pi^2}{3} - \pi\omega_i + \omega_i^2 \right) E(\varepsilon_t^2) < \infty. \end{aligned} \quad (2.26)$$

Et

$$\begin{aligned} \frac{\partial^2 \varepsilon_t}{\partial d_i^2} &= \log^2(1 - 2\eta_i B + B^2) \varepsilon_t \\ &= 4 \left(\sum_{k=1}^{\infty} \frac{\cos(k\omega_i)}{k} B^k \right)^2 \varepsilon_t. \end{aligned} \quad (2.27)$$

Comme le processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ est stationnaire, on déduit alors que :

$$E \left(\frac{\partial^2 \varepsilon_t}{\partial d \partial d'} \right) < \infty.$$

De plus, $E[(\partial^2 \varepsilon_t) / (\partial d \partial d') | F_{t-1}] = 0$.

Nous montrons ainsi, sous les hypothèses (H_0) , (H_2) - (H_3) et (H_6) - (H_7) , que $E \left(\frac{\partial^2 \ell_t}{\partial \delta \partial \delta'} \right)$, $E \left(\frac{\partial^2 \ell_t}{\partial \delta \partial \gamma'} \right)$ et $E \left(\frac{\partial^2 \ell_t}{\partial \gamma \partial \gamma'} \right)$ sont finies. De plus, comme $(\varepsilon_t)_{t \in \mathbb{Z}}$ a une distribution symétrique alors $E \left[\frac{\partial^2 \ell_t}{\partial \gamma \partial \delta'} \right] = 0$.

Sous les hypothèse (H_0) , (H_2) - (H_3) , les processus $(X_t)_{t \in \mathbb{Z}}$ et $(\varepsilon_t)_{t \in \mathbb{Z}}$ sont strictement stationnaires et ergodiques. D'après le théorème d'ergodicité, nous obtenons alors :

$$-\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \ell_t}{\partial \gamma \partial \gamma'} \xrightarrow{p.s} \Omega_\gamma, \text{ quand } n \rightarrow \infty \quad (2.28)$$

et

$$-\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \ell_t}{\partial \delta \partial \delta'} \xrightarrow{p.s} \Omega_\delta, \text{ quand } n \rightarrow \infty \quad (2.29)$$

Enfin, en suivant la preuve des lemmes 3.1-3.3 de Weiss (1986), nous déduisons que Ω_γ et Ω_δ sont des matrices définies positives. D'où le lemme 2.3.1.

Nous montrons maintenant le théorème 2.3.1

2. Nous commençons par établir les conditions 1) du théorème 2.3.1.

Pour montrer l'existence et la consistance de $\hat{\omega}$, nous allons utiliser un résultat dû à Weiss (1986). Pour cela nous avons besoin d'établir les conditions suivantes (voir Basawa, Feign et Heyde (1976)) :

$$- \frac{1}{n} \sum \frac{\partial \ell_t(\omega_0)}{\partial \omega} \xrightarrow{P} 0;$$

- $\exists M(\omega_0) > 0$ non aléatoire tel que

$$\forall \varepsilon > 0, P \left\{ -\frac{1}{n} \sum \frac{\partial^2 \ell_t(\omega_0)}{\partial \omega \partial \omega'} \geq M(\omega_0) \right\} > 1 - \varepsilon \quad \forall n > n_1(\varepsilon);$$

- $\exists M < \infty$ tel que $E \left[\left| \frac{\partial^3 \ell_t(\omega)}{\partial \omega_i \partial \omega_j \partial \omega_k} \right| \right] < M \quad \forall \omega \in \Theta$.

Ces trois conditions assurent l'existence et la consistance des estimateurs du CSS.

A partir des équations (2.15) et (2.16), nous avons $E \left[\frac{\partial \ell_t}{\partial \omega} \right]_{\omega=\omega_0} = 0$. Sous les hypothèses (H_0) , (H_2) - (H_3) , les processus $(X_t)_{t \in \mathbb{Z}}$ et $(\varepsilon_t)_{t \in \mathbb{Z}}$ sont strictement stationnaires et ergodiques. En appliquant le théorème d'ergodicité, nous obtenons la condition (a).

Le lemme 2.3.1, nous dit que Ω_0 est définie positive, alors pour toute constante $c \neq 0$, nous avons :

$$\frac{1}{n} \sum_{t=1}^n c' \frac{\partial^2 \ell_t(\omega_0)}{\partial \omega \partial \omega'} c \xrightarrow{p.s} -c' \Omega_0 c.$$

Ainsi, pour tout c donné, si nous posons $0 < \Delta(c) < \frac{c}{2} \Omega_0$, alors, pour tout $\varepsilon > 0$, nous pouvons trouver un $n_1 = n_1(\varepsilon)$ tel que pour tout $n > n_1$,

$$P \left\{ \left| \frac{1}{n} \sum_{t=1}^n c' \frac{\partial^2 \ell_t(\omega_0)}{\partial \omega \partial \omega'} c + c' \Omega_0 c \right| < \Delta(c) \right\} > 1 - \varepsilon.$$

Alors

$$P \left\{ -\frac{1}{n} \sum_{t=1}^n c' \frac{\partial^2 \ell_t}{\partial \omega \partial \omega'} c > \frac{c' \Omega_0 c}{2} \right\} > 1 - \varepsilon$$

pour tout $n > n_1$, et la condition (b) est ainsi satisfaite.

Pour montrer la condition (c), nous pouvons remarquer que

$$\begin{aligned} \frac{\partial^3 \varepsilon_t}{\partial d_i \partial d_j \partial d_k} &= \log(1 - 2\eta_i B + B^2) \log(1 - 2\eta_j B + B^2) \log(1 - 2\eta_k B + B^2) \varepsilon_t \\ &= \left(\sum_{k_1, k_2, k_3=1}^{\infty} \frac{\cos(\omega_i k_1)}{k_1} \frac{\cos(\omega_j k_2)}{k_2} \frac{\cos(\omega_k k_3)}{k_3} \right) \varepsilon_{t-k_1-k_2-k_3}, \end{aligned}$$

et

$$\begin{aligned} E \left(\frac{\partial^3 \varepsilon_t}{\partial d_i \partial d_j \partial d_k} \right)^2 &= E \left(\sum_{k_1, k_2, k_3=1}^{\infty} \frac{\cos(\omega_i k_1) \cos(\omega_j k_2) \cos(\omega_k k_3)}{k_1 k_2 k_3} \varepsilon_{t-k_1-k_2-k_3} \right)^2 \\ &= \sum_{k_1, k_2, k_3=1}^{\infty} \frac{\cos^2(k_1 \omega_i)}{k_1^2} \frac{\cos^2(k_2 \omega_j)}{k_2^2} \frac{\cos^2(k_3 \omega_k)}{k_3^2} E(\varepsilon_{t-k_1-k_2-k_3}^2) \\ &= 8 \left(\frac{\pi}{3} - \omega_i \pi + \omega_i^2 \right) \left(\frac{\pi}{3} - \omega_j \pi + \omega_j^2 \right) \left(\frac{\pi}{3} - \omega_k \pi + \omega_k^2 \right) E(\varepsilon_t^2). \end{aligned}$$

Nous obtenons :

$$E \left(\frac{\partial^3 \varepsilon_t}{\partial d_i \partial d_j \partial d_k} \right)^2 < \infty \quad (2.30)$$

En calculant le différentiel par rapport à ω de $\frac{\partial^2 \ell_t}{\partial \omega \partial \omega'}$ et en utilisant le résultat 2.30 nous obtenons la condition (c).

Enfin, en utilisant le développement limité de Taylor au premier ordre de l'expression $\frac{\partial L_G}{\partial \omega}$ au voisinage du paramètre ω_0 et les conditions (a)-(b) et (c), on obtient l'existence et la consistance de l'estimateur des moindres carrés conditionnels ω .

3. Nous montrons maintenant les conditions 2) du théorème 2.3.1

D'après le lemme 2.3.1, nous avons le résultat suivant :

$$\frac{1}{n} \sum_{t=1}^n \left(\frac{\partial \ell_t}{\partial \omega} \frac{\partial \ell_t}{\partial \omega'} \right)_{\omega=\omega_0} \xrightarrow{p.s} \Omega_0.$$

Si nous posons $S_n = \sum_{t=1}^n b_0 \frac{\partial \ell_t}{\partial \omega} |_{\omega=\omega_0}$, où $b_0 \neq 0$ est une constante arbitraire, alors, S_n est une martingale et $\frac{1}{n} E(S_n^2) = b_0^2 \Omega_0 > 0$. Comme

les processus $(X_t)_{t \in \mathbb{Z}}$ et $(\varepsilon_t)_{t \in \mathbb{Z}}$ sont ergodiques, nous en déduisons que :

$$\left[\frac{1}{n} E(S_n^2) \right]^{-1} \left[\frac{1}{n} E(S_n^2 | \mathcal{F}_{n-1}) \right] \xrightarrow{p.s} 1. \quad (2.31)$$

A partir du théorème central limite de Stout (1974), nous avons :

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \ell_t}{\partial \omega} \xrightarrow{D} \mathbb{N}(0, \Omega_0). \quad (2.32)$$

D'autre part :

$$\frac{1}{n} \sum_{t=1}^n c' \left[\frac{\partial^2 \ell_t}{\partial \omega \partial \omega'} \right] c \xrightarrow{p.s} -c' \Omega_0 c. \quad (2.33)$$

Maintenant, en utilisant le développement limité de Taylor au premier ordre de l'expression $\frac{\partial L_G}{\partial \omega}$ au voisinage du paramètre ω_0 et les résultats des équations (2.31)-(2.33), nous obtenons

$$\sqrt{n}(\hat{\omega}_n - \omega_0) \xrightarrow{D} \mathbb{N}(0, \Omega_0^{-1}).$$

D'où le théorème 2.3.1

■

2.3.1.2 Estimation des paramètres par CSS pour un bruit Student

Dans le cas où la densité conditionnelle du bruit $(\varepsilon_t)_{t \in \mathbb{Z}}$ suit une loi de Student t à l degrés de liberté, alors la log-vraisemblance conditionnelle que nous devons maximiser sur \mathcal{F}_0 s'écrit :

$$\begin{aligned} L_S(\omega) = & n \left[\log \Gamma \left\{ \frac{(l+1)}{2} \right\} - \log \Gamma \left(\frac{l}{2} \right) - \frac{1}{2} \log(l-2) \right] \\ & - \frac{1}{2} \sum_{t=1}^n \left\{ \log(\sigma_t^2) + (l+1) \left[\log \left(1 + \frac{\varepsilon_t^2}{\sigma_t^2(l-2)} \right) \right] \right\}. \end{aligned} \quad (2.34)$$

L'ensemble des résultats présentés au paragraphe 2.3.1.1 dans le cadre Gaussien se généralisent sans difficulté si $(\varepsilon_t)_{t \in \mathbb{Z}}$ suit une loi de Student conditionnellement à \mathcal{F}_{t-1} . Nous obtenons le résultat suivant :

Théorème 2.3.2 *Soit un processus $(X_t)_{t \in \mathbb{Z}}$ défini par les équations (2.2)-(2.13), avec k entier quelconque. Supposons que $(\varepsilon_t)_{t \in \mathbb{Z}}$ suive une loi de Student. Etant donné un échantillon de taille n et les hypothèses (H_0) , (H_2) - (H_3) et (H_6) - (H_7) , l'estimateur des moindres carrés conditionnels de ω obtenu en maximisant la log vraisemblance (2.34) a les propriétés suivantes :*

1. $\hat{\omega}_n$ existe et satisfait l'équation

$$\frac{\partial L_S(\omega)}{\partial \omega} = 0,$$

et $\hat{\omega}_n \xrightarrow{P} \omega_0$ quand $n \rightarrow \infty$.

2. $\sqrt{n}(\hat{\omega} - \omega) \xrightarrow{D} \mathbb{N}(0, \Omega_0^{-1})$ quand $n \rightarrow \infty$, où $\Omega_0 = \text{diag}(\Omega_{\gamma_0}, \Omega_{\delta_0},)$

avec

$$\Omega_{\gamma_0} = E \left(\left[\frac{1}{\sigma_t^2} \frac{\partial \varepsilon_t}{\partial \gamma} \frac{\partial \varepsilon_t}{\partial \gamma'} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma'} \right] \right),$$

et

$$\Omega_{\delta_0} = E \left(\left[\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta} \frac{\partial \sigma_t^2}{\partial \delta'} \right] \right).$$

3. Des estimateurs consistants de Ω_{γ_0} et Ω_{δ_0} sont donnés par :

$$\hat{\Omega}_{\gamma} = \frac{1}{n} \sum_{t=1}^n \left[\frac{1}{\sigma_t^2} \frac{\partial \varepsilon_t}{\partial \gamma} \frac{\partial \varepsilon_t}{\partial \gamma'} + \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \gamma} \frac{\partial \sigma_t^2}{\partial \gamma'} \right],$$

et

$$\hat{\Omega}_{\delta} = \frac{1}{n} \sum_{t=1}^n \left[\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \delta} \frac{\partial \sigma_t^2}{\partial \delta'} \right].$$

Nous remarquons que ce Théorème 2.3.2 se démontre très facilement en utilisant les mêmes étapes que celles développées dans la preuve du Théorème 2.3.1.

2.3.2 Estimation des paramètres par la méthode de Whittle

Dans ce paragraphe, nous développons une méthode séquentielle d'estimation des paramètres du processus GIGARCH à k facteurs défini par les équations

(2.2)-(2.13), basée sur une extension de la méthode du pseudo-maximum de vraisemblance de Whittle dans le cas où les paramètres de localisation sont connus, (voir Ferrara et Guégan (2001a et b) et Ferrara (2000b)).

Nous supposons alors que X_1, \dots, X_n est une trajectoire issue du processus défini par les équations (2.2)-(2.13) avec k quelconque. Le processus $(X_t)_{t \in \mathbb{N}}$ est supposé, centré stationnaire et inversible. Nous notons $\psi_0 = (\alpha_0, d_0)$ la vraie valeur des paramètres à estimer, où $\alpha_0 = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ et $d_0 = (d_1, \dots, d_k)$. Nous allons considérer ici une méthode d'estimation en deux étapes. Cela signifie que dans un premier temps, on estime les paramètres de longue mémoire d_0 et de mémoire courte α_0 , puis les paramètres du modèle GARCH sont estimés, par maximum de vraisemblance de Whittle, à partir du carré des résidus. La méthode est la suivante :

Notons $\hat{\psi}_n = (\hat{\alpha}_n, \hat{d}_n)$ l'estimateur de Whittle de ψ_0 obtenu en minimisant la vraisemblance de Whittle $L_W(Y, \psi)$ définie par :

$$L_W(Y, \psi) = \sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} \frac{I_n(X, \omega_j)}{g(\omega_j, \psi)}, \quad (2.35)$$

où

$$g(\omega_j, \psi) = \left| \frac{\theta(e^{i\omega_j})}{\phi(e^{i\omega_j})} \right|^2 \prod_{i=1}^k |2(\cos(\omega_j) - \eta_i)|^{-2d_i}, \quad (2.36)$$

et $I_n(X, \omega_j)$ est le périodogramme du processus $(X_t)_{t \in \mathbb{Z}}$ donné par l'équation 1.20.

Les propriétés des estimateurs $\hat{\psi}_n$ de ψ_0 sont résumées dans le Théorème 2.3.3 suivant.

Théorème 2.3.3 *Soit $(X_t)_{t \in \mathbb{Z}}$ un processus GIGARCH Gaussien à k facteurs défini par les équations (2.2)-(2.13) stationnaire et inversible, avec k fixé. Nous supposons que les fréquences de localisation, $\eta_i, i = 1, \dots, k$, sont connues. Si $n \rightarrow \infty$, alors les estimateurs $\hat{\alpha}_n$ et \hat{d}_n de α et d sont asymptotiquement indépendants, et*

$$\hat{\psi}_n \xrightarrow{p.s} \psi.$$

De plus,

$$n^{\frac{1}{2}} (\hat{\alpha}_n - \alpha_0) \rightarrow \mathbb{N}(0, 4\pi V(\alpha)^{-1}),$$

où $V(\alpha)$ est la matrice de dimension $(p+q) \times (p+q)$ de terme général :

$$V(\alpha)_{ij} = \int_{-\pi}^{\pi} g^2(\omega, \psi) \frac{\partial g^{-1}(\omega, \psi)}{\partial \alpha_i} \frac{\partial g^{-1}(\omega, \psi)}{\partial \alpha_j} d\omega,$$

et

$$n^{\frac{1}{2}} (\hat{d}_n - d_0) \rightarrow \mathbb{N}(0, 4\pi V(d)^{-1}),$$

où $V(d)$ est la matrice de dimension $k \times k$ de terme général :

$$V(d)_{ij} = \int_{-\pi}^{\pi} \log \left| 4 \sin \frac{(\omega - \omega_i)}{2} \sin \frac{(\omega + \omega_i)}{2} \right| \log \left| 4 \sin \frac{(\omega - \omega_j)}{2} \sin \frac{(\omega + \omega_j)}{2} \right| d\omega.$$

Le Théorème 2.3.3 est un corollaire du Théorème 2.3 de Hosoya (1997) car ici les paramètres de localisation sont supposés connus. Cette méthode d'estimation en deux étapes s'appuie sur le fait que la densité spectrale du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$, $f(\omega)$ est bornée et ne dépend pas du paramètre ψ .

Comme le carré d'un processus $GARCH(r, s)$ peut être approché par un processus $ARMA(\max(r, s), s)$, (Bollerslev (1986)), nous allons maintenant supposer que le carré du processus résiduel $(\varepsilon_t)_{t \in \mathbb{Z}}$ vérifie $\forall t$:

$$\varepsilon_t^2 - \sum_{i=1}^{\max(r, s)} c_i \varepsilon_{t-i}^2 = a_0 + v_t - \sum_{j=1}^s b_j v_{t-j}, \quad (2.37)$$

où $c_i = a_i + b_i$. Si $i \in (s, r]$, nous posons $b_i = 0$, et si $i \in (r, s]$ nous posons $a_i = 0$ et $\forall t$, $v_t = \varepsilon_t^2 - \sigma_t^2$.

L'estimateur de Whittle $\hat{\beta}$ des paramètres du processus introduit en (2.37), noté $\beta = (a_0, a_1, \dots, a_r, b_1, \dots, b_s)$, s'obtient en minimisant sur le compact $\Omega \subset \mathbb{R}^{r+s+1}$ la vraisemblance de Whittle définie par :

$$L_W(\beta) = \frac{1}{n} \sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} \frac{I_n(X, \omega_j)}{f(\omega_j, \beta)}, \quad (2.38)$$

où

$$f(\omega_j, \beta) = \frac{\sigma^2}{2\pi} \left| \frac{\psi(e^{-i\omega_j})}{\varphi(e^{-i\omega_j})} \right|^2, \quad (2.39)$$

avec $\sigma^2 = E(v_t^2)$.

Les propriétés des estimateurs des paramètres hétéroscédastiques β sont données dans le Théorème 2.3.4 ci-dessous. Nous précisons tout d'abord quelques hypothèses.

- (H_8) : Pour tout $t = 0, \pm 1, \dots$, le processus $(\xi_t)_{t \in \mathbb{Z}}$ qui intervient dans l'équation (2.2), est strictement stationnaire, ergodique, de moments d'ordre J finis tels que :

$$E(\xi_t | \mathcal{F}_{t-1}) = 0 \text{ p.s.},$$

$$E(\xi_t^2 | \mathcal{F}_{t-1}) = 1 \text{ p.s.},$$

$$E(\xi_t^{2j} | \mathcal{F}_{t-1}) = v_{2j} \text{ p.s. } j = 2, \dots, \frac{J}{2},$$

où v_{2j} est une constante et où

$$|v_J|^{\frac{2}{J}} \left(\sum_{i=1}^r a_i + \sum_{i=1}^s b_i \right) < 1.$$

Pour obtenir la consistance de l'estimateur $\hat{\beta}$, nous utilisons $J = 4$, et pour la normalité asymptotique, $J = 8$.

Les hypothèses ci-dessous (H_9) et (H_{10}) sont techniques et interviennent pour généraliser la démonstration du Théorème 2.1 de Giraitis et Leipus (2001).

- (H_9) : La densité spectrale $f(\omega, \beta)$ du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ défini en (2.2) vérifie les propriétés suivantes :

1. Pour tout β , $\int_{-\pi}^{\pi} \log f(\omega, \beta) d\omega = 0$.
2. $f(\omega, \beta)^{-1}$ est continue en $(\omega, \beta) \in [-\pi, \pi] \times \Omega$.
3. $\mu_L(\{\omega; f(\omega, \beta) \neq f(\omega, \beta_0)\}) \geq 0$, pour tout $\beta \in \Omega$. Ici μ_L représente la mesure de Lebesgue.

- (H_{10}) : Enfin,

1. β_0 appartient à l'intérieur de Ω .
2. Quelque soit β appartenant au voisinage de β_0 , on suppose que $\frac{\partial f^{-1}(\omega, \beta_0)}{\partial \beta}$ et $\frac{\partial^2}{\partial \beta \partial \beta'} f^{-1}(\omega, \beta_0)$ existent et sont continues en ω et β .
3. $\frac{\partial f^{-1}}{\partial \beta}(\omega, \beta_0)$ est M -lipschitzienne avec $M > \frac{1}{2}$.

4. La matrice

$$W = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \log f(\omega, \beta_0)}{\partial \beta} \frac{\partial \log f(\omega, \beta_0)}{\partial \beta'} d\omega$$

est non singulière.

Théorème 2.3.4 *Soit $(X_t)_{t \in \mathbb{Z}}$ un processus GIGARCH Gaussien à k facteurs défini par les équations (2.2)-(2.13), avec k fixé. Supposons que les hypothèses (H_0) , (H_2) - (H_3) , (H_6) - (H_8) (avec $J = 4$) et (H_9) sont vérifiées, alors :*

$$\hat{\beta} \xrightarrow{P} \beta \text{ quand } n \rightarrow \infty.$$

Preuve : Elle est semblable à celle du Théorème 2.1 de Giraitis et Robinson (2001).

Théorème 2.3.5 *Soit $(X_t)_{t \in \mathbb{Z}}$ un processus GIGARCH Gaussien à k facteurs défini par les équations (2.2)-(2.13), avec k fixé. Supposons que les hypothèses (H_0) , (H_2) - (H_3) , (H_6) - (H_8) (avec $J = 8$), et (H_9) - (H_{10}) sont vérifiées, alors :*

$$n^{\frac{1}{2}}(\hat{\beta} - \beta_0) \xrightarrow{d} \mathbb{N}(0, 2W^{-1} + W^{-1}VW^{-1}),$$

où

$$V = \frac{2\pi}{\sigma^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\partial f(\omega, \beta_0)^{-1}}{\partial \beta} \frac{\partial f(\omega, \beta_0)^{-1}}{\partial \beta'} h(\omega, -\omega, \omega) d\omega d\omega,$$

et

$$h(\omega, \omega, v) = \frac{1}{(2\pi)^3} \sum_{j,k,l=-\infty}^{+\infty} e^{-ij\omega - ik\omega - ilv} \text{Cum}(X_0, X_j, X_k, X_l),$$

où Cum est le cumulants d'ordre 4 du processus $(X_t)_{t \in \mathbb{Z}}$.

La preuve de ce Théorème est semblable à celle du Théorème 2.2 de Giraitis et Robinson (2001).

Remarquons que ces résultats sont encore valables pour des processus non Gaussiens pourvu que les hypothèses (H_9) et (H_{10}) soient vérifiées.

2.4 Simulations numériques

Nous présentons maintenant un ensemble de simulations pour valider les méthodes d'estimation présentées ci-dessus. Pour ce faire, nous simulons, en utilisant une généralisation des méthodes proposées dans la section 1.2.3, des processus longue mémoire généralisés à un facteur ($k = 1$) avec une variance conditionnelle $(\sigma_t^2)_t$ qui suit un processus $ARCH(1)$. Nous simulons, par méthode de Monte Carlo, des échantillons de taille $n = 300$ puis $n = 500$ et $n = 1000$. Nous utilisons deux types de bruits : un bruit $ARCH(1)$ -Gaussien centré réduit et un processus $ARCH(1)$ -Student à $l = 5$ degrés de liberté. Les modèles simulés sont définis par les équations suivantes :

$$(I - 2\eta B + B^2)^d X_t = \varepsilon_t, \quad (2.40)$$

où

$$\varepsilon_t = \sigma_t \xi_t, \quad (2.41)$$

et

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2. \quad (2.42)$$

Le processus $(\xi_t)_{t \in \mathbb{Z}}$ est un processus bruit blanc fort Gaussien ou Student à l degrés de liberté. Pour chaque plan de simulations, nous effectuons $M = 500$ réplifications. Nous choisissons pour tous ces modèles $\eta = \cos\left(\frac{\pi}{6}\right)$. Les résultats des estimations par les différentes méthodes sont résumés dans les tableaux A.1 à A.6, voir annexe A. Dans ces tableaux, les trois premières colonnes donnent les valeurs théoriques des paramètres utilisés pour simuler les modèles. Dans les colonnes suivantes, nous donnons les paramètres estimés par les différentes méthodes. Les valeurs entre parenthèses représentent les MAE (*Mean Absolute Error*) des estimateurs et celles entre crochets sont les RMSE (*Root Mean Square Error*). Nous avons utilisé, pour effectuer nos simulations, une extension des méthodes de simulations numériques développées dans Ferrara (2000b), voir aussi, Beran (1994).

Dans la table A.1, nous donnons les résultats des estimations pour un échantillon de petite taille $n = 300$. Dans le cas d'un modèle simulé à partir d'un bruit Gaussien, les résultats sont satisfaisants quelle que soit la méthode utilisée CSS ou Whittle. Nous pouvons néanmoins noter quelques différences. Les résultats sont meilleurs par CSS quand $d = 0.25$ et $d = 0.3$. Par contre pour $d = 0.35$, ils sont meilleurs par la méthode Whittle. Les estimations des

paramètres de mémoire courte ne sont pas très bonnes quelles que soient les méthodes considérées. Nous retrouvons un fait souvent noté dans la littérature : l'estimation des paramètres de mémoire courte est plus sensible dès que l'on est en présence de mémoire longue. Dans le cas d'un modèle simulé à partir d'un bruit Student, les résultats donnés dans la table A.2, pour les paramètres d nous semblent satisfaisants. Les mêmes remarques que précédemment s'appliquent pour l'estimation des paramètres de mémoire courte hétéroscédastiques.

Quand nous augmentons de manière significative la taille de l'échantillon ($n = 1000$), les résultats ne s'améliorent pas de manière sensible, en particulier quand nous utilisons la méthode CSS. En revanche les résultats semblent globalement plus satisfaisants avec la méthode de Whittle, surtout pour l'estimateur des paramètres de mémoire courte. En cela l'approche nous semble très intéressante.

D'un point de vue empirique, il nous semble donc plus adapté d'utiliser la méthode de Whittle pour estimer les paramètres d'un modèle GIGARCH à k facteurs. D'une part, cette méthode est moins coûteuse en temps de calculs que la méthode des moindres carrés conditionnels. D'autre part, elle nous semble plus appropriée car il est difficile de montrer empiriquement que la loi conditionnelle d'un bruit non observé est symétrique. Enfin, nous savons que souvent la loi non conditionnelle des séries financières a une distribution plus leptokurtique que la distribution Gaussienne et que ceci est souvent due à la présence d'effet ARCH.

2.5 Prédiction avec un processus GIGARCH à k facteurs

Dans cette section, nous donnons des résultats nouveaux concernant la prédiction à la fois en moyenne et en variance, par un processus GIGARCH à k facteurs défini par les équations (2.2)-(2.13). Nous nous intéressons particulièrement au prédicteur de moindres carrés et nous fournissons l'intervalle de confiance de l'erreur de prédiction sous certaines conditions. Ainsi, dans la suite nous supposons que le processus $(\xi_t)_{t \in \mathbb{Z}}$ est un bruit blanc fort. Nous ne spécifions pas pour l'instant la distribution de ce bruit dans la plupart

des résultats obtenus dans ce paragraphe. En vue de donner un intervalle de confiance pour l'erreur de prévision, nous supposons que le processus $(\xi_t)_{t \in \mathbb{Z}}$ suit une loi Normale.

2.5.1 Prévision en moyenne

Soit X_1, \dots, X_n des observations issues d'un processus GIGARCH à k facteurs défini par les équations (2.2)-(2.13). Nous nous intéressons au prédicteur de moindres carrés à l'horizon $h > 0$ d'un processus GIGARCH à k facteurs supposé stationnaire, inversible et causal. Nous notons $\hat{X}_t(h)$ le prédicteur à l'horizon h ($h > 0$). Le prédicteur de moindres carrés est obtenu en minimisant l'erreur quadratique moyenne de prévision. Dans toute la suite, nous supposons que les paramètres du modèle défini par les équations (2.2)-(2.13) sont connus.

Nous obtenons alors dans le Théorème 2.5.1 suivant l'expression du prédicteur optimal au temps t pour l'horizon h :

Théorème 2.5.1 *Soit $(X_t)_{t \in \mathbb{Z}}$ un processus GIGARCH à k facteurs centré stationnaire inversible et causal défini par les équations (2.2)-(2.13). Alors le prédicteur des moindres carrés à l'horizon h , pour $h \geq 1$, noté $\hat{X}_t(h)$, est donné par l'expression suivante :*

$$\hat{X}_t(h) = - \sum_{j=1}^{h-1} \alpha_j(d, \nu, \phi, \theta) \hat{X}_t(h-j) - \sum_{j \geq 0} \alpha_{j+h}(d, \nu, \phi, \theta) X_{t-j}, \quad (2.43)$$

où les poids $(\alpha_j(d, \nu, \phi, \theta))_{j \in \mathbb{Z}}$ sont donnés dans la Proposition 1.2.1.

La preuve de ce Théorème est similaire à celle de la Proposition 2 dans Ferrara et Guégan (2001). Ce résultat peut être généralisé comme dans Ferrara et Guégan (2001) (Proposition 4) au cas où les paramètres du modèle ne sont pas connus et sont estimés par la méthode de pseudo-maximum de vraisemblance de Whittle.

2.5.2 Prévision en variance

Soit \mathcal{F}_t la σ -algèbre engendré par $\{\varepsilon_0, \dots, \varepsilon_t\}$. L'erreur de prévision du prédicteur de moindres carrés à l'horizon h , notée $e_{t,h}$, pour $h > 0$ est donnée

par l'équation suivante :

$$\begin{aligned} e_{t,h} &= X_{t+h} - \hat{X}_t(h) \\ &= \sum_{j=0}^{h-1} \beta_j(d, \nu, \phi, \theta) \varepsilon_{t+h-j}. \end{aligned} \quad (2.44)$$

où les poids $\beta_j(d, \nu, \phi, \theta)$ sont donnés dans la Proposition 1.2.1.

Preuve : Il s'agit d'une extension de la Proposition 5.1 démontrée dans Ferrara et Guégan (2001a).

Nous allons nous intéresser à la variance conditionnelle de l'erreur de prévision $e_{t,h}$. Elle peut s'écrire de la manière suivante :

$$E(e_{t,h}^2 | \mathcal{F}_t) = \sum_{j=0}^{h-1} \beta_j^2(d, \nu, \gamma) E(\varepsilon_{t+h-j}^2 | \mathcal{F}_t). \quad (2.45)$$

Nous remarquons que la variance conditionnelle de l'erreur de prévision est fonction du prédicteur de moindres carrés du processus $(\varepsilon_t^2)_{t \in \mathbb{Z}}$. Dans le Théorème 2.5.2, nous fournissons l'expression de la variance conditionnelle de l'erreur de prévision du prédicteur optimal de moindres carrés d'un processus GIGARCH à k facteurs.

Théorème 2.5.2 *Soit $e_{t,h}$ l'erreur de prévision du prédicteur de moindres carrés à l'horizon h définie par l'équation (2.44), alors*

1. *la variance conditionnelle de $e_{t,h}$, notée $\text{var}_t(e_{t,h})$, est donnée par :*

$$\begin{aligned} \text{var}_t(e_{t,h}) &= \sum_{i=1}^h \beta_{h-i}^2(d, \nu, \gamma) \omega_i \\ &+ \sum_{k=1}^h \beta_{h-k}^2(d, \nu, \gamma) \left[\sum_{i=0}^{s-1} \delta_{i,k} h_{t-i} + \sum_{i=0}^{m-1} \rho_{i,k} \varepsilon_{t-i}^2 \right]. \end{aligned} \quad (2.46)$$

où $m = \max(r, s)$, $\omega_h = e_1' (I + \Gamma + \dots + \Gamma^{h-1}) e_1 \hat{a}_0$, $\delta_{j,h} = -e_1' \Gamma^h e_{m+j+1}$ si $j = 0, \dots, s-1$, $\rho_{j,h} = e_1' \Gamma^h (e_{j+1} + e_{m+j+1})$ si $j = 0, \dots, s-1$, $\rho_{j,h} = e_1' \Gamma^h e_{j+1}$ if $j = s, \dots, m-1$. Ici, $\{e_j\}_{j=1}^{m+s}$ est la base canonique de \mathbb{R}^{m+s} et la matrice Γ est telle que sa première ligne est égale à $(a_1 + b_1, \dots, a_m + b_m, -b_1, \dots, -b_s)$,

et les autres éléments sont nuls sauf ceux de la première sous-diagonale qui sont tous égaux à un.

2. Nous supposons maintenant que le processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ suit un modèle $GARCH(r, s)$ Gaussien. Alors nous obtenons

$$\mathbb{P} \left\{ \frac{|X_{t+h} - \hat{X}_t(h)|}{\sqrt{\text{var}(e_{t,h})}} < t_{\frac{\alpha}{2}} \mid \mathcal{F}_t \right\} = 1 - \alpha \quad (2.47)$$

où $t_{\frac{\alpha}{2}}$ est le quantile au risque α d'une loi Normale centrée réduite et $MSE(h)$ est définie par l'équation (2.46)

Preuve :

1. D'après Bollerslev (1988), le carré du processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ défini par l'équation (2.2) peut être réécrit sous la forme suivante :

$$\varepsilon_t^2 = a_0 + \sum_{i=1}^m (a_i + b_i) \varepsilon_{t-i}^2 + \nu_t - \sum_{j=1}^s b_j \nu_{t-j}, \quad (2.48)$$

où $m = \max(r, s)$, $a_i = 0$ if $i > s$ and $b_j = 0$ if $j > r$. Le processus $(\nu_t)_{t \in \mathbb{Z}}$ est défini par $\nu_t = \varepsilon_t^2 - h_t$, et forme une série indépendante et identiquement distribuée. L'équation (2.48) suggère que le processus $(\varepsilon_t^2)_{t \in \mathbb{Z}}$ suit un modèle ARMA (m, s) . Alors le prédicteur de moindres carrés à l'horizon h du processus $(\varepsilon_t^2)_{t \in \mathbb{Z}}$ est donné par l'équation (2.49) suivante.

$$\begin{aligned} E(\varepsilon_{t+h}^2 \mid \mathcal{F}_t) &= E(h_{t+h} \mid \mathcal{F}_t) \\ &= \omega_h + \sum_{j=0}^{s-1} \delta_{j,h} h_{t-j} + \sum_{j=0}^{m-1} \rho_{j,h} \varepsilon_{t-j}^2, \end{aligned} \quad (2.49)$$

où ω_h , e_j , Γ , $\delta_{j,h}$, et $\rho_{j,h}$ sont donnés dans le Théorème 2.5.1. En combinant les équations (2.45) et (2.49), nous obtenons l'expression de l'erreur quadratique conditionnelle du prédicteur de moindres carrés d'un processus GIGARCH à k facteurs.

2. Si le processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ admet une distribution conditionnelle normale, alors le processus $(e_{t,h})_{t \in \mathbb{Z}}$ introduit dans l'équation (2.44) admet aussi la même loi. De plus, nous montrons très facilement que :

$$E(e_{t,h} \mid \mathcal{F}_t) = 0,$$

et $var_t(e_{t,h})$ est donnée par l'équation (2.46). D'où, l'expression (2.47) du Théorème 2.5.2 s'en suit.

■

2.6 Conclusions et perspectives

Dans ce chapitre, nous nous sommes particulièrement intéressés à la famille des processus longue mémoire généralisés à variance conditionnelle non constante. Nous avons étudié les questions d'estimation des paramètres et de prévision pour cette famille processus. A cet effet, deux méthodes d'estimation des paramètres ont été proposées dans ce travail. La première est basée sur la somme des carrés conditionnels alors que la seconde est une extension de la méthode de pseudo-maximum de vraisemblance de Whittle. Dans toutes ces deux méthodes, nous avons supposé que les G-fréquences sont connues. Nous avons alors montré les propriétés asymptotiques (consistance et normalité asymptotique) de chacun des estimateurs obtenus par ces méthodes. Afin de comparer les vitesses de convergence, un plan de simulations numériques est effectué. Les deux méthodes ont fourni des résultats satisfaisants. Cependant, la méthode de Whittle est plus simple à mettre en oeuvre et converge plus vite que la méthode CSS. Nous avons aussi donné l'expression analytique du prédicteur des moindres carrés. Enfin, sous l'hypothèse de gaussianité, nous avons fourni l'intervalle de confiance de l'erreur de prévision.

D'autres pistes restent à explorer pour cette famille de processus. Par exemple, dans cette étude, nous nous sommes restreints à l'estimation des paramètres dans le cas où les G-fréquences sont supposées connues. On pourrait s'intéresser à l'estimation simultanée des paramètres du processus et aux propriétés asymptotiques de cet estimateur. De plus, plusieurs séries présentent un caractère non stationnaire. Il serait alors très intéressant de s'investir sur l'estimation des paramètres ainsi que les méthodes de prévision. Il serait aussi bien de relaxer l'hypothèse de gaussianité pour déterminer l'intervalle de confiance de l'erreur de prévision. La prévision de processus GIGARCH à k facteurs non gaussiens pourrait aussi faire l'objet d'autres travaux.

Chapitre 3

Les Processus Longue Mémoire Généralisés Multivariés

3.1 Introduction

Les processus longue mémoire multivariés sont très peu étudiés dans la littérature. Hosoya (1996) a proposé une méthode de quasi-maximum de vraisemblance (QMLE) et un test QLR (Quasi-likelihood ratio). Lobato (1999) a proposé une méthode d'estimation à deux temps des processus longue mémoire multivariés. Il a établi des résultats de normalité asymptotique de l'estimateur et a appliqué son modèle sur des données de taux de change. Callabero et Jewson (2002) ont modélisé des séries journalières de températures pour l'évaluation des dérivés climatiques. Enfin, Nielsen (2004) a introduit une procédure d'estimation des paramètres d'un processus FARIMA multivarié basée sur la méthode du maximum de vraisemblance. Particulièrement, il a montré l'existence d'une solution locale et a établi la normalité asymptotique de cet estimateur sous l'hypothèse de gaussianité.

Si l'estimation des paramètres d'un processus longue mémoire multivarié a fait l'objet de plusieurs études, leur utilisation pour la prévision est très peu développée dans la littérature. Dans ce chapitre, nous considérons une extension au cas multivarié des processus longue mémoire généralisés, introduits par Woodward *et al.* (1998). Ce modèle nous permet de capturer les corrélations entre plusieurs séries. Nous examinons ici les propriétés probabilistiques et statistiques du modèle. Une procédure pratique d'estimation des paramètres est proposée. Enfin, nous étudions les capacités prédictives du modèle en fournissant l'expression analytique du prédicteur de moindres carrés.

Ce chapitre est organisé comme suit. Dans le deuxième paragraphe, nous introduisons le modèle longue mémoire généralisé multivarié. Dans le troisième paragraphe, nous établissons les propriétés probabilistiques du processus. Le paragraphe 4 contient une procédure d'estimation à deux temps des paramètres du modèle. Dans le cinquième paragraphe, nous fournissons l'expression analytique du prédicteur des moindres carrés et nous donnons l'intervalle de confiance de l'erreur de prévision.

3.2 Les processus GARMA à k -facteurs multivariés

Dans ce paragraphe, nous considérons les processus longue mémoire généralisés multivariés, dénommés MVGARMA à k facteurs processus. Nous étudions ensuite les conditions de stationnarité et d'inversibilité du processus. Enfin, nous présentons une procédure pratique d'estimation des paramètres du processus longue mémoire généralisé multivarié.

3.2.1 Présentation du modèle

Les modèles longue mémoire multivariés (VARFIMA) ont été étudiés dans les articles Baillie et Chung (2002) ou Davidson (2002). Nous introduirons dans cette section les processus longue mémoire généralisés multivariés qui sont une généralisation des processus VARFIMA. Pour ce type de modèle, les corrélations croisées entre les séries sont capturées par la partie mémoire courte. En d'autres termes, nous supposons que toute la dépendance temporelle entre les séries est à court terme. Sans perte de généralités, nous considérons dans ce travail le cas de processus bivarié. Nous avons alors la Définition 3.2.1 suivante :

Définition 3.2.1 Soit $\xi_t = (\xi_{1t}, \xi_{2t})'$ un bruit blanc vectoriel. Alors le processus vectoriel $X_t = (X_{1t}, X_{2t})'$ de moyenne $\mu = (\mu_1, \mu_2)'$ suit un modèle longue mémoire généralisé s'il vérifie l'équation suivante :

$$D(B)(X_t - \mu) = \varepsilon_t, \quad \text{pour tout } t = 1, 2, \dots \quad (3.1)$$

où $D(B) = \text{diag} \left[\prod_{i=1}^{k_1} (I - 2\nu_i B + B^2)^{d_i}, \prod_{i=1}^{k_2} (I - 2\eta_i B + B^2)^{\delta_i} \right]$ avec B est l'opérateur retard, $|\nu_i| \leq 1$ pour $i = 1, \dots, k_1$ et $|\eta_i| \leq 1$ pour $i = 1, \dots, k_2$ sont respectivement les fréquences de Gegenbauer des processus $(X_{1t})_{t \in \mathbb{Z}}$ et $(X_{2t})_{t \in \mathbb{Z}}$. On note $d_i, i = 1, \dots, k_1$ et $\delta_i, i = 1, \dots, k_2$ les paramètres longue mémoire respectifs des processus $(X_{1t})_{t \in \mathbb{Z}}$ et $(X_{2t})_{t \in \mathbb{Z}}$. Ici, k_1 et k_2 sont des entiers positifs. Le processus vectoriel $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ suit un modèle VARMA défini par

$$\Phi(B)\varepsilon_t = \Theta(B)\xi_t, \quad (3.2)$$

où $\Phi(B)$ et $\Theta(B)$ sont des polynômes vectoriels d'ordre p et q respectifs définis par $\Phi(B) = I - \Phi_1 B - \dots - \Phi_p B^p$ et $\Theta(B) = I - \Theta_1 B - \dots - \Theta_q B^q$ avec $\Phi_i, i = 1, \dots, p$ et $\Theta_i, i = 1, \dots, q$ sont des matrices carrées d'ordre deux.

3.2.2 Propriétés probabilistiques

Dans cette partie, nous nous intéressons aux conditions de stationnarité et d'inversibilité du processus MVGARMA introduit dans la Définition 3.2.1. Nous introduirons d'abord un certain nombre d'hypothèses qui nous permettent d'assurer la stationnarité et l'inversibilité du processus.

- (H_{11}) : Supposons que les polynômes $\Delta(B) = \det(\Phi(B))$ et $\Delta_1(B) = \det(\Theta(B))$ n'ont pas de racines communes et ont toutes leurs racines en dehors du cercle unité.
- (H_{12}) : Soit $d_j \neq 0$ tel que $|d_j| \leq \frac{1}{2}$ si $|\nu_j| \leq 1$ ou $|d_j| \leq \frac{1}{4}$ si $|\nu_j| = 1$ pour $j = 1, \dots, k_1$.
- (H_{13}) : Soit $\delta_j \neq 0$ tel que $|\delta_j| \leq \frac{1}{2}$ si $|\eta_j| \leq 1$ ou $|\delta_j| \leq \frac{1}{4}$ si $|\eta_j| = 1$ pour $j = 1, \dots, k_2$.

Dans la Proposition 3.2.1 suivante, nous énonçons les propriétés de stationnarité et d'inversibilité. Nous donnons aussi les formes moyenne mobile infinie et autorégressive infinie du processus MVGARMA à k facteurs. Avant d'énoncer la Proposition 3.2.1, nous introduirons d'abord les polynômes $\Delta(B)$, $\Delta_1(B)$ et $\Gamma_i(B)$ pour $i = 1, \dots, 8$.

Soient $\phi(B)$ et $\theta(B)$ deux polynômes d'ordre respectifs p et q , définis par $\phi(B) = \phi_0 I - \sum_{i=1}^p \phi_i B^i$ et $\theta(B) = \theta_0 I - \sum_{i=1}^q \theta_i B^i$. Nous allons calculer les coefficients du produit des polynômes $\Gamma(B)$, défini par $\Gamma(B) = \phi(B)\theta(B)$. Nous obtenons alors

$$\Gamma(B) = \gamma_0 I - \sum_{i=1}^{p+q} \gamma_i B^i \quad (3.3)$$

avec

$$\begin{cases} \gamma_0(\phi, \theta) = \phi_0 \theta_0 \\ \gamma_j(\phi, \theta) = \theta_0 \phi_j - \sum_{i=1}^{j-1} \phi_i \theta_{j-i} + \phi_0 \theta_j, \quad \forall j \geq 1. \end{cases} \quad (3.4)$$

A partir des équations (3.3) et (3.4), nous pouvons déterminer les coefficients des polynômes suivants :

$$\Gamma_1(B) = \Phi_{22}(B)\Theta_{11}(B), \quad (3.5)$$

$$\Gamma_2(B) = \Phi_{22}(B) \Theta_{12}(B), \quad (3.6)$$

$$\Gamma_3(B) = \Phi_{12}(B) \Theta_{22}(B), \quad (3.7)$$

$$\Gamma_4(B) = \Phi_{12}(B) \Theta_{21}(B), \quad (3.8)$$

$$\Gamma_5(B) = \Phi_{11}(B) \Theta_{21}(B), \quad (3.9)$$

$$\Gamma_6(B) = \Phi_{11}(B) \Theta_{22}(B), \quad (3.10)$$

$$\Gamma_7(B) = \Phi_{21}(B) \Theta_{11}(B), \quad (3.11)$$

$$\Gamma_8(B) = \Phi_{21}(B) \Theta_{12}(B), \quad (3.12)$$

$$\Delta(B) = \Phi_{11}(B) \Phi_{22}(B) - \Phi_{12}(B) \Phi_{21}(B), \quad (3.13)$$

et

$$\Delta_1(B) = \Theta_{11}(B) \Theta_{22}(B) - \Theta_{12}(B) \Theta_{21}(B). \quad (3.14)$$

Proposition 3.2.1 1. *Sous les Hypothèses (H_{11}) - (H_{13}) , le processus $(X_t)_{t \in \mathbb{Z}}$ donné dans la Définition 3.2.1 est stationnaire, inversible et causal.*

2. *Sa forme moyenne mobile infinie est :*

$$\begin{aligned} X_{1t} &= \sum_{j \geq 0} \beta_j(d, \nu, \Gamma_1 - \Gamma_4, \Delta) \xi_{t-j} \\ &+ \sum_{j \geq 0} \beta_j(d, \nu, \Gamma_2 - \Gamma_3, \Delta) \xi_{2t-j}, \end{aligned} \quad (3.15)$$

et

$$\begin{aligned} X_{2t} &= \sum_{j \geq 0} \beta_j(\delta, \eta, \Gamma_5 - \Gamma_7, \Delta) \xi_{1t-j} \\ &+ \sum_{j \geq 0} \beta_j(\delta, \eta, \Gamma_6 - \Gamma_8, \Delta) \xi_{2t-j}, \end{aligned} \quad (3.16)$$

Les poids $(\beta_j)_{j \in \mathbb{N}}$ sont définis par l'expression suivante :

$$\left\{ \begin{array}{l} \beta_j(d, \nu, \phi, \theta) = \theta_0, \\ \beta_j(d, \nu, \phi, \theta) = \theta_0 \psi_j(d, \nu) + \sum_{i=1}^{\min(j,p)} \phi_i \beta_{j-i}(d, \nu, \phi, \theta) \\ \quad - \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu), \quad \forall j \geq 1. \end{array} \right. \quad (3.17)$$

avec les coefficients $(\psi_j(d, \nu))_{j \in \mathbb{N}}$ donnés dans l'équation (1.10).

3. La forme autorégressive infinie du processus vectoriel $(X_t)_{t \in \mathbb{Z}}$ est donnée par :

$$\begin{aligned} \xi_{1t} &= \sum_{j \geq 0} \alpha_j (d, \nu, \Gamma_6 - \Gamma_8, \Delta_1) X_{1t-j} \\ &+ \sum_{j \geq 0} \alpha_j (\delta, \eta, \Gamma_3 - \Gamma_2, \Delta_1) X_{2t-j} \end{aligned} \quad (3.18)$$

et

$$\begin{aligned} \xi_{2t} &= \sum_{j \geq 0} \alpha_j (d, \nu, \Gamma_7 - \Gamma_5, \Delta_1) X_{1t-j} \\ &+ \sum_{j \geq 0} \alpha_j (\delta, \eta, \Gamma_1 - \Gamma_4, \Delta_1) X_{2t-j} \end{aligned} \quad (3.19)$$

où les poids $(\alpha_j)_{j \in \mathbb{N}}$ sont définis dans l'équation suivante :

$$\left\{ \begin{array}{l} \alpha_0 (d, \nu, \phi, \theta) = \phi_0 \\ \alpha_j (d, \nu, \phi, \theta) = \phi_0 \pi_j (d, \nu) + \sum_{i=1}^{\min(j,q)} \theta_i \alpha_{j-i} (d, \nu, \phi, \theta) \\ \quad - \sum_{i=1}^{\min(j,p)} \phi_i \psi_{j-i} (d, \nu), \quad \forall j \geq 1. \end{array} \right. \quad (3.20)$$

Les coefficients $(\pi_j (d, \nu))_{j \in \mathbb{N}}$ ont été donnés dans l'équation (1.9).

Preuve :

Si toutes les racines du polynôme $\Delta (B)$ sont en dehors du cercle unité, nous obtenons alors

$$D(B) X_t = \Phi^{-1}(B) \Theta(B) \xi_t, \quad (3.21)$$

avec

$$\Phi^{-1} = \begin{bmatrix} \Phi_{22}(B) & -\Phi_{12}(B) \\ -\Phi_{21}(B) & \Phi_{11}(B) \end{bmatrix}, \quad (3.22)$$

et

$$\Phi^{-1}(B) \Theta(B) = \Delta^{-1}(B) \begin{bmatrix} \Gamma_1 - \Gamma_4 & \Gamma_2 - \Gamma_3 \\ \Gamma_5 - \Gamma_7 & \Gamma_6 - \Gamma_8 \end{bmatrix}. \quad (3.23)$$

D'où, nous avons

$$\begin{aligned} X_{1t} &= \sum_{j \geq 0} \beta_j (d, \nu, \Gamma_1 - \Gamma_4, \Delta) \xi_{1t} \\ &+ \sum_{j \geq 0} \beta_j (d, \nu, \Gamma_2 - \Gamma_3, \Delta) \xi_{2t}, \end{aligned} \quad (3.24)$$

et

$$\begin{aligned} X_{2t} &= \sum_{j \geq 0} \beta_j (\delta, \eta, \Gamma_5 - \Gamma_7, \Delta) \xi_{1t} \\ &+ \sum_{j \geq 0} \beta_j (\delta, \eta, \Gamma_6 - \Gamma_8, \Delta) \xi_{2t}, \end{aligned} \quad (3.25)$$

Si toutes les racines du polynôme $\Theta(B)$ sont en dehors du cercle unité alors nous obtenons la relation suivante :

$$\Theta^{-1}(B) \Phi(B) = \Delta_1^{-1}(B) \begin{bmatrix} \Gamma_6 - \Gamma_8 & \Gamma_3 - \Gamma_2 \\ \Gamma_7 - \Gamma_5 & \Gamma_1 - \Gamma_4 \end{bmatrix}, \quad (3.26)$$

et nous obtenons

$$\begin{aligned} \xi_{1t} &= \frac{\Gamma_6(B) - \Gamma_8(B)}{\Delta_1(B)} \prod_{i=1}^{k_1} (I - 2\nu_i B + B^2)^{d_i} X_{1t} \\ &+ \frac{\Gamma_3(B) - \Gamma_2(B)}{\Delta_1(B)} \prod_{i=1}^{k_2} (I - 2\eta_i B + B^2)^{\delta_i} X_{2t}, \end{aligned} \quad (3.27)$$

et

$$\begin{aligned} \xi_{2t} &= \frac{\Gamma_7(B) - \Gamma_5(B)}{\Delta_1(B)} \prod_{i=1}^{k_1} (I - 2\nu_i B + B^2)^{d_i} X_{1t} \\ &+ \frac{\Gamma_1(B) - \Gamma_4(B)}{\Delta_1(B)} \prod_{i=1}^{k_2} (I - 2\eta_i B + B^2)^{\delta_i} X_{2t}. \end{aligned} \quad (3.28)$$

Alors

$$\begin{aligned} \xi_{1t} &= \sum_{j \geq 0} \alpha_j (d, \nu, \Gamma_6 - \Gamma_8, \Delta_1) X_{1t-j} \\ &+ \sum_{j \geq 0} \alpha_j (\delta, \eta, \Gamma_3 - \Gamma_2, \Delta_1) X_{2t-j}, \end{aligned} \quad (3.29)$$

et

$$\begin{aligned}\xi_{2t} &= \sum_{j \geq 0} \alpha_j(d, \nu, \Gamma_7 - \Gamma_5, \Delta_1) X_{1t-j} \\ &+ \sum_{j \geq 0} \alpha_j(\delta, \eta, \Gamma_1 - \Gamma_4, \Delta_1) X_{2t-j}.\end{aligned}\quad (3.30)$$

■

3.3 Procédure d'estimation des paramètres

Dans cette section, nous proposons une procédure pratique à deux étapes pour l'estimation des paramètres d'un processus longue mémoire généralisé multivarié. Cette procédure est décrite comme suit :

1. Dans la première étape, nous estimons séparément les paramètres longue mémoire de chaque processus $(X_{it})_{t \in \mathbb{Z}}$, $i = 1, 2$. Pour ce faire, nous utilisons la méthode de pseudo-maximum de vraisemblance de Whittle dans le cas où les G-fréquences sont connues (nous renvoyons à Ferrara et Guégan (2001)).
2. Nous utilisons les valeurs estimées des paramètres d et δ pour récupérer les résidus de chaque série. Puis, nous utilisons le filtre de Gegenbauer défini, par exemple pour le processus $(X_{jt})_{t \in \mathbb{Z}}$, par :

$$\prod_{i=1}^{k_j} (I - 2\nu_i B + B^2)^{d_i} (X_{jt} - \mu_j) = \sum_{k=0}^t \pi_k(d, \nu) X_{jt-k}. \quad (3.31)$$

3. La dernière étape consiste à caler un modèle du type VARMA sur ces résidus. Nous remarquons que cette partie permet de capturer l'interdépendance entre les séries.

A des fins d'application, nous nous intéressons à la prévision des processus MVGARMA à k facteurs.

3.4 Prévision

Dans ce paragraphe, nous étudions les capacités prédictives du modèle MVGARMA. Nous nous intéressons d'abord aux résultats théoriques du pré-

dicteur des moindres carrés à l'horizon $h \geq 1$ pour un tel processus. Nous donnons l'expression analytique du prédictor des moindres carrés et l'écart moyen au carré de l'erreur de prévision dans le cas où les paramètres du modèle sont connus. Soit $(X_t)_{t \in \mathbb{Z}}$ un processus longue mémoire généralisé bivarié stationnaire, inversible et causal. Nous supposons que le processus $(\xi_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien bivarié. Nous savons que, d'après Priestley (1981), le prédictor des moindres carrés du processus $(X_t)_{t \in \mathbb{Z}}$, noté $\hat{X}_t(h)$, est défini par :

$$\hat{X}_t(h) = \mathbb{E}(X_{t+h} | \mathcal{F}_t), \quad (3.32)$$

où $\mathcal{F}_t = \sigma(X_{1s}, X_{2s}, \xi_{1s}, \xi_{2s}, s \leq t)$.

Nous donnons dans le Théorème 3.4.1 suivant, l'expression analytique du prédictor des moindres carrés du processus $(X_t)_{t \in \mathbb{Z}}$.

Théorème 3.4.1 *Soit $h \geq 1$. Si $(X_t)_{t \in \mathbb{Z}} = (X_{1t}, X_{2t})'_{t \in \mathbb{Z}}$ est un processus longue mémoire généralisé bivarié gaussien centré, stationnaire, inversible et causal, donné dans la Définition 3.2.1. Alors le prédictor des moindres carrés est défini par les expressions suivantes :*

$$\begin{aligned} \hat{X}_{1t}(h) &= - \sum_{j=1}^{h-1} \alpha_j(d, \nu, \Gamma_6 - \Gamma_8, \Delta_1) \hat{X}_{1t}(h-j) \\ &\quad - \alpha_j(\delta, \eta, \Gamma_3 - \Gamma_2, \Delta_1) \hat{X}_{2t}(h-j) \\ &\quad - \sum_{j \geq 0} \alpha_{j+h}(d, \nu, \Gamma_6 - \Gamma_8, \Delta_1) X_{1t-j} \\ &\quad - \alpha_{j+h}(\delta, \eta, \Gamma_3 - \Gamma_2, \Delta_1) X_{2t-j}, \end{aligned} \quad (3.33)$$

et

$$\begin{aligned} \hat{X}_{2t}(h) &= - \sum_{j=1}^{h-1} \alpha_j(\delta, \eta, \Gamma_1 - \Gamma_4, \Delta_1) \hat{X}_{2t}(h-j) \\ &\quad - \alpha_j(d, \nu, \Gamma_7 - \Gamma_5, \Delta_1) \hat{X}_{1t}(h-j) \\ &\quad - \sum_{j \geq 0} \alpha_{j+h}(\delta, \eta, \Gamma_1 - \Gamma_4, \Delta_1) X_{2t-j} \\ &\quad - \alpha_{j+h}(d, \nu, \Gamma_7 - \Gamma_5, \Delta_1) X_{1t-j}, \end{aligned} \quad (3.34)$$

où les poids $(\alpha_k(d, \nu, \phi, \theta))_{k \in \mathbb{N}}$ sont donnés dans l'équation (3.20), les polynômes Γ_1 à Γ_8 sont définis par les équations (3.5)-(3.12) et les polynômes Δ et Δ_1 sont donnés respectivement dans les équation (3.13) et (3.14).

Preuve :

A partir de l'équation (3.18), nous avons :

$$\begin{aligned}\xi_{1t+h} &= \sum_{j \geq 0} \alpha_j(d, \nu, \Gamma_6 - \Gamma_8, \Delta_1) X_{1t+h-j} \\ &+ \sum_{j \geq 0} \alpha_j(\delta, \eta, \Gamma_3 - \Gamma_2, \Delta_1) X_{2t+h-j}.\end{aligned}\quad (3.35)$$

Nous montrons facilement que $\alpha_0(d, \nu, \Gamma_6 - \Gamma_8, \Delta_1) = 1$ et $\alpha_0(\delta, \eta, \Gamma_3 - \Gamma_2, \Delta_1) = 0$.

On en déduit :

$$\begin{aligned}\xi_{1t+h} &= X_{1t+h} + \sum_{j=1}^{h-1} \alpha_j(d, \nu, \Gamma_6 - \Gamma_8, \Delta_1) X_{1t+h-j} \\ &+ \sum_{j \geq 0} \alpha_{j+h}(d, \nu, \Gamma_6 - \Gamma_8, \Delta_1) X_{1t-j} \\ &+ \sum_{j=1}^{h-1} \alpha_j(\delta, \eta, \Gamma_3 - \Gamma_2, \Delta_1) X_{2t+h-j} \\ &+ \sum_{j \geq 0} \alpha_{j+h}(\delta, \eta, \Gamma_3 - \Gamma_2, \Delta_1) X_{2t-j},\end{aligned}\quad (3.36)$$

en prenant l'espérance conditionnelle de ξ_{1t+h} conditionnellement à \mathcal{F}_t . Et, nous obtenons :

$$\begin{aligned}\hat{X}_{1t}(h) &= E(X_{1t+h} | \mathcal{F}_t) \\ &= - \sum_{j=1}^{h-1} \alpha_j(d, \nu, \Gamma_6 - \Gamma_8, \Delta_1) \hat{X}_{1t}(h-j) \\ &\quad - \sum_{j \geq 0} \alpha_{j+h}(d, \nu, \Gamma_6 - \Gamma_8, \Delta_1) X_{1t-j} \\ &\quad - \sum_{j=1}^{h-1} \alpha_j(\delta, \eta, \Gamma_3 - \Gamma_2, \Delta_1) \hat{X}_{2t}(h-j) \\ &\quad - \sum_{j \geq 0} \alpha_{j+h}(\delta, \eta, \Gamma_3 - \Gamma_2, \Delta_1) X_{2t-j}.\end{aligned}\quad (3.37)$$

En utilisant l'équation (3.19), nous pouvons très facilement montrer que :

$$\begin{aligned}
\hat{X}_{2t}(h) &= - \sum_{j=1}^{h-1} \alpha_j (\delta, \eta, \Gamma_1 - \Gamma_4, \Delta_1) \hat{X}_{2t}(h-j) \\
&\quad - \sum_{j \geq 0} \alpha_{j+h} (\delta, \eta, \Gamma_1 - \Gamma_4, \Delta_1) X_{2t-j} \\
&\quad - \sum_{j=1}^{h-1} \alpha_j (d, \nu, \Gamma_7 - \Gamma_5, \Delta_1) \hat{X}_{1t}(h-j) \\
&\quad - \sum_{j \geq 0} \alpha_{j+h} (d, \nu, \Gamma_7 - \Gamma_5, \Delta_1) X_{1t-j}. \tag{3.38}
\end{aligned}$$

■

Dans le Théorème 3.4.2 suivant, nous établissons les intervalles de confiance des erreurs de prévision pour le processus longue mémoire généralisé bivarié Gaussien introduit dans la Définition 3.2.1.

Théorème 3.4.2 *Soit $(X_t)_{t \in \mathbb{Z}} = (X_{1t}, X_{2t})'_{t \in \mathbb{Z}}$ un processus longue mémoire généralisé bivarié gaussien centré, stationnaire, inversible et causal donné dans la Définition 3.2.1. Nous supposons que les innovations $\xi_t = (\xi_{1t}, \xi_{2t})'$ suivent un bruit blanc gaussien de moyenne nulle et tel que $\text{var}(\xi_{1t}) = \sigma_1^2$, $\text{var}(\xi_{2t}) = \sigma_2^2$ et $\text{cov}(\xi_{1t}, \xi_{2t}) = \text{cov}(\xi_{2t}, \xi_{1t}) = 0$. Alors, nous avons :*

$$\mathbb{P} \left\{ \frac{|X_{1t+h} - \hat{X}_{1t}(h)|}{\sqrt{MSE_1(h)}} < t_{\frac{\alpha_1}{2}} \right\} = 1 - \alpha_1 \tag{3.39}$$

et

$$\mathbb{P} \left\{ \frac{|X_{2t+h} - \hat{X}_{2t}(h)|}{\sqrt{MSE_2(h)}} < t_{\frac{\alpha_2}{2}} \right\} = 1 - \alpha_2 \tag{3.40}$$

où $t_{\frac{\alpha}{2}}$ est le quantile au risque α d'une loi normale centrée réduite et $MSE_1(h)$ ainsi que $MSE_2(h)$ sont définies par les équations suivantes

$$\begin{aligned}
MSE_1(h) &= \sigma_1^2 \sum_{j=0}^{h-1} \beta_j^2 (d, \nu, \Gamma_1 - \Gamma_4, \Delta) \\
&\quad + \sigma_2^2 \sum_{j=0}^{h-1} \beta_j^2 (d, \nu, \Gamma_2 - \Gamma_3, \Delta), \tag{3.41}
\end{aligned}$$

et

$$\begin{aligned} MSE_2(h) &= \sigma_1^2 \sum_{j=0}^{h-1} \beta_j^2(\delta, \eta, \Gamma_5 - \Gamma_7, \Delta) \\ &+ \sigma_2^2 \sum_{j=0}^{h-1} \beta_j^2(\delta, \eta, \Gamma_6 - \Gamma_8, \Delta). \end{aligned} \quad (3.42)$$

Les poids $(\beta_j(d, \nu, \phi, \theta))_{j \in \mathbb{N}}$ sont définis dans l'équation (3.17).

Preuve :

A partir de l'équation (3.15), nous avons :

$$\begin{aligned} X_{1t+h} &= \sum_{j \geq 0} \beta_j(d, \nu, \Gamma_1 - \Gamma_4, \Delta) \xi_{1t+h-j} \\ &+ \sum_{j \geq 0} \beta_j(d, \nu, \Gamma_2 - \Gamma_3, \Delta) \xi_{2t+h-j}. \end{aligned} \quad (3.43)$$

En prenant l'espérance de X_{1t+h} conditionnellement à \mathcal{F}_t , nous obtenons

$$\begin{aligned} \hat{X}_{1t}(h) &= E(X_{1t+h} | \mathcal{F}_t) \\ &= \sum_{j \geq h} \beta_j(d, \nu, \Gamma_1 - \Gamma_4, \Delta) \xi_{1t+h-j} \\ &+ \sum_{j \geq h} \beta_j(d, \nu, \Gamma_2 - \Gamma_3, \Delta) \xi_{2t+h-j}. \end{aligned} \quad (3.44)$$

Nous posons $e_{1t}(h) = X_{1t+h} - \hat{X}_{1t}(h)$, alors nous avons le résultat suivant :

$$\begin{aligned} e_{1t}(h) &= \sum_{j=0}^{h-1} \beta_j(d, \nu, \Gamma_1 - \Gamma_4, \Delta) \xi_{1t+h-j} \\ &+ \sum_{j=0}^{h-1} \beta_j(d, \nu, \Gamma_2 - \Gamma_3, \Delta) \xi_{2t+h-j}. \end{aligned} \quad (3.45)$$

Les processus $(\xi_{1t})_{t \in \mathbb{Z}}$ et $(\xi_{2t})_{t \in \mathbb{Z}}$ étant indépendants, nous pouvons donc conclure que $e_{1t}(h) \sim \mathbb{N}\left(0, \sqrt{MSE_1(h)}\right)$, avec $MSE_1(h)$ défini par :

$$\begin{aligned} MSE_1(h) &= \sigma_1^2 \sum_{j=0}^{h-1} \beta_j^2(d, \nu, \Gamma_1 - \Gamma_4, \Delta) \\ &+ \sigma_2^2 \sum_{j=0}^{h-1} \beta_j^2(d, \nu, \Gamma_2 - \Gamma_3, \Delta). \end{aligned} \quad (3.46)$$

Par analogie, nous pouvons montrer, en utilisant l'équation (3.16), que le processus $(e_{2t}(h))_{t \in \mathbb{Z}}$ défini par $e_{2t}(h) = X_{2t+h} - \hat{X}_{2t}(h)$ admet une distribution normale de moyenne nulle et de variance notée $MSE_2(h)$. La variance est donnée par l'équation suivante :

$$\begin{aligned} MSE_2(h) &= \sigma_1^2 \sum_{j=0}^{h-1} \beta_j^2(\delta, \eta, \Gamma_5 - \Gamma_7, \Delta) \\ &+ \sigma_2^2 \sum_{j=0}^{h-1} \beta_j^2(\delta, \eta, \Gamma_6 - \Gamma_8, \Delta). \end{aligned} \quad (3.47)$$

■

3.5 Conclusions et perspectives

Dans ce chapitre, nous avons introduit les processus longue mémoire généralisés multivariés (MVGARMA à k facteurs). Nous avons étudié les conditions de stationnarité, d'inversibilité et de causalité de ce processus. A des fins d'application, nous avons proposé une procédure pratique d'estimation des paramètres du processus, et nous avons donné des résultats théoriques pour la prévision. Nous avons fourni l'expression analytique du prédicteur de moindres carrés des processus MVGRMA à k facteurs bivariés. Sous l'hypothèse de gaussianité, nous avons déterminé l'intervalle de confiance de l'erreur de prévision.

Beaucoup de questions restent encore sans réponses pour ce type de processus. Il serait très intéressant de proposer une méthode théorique d'estimation des paramètres de ce modèle. Par exemple, la méthode de pseudo-maximum de vraisemblance de Whittle utilisé pour les processus VARFIMA (processus longue mémoire multivarié) pourrait être étendue. On pourrait aussi tenter de relaxer l'hypothèse de gaussianité et alors étudier la prévision de processus MVGARMA à k facteurs non gaussiens. On pourrait étudier le cas des processus longue mémoire multivariés non Gaussiens ou à variances hétéroscédastiques.

Chapitre 4

Application aux prix spot de marchés de l'électricité

Ce chapitre correspond en partie aux deux articles suivants :

"A k -factor GIGARCH process : estimation and application to electricity market spot prices" (2004), *IEEE Proceedings of the 8th International Conference on Probabilistic Methods Applied to Power Systems*, Iowa State University, Ames, Iowa : 1-7.

"Forecasting electricity market spot prices using a k -factor Multivariate GARMA process." (en cours de rédaction)

4.1 Introduction

Le comportement de longue mémoire, tel qu'il a été décrit par Mandelbrot et Van Ness (1968), puis repris par Granger (1980), et ensuite modélisé à l'aide de différents modèles, est présent dans de nombreuses données, appartenant à des domaines aussi variés que l'hydrologie, la climatologie, la médecine ou la finance. Ce n'est que très récemment que l'on s'est intéressé à cette approche pour la modélisation des prix spot de l'électricité. La modélisation de ces prix a vu récemment un développement important avec la mise en place d'institutions comme Powernext sur la place financière française par exemple. Une entreprise comme EDF en France est particulièrement intéressée par des développements robustes sur l'évolution de ces prix. Néanmoins quelques travaux existent sur le sujet, voir par exemple Madala et *al.* (1997), Schwarz et *al.* (2002) et Barrieu et El Karoui (2002). Les approches sont variées, mais sont toutes basées sur des modélisations en temps continu. L'étude empirique de ces prix montre que ceux-ci présentent des comportements de longue mémoire, associés à des phénomènes de saisonnalité ayant des phases plus ou moins fixes. Nous observons aussi sur ces prix de l'électricité des pics de prix, un phénomène de retour à la moyenne et de volatilité non constante. D'autre part ces prix, comme tous les prix, correspondent à des chroniques établies à partir du temps discret. Il est donc apparu naturel de modéliser ces séries de prix par des modèles permettant de prendre en compte ces différentes caractéristiques reposant sur une modélisation développée en temps discret.

De plus, en finance, la prévision des prix des commodités, en particulier les prix de l'électricité, sont cruciales. En effet, depuis l'ouverture des marchés de l'électricité, les prévisions de prix d'électricité peuvent aider les producteurs aussi bien que les consommateurs à évaluer leurs contrats bilatéraux. Ces prévisions peuvent aussi être utiles pour la mesure de risque, l'aide au trading ou l'évaluation de position. Comme tous les prix de commodités, la prévision des prix de l'électricité peut être basée sur une analyse pertinente du spot. Nous considérons dans cette étude les prix spot de l'électricité des marchés français et allemand. Les différentes caractéristiques citées ci-dessus rendent la modélisation et la prévision des prix spot de l'électricité très complexes.

Nous nous proposons d'utiliser dans ce chapitre deux approches différentes pour prendre en compte ces différentes caractéristiques. La première approche est univariée et nous faisons appel aux processus longue mémoire généralisés à variance conditionnelle hétéroscédastique. La seconde approche est multivariée et est basée sur les processus longue mémoire multivariés. Le chapitre est alors organisé comme suit. Dans le premier paragraphe, nous présentons les données. Nous présentons en détails dans le second paragraphe les différents modèles proposés pour analyser ces données. Enfin, nous étudions puis comparons les capacités prédictives de chacun des modèles proposés. Dans le dernier paragraphe, nous proposons quelques perspectives.

4.2 Présentation des données

Nous considérons ici les séries horaires de prix spot de l'électricité des marchés "day-ahead" allemand EEX (*European Energy eXchange*) et français Powernext, notées respectivement par $\{S_{1t}\}_{t=1}^n$ et $\{S_{2t}\}_{t=1}^n$, sur la période allant du 01/01/2002 au 31/12/2004, soit un total de $n = 28968$ données. En Allemagne, le marché de l'électricité LPX (*Leibzig Power eXchange*) fut ouvert en juin 2000. Depuis lors, il s'est établi comme le marché spot de l'électricité le plus liquide dans ce pays. Deux ans plus tard (juin 2002), LPX fusionne avec EEX (*European Energy eXchange*) à Frankfurt, pour former un des plus grands marchés de l'électricité du continent européen avec des volumes mensuels échangés dépassant régulièrement 1.5 millions de MWh. Le marché français, Powernext est, quant à lui, opérationnel depuis l'automne 2001. Dans ces marchés, les participants soumettent des offres et des demandes la veille pour les 24h du lendemain. Les fixations des prix et des volumes échangés sont ainsi établies la veille pour chaque heure du lendemain en confrontant l'offre et la demande. Ensuite, les marchés nominent les engagements physiques des membres responsables d'équilibre au TSO (*Transmission System Operators*) et informent les enchérisseurs des prix et des volumes à délivrer et à recevoir. Les contrats peuvent être signés 7j/j en France et 5j/7j en Allemagne, toute l'année y compris les jours fériés.

La figure 4.1 représente les prix spot horaires en moyenne sur la période allant du 01/01/2002 au 31/12/2004. Nous remarquons que les prix spot de l'électricité des marchés allemand et français ont un comportement similaire. Nous constatons une alternance de décroissance et de croissance des prix. Nous avons une première décroissance des prix jusqu'à 5h, une seconde de

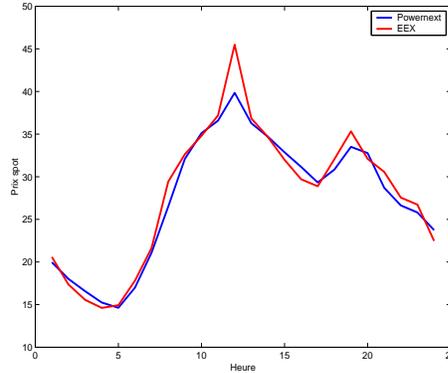


FIG. 4.1 – Prix spot horaires moyens des marchés français et allemand pour la période allant du 01/01/2002 au 31/12/2004

12h à 17h et puis une dernière de 19h à la fin de la journée. Nous observons une croissance des prix spot d'abord de 5h à midi et puis de 17h à 19h. Nous notons aussi la présence de deux pics de prix : un premier pic de prix observé à midi qui est plus élevé que le second observé à 19h. Nous constatons aussi que les prix spot du marché allemand sont en moyenne plus élevés que ceux du marché français. Sur la figure 4.2, nous présentons l'évolution des prix spot aussi bien du marché allemand que celui du marché français et un zoom sur un mois (01/03/2002 au 31/03/2002) permettant de mieux voir certaines caractéristiques des prix spot. Nous notons la présence de pics de prix aussi bien sur le marché allemand que français. Par exemple en France, les pics de prix (environ 1000 euros/MWh) les plus élevés sont observés le 11/08/2003 de 10h à 16h alors qu'en Allemagne, le pic le plus élevé est observé le 07/01/2003 à 19h. Pour la France, les pics de prix ont été observés pendant la canicule de 2003 alors qu'en Allemagne cela peut s'expliquer par une vague de froid importante. De plus, nous notons que les prix les plus élevés sont souvent apparus entre 12h et 14h pour l'Allemagne et entre 10h et 16h pour la France. Sur le zoom (figure 4.3), nous notons sur l'évolution des prix spot une présence de cycles journalier et hebdomadaire. Il apparaît aussi clairement que l'évolution des deux séries (prix spot allemand et prix spot français) est quasiment similaire sur cette période.

Dans la suite, pour avoir des séries de distributions approximativement symétriques, nous considérons les séries de prix spot en logarithme des marchés

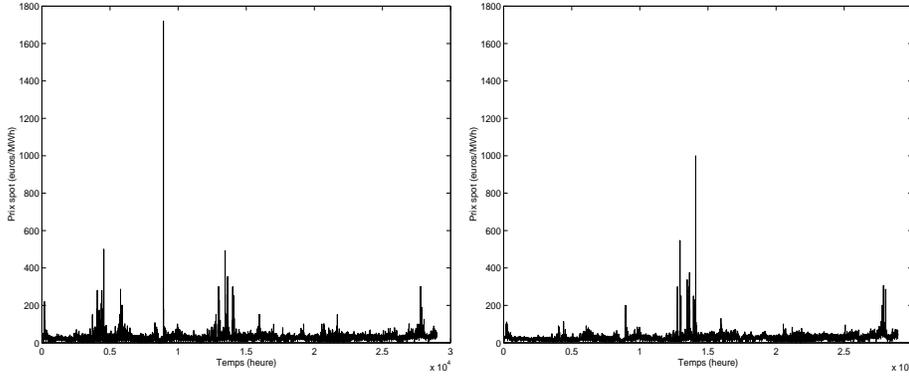


FIG. 4.2 – Evolution des prix spot horaires des marchés EEX (en haut à gauche), Powernext (en haut à droite) pour la période du 01/01/2002 au 31/12/2004.

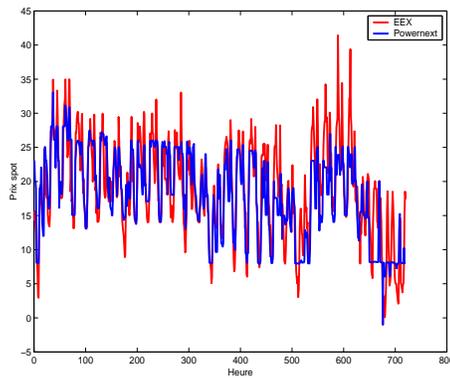


FIG. 4.3 – Zoom sur un mois du 1/03/2002 au 30/03/2002 (en bas).

allemand et français, notées respectivement par $X_{1t} = \log(S_{1t})$ et $X_{2t} = \log(S_{2t})$. Nous décomposons les séries en deux parties : la première partie allant du 01/01/2002 au 31/12/2003 pour l'estimation de ces paramètres des modèles proposés et la seconde partie allant du 01/01/2004 au 31/12/2004 pour la validation des modèles. Sur la figure 4.4, nous représentons l'évolution des prix spot d'électricité des deux marchés sur la période étudiée. Les trajectoires de ces prix semblent être stationnaires en moyenne. Cependant, nous notons clairement que ces deux séries sont non stationnaires en variance. Nous résumons dans le tableau 4.1 les quatre premiers moments des séries considérées. Nous remarquons que les prix spot du marché allemand sont plus dispersés que ceux du marché français. De plus les moments d'ordre

trois et quatre laissent conclure que ces séries sont dissymétriques et sont à queues épaisses. Ces hypothèses sont confirmées par les distributions des séries étudiées (voir figure 4.5). Nous nous intéressons par la suite aux pro-

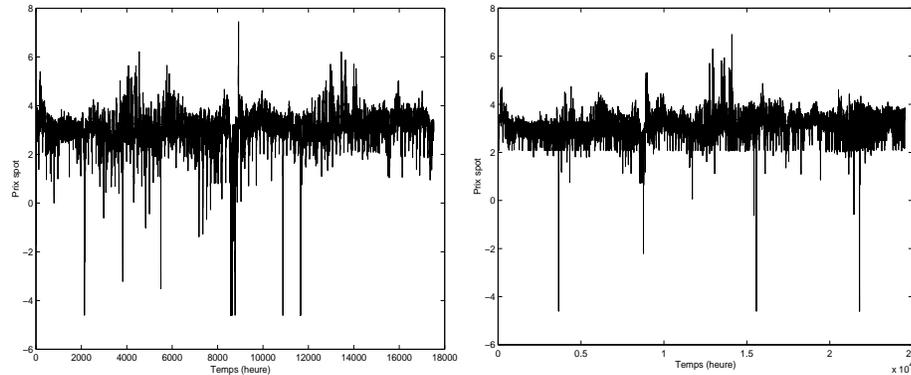


FIG. 4.4 – Evolution à l'échelle logarithmique des prix spot horaires des marchés EEX (en haut à gauche), Powernext (en haut à droite) pour la période du 01/01/2002 au 31/12/2003.

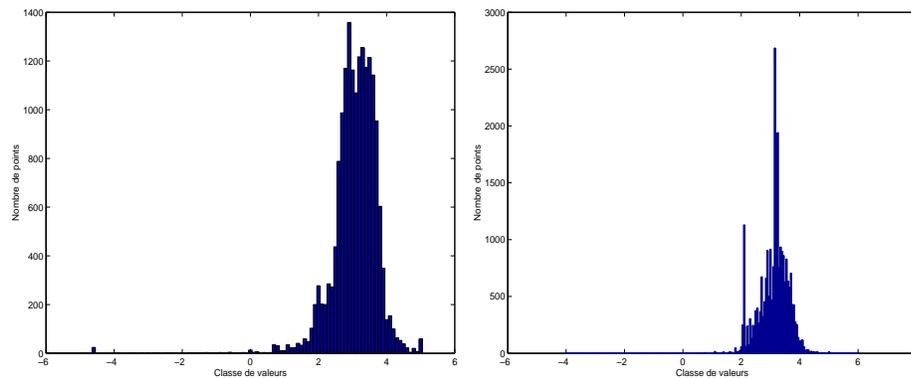


FIG. 4.5 – Histogramme des prix spot horaires pour la période du 01/01/2002 au 31/12/2004 des marchés EEX (en haut à gauche), Powernext (en haut à droite).

priétés statistiques des deux séries à étudier. Nous regardons alors les FAC (Fonctions d'autocorrélation) et les périodogrammes (voir figure 4.6 et 4.7). Sur les FAC, nous observons une décroissance très lente (hyperbolique) vers zéro. De plus, nous notons la présence de pics aux retards multiples de 24h (un jour) et de 168h (une semaine). Cette observation laisse conclure que les

Statistique	EEX	Powernext
Moyenne	3.072	3.065
Variance	0.480	0.324
Skewness	-2.819	-1.388
Kurtosis	27.787	21.722

TAB. 4.1 – Statistiques descriptives des prix spot des marchés allemands et français sur la période du 01/01/2002 au 31/12/2003.

séries étudiées sont cycliques avec des périodes de 24h et de 168h. Les périodogrammes sont non bornés. Nous notons sur chacun d'eux, trois explosions correspondant respectivement aux période de 168h, 24h et 12h.

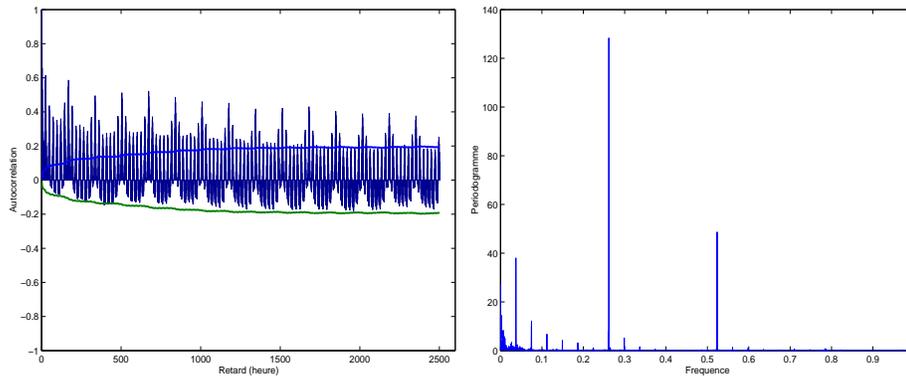


FIG. 4.6 – FAC et Périodogramme des prix spot d'électricité à l'échelle logarithmique du marché allemand EEX.

L'ensemble des remarques précédentes :

- distribution non gaussienne,
- volatilité non constante,
- comportement de type longue mémoire observé sur le périodogramme et sur la fonction d'autocorrélation,
- présence de cycles plus ou moins fixes,

nous conduit à proposer dans le cadre univarié une modélisation de la série $(X_t)_t$ par un processus GARMA à trois facteurs non gaussien avec un bruit à variance non constante (GIGARCH à 3 facteurs). D'autre part, l'interconnexion des marchés européens peut entraîner un phénomène de causalité entre ces marchés. Cette observation nous pousse à considérer une deuxième

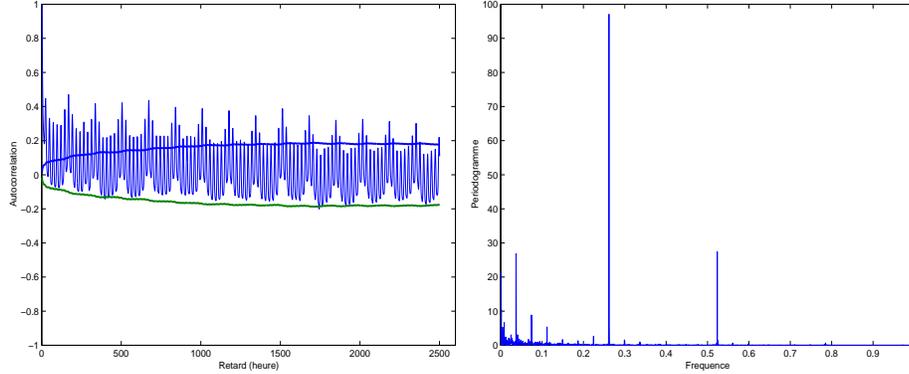


FIG. 4.7 – FAC et Périodogramme des prix spot d'électricité à l'échelle logarithmique du marché français Powernext.

classe de modèles : les modèles longue mémoire généralisés multivariés (MVGARMA à k facteurs).

4.3 Approche univariée

Dans cette section, nous nous intéressons à une modélisation univariée de chacune des séries de marché de l'électricité étudiées. Nous nous proposons de les modéliser par des processus GIGARCH à k facteurs.

Dans les paragraphes 2.3.1 et 2.3.2 nous avons introduit deux méthodes d'estimation des paramètres caractérisant les processus GIGARCH à k facteurs. Ici nous utiliserons la méthode de Whittle développée dans la section 2.3.2. Ce choix est motivé par le fait que les distributions des séries étudiées sont dissymétriques, voir la figure 4.5. Nous rappelons que l'hypothèse de symétrie de la distribution est très importante dans l'estimation des paramètres par la méthode CSS. Nous précisons maintenant les différentes étapes utilisées pour modéliser les séries de prix spot des marchés allemand et français.

Nous ajustons d'abord sur les séries X_{1t} et X_{2t} des processus Gegenbauer à trois facteurs prenant en compte les trois explosions principalement observées sur les périodogrammes, définis par :

$$\prod_{i=1}^3 (I - 2 \cos(\lambda_i) B + B^2)^{d_i} (X_{jt} - \bar{X}_{jt}) = \varepsilon_{jt}, \quad j = 1, 2, \quad (4.1)$$

où \bar{X}_{jt} est la moyenne empirique de X_{jt} . Les paramètres sont estimés à partir de la méthode de Whittle et sont résumés dans le tableau 4.2. Le processus

Paramètres	Powernext	EEX
λ_1	0.2618	0.2618
λ_2	0.0374	0.0374
λ_3	0.5236	0.5236
d_1	0.1223	0.1232
d_2	0.2615	0.2216
d_3	0.1021	0.1243
$\sigma_{\varepsilon_{it}}^2$	0.0542	0.0753

TAB. 4.2 – Paramètres longue mémoire estimés pour le modèle (4.1) pour les séries log prix $(X_{it})_t$ $i = 1, 2$.

$(X_t)_t = (X_{1t}, X_{2t})_t$ défini par l'équation (4.1) est alors stationnaire et inversible. Il admet donc une représentation unique sous forme autorégressive infinie qui fournit l'expression des résidus notés ici $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})$. Les trajectoires des processus $(\varepsilon_{1t})_t$ et $(\varepsilon_{2t})_t$ sont données sur la figure 4.8. Elles présentent des phases de tranquillité suivies de phases de fortes agitations (volatilité non constante). L'objectif est de modéliser les processus $(\varepsilon_{it})_t$,

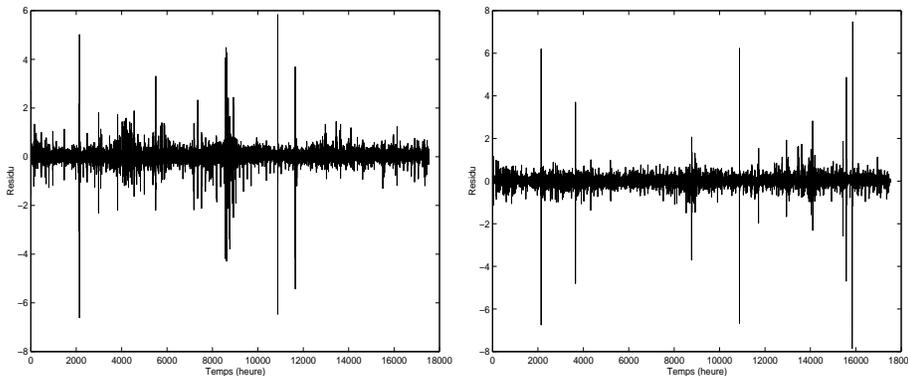


FIG. 4.8 – Évolution des processus résiduels $(\varepsilon_{it})_t$, $i = 1, 2$, issus de la modélisation (4.1).

$i = 1, 2$ par un modèle $\text{ARMA}(p, q) - \text{GARCH}(r, s)$. Nous commençons par regarder la distribution des séries ainsi que leurs caractéristiques statistiques. Nous résumons dans le tableau 4.3 suivant les quatre premiers moments des séries $(\varepsilon_{it})_t$, $i = 1, 2$.

Statistiques	EEX	Powernext
Moyenne	-1.810310^{-5}	$-2.3301 \cdot 10^{-5}$
Variance	0.0771	0.0555
Skewness	-1.874	-1.321
Kurtosis	110.0710	303.9730

TAB. 4.3 – Statistiques descriptives des résidus $(\varepsilon_{it})_t$, $i = 1, 2$, issus de la modélisation (4.1).

Le moment d'ordre trois des séries est non nul et négatif. Nous pouvons alors conclure que les distributions sont dissymétriques à gauche. Les moments d'ordre 4 (kurtosis) sont très écartés de celui d'une loi Normale. Nous en déduisons alors que les distributions des résidus ne sont pas gaussiennes. La figure 4.9 semble confirmer cette affirmation. L'étude des FAC et des PACF

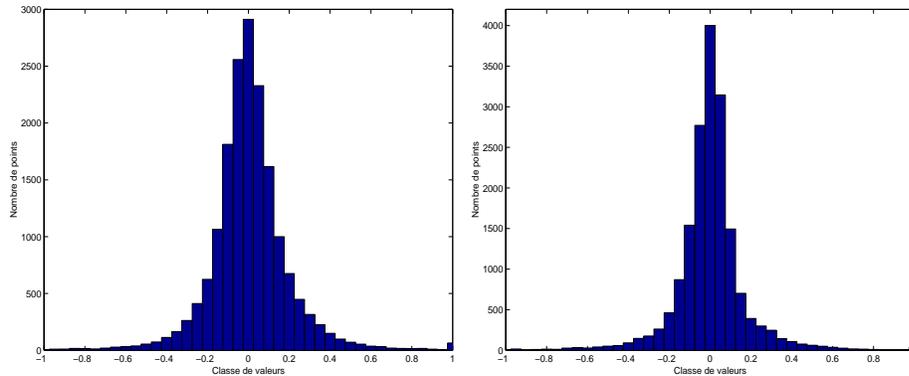


FIG. 4.9 – Histogramme des processus résiduels $(\varepsilon_{it})_t$, $i = 1, 2$, issus de la modélisation (4.1).

(*Partial Autocorrelation Function*) (voir figures 4.10 et 4.11) des processus $(\varepsilon_{1t})_t$ et $(\varepsilon_{2t})_t$ nous conduisent à privilégier un processus à mémoire courte saisonnier, car les FAC ainsi que les PACF présentent des pics significativement non nuls aux retards multiples de 24h et de 168h. Plusieurs modèles ont été ainsi testés. Nous retenons en définitif les deux modèles suivants :

1. Pour le marché allemand :

$$\phi(B) \varepsilon_{1t} = \theta(B) \xi_{1t}, \quad (4.2)$$

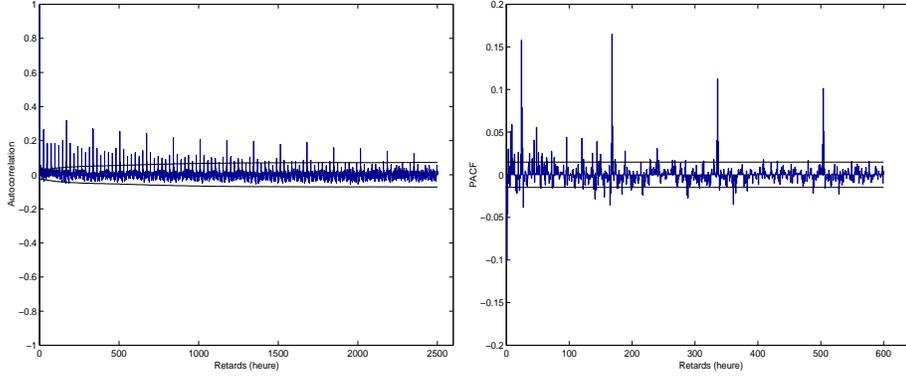


FIG. 4.10 – FAC et PACF des processus $(\varepsilon_{1t})_t$, issus de la modélisation (4.1).

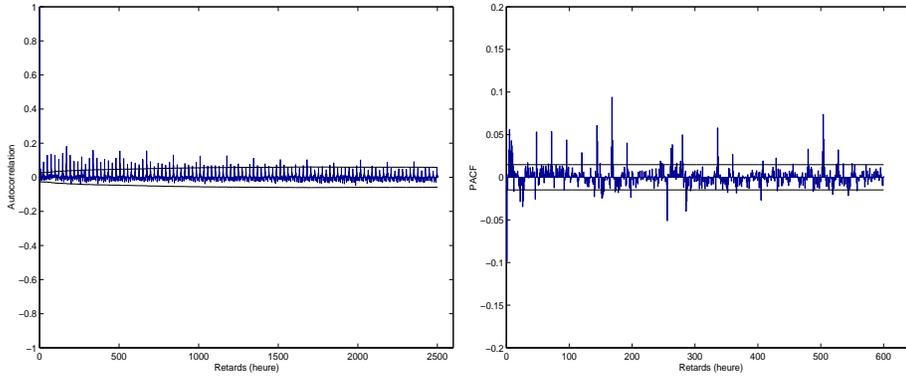


FIG. 4.11 – FAC et PACF du processus $(\varepsilon_{2t})_t$, issus de la modélisation (4.1).

avec $\phi(B) = (I - \phi_1^{(1)}B)(I - \phi_2^{(1)}B^{24})(I - \phi_3^{(1)}B^{168})$ et $\theta(B) = (I - \theta_1^{(1)}B)(I - \theta_2^{(1)}B^{24})(I - \theta_3^{(1)}B^{168})$.

Ici,

$$\xi_{1t} = \sqrt{h_{1t}}\eta_{1t}, \text{ avec } h_{1t} = a_0^{(1)} + a_1^{(1)}\xi_{1t-1}^2 + a_2^{(1)}\xi_{1t-2}^2 + a_3^{(1)}\xi_{1t-3}^2. \quad (4.3)$$

2. Pour le marché français :

$$(I - \phi_2^{(2)}B^{24})\varepsilon_{2t} = (I - \theta_2^{(2)}B^{24})\xi_{2t}, \quad (4.4)$$

où

$$\xi_{2t} = \sqrt{h_{2t}}\eta_{2t}, \text{ avec } h_{2t} = a_0^{(2)} + a_1^{(2)}\xi_{2t-1}^2. \quad (4.5)$$

Paramètres	Powernext	EEX
$\hat{\phi}_1$	-	0.110
$\hat{\phi}_2$	0.982	0.348
$\hat{\phi}_3$	-	0.926
$\hat{\theta}_1$	-	0.226
$\hat{\theta}_2$	0.925	0.169
$\hat{\theta}_3$	-	0.768
\hat{a}_0	0.025	0.0092
\hat{a}_1	0.6088	0.6257
\hat{a}_2	-	0.3563
\hat{a}_3	-	0.1266

TAB. 4.4 – Estimation des paramètres pour les modèles 4.2-4.3 et 4.4-4.5.

Les paramètres estimés des modèles définis par les équations (4.2)-(4.3) et (4.4)-(4.5) sont résumés dans le tableau 4.4.

Les FAC empiriques des processus $(\eta_{it})_t$, $i = 1, 2$, représentées par la figure 4.12, conduisent à penser que les bruits $(\eta_{it})_t$, $i = 1, 2$, sont blancs. Ainsi, nous pouvons conclure que les modèles retenus sont satisfaisants et semblent bien prendre en compte les caractéristiques observées sur les séries de prix spot de marché de l'électricité allemand et français.

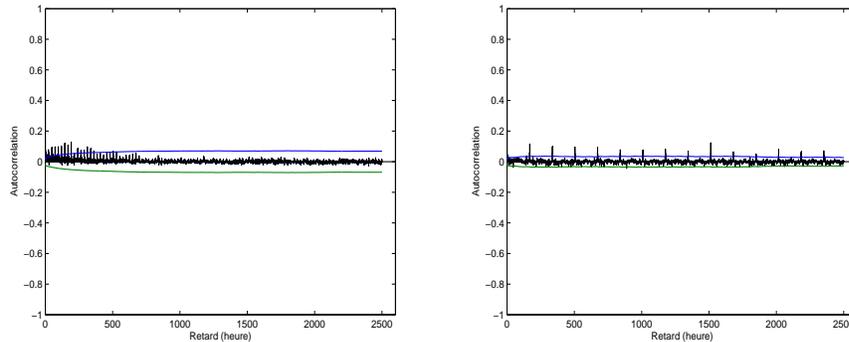


FIG. 4.12 – FAC des processus $(\xi_{it})_t$, $i = 1, 2$.

Les modèles (M_1) et (M_2) respectivement associés au log du prix spot $(X_{1t})_t$ du marché allemand et $(X_{2t})_t$ du marché français sur la période du 01/01/2002 au 31/12/2003 sont respectivement donnés par les équations

tions suivantes :

$$M_1 : \phi^{(1)}(B) \prod_{i=1}^3 (I - 2\nu_i B + B^2)^{d_i^{(1)}} (X_{1t} - 3.072) = \theta^{(1)}(B) \varepsilon_{1t}, \quad (4.6)$$

où,

$$\phi^{(1)}(B) = (I - \phi_1^{(1)} B) (I - \phi_2^{(1)} B^{24}) (I - \phi_3^{(1)} B^{168}), \quad (4.7)$$

$$\theta^{(1)}(B) = (I - \theta_1^{(1)} B) (I - \theta_2^{(1)} B^{24}) (I - \theta_3^{(1)} B^{168}), \quad (4.8)$$

$$\varepsilon_{1t} = \sqrt{h_{1t}} \eta_{1t}, \text{ avec } h_t = a_0^{(1)} + \sum_{i=1}^3 a_i^{(1)} \xi_{1t-i}^2, \quad (4.9)$$

et

$$M_2 : \phi^{(2)}(B) \prod_{i=1}^3 (I - 2\nu_i B + B^2)^{d_i^{(2)}} (X_{2t} - 3.065) = \theta^{(2)}(B) \varepsilon_{2t}, \quad (4.10)$$

avec $\phi^{(2)}(B) = (I - \phi_2^{(2)} B^{24})$, $\theta^{(2)}(B) = (I - \theta_2^{(2)} B^{24})$. Le processus $(\varepsilon_{2t})_{t \in \mathbb{Z}}$ vérifie l'équation (4.11) suivante :

$$\varepsilon_{2t} = \sqrt{h_{2t}} \eta_{2t}, \text{ avec } h_{2t} = a_0^{(2)} + a_1^{(2)} \xi_{2t-i}^2. \quad (4.11)$$

4.4 Approche multivariée

Dans cette section, nous étudions la modélisation jointe des deux marchés. En effet, les marchés européens de l'électricité étant interconnectés, il nous paraît alors intéressant de les modéliser conjointement.

Nous proposons à la suite de modéliser la série $(X_t)_t$ par un processus MVGARMA à k facteurs. Pour ce faire, nous utilisons la procédure pratique d'estimation décrite dans la section 3.3. La partie longue mémoire, définie par l'équation (4.12), est modélisée de façon analogue au cas univarié (voir section 4.3). Le modèle se formalise alors comme suit :

$$D(B) X_t = \varepsilon_t, \quad (4.12)$$

où

$$D(B) = \text{diag} \left[\prod_{i=1}^3 (I - 2\nu_i^1 B + B^2)^{d_i^{(1)}}, \prod_{i=1}^3 (I - 2\nu_i^2 B + B^2)^{d_i^{(2)}} \right].$$

Les paramètres estimés sont résumés dans le tableau 4.2. Le processus $(X_t)_t$ est stationnaire et inversible alors nous nous servons de l'équation (3.31) pour récupérer les résidus $(\varepsilon_t)_t$ issus du modèle défini par l'équation (4.12). Dans la suite, nous cherchons à modéliser les résidus $(\varepsilon_t)_t$ par un modèle VARMA. Les FAC et PACF des résidus $(\varepsilon_t)_t$ sont présentées sur les figures 4.10 et 4.11. Sur la figure 4.13, nous représentons les autocorrélations croisées des processus $(\varepsilon_{1t})_t$ et $(\varepsilon_{2t})_t$. Nous observons la présence de pics significativement non nuls aux retards 1, $24k$ et $168k$, $k = 1, 2, \dots$.

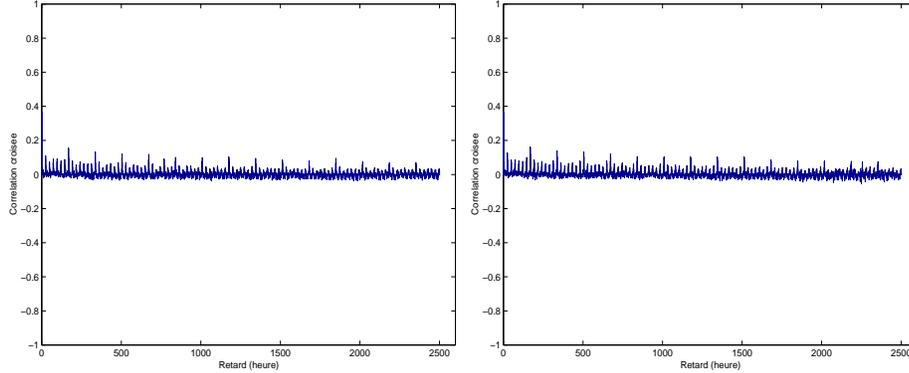


FIG. 4.13 – FAC croisées des processus $(\varepsilon_{it})_t$, $i = 1, 2$.

Ainsi, plusieurs modèles VAR ont été testé sur les résidus $(\varepsilon_t)_t$ et finalement le modèle retenu se formalise de cette manière :

$$\Phi(B) \varepsilon_t = \xi_t, \quad (4.13)$$

avec

- (a) . $\Phi(B) = I - \Phi_{24}B^{24} - \Phi_{48}B^{48} - \Phi_{72}B^{72} - \Phi_{96}B^{96} - \Phi_{168}B^{168} - \Phi_{336}B^{336} - \Phi_{504}B^{504}$,
- (b) . $\xi_t = (\xi_{1t}, \xi_{2t})' \sim i.i.d.\mathcal{N}(0, \Sigma)$,
- (c) . $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$ avec Σ symétrique et définie positive.

Remarque 4.4.1 *Un modèle VARMA pourrait être utilisé pour modéliser le processus $(\varepsilon_t)_t$ mais des études empiriques ont montré que les modèles VAR étaient satisfaisants pour décrire la dépendance temporelle souvent observée sur les données.*

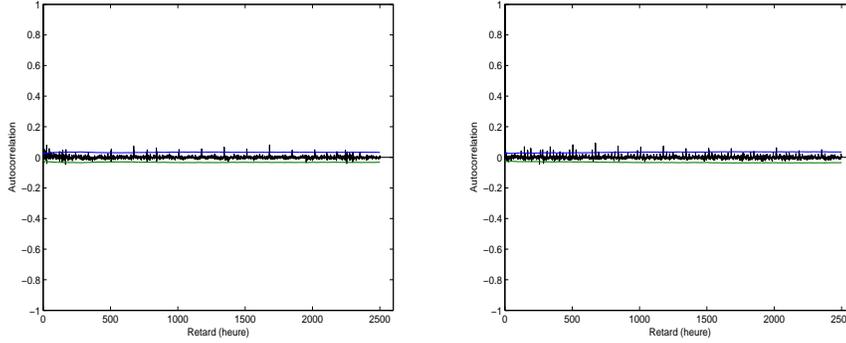


FIG. 4.14 – FAC des processus $(\xi_{it})_t$, $i = 1, 2$.

Sur la figure 4.14, nous représentons les FAC empiriques des processus $(\xi_{it})_t$, $i = 1, 2$. De celles-ci, nous concluons que les processus $(\xi_{it})_t$, $i = 1, 2$ semblent être des bruits blancs. Le modèle multivarié proposé semble alors satisfaisant et capture suffisamment les caractéristiques observées sur ces marchés spot. Les résultats des paramètres estimés du modèle retenu sont donnés dans le tableau 4.5. Nous remarquons que tous les paramètres sont significativement non nuls.

4.5 Comparaison des résultats

Dans cette section, nous nous intéressons aux capacités prédictives des modèles présentés dans les sections 4.3 et 4.4. Nous considérons la période allant du 01/01/2004 au 31/12/2004 comme période de validation. Nous comparons ainsi les performances de prévision des modèles en utilisant le RMSE (*Root Mean Squares Error*) défini par l'équation suivante :

$$RMSE_k^i = \sqrt{\frac{1}{M} \sum_{j=1}^M \left(X_{t+j+k}^i - \hat{X}_{t+j}^i(k) \right)^2}, \quad k = 1h, \dots, 720h, \quad i = 1, \dots, 24, \quad (4.14)$$

Variables	Paramètres	Paramètres estimés	Ecart type
	ϕ_1^{11}	-0.0647	0.0079
	ϕ_{24}^{11}	0.0783	0.0079
	ϕ_{24}^{12}	0.0842	0.0079
	ϕ_{24}^{22}	0.0183	0.0088
	ϕ_{48}^{11}	0.0717	0.0079
	ϕ_{48}^{12}	0.0934	0.0080
	ϕ_{48}^{22}	0.0252	0.0089
	ϕ_{72}^{11}	0.0745	0.0080
	ϕ_{72}^{12}	0.0770	0.0079
	ϕ_{72}^{21}	0.0272	0.0088
	ϕ_{96}^{11}	0.0382	0.0068
	ϕ_{96}^{12}	0.0634	0.0070
	ϕ_{96}^{21}	-0.0528	0.0075
	ϕ_{96}^{22}	0.1516	0.0078
	ϕ_{168}^{21}	0.0554	0.0079
	ϕ_{168}^{22}	0.0495	0.0078
	ϕ_{336}^{21}	0.0539	0.0072
	ϕ_{336}^{22}	0.1826	0.0079
	ϕ_{504}^{11}	0.0329	0.0072
	ϕ_{504}^{12}	0.0264	0.0070
	ϕ_{504}^{21}	0.1147	0.0079
	ϕ_{504}^{22}	0.1191	0.0078
EEX	σ_1^2	0.0496	
Powernext	σ_2^2	0.0597	
	σ_{12}	0.0172	

TAB. 4.5 – Paramètres estimés pour le modèle VAR défini par 4.13

où M est la taille de l'échantillon de validation des modèles et $\hat{X}_t^i(k)$ est le prédicteur de moindres carrés à l'instant t et à l'horizon k de l'heure i de X_{t+h}^i . Ce critère est utilisé dans ce travail car nous nous intéressons particulièrement à mesurer la variation de l'erreur quadratique. De plus, nous comparons ces résultats avec un modèle de référence supposant que la meilleure prévision d'une heure donnée est la moyenne empirique de l'heure considérée sur toute la période de validation. Ce qui nous ramène alors à comparer l'écart moyen quadratique de l'erreur de prévision avec un écart type.

Néanmoins, d'autres mesures de performances de prévision existent. Par exemple, Slama (2000) a utilisé le MAPE (*Mean Absolute Percentage Error*) alors que Bakitzis *et al.* (1996) ont utilisé le MAE (*Mean Absolute Error*).

Pour chacun des modèles proposés dans ce travail, nous effectuons des prévisions allant d'un jour à un mois. Cependant, ces prévisions sont faites sur les prix spot pris à l'échelle logarithmique. Pour fournir des résultats de prévision dans l'échelle linéaire, nous avons besoin de calculer l'espérance conditionnelle de l'exponentiel des séries considérés. Ce problème étant très complexe, nous supposons alors ici que les résidus issus des modèles proposés suivent une loi Normale pour contourner toutes les difficultés liées à la loi du bruit. Le prédicteur des moindres carrés du prix spot au temps n et à l'horizon h est alors calculé à partir de l'équation suivante :

$$\hat{S}_n(h) = \exp \left[\hat{X}_n(h) + \frac{1}{2} \text{var}(\hat{e}_{n,h}) \right]. \quad (4.15)$$

Les figures 4.15-4.17 et B.1-B.8 représentent l'ensemble des résultats de prévision en $J + 30$ obtenus à partir des modèles développés dans les sections 4.3 et 4.4. Sur les figures 4.15 et 4.16, nous présentons les résultats de prévisions de deux mois spécifiques : le premier allant du 01/12/2004 au 31/12/2004 et le second du 01/07/2004 au 31/07/2004. Nous remarquons, à partir de ces figures, que les résultats de prévision fournis par les deux modèles sont satisfaisants aussi bien pour le marché français que pour le marché allemand. Cependant, le modèle multivarié semble en général donner des résultats plus satisfaisants que le modèle univarié pour les deux marchés considérés. Pour confirmer ou infirmer cette assertion, nous calculons les RMSE heure par heure pour

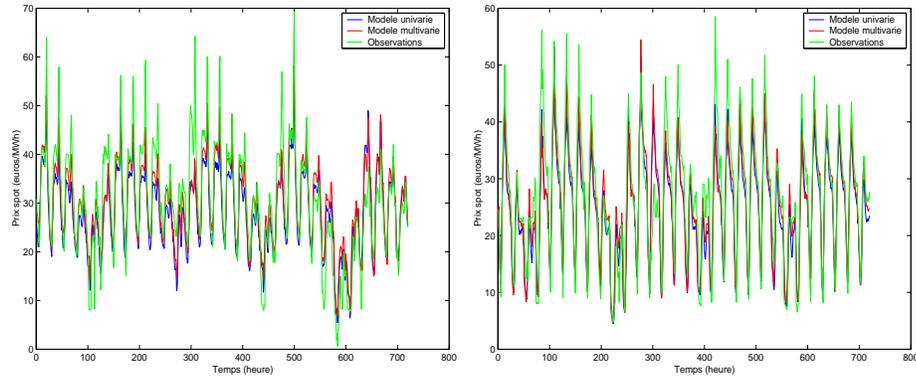


FIG. 4.15 – Résultats de prévision en $J+30$ de prix spot d'électricité du marché allemand pour la période du 01/12/2004 au 31/12/2004 et du 01/07/2004 au 31/07/2004.

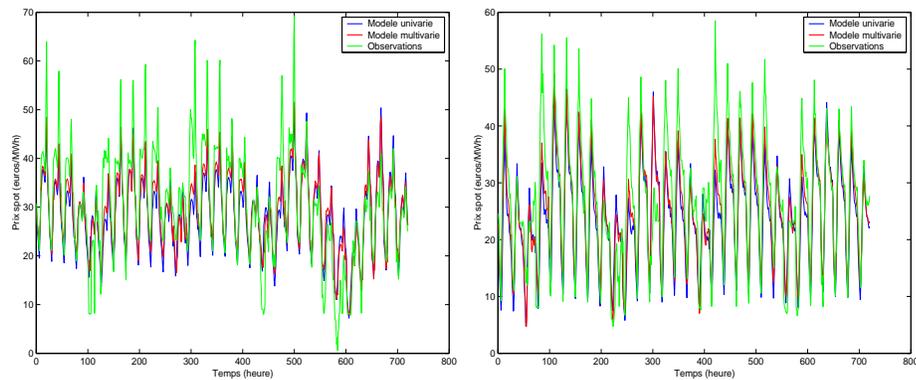


FIG. 4.16 – Résultats de prévision en $J + 30$ de prix spot du marché français pour la période du 01/12/2004 au 31/12/2004 et du 01/07/2004 au 31/07/2004.

un horizon de prévision d'un mois et ceci pour toute l'année 2004. En effet, les marchés français et allemand étant des marchés "day-ahead" alors la comparaison d'une heure particulière d'un jour avec une autre n'a pas trop de sens.

Ces RMSE révèlent que pour le marché allemand, respectivement pour le marché français, le modèle multivarié fournit des résultats plus satisfaisants que le modèle univarié jusqu'à un horizon de 19 jours, respectivement 24 jours, (voir figure 4.17). Ils croissent en fonction de l'horizon de prévision et varient en moyenne pour le marché allemand entre 6.40 euros/MWh et 8.18 euros/MWh en univarié et entre 6 euros/MWh et 8.44 euros/MWh en multivarié. Pour le marché français, ils croissent

entre 6.77 euros/MWh et 8.50 euros/MWh en univarié et entre 5.75 euros/MWh et 8.58 euros/MWh en multivarié. Enfin, sur les figures B.1-B.8 données en annexe B, on représente les RMSE heure par heure pour un horizon de prévision allant jusqu'à un mois. Cependant, les modèles proposés ne permettent pas de bien prévoir les jour fériés (par exemple le 14/07/2004 en France).

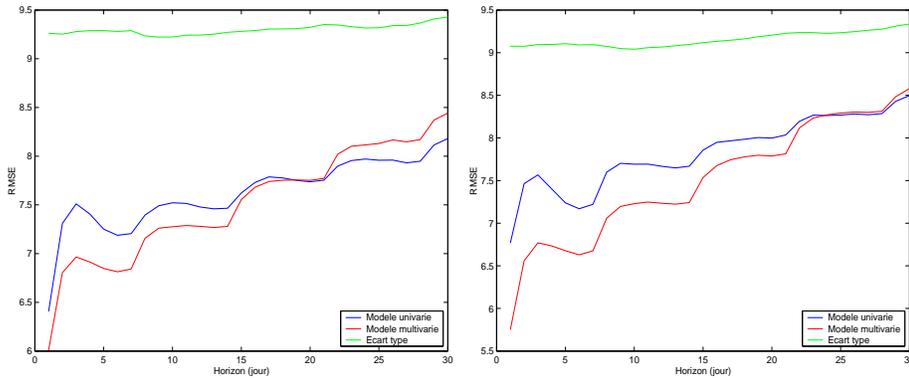


FIG. 4.17 – RMSE en moyenne (EEX à gauche et Powernext à droite) pour une prévision en $J + 30$.

4.6 Conclusions et perspectives

Dans ce chapitre, nous nous sommes intéressés à l'application des processus étudiés précédemment sur les données de prix spot de marchés de l'électricité. Nous avons considéré l'étude des marchés spot français et allemand. Pour chacun des marchés, deux modèles sont proposés : un modèle univarié (GIGARCH à k facteurs) et un modèle permettant de modéliser conjointement les deux marchés (MVGARMA à k facteurs). Nous nous sommes penchés sur les capacités prédictives de ces modèles. Pour cela, nous avons donné des prévisions allant jusqu'à un horizon d'un mois avec ces deux modèles. L'étude a montré que les modèles proposés reproduisent bien les caractéristiques observées sur les séries de prix spot français et allemand. En effet, les bruits issus des modèles semblent être tous blancs. Les résultats de prévision obtenus avec ces modèles sont satisfaisants. Cependant, pour le marché français le modèle multivarié est de loin meilleur que le modèle univarié.

Les modèles proposés ne permettent pas de bien prévoir les jours fériés et les pics de prix. Il serait alors intéressant de trouver d'autres modèles pouvant aussi prendre en compte ces deux composantes. Dans nos applications, nous n'avons considéré que les marchés spot français et allemand. On pourrait aussi appliquer ces modèles à d'autres marchés spot européens ou même à d'autres séries financières. Des études ont montré que le prix spot dépend fortement de la demande en électricité. Il serait intéressant d'appliquer le modèle multivarié par exemple aux séries prix et consommation ou prix et température ou prix, température et consommation. Une autre approche, consistant à décomposer les séries heure par heure (ce qui donne 24 séries) et de modéliser indépendamment chacune d'elle, pourrait être testée.

Conclusion générale

Notre travail consiste à l'utilisation des processus longue mémoire pour la modélisation et la prévision des prix spot de marché de l'électricité. D'une part, pour prendre en compte à la fois les phénomènes de dépendance à long terme, de cycles plus ou moins fixes et de volatilité non constante observés sur les prix spot de marché de l'électricité, nous avons introduit les modèles longue mémoire généralisés conditionnellement hétéroskédastiques. A des fins d'application, nous nous sommes intéressés aux propriétés statistiques de ces processus. Nous avons alors proposé deux méthodes d'estimation des paramètres : une première méthode basée sur les moindres carrés conditionnels et une seconde méthode basée sur la méthode de pseudo-maximum de vraisemblance de Whittle. Pour chacune des méthodes, nous avons étudié les propriétés asymptotiques des estimateurs associés : consistance et normalité asymptotique. Afin de comparer les propriétés asymptotiques des estimateurs, un plan de simulations numériques est effectué. Les résultats obtenus sont tous satisfaisants. Cependant la méthode de Whittle est privilégiée car le méthode CSS impose des hypothèses un peu contraignantes sur la distribution du bruit. Pour évaluer la capacité prédictive de ces modèles, nous avons établi les expressions analytiques des prédicteurs de moindres carrés en moyenne et en variance. Sous l'hypothèse de gaussianité, nous avons donné l'intervalle de confiance de l'erreur de prévision.

D'autre part, parce que les marchés européens de l'électricité sont interconnectés, nous avons introduit les processus longue mémoire multivariés pour modéliser conjointement les prix spot de l'électricité de ces marchés. Nous nous sommes particulièrement intéressés aux processus MVGARMA à k facteurs. Nous avons proposé une procédure

pratique d'estimation des paramètres du modèle et nous avons étudié les conditions de stationnarité et d'inversibilité du processus. Le problème principal qu'on cherchait à résoudre est l'apport des modèles longue mémoire multivariés sur la prévision de prix spot de l'électricité de marchés européens. Pour ce faire, nous avons fourni l'expression analytique du prédicteur de moindres carrés d'un processus MVGARMA à k facteurs. Nous avons aussi donné, sous l'hypothèse de gaussianité, l'intervalle de confiance de l'erreur de prévision.

Les modèles proposés ont permis de prendre en compte presque l'ensemble des caractéristiques observées sur les prix spot de marché de l'électricité. En effet, les résidus issus des modèles proposés semblent être blancs et de distributions non gaussiennes. Afin de comparer les modèles pour chacun des marchés, nous avons évalué leurs capacités prédictives. Les résultats de prévision obtenus sont dans l'ensemble satisfaisants pour tous les modèles. Cependant, les résultats de prévision du marché français ainsi que du marché allemand avec le modèle multivarié sont dans l'ensemble meilleurs que ceux obtenus avec le modèle univarié.

Bibliographie

- [1] Arteche, J., Robinson, J.M. (2000) Semiparametric inference in seasonal and cyclical long memory processes, *Journal of Time Series Analysis*, **21**(1), 1-25.
- [2] Atkins, F.J., Chen, J. (2002) Some statistical properties of deregulated electricity prices in Alberta, *Discussion Paper*, 2002-06 University of Calagary, Canada.
- [3] Baillie, R.T., Bollerslev, T. (1992) Prediction in dynamic models with time-dependent conditional variances, *Journal of Econometrics*, **52**, 91-113.
- [4] Baillie, R.T., Chung, C.-F., Tieslau, M.A.(1996) Analyzing Industrialized Countries Inflation by the Fractionally Integrated ARFIMA-GARCH Model, *Journal of Applied Econometrics*, **11**, 23-40.
- [5] Baillie, R.T., Chung, C-F. (2002) Modeling and forecasting from trend-stationary long memory models with applications to climatology, *International Journal Forecasting*, **18**, 215-226.
- [6] Bakirtzis, A.G., Petridis, V., Kiartzis, S.J., Alexiadis, M.C., Masis, A.H. (1996) A neural network short term load forecasting model for the Greek power system, *IEEE Transactions on Power Systems*, **11**, 858-863.
- [7] Barkoulas, J.T., Baum, C.F. (1997) Fractional differencing modeling and forecasting of eurocurrency, *Journal of Financial Research*, **20**(3), 355-372.
- [8] Barquin, J. (2004) A new model for electricity price series modeling and forward and volatility curves computation, *IEEE Proceedings of the 8th International Conference on Probabilistic Methods*

Applied to Power Systems, Iowa State University, Ames, Iowa, 1-7.

- [9] Barrieu, P., El Karoui, N. (2002) Reinsuring Climatic Risk Using Optimally Designed Weather Bonds, *Preprint HEC*.
- [10] Basawa, I. V., Feign, P. D., Heyde, C. C. (1976) Asymptotic Properties of Maximum Likelihood Estimators for Stochastic Processes, *Sankhya, Ser. A*, **38**, 259-270.
- [11] Benini, M., Marracci, M., Pelacchi, P., Venturini, A. (2002) Day-Ahead Market Price Volatility Analysis in Deregulated Electricity Markets, *Proc. 2002 IEEE Power Engineering Society Summer Meeting, Chicago, IL, United States* **3**, 1354-1359.
- [12] Beran, J. (1994) Statistics for Long memory Processes, *Chapman and Hall*, London.
- [13] Bertelli, R.J., Caporin, M. (2002) A note on calculating autocovariances of long memory processes, *Journal of Time Series Analysis*, **23**(5), 503-508.
- [14] Bhansali, R.J., Kokoszka, P.S. (2003) Prediction of long-memory time series, *Theory and applications of long-range dependence*, 335-367, Birkhäuser Boston, Boston, MA.
- [15] Bhardwaj, G., Swanson, N.R. (2005) An empirical investigation of the usefulness of ARFIMA models for predicting macroeconomic and financial time series, *Journal of Econometrics*, in press.
- [16] Bisaglia, L. (1998) Processi a memoria lunga : problemi di stima, identificazione e previsione, Dottora di Ricerca in Statistica, Ciclo X, Università degli Studi di Padova.
- [17] Bisaglia, L. Guégan, D. (1998) A comparison of techniques of estimation in long memory processes : application to intra-day data, *Computational Statistics and Data Analysis*, **27**, 61-81.
- [18] Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* **31**, 307-327.
- [19] Bollerslev, T. (1987) A conditional heteroskedastic time series model for speculative prices and rates of return, *Review of Economic and Statistics* **69**, 542-574.

- [20] Bollerslev, T. (1988) On the correlation structure for the generalized autoregressive conditional heteroskedastic process, *Journal of Time Series Analysis* **9**, 121-131.
- [21] Bondon, P., Palma, W. (2004) Prediction of strongly dependent time series, *Preprint*.
- [22] Breipohl, A.M. (2002) Electricity price forecasting model, *IEEE Power Engineering Society Winter Meeting*, **2**, 963-966.
- [23] Brillinger, D. (1964) The asymptotic behaviour of Tuckey's general method of setting approximate confidence limits (the jackknife) when applied to maximum likelihood estimates, *Revue de l'Institution Internationale de Statistique*, **32**, 202-206.
- [24] Brockwell, P.J., Davis, R.A. (1987) Time series : theory and methods, Springer-verlag, New York.
- [25] Brodsky, J., Hurvich, C.M. (1999) Multi-step forecasting for long-memory processes, *Journal of Forecasting* **18**, 59-75.
- [26] Bunn, D.W. (2003) Modelling prices in competitive electricity market, Wiley.
- [27] Bunn, D.W., Karakatsani, N. (2003) Forecasting electricity prices, *Working Paper*, London Business School.
- [28] Burger, M., Klar, B., Müller, A., Schindlmayr, G. (2004) A spot market model for pricing derivatives in electricity markets, *Quantitative Finance*, **4**, 109-122.
- [29] Caballero, R., Jewson, S.P. (2002) Multivariate long-memory modeling for daily surface air temperatures and the valuation of weather derivative portfolios, *Preprint University of Copenhagen Denmark 2002*, 1-25.
- [30] Caporin, M. (2002) Long memory conditional heteroskedastic and second order causality, *Università di Venezia*, Ph.D. Dissertation.
- [31] Chambers, M.J. (1995) The simulation of random vector time series with given spectrum, *Math. Comput. Modelling*, **22**(2), 1-6.
- [32] Chung, C.F. (1996a) Estimating a Generalized Long Memory Process, *Journal of Econometrics*, **73**, 237-259.
- [33] Chung, C.-F. (1996b) A Generalized Fractionally Integrated ARMA Process, *Journal of Time Series Analysis*, **17**(2), 111-140.

- [34] Chen, C.W.S., Yu, T.H.K. (2005) Long-term dependence with asymmetric conditional heteroscedasticity in stock return, *Physica A*, in press.
- [35] Coifman, R.R., Meyer, Y., Whickerhauser, M.V. (1992) Wavelet analysis and signal processing, *In Wavelets and their applications, Boston*, 153-178.
- [36] Collet, J.J. (2003) Les processus longue mémoire : prévisions, estimations et valeurs extrêmes, *Ph.D Thesis*, Université de Reims, France.
- [37] Collet, J.J., Guégan, D., Valdes-Sosa, P. (2003) How shall we determine the number and the location of the Gegenbauer frequencies? An empirical approach, *Note de Recherche MORA 2003-09, Ecole Normale Supérieure de Cachan, France*, 1-31.
- [38] Collet, J.J., Fadili, J.M. (2005) Simulation of Gegenbauer processes using wavelet packets, *Preprint*
- [39] Contreras, J., Espínola, R., Nogales, F.J., Conejo, A.J. (2003) ARIMA models to predict next-day electricity prices, *IEEE Trans Power Systems* **18**(3), 1014-1020.
- [40] Crato, N., de Lima, P.J.F. (1994) Long-range dependence in the conditional variance of stock returns, *Economics Letters*, **45**, 281-285.
- [41] Crato, N., Ray, B.K. (1996) Model selection and forecasting for long-range dependent processes, *Journal of Forecasting*, **15**, 107-125.
- [42] Dahlgren, R.W., Liu, C-C., Lawarree, J. (2001) Volatility in the California Power Market : Source, Methodology and Recommendations, *IEE Proc.-Gener. Transm. Distrib.*, **148**(2),189-193.
- [43] Darbellay, G.A., Slama, M. (2000) Forecasting the short-term demand for electricity. Do the neural networks stand a better chance? *International Journal of Forecasting*, **16**, 71-83.
- [44] Davidson, J. (2002) Testing for fractional cointegration. *Journal of Time Series Analysis*, in press.
- [45] Ding, Z., Granger, C.W.J., Engle, R.F. (1993) A long memory property of stock market and a new model, *Journal of Empirical Finance*, **1**, 83-106.

- [46] Diongue, A.K., Guégan, D. (2004) Estimating parameters of a k -factor GIGARCH process. *C. R. Acad. Sci. Paris*, **I 339**(6), 435-440.
- [47] Diongue, A.K., Guégan, D., Vignal, B. (2004a) A k -factor GIGARCH process : estimation and application on electricity market spot prices, *IEEE Proceedings of the 8th International Conference on Probabilistic Methods Applied to Power Systems, Iowa State University, Ames, Iowa*, 1-7.
- [48] Diongue, A.K., Guegan, D., Vignal, B. (2004b) Forecasting electricity spot market prices with a k -factor GIGARCH process, *Note de Recherche MORA 2004-09, Ecole Normale Supérieure de Cachan, France*, 1-31.
- [49] Diongue, A.K., Guégan, D. (2005) Forecasting electricity market spot prices with a k -factor MGARMA process *en cours de rédaction*.
- [50] Dolado, J.J., Marmol, F. (2004) Asymptotic Inference Results for Multivariate Long-Memory Processes, *Econometrics Journal*, **7**(1), 168-190.
- [51] Espínola, R., Contreras, J., Nogales, F.J., Conejo, A.J. (2002) Day-ahead electricity price forecasting based on time series models : a comparison, *Proceedings of the 14th PSCC*, Sevilla, 24-28.
- [52] Eydeland, A., Geman, H. (1999) Fundamentals of electricity derivative pricing, In *Jameson, R. (ed.), Energy Modelling and Management of Uncertainty*, Risk Books.
- [53] Fei, L., Yaan, L., Qing, W., Yunguo, L. (2000) Electricity power system load forecast and electricity price forecast, *Relay*, **28**(10), 31-33 (in chinese).
- [54] Ferrara, L., Guégan, D. (1999) Estimation and application of Gegenbauer processes, *Working Paper n 9927*, CREST-INSEE.
- [55] Ferrara, L. (2000) Processus Longue Mémoire Généralisé : Estimation, Prévision et Applications, *Ph.D Thesis*, Paris XIII, France.
- [56] Ferrara, L., Guégan, D. (2000) Forecasting financial time series with generalized long memory processes, In *Advances in Quantitative Asset Management*, Vol. 1, Davis C (eds) ; Kluwer, Amsterdam. 319-342.

- [57] Ferrara, L., Guégan, D. (2001a) Forecasting with k -factor Gegenbauer processes : Theory and Applications, *Journal of Forecasting*, **20**, 1-21.
- [58] Ferrara, L., Guégan, D. (2001b) Comparison of Parameter Estimation Methods in Cyclical Long Memory Time Series, *Developments in Forecasts Combination and Portfolio Choice*, eds C. Dunis and J. Timmerman, Chapter 8, J. Wiley.
- [59] Fleten, S.E., Wallace, S.W. (1998) Power scheduling with forward contracts, *Proceedings of the Nordic MPS*, Molde, Norway.
- [60] Fossa, A. (2001) The California Electricity Market Analysis during the Summer 2000 : Application of stochastic Techniques, *Ph.D dissertation*, Dipartimento de Ingegneria Elettrica, Università degli Studi de Pavia.
- [61] Fosso, O.B., Gjelsvik, A., Haugstad, A., Birger, M., Wangensteen, I. (1999) Generation Scheduling in a Deregulated System. The Norwegian case, *IEEE Trans. Power Systems*, **14**(1), 75-81.
- [62] Franses, P.H, Ooms, M. (1997) A periodic long memory model for quarterly UK inflation, *International Journal of Forecasting*, **13**,117-126.
- [63] Francisco, N.J., Javier, C. *et al.* (2002) Forecasting next-day electricity prices by time series models, *IEEE Transactions on Power Systems*, **17**(2), 342-348.
- [64] Gao, F., Guan, X., Cao, X.-R., Papalexopoulos, A. (2000) Forecasting power market clearing price and quantity using a neural network method, In *Proceedings of the Power Engineering Summer Meeting*, Seattle, July 2000.
- [65] Geweke, J., Porter-Hudak, S. (1983) The estimation and application of long-memory time series models, *Journal of Time Series Analysis*, **4**, 221-228.
- [66] Giraitis, L., Leipus, R. (1995) A Generalized Fractionally Differencing Approach in Long Memory Modelling, *Lithuanian Mathematical Journal*, **35**, 65-81.
- [67] Giraitis, L., Robinson, P.M. (2001) Whittle Estimation of ARCH models, *Econotric Theory*, **17**, 608-631.

- [68] Granger, C.W.J. (1969) Investigating Causal relations by Econometric Models and Cross-Spectral Methods, *Econometrica*, **37**, 434-438.
- [69] Granger, C.W.J. (1980) Long Memory Relationships and the Aggregation of Dynamic models, *Journal of Econometrics*, **14**, 227-238.
- [70] Gray, H.L., Zhang, N.-F., Woodward, W.A. (1989) On generalized fractional processes, *Journal of Time Series Analysis*, **10**, 233-257.
- [71] Guégan, D. (1994) Séries chronologiques non linéaires à temps discret, Economica, Paris.
- [72] Guégan, D. (2000) A New Model : The k -factor GIGARCH Process. *Journal of Signal Processing*, **4**, 265-271.
- [73] Guégan, D. (2003) A Prospective Study of the k -factor Gegenbauer Process with Heteroscedastic Errors and an Application to Inflation Rates, *Finance India* **17**(1), 165-197.
- [74] Guégan, D. (2004) How can we define the concept of long memory ? An econometric survey, *Econometric Society*, 1-42.
- [75] Hannan, E.J. (1970) Multiple time series, New York, Wiley.
- [76] Heyde, C.C., Gay, H.L. (1993) Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence, *Stochastic Processes and their Applications*, **45**, 169-182.
- [77] Hosking, J.R.M. (1981) Fractional differencing, *Biometrika*, **68**(1), 165-176.
- [78] Hosoya, Y. (1996) The quasi-likelihood approach to statistical inference on multiple time-series with long-range dependence, *Journal of Econometrics*, **73**, 217-236.
- [79] Hosoya, Y. (1997) A Limit Theory of Long Range Dependence and Statistical Inference in Related Model, *Annals of Statistics*, **25**, 105-137.
- [80] Huang, R., Chongqing, K., Qing, X. (2000) System marginal price forecasting in electricity market, *Automation of electric Power System*, **24**(25), 9-12.

- [81] Kim, C.-I., Yu, I.-K., Song, Y.H. (2002) Prediction of system marginal price of electricity using wavelet transform analysis, *Energy Conversion and Management*, **43**, 1839-1851.
- [82] Knittel, C.R., Roberts, M. (2001) An empirical examination of deregulated electricity prices, *Power Working Paper*, PWP-087.
- [83] Koreneff, G., Seppala, A., Lehtonen, M., Kekkonen, V., Laitinen, E., Hakli, J., Antila, E. (1998) Electricity spot price forecasting as a part of energy management in de-regulated power market, *Proceedings of EMPD*, **1**, 223-228.
- [84] Li, C.W., Li, W.K. (1996) On a double-threshold autoregressive with heteroskedasticity times series model, *Journal of Applied Econometrics*, **11**, 253-274.
- [85] Ling, S., Li, W. K. (1997) On Fractionally Integrated Autoregressive Moving-Average Time Series Models with Conditional Heteroscedasticity, *Journal of the American Statistical Association*, **92**(439), 1184-1194.
- [86] Lobato, I.N. (1999) A semiparametric two-step estimator in a multivariate long memory model, *Journal of Econometrics*, **90**, 129-153.
- [87] Madala, G.S., Trost, R.P., Li, H., Joutz, F. (1997) Estimation of Short-run and Long-run Elasticities of Energy Demand from Panel Data using Shrinkage Estimators, *Journal of Business and Economic Statistics*, **15**, 90-100.
- [88] Magnus, W., Oberhettinger, F., Soni, R.P. (1966) *Formulas and Theorems for the special functions of mathematical physics*, 3rd edn. Springer-Verlag, New York.
- [89] Mandelbrot, B.B., Van Ness, J.W. (1968) Fractional Brownian Motions, Fractional Noises and Applications, *SIAM Review*, **10**, 422-437.
- [90] Miller, R.G., (1974) The jackknife - a review, *Biometrika*, **61**, 1-15.
- [91] Mount, T. (2001) Market Power and Price Volatility in Restructured Markets for Electricity, *Decision Support Systems* **30**, 311-325.
- [92] Ni, E., Luh, P.B. (2001) Forecasting power market clearing price and its discrete PDF using a Bayesian-based classification method, *IEEE Power Engineering Society Winter Meeting*, **3**, 1518-1523.

- [93] Nicolaisen, J.D., Richter, C.W., Sheblé, G.B. (2000) Price signal analysis for competitive electric generation companies, *Proc. of the Conference on Electric Utility Deregulation and Restructuring and Power Technologies 2000, City University, London, U.K.*, 66-71.
- [94] Nielsen, M.O. (2004) Efficient Inference in Multivariate Fractionally Integrated Time Series Models, *Econometrics Journal*, **7**(1), 63-97.
- [95] Noakes, D.J., Hipel, K.W., Mcleod, A.I., Jimenez, C., Yakowitz, C. (1988) Forecasting annual geophysical time series, *International Journal of Forecasting*, **4**, 103-115.
- [96] Nogales, F.J., Conteras, J., Conejo, A.J., Espínola, R. (2002) Forecasting next-day electricity prices by time series models, *IEEE Trans. Power Systems*, **17**(2), 342-348.
- [97] Olsson, M., Söder, L. (2004) Generation of regulating power price scenarios, *IEEE Proceedings of the 8th International Conference on Probabilistic Methods Applied to Power Systems, Iowa State University, Ames, Iowa*, 1-7.
- [98] Palma, W., Chan, N.H. (1997) Estimation and forecasting of long-memory processes with missing values, *Journal of forecasting*, **16**, 395-410.
- [99] Porter-Hudak, S. (1990) An application to the seasonal fractionally differenced model to the monetary aggregates, *Journal of the American Statistical Association*, **85**(410), 338-344.
- [100] Priestley, M.B. (1981) Spectral Analysis of Time Series, *Academic Press*, New York.
- [101] Rainville, E.D. (1960) Special Functions, *Mac Millan*, New York.
- [102] Ramsay, B., Wang, A.J. (1998) An electricity spot price estimator with particular reference to weekends and public holidays, *Neurocomputing*, **23**(1-3), 47-57.
- [103] Ravishanker, N., Ray, B.K. (2002) Bayesian prediction for vector ARFIMA process, *International Journal of Forecasting*, **18**, 204-214.
- [104] Ray, B.K. (1993a) Modeling long memory processes for optimal long -range prediction, *Journal of Times Series Analysis*, **14**, 511-526.

- [105] Ray, B.K. (1993b) Long-range forecasting of IBM product revenues using a seasonal fractionally differenced ARMA model, *Journal of International Forecasting*, **9**, 255-269.
- [106] Reeds, J. (1978) Jackknifing maximum likelihood estimates, *Annals of Statistics*, **6**(4), 727-739.
- [107] Robinson, P.M. (1995) Gaussian semiparametric estimation of long range dependence, *Annals of Statistics*, **23**, 1630-1661.
- [108] Sansom, D.C., Saha, T.K. (1999) Neural networks for forecasting electricity pool price in a deregulated electricity supply industry, *AUPEC/EECON'99*, Darwin, Australia.
- [109] Schwarz, P.M., Taylor, T.N., Birmingham, M., Dardan, S.L. (2002) Industrial Response to Electricity Real-time prices : Short-run and Long-run, *Economic Inquiry*, **40**, 597-610.
- [110] Sethuraman, S., Basawa, I.V. (1997) The asymptotic distribution of the maximum likelihood estimator for a vector time series model with long memory dependence, *Statistics and Probability Letters*, **31**(4), 285-293.
- [111] Skantze, P., Gubina, A., Ilic, M. (2000) Bid-based stochastic model for electricity prices : the impact of fundamental drivers on market dynamics, *Technical Report Energy Laboratory*, Massachusetts Institute of Technology, URL. <http://lfee.mit.edu/publications/PDF/el00.004.pdf>.
- [112] Smith, J., Yadav, S. (1994) Forecasting cost incurred from unit differencing fractionally integrated processes, *International Journal of Forecasting*, **10**, 507-514.
- [113] Sowell, F. (1987) Maximum likelihood estimation of fractionally integrated time series models, *Discussion Paper 87-07*, Department of Economics, Duke University.
- [114] Sowell, F (1989) A decomposition of block toeplitz matrices with applications to vector time series, *Discussion paper*, **GSIA**, Carnegie Mellon University.
- [115] Sowell, F. (1992) Maximum likelihood estimation of stationary univariate fractionally integrated time series models, *Journal of Econometrics*, **53**, 165-188.

- [116] Stout, W. F. (1974) Almost Sure Convergence, New York : Academic press.
- [117] Sutcliffe, A. (1994) Time series forecasting using fractional differencing, *Journal of Forecasting*, **13**, 383-393.
- [118] Szegő, G. (1959) *Orthogonal polynomials*, AMS Colloquium Publications, vol. 23, Amer. Math. Soc. : New York.
- [119] Szkuta, B.R., Sanabria, .A., Dillon, T.S. (1999) Electricity price short term forecasting using Artificial Neural Networks, *IEEE Trans. Power Systems*, **14**(3), 851-857.
- [120] Tipping, J.P., Read, E.G., McNickle, D.C. (2004) The incorporation of hydro storage into a spot price model for the New Zealand, *Working Paper*, University of Canterbury.
- [121] Valenzuela, J., Mazumdar, M. (2001) On the computation of the Probability Distribution of the spot market price in a Deregulated electricity market, *Proc. of the 22nd International Conference on Power Industry Computer Applications, PICA '2001 Innovative Computing for Power - Electric Energy Meets the Market*, 268-271.
- [122] Vehviläinen, I., Pyykkönen, (2005) Stochastic factor model for electricity spot price - the case of the Nordic market, *Energy Economics*, in press.
- [123] Velasco, C., Robinson, P.M. (2000) Whittle pseudo-maximum likelihood estimation for nonstationary time series, *Journal of American Statistical Association*, **95**(452), 1229-1243.
- [124] Wang, A., Ramsay, B. (1997) Prediction of system marginal price in the U.K. power pool using neural networks, *IEEE Proceedings of the 1997 International Conference on Neural Networks*, **4**, 2116-2120.
- [125] Weiss, A. (1986) Asymptotic Theory for ARCH Models : Estimation and Testing, *Econometric Theory*, **2**, 107-131.
- [126] Whitcher, B. (2001) Simulating gaussian stationary processes with unbounded spectra, *Journal of Computational and Graphical Statistics*, **10**, 112-134.
- [127] Whittle, P. (1951) Hypothesis testing in time series analysis, Hafner, New York.

- [128] Woodward, W.A., Cheng, Q.C., Gray, H.L (1998) A k -factor GARMA long-memory model, *Journal of Time Series Analysis*, **19**(5), 485-504.
- [129] Yajima, Y. (1996) Estimation of the frequency of unbounded spectral densities, *Proceedings of the Business and Econ. Statist. Section of ASA*, 4-7.
- [130] Yajima, Y. (1995) Estimation of the Frequency of unbounded Spectral Densities, *Discussion Paper Series, Faculty of Economics, University of Tokyo*.
- [131] Ying, Y.H., Hsino, C.-Y. (2001) Locational marginal price forecasting in deregulated electric markets using a recurrent neural network, In *IEEE Power Engineering Society Winter Meeting*, **2**, 539-544.
- [132] Zhang, P., Guan, X. (2000) Fuzzy modeling for electrical market price forecasting, *Proceedings of the 3rd World Congress on Intelligent Control and Automation*, **3**, 2262-2266.

Annexes

Dans cette partie, nous présentons deux annexes. Dans l'annexe A, nous avons les résultats des simulations de Monte Carlo permettant de comparer les méthodes d'estimation des paramètres de processus GI-GARCH à k facteurs. Dans l'annexe B les résultats des RMSE heure par heure de prévision sont donnés. Ces RMSE ont été calculés sur toute l'année 2004 et pour un horizon d'un mois. Des figures B.1 à B.4, (respectivement figures B.5 à B.8), nous avons les RMSE des 24h du premier jour du mois jusqu'à ceux des 24h du dernier jour du mois des prévisions de prix spot de l'électricité du marché français (respectivement du marché allemand).

Annexe A

Comparaison des méthodes d'estimation

Vraie valeur		CSS			Whittle			
d	a_1	\hat{d}	\hat{a}_0	\hat{a}_1	\hat{d}	\hat{a}_0	\hat{a}_1	
0.25	0.6	0.4	0.253	0.661	0.339	0.206	0.674	0.326
			(0.003)	(0.061)	(0.061)	(0.044)	(0.074)	(0.074)
			[0.017]	[0.069]	[0.069]	[0.049]	[0.136]	[0.136]
0.3	0.2	0.3	0.290	0.168	0.411	0.261	0.213	0.253
			(0.010)	(0.032)	(0.111)	(0.039)	(0.013)	(0.047)
			[0.035]	[0.032]	[0.112]	[0.048]	[0.032]	[0.112]
0.35	0.8	0.5	0.286	0.866	0.459	0.323	0.994	0.378
			(0.064)	(0.066)	(0.042)	(0.027)	(0.194)	(0.122)
			[0.021]	[0.079]	[0.049]	[0.057]	[0.273]	[0.171]

TAB. A.1 – Paramètres estimés pour un modèle GIGARCH Gaussien centré réduit défini par (2.40)-(2.42) ($n = 300$).

Vraie valeur		CSS			Whittle			
d	a_0	a_1	\hat{d}	\hat{a}_0	\hat{a}_1	\hat{d}	\hat{a}_0	\hat{a}_1
0.25	0.6	0.4	0.274 (0.024)	0.418 (0.182)	0.582 (0.182)	0.203 (0.047)	0.644 (0.044)	0.356 (0.044)
			[0.028]	[0.016]	[0.016]	[0.082]	[0.158]	[0.158]
0.3	0.2	0.3	0.295 (0.005)	0.163 (0.037)	0.429 (0.129)	0.26 (0.040)	0.203 (0.003)	0.288 (0.012)
			[5.88 10 ⁻⁴]	[0.001]	[0.005]	[0.066]	[0.041]	[0.144]
0.35	0.8	0.5	0.275 (0.075)	0.707 (0.093)	0.558 (0.058)	0.318 (0.032)	0.998 (0.198)	0.377 (0.123)
			[0.043]	[0.028]	[0.017]	[0.092]	[0.323]	[0.202]

TAB. A.2 – Paramètres estimés pour un modèle GIGARCH Student à 5 degrés de liberté défini par (2.40)-(2.42) ($n = 300$).

Vraie valeur		CSS			Whittle			
d	a_0	a_1	\hat{d}	\hat{a}_0	\hat{a}_1	\hat{d}	\hat{a}_0	\hat{a}_1
0.25	0.6	0.4	0.311 (0.061) [0.027]	0.708 (0.108) [0.089]	0.292 (0.108) [0.090]	0.296 (0.046) [0.042]	0.64 (0.04) [0.119]	0.36 (0.04) [0.119]
0.3	0.2	0.3	0.257 (0.044) [0.050]	0.149 (0.051) [0.022]	0.477 (0.177) [0.077]	0.356 (0.056) [0.035]	0.174 (0.026) [0.027]	0.39 (0.09) [0.096]
0.35	0.8	0.5	0.364 (0.014) [0.035]	0.782 (0.018) [0.130]	0.511 (0.011) [0.081]	0.418 (0.068) [0.045]	0.912 (0.112) [0.241]	0.43 (0.07) [0.151]

TAB. A.3 – Paramètres estimés pour un modèle GIGARCH Gaussien centré réduit défini par (2.40)-(2.42) ($n = 500$).

Vraie valeur		CSS		Whittle			
d	a_1	\hat{d}	\hat{a}_0	\hat{a}_1	\hat{d}	\hat{a}_0	\hat{a}_1
0.25	0.6	0.4	0.239 (0.011) [0.007]	0.474 (0.126) [0.019]	0.526 (0.126) [0.019]	0.297 (0.047) [0.068]	0.639 (0.039) [0.146]
0.3	0.2	0.3	0.338 (0.038) [0.050]	0.139 (0.061) [0.012]	0.514 (0.214) [0.043]	0.359 (0.059) [0.063]	0.310 (0.010) [0.136]
0.35	0.8	0.5	0.326 (0.024) [0.012]	0.672 (0.128) [0.004]	0.58 (0.080) [0.002]	0.422 (0.072) [0.092]	0.398 (0.102) [0.305]

TAB. A.4 – Paramètres estimés pour un modèle GIGARCH Student à 5 degrés de liberté défini par (2.40)-(2.42) ($n = 500$).

Vraie valeur		CSS			Whittle			
d	a_0	a_1	\hat{d}	\hat{a}_0	\hat{a}_1	\hat{d}	\hat{a}_0	\hat{a}_1
0.25	0.6	0.4	0.315	0.602	0.398	0.297	0.643	0.357
			(0.065)	(0.002)	(0.002)	(0.047)	(0.043)	(0.043)
0.3	0.2	0.3	[0.025]	[0.069]	[0.069]	[0.028]	[0.097]	[0.97]
			0.243	0.170	0.405	0.355	0.204	0.285
0.35	0.8	0.5	(0.057)	(0.030)	(0.105)	(0.055)	(0.004)	(0.015)
			[0.030]	[0.014]	[0.049]	[0.024]	[0.024]	[0.084]
			0.310	0.986	0.384	0.416	0.937	0.415
			(0.040)	(0.186)	(0.116)	(0.066)	(0.137)	(0.085)
			[0.022]	[0.208]	[0.1302]	[0.034]	[0.202]	[0.127]

TAB. A.5 – Paramètres estimés pour un modèle GIGARCH Gaussien centré réduit défini par (2.40)-(2.42) ($n = 1000$).

Vraie valeur		CSS			Whittle			
d	a_0	a_1	\hat{d}	\hat{a}_0	\hat{a}_1	\hat{d}	\hat{a}_0	\hat{a}_1
0.25	0.6	0.4	0.255 (0.005)	0.460 (0.140)	0.540 (0.140)	0.295 (0.045)	0.619 (0.018)	0.382 (0.018)
			[0.031]	[0.018]	[0.018]	[0.061]	[0.143]	[0.143]
0.3	0.2	0.3	0.298 (0.002)	0.143 (0.057)	0.499 (0.199)	0.356 (0.056)	0.193 (0.007)	0.324 (0.024)
			[0.065]	[0.027]	[0.095]	[0.047]	[0.037]	[0.129]
0.35	0.8	0.5	0.311 (0.039)	0.673 (0.127)	0.579 (0.079)	0.419 (0.069)	0.942 (0.142)	0.411 (0.089)
			[0.013]	[0.005]	[0.008]	[0.079]	[0.256]	[0.160]

TAB. A.6 – Paramètres estimés pour un modèle GIGARCH Student à 5 degrés de liberté défini par (2.40)-(2.42) ($n = 1000$).

Annexe B

RMSE heure par heure

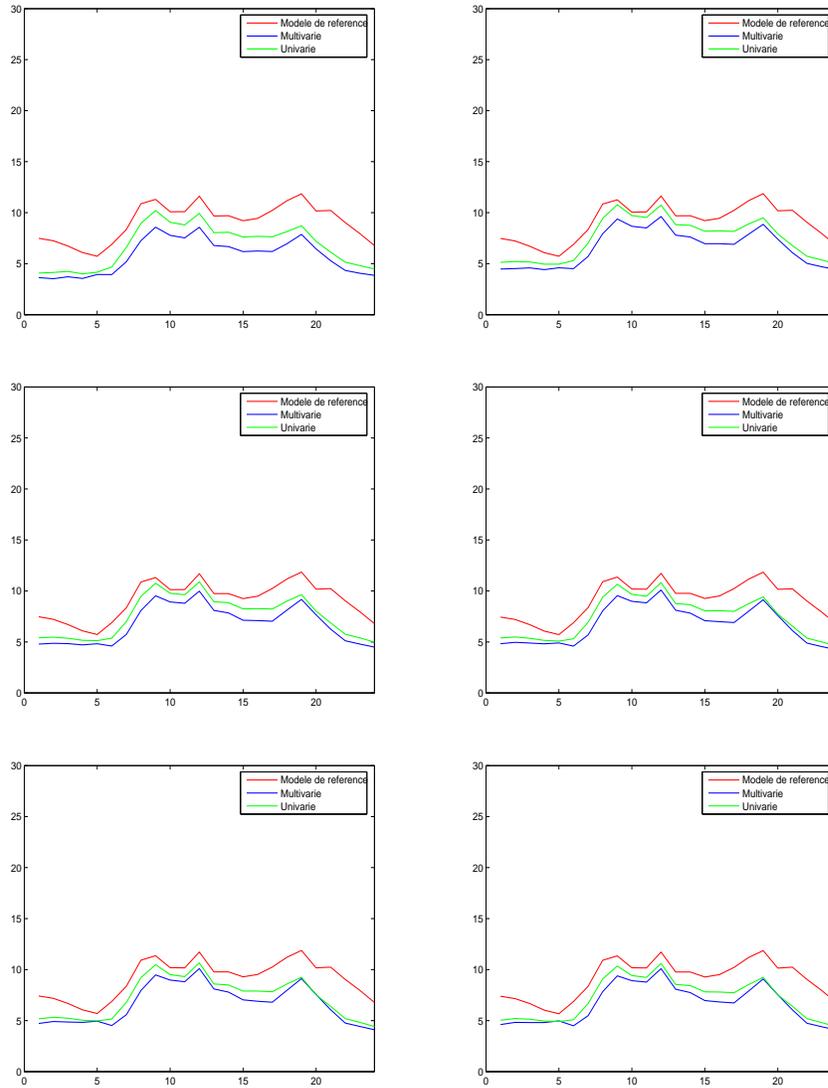


FIG. B.1 – RMSE heure par heure des prix spot français pour un horizon de prévision d'un mois.

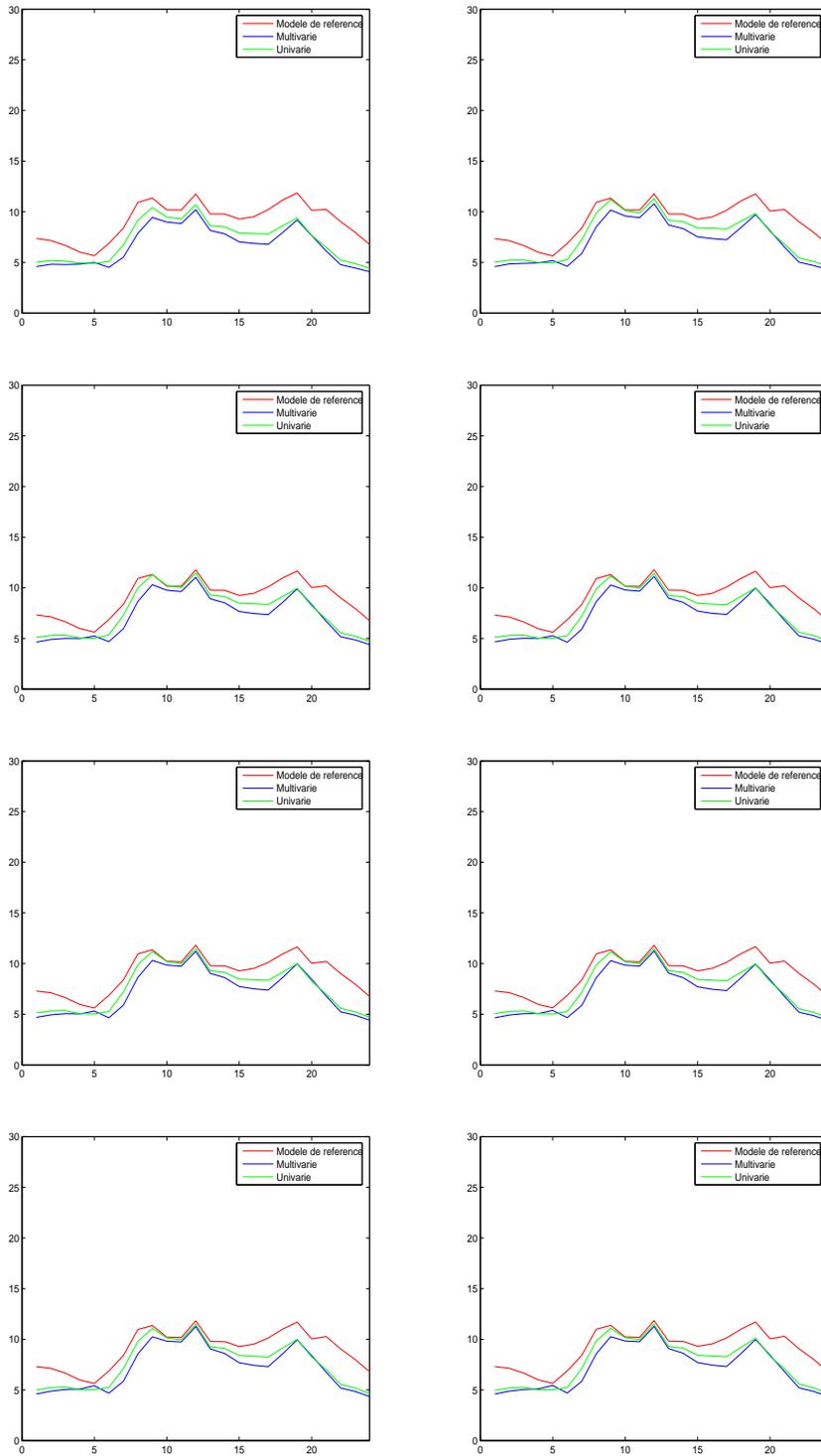
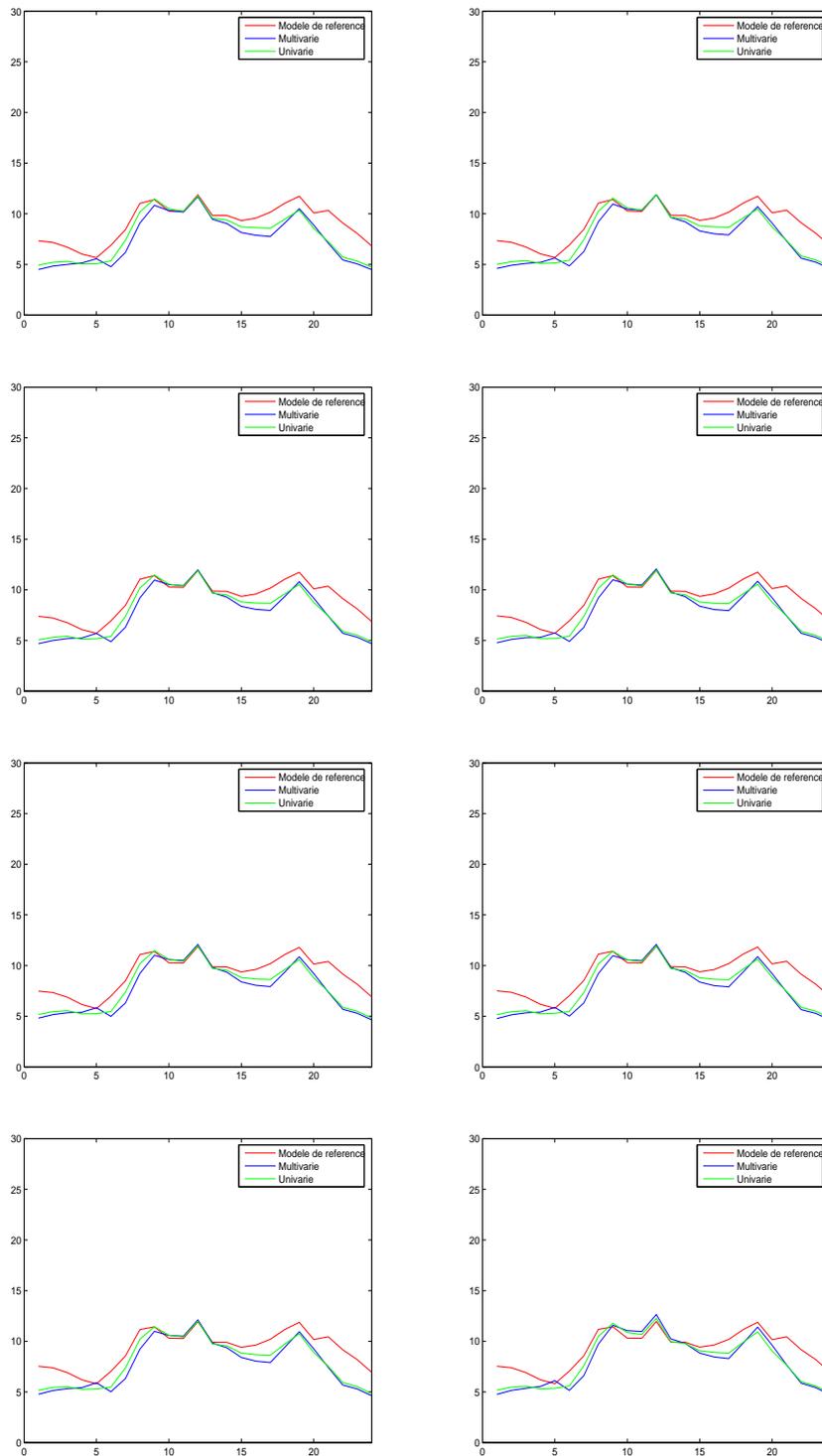


FIG. B.2 – RMSE heure par heure des prix spot français pour une prévision en $J + 30$.

FIG. B.3 – RMSE heure par heure des prix spot français pour une prévision en $J + 30$.

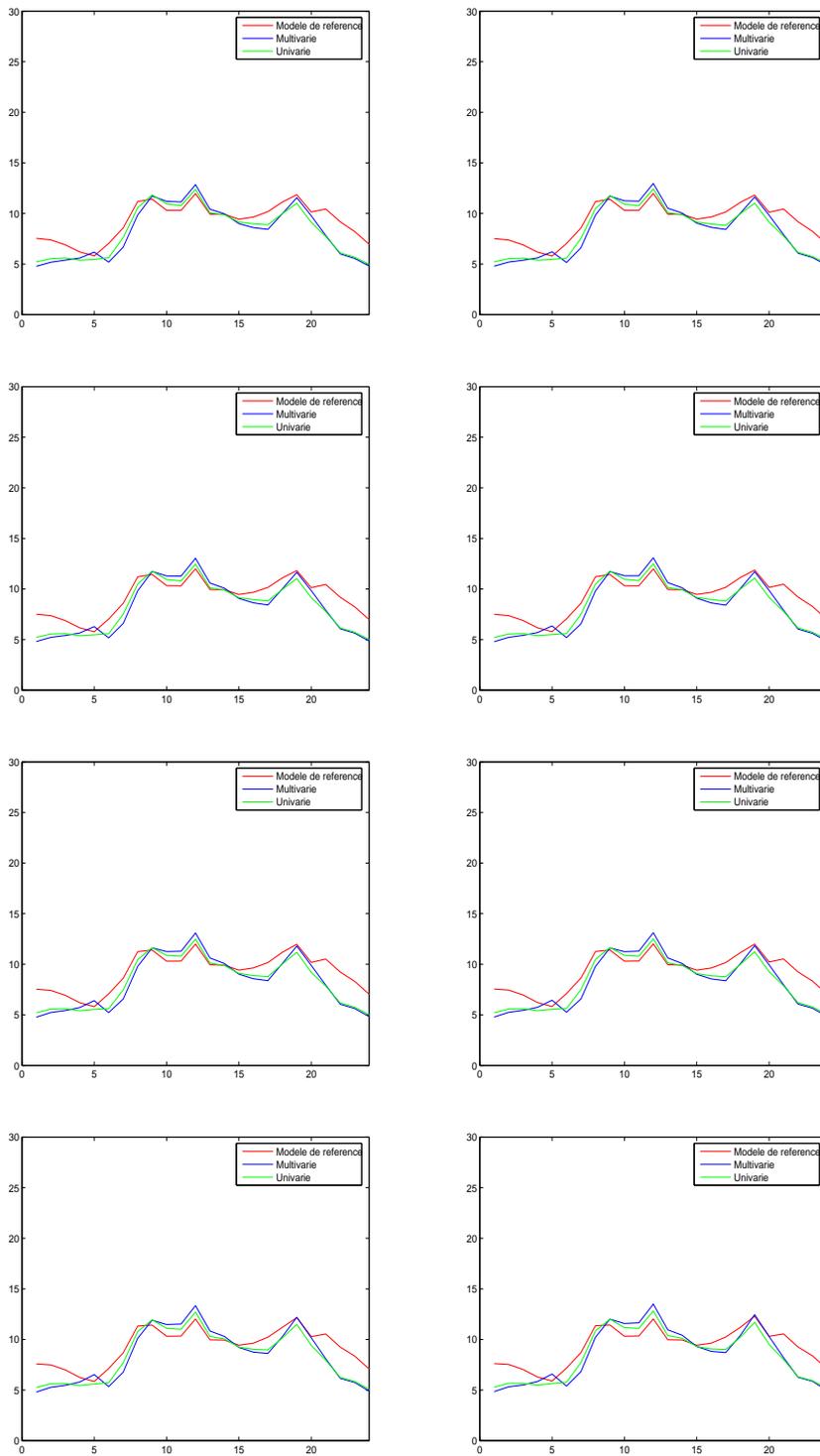
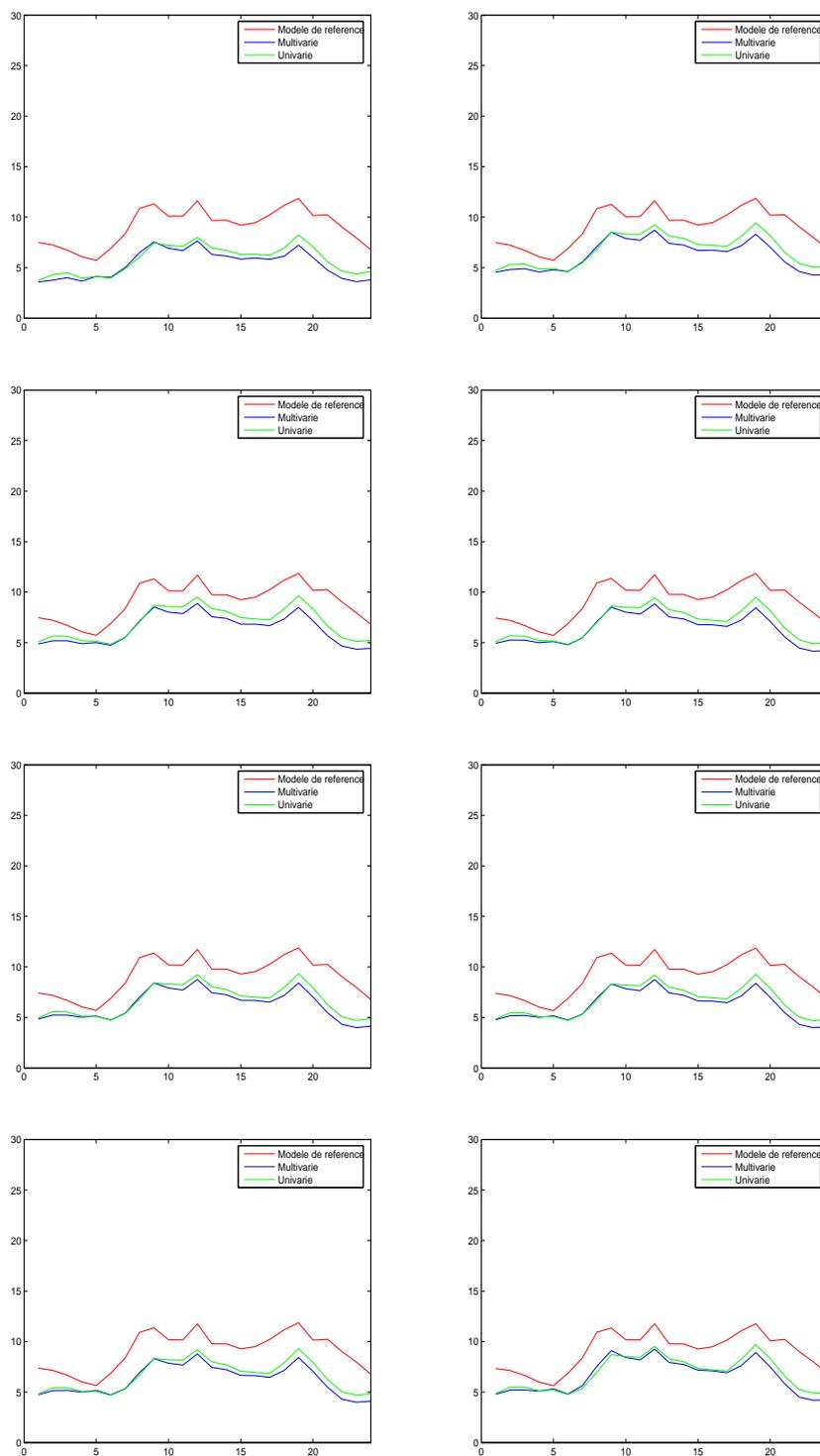


FIG. B.4 – RMSE heure par heure des prix spot français pour une prévision en $J + 30$.

FIG. B.5 – RMSE heure par heure des prix spot allemand pour une prévision en $J + 30$.

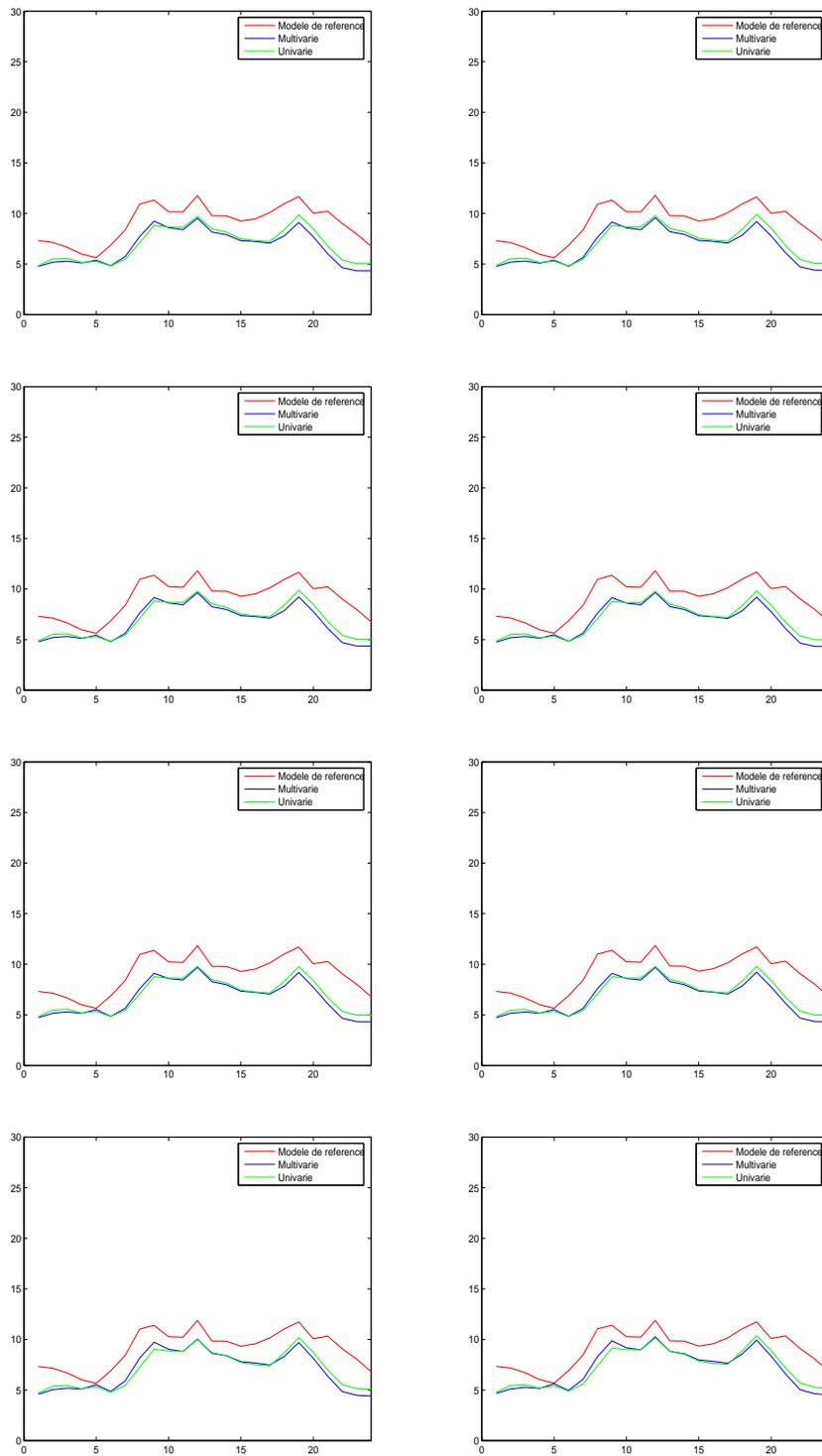
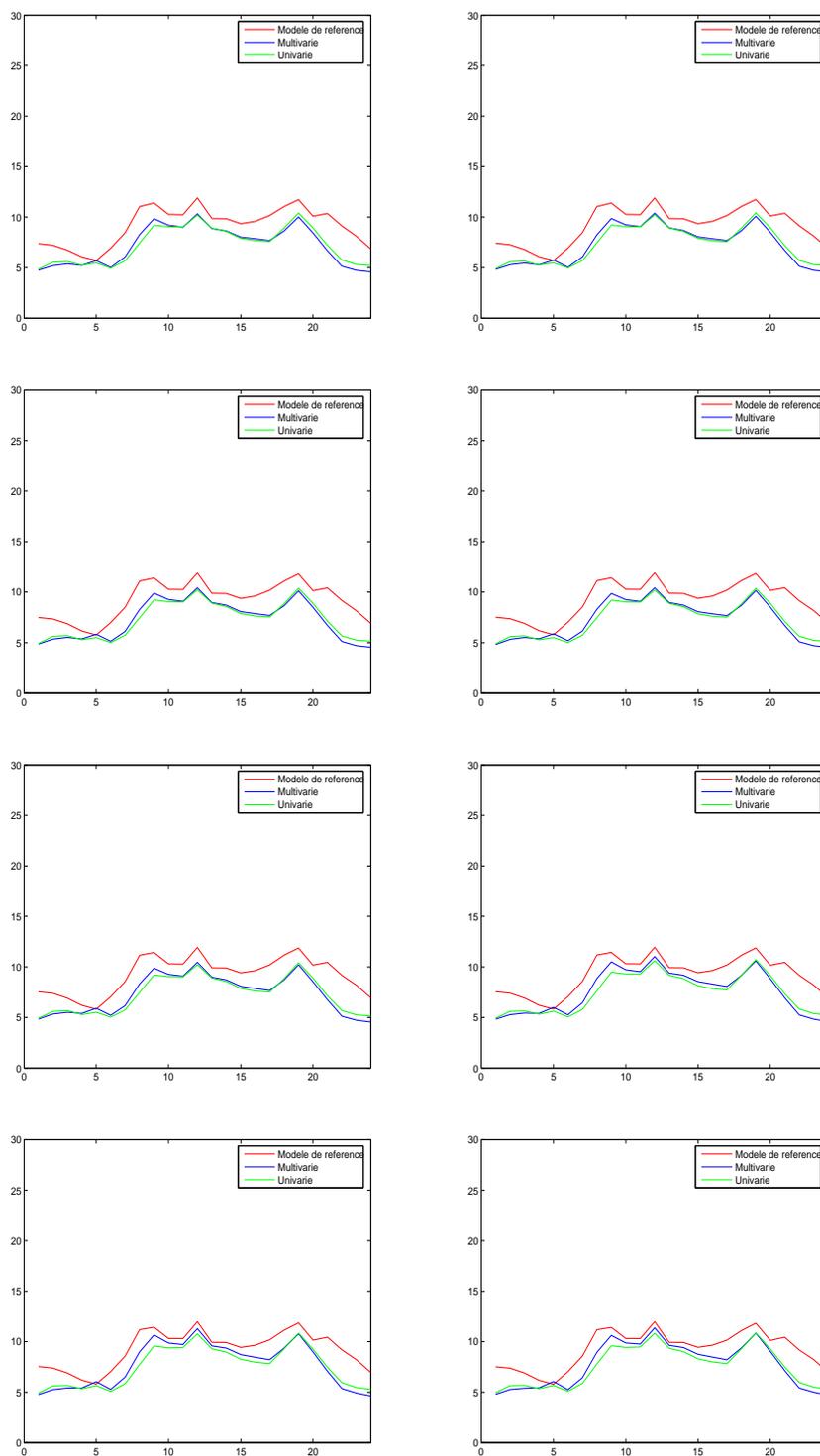


FIG. B.6 – RMSE heure par heure des prix spot allemand pour une prévision en $J + 30$.

FIG. B.7 – RMSE heure par heure des prix spot allemand pour une prévision en $J + 30$.

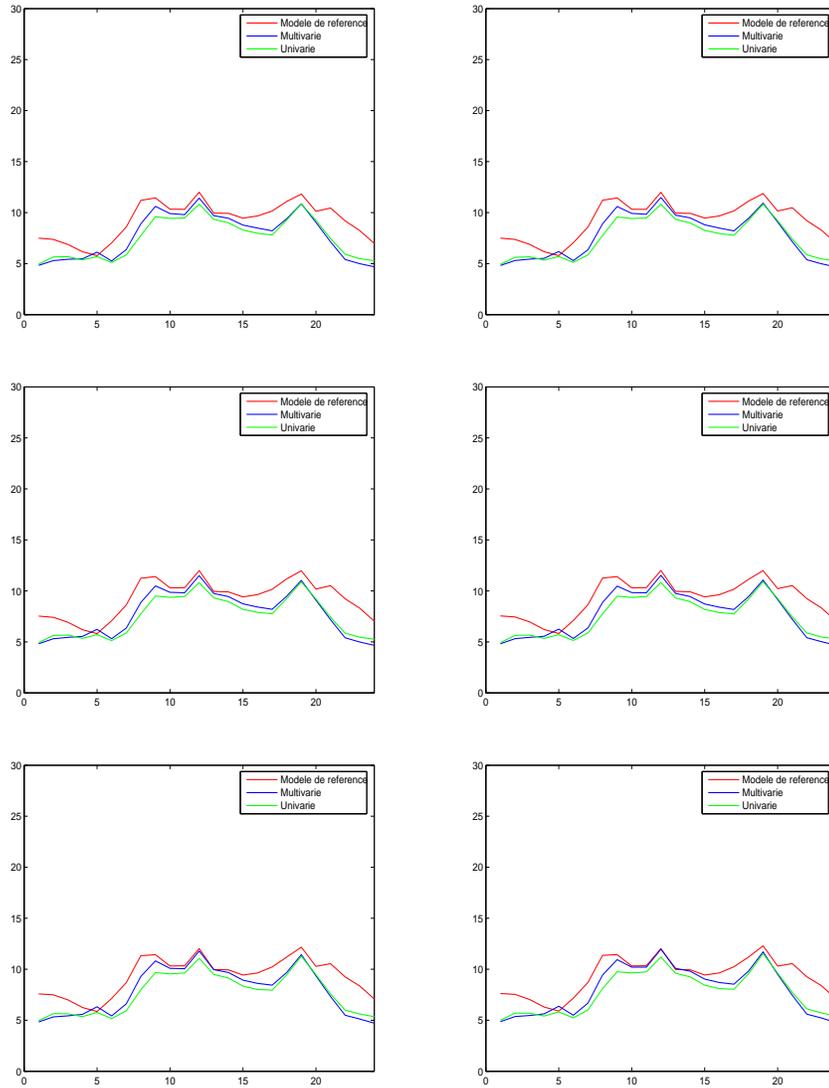


FIG. B.8 – RMSE heure par heure des prix spot allemand pour une prévision en $J + 30$.

Résumé

Plusieurs données de marchés financiers, telles que les prix spot de marchés européens de l'électricité interconnectés, présentent de la longue mémoire, au sens de la décroissance hyperbolique des autocorrélations combinée avec un phénomène d'hétéroskédasticité et de cycles périodiques ou non. Pour modéliser de tels comportements, nous introduisons d'une part les processus GIGARCH à k facteurs et nous proposons deux méthodes d'estimation des paramètres. Nous développons les propriétés asymptotiques des estimateurs de chacune des méthodes. De plus, afin de comparer les propriétés asymptotiques des estimateurs, des simulations de Monté Carlo sont effectuées. D'autre part, nous proposons un modèle longue mémoire généralisé multivarié (MVGARMA à k facteurs) pour modéliser conjointement deux marchés européens de l'électricité interconnectés. Nous donnons une procédure pratique d'estimation des paramètres. Pour la prévision, nous fournissons les expressions analytiques des prédicteurs de moindres carrés pour les modèles proposés et les intervalles de confiance des erreurs de prévision. Enfin, nous appliquons ces deux modèles sur les prix spot de l'électricité des marchés français et allemand et nous comparons leurs capacités prédictives.

Mots-clés : Prix spot de l'électricité ; Mémoire longue ; Hétéroskédasticité ; Prédicteur de moindres carrés ; Multivarié.

Titre en anglais

Multivariate Long Memory Processes : applications to the EDF producer problematic in the context of the european electricity market liberalization

Abstract

Certain crucial financial time series, such as the interconnected european electricity market spot prices, exhibit long-memory, in the sense of slowly-decaying correlations combined with heteroskedasticity and periodic or none cycles. In modeling such behavior, we consider on one hand, the k -factor GIGARCH process and additionally propose two methods to address the related parameter estimation problem. In each method, we explore the asymptotic theory for estimation. Moreover, the asymptotic properties are validated and compared via Monte Carlo simulations. On the other hand, we introduce a new multivariate long-memory generalized model (k -factor MVGARMA) in order to model interconnected european electricity market spot prices. We suggest a practical framework to address the parameter estimation problem. We investigate the analytical expressions of the least squares predictors for the two proposed models and their confidence intervals. To finish, we apply the two proposed models to the french and german electricity market spot prices and a comparison is made between their forecasting abilities.

Keywords : Electricity spot prices ; Long-memory ; Heteroskedasticity ; Least squares predictor ; Multivariate.

Adresses

ENS Cachan IDHE-MORA, UMR CNRS 8533, 61, avenue du président Wilson 94231 Cachan cedex ou EDF R&D, 1 avenue du général de gaulle 92141 Clamart cedex.
E.mail : diongue@idhe.ens-cachan.fr