

2 m 11. 3145. 6

v. 003
11497 352

Université de Montréal

Tests d'adéquation en séries chronologiques

par

Lamya Souktani

Département de mathématiques et de statistique

Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures

en vue de l'obtention du grade de

Maître ès sciences (M.Sc.)
en Statistique

décembre 2003



QA

3

U54

2004

V.003

Direction des bibliothèques

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

Tests d'adéquation en séries chronologiques

présenté par

Lamya Souktani

a été évalué par un jury composé des personnes suivantes :

Martin Bilodeau

(président-rapporteur)

Roch Roy

(directeur de recherche)

Louis G. Doray

(membre du jury)

Mémoire accepté le:

11 février 2004

REMERCIEMENTS

Je tiens à exprimer ma reconnaissance envers les personnes qui m'ont aidé à réaliser ce travail.

La première personne que je tiens à remercier est mon directeur de recherche M. Roch Roy pour ses conseils, sa patience et son support financier. Je veux lui témoigner ici toute mon estime.

J'ai bénéficié d'un soutien informatique de la part de Claude Beauchamp et Zahia Salmi. Je les remercie pour l'aide qu'ils m'ont apportée.

Ce travail de recherche a bénéficié d'une bourse octroyée par la Faculté des études supérieures et d'un soutien financier de la part du Conseil de Recherche en Sciences Naturelles et en Génie du Canada (CRNSG) et du réseau de centres d'excellences sur les mathématiques des technologies de l'information et des systèmes complexes (MATISC), par le biais de subventions à mon directeur de recherche.

Ce travail aurait été certainement pénible si je n'avais pas bénéficié du soutien de ma famille. Je remercie ma fille Oumaima pour sa patience et mon mari Nouredine pour son énorme soutien.

SOMMAIRE

Les tests usuels utilisés pour la validation d'un modèle estimé sont ceux de Box et Pierce (1970) et de Ljung et Box (1976) connus sous le nom de tests *portemanteau*. Ces derniers, qui sont basés sur le carré des autocorrélations échantillonales des résidus, permettent de détecter une dépendance, si elle existe, dans la série des résidus. Cependant, dans certaines situations, plusieurs autocorrélations successives sont petites et de même signe, ce qui porte à douter de l'adéquation du modèle estimé. Les tests de Box et Pierce et de Ljung et Box ne peuvent pas prendre en compte cette particularité et conduisent dans la plupart du temps au non rejet du modèle estimé. Un test asymptotique a été développé par Koch et Yang (1986) pour tester l'indépendance de deux séries chronologiques qui permet de prendre en compte des comportements particuliers des corrélations croisées. Dans ce mémoire, nous reprenons l'idée de ces auteurs mais dans le contexte des tests d'ajustement. Ainsi, nous utilisons ce test pour détecter des patrons particuliers dans des autocorrélations résiduelles successives et nous le comparons au test spectral de Hong (1996a). Nous décrivons les principales propriétés du test de Koch et Yang modifié et de celui de Hong et nous étudions leurs niveaux et puissance à l'aide de simulations de Monte Carlo. Finalement, nous appliquons ces deux tests à deux séries réelles.

Mots clés : Processus linéaires, modèles *ARMA*, tests d'ajustement, méthode de Monte Carlo.

SUMMARY

The usual tests used for the validation of fitted models are those of Box and Pierce (1970) and Ljung and Box (1976) known under the name of *portmanteau* tests. These tests use the sample square residual autocorrelations to detect the existence of dependence in a series of residuals. However, in certain situations, several successive autocorrelations are small and have the same sign, bringing into doubt the adequacy of the fitted model. The tests of Box and Pierce and Ljung and Box cannot take into account this characteristic and lead most of the time to non rejection of the fitted model. An asymptotic test was developed by Koch and Yang (1986) to test independence of two time series that takes into account particular patterns in the cross-correlation function. In this Master thesis, we take up the idea of these authors, but in the context of goodness-of-fit tests. We will use these tests to detect particular patterns in successive residual autocorrelations and will compare it with the spectral test of Hong (1996a). We describe the principal properties of these two tests and study their levels and powers using Monte Carlo simulations. Finally, we apply the two tests to two real series.

Keywords : Linear process, *ARMA* models, goodness-of-fit tests, Monte Carlo methods.

TABLE DES MATIÈRES

| | |
|---|------|
| Remerciements | i |
| Sommaire | ii |
| Summary | iii |
| Liste des figures | vi |
| Liste des tableaux | viii |
| Introduction | 1 |
| Chapitre 1. Préliminaires | 5 |
| 1.1. Processus stationnaires..... | 5 |
| 1.2. Estimation des paramètres d'un processus stationnaire..... | 8 |
| 1.2.1. Estimation de la moyenne et des autocovariances..... | 8 |
| 1.2.2. Estimation des autocorrélations..... | 9 |
| 1.3. Processus ARMA..... | 11 |
| 1.4. Le test de Box et Pierce : Description et propriétés..... | 12 |
| Chapitre 2. Le test de Koch et Yang modifié | 22 |
| 2.1. Description du test | 22 |
| 2.2. Loi asymptotique de la statistique de Koch et Yang modifiée..... | 23 |

| | |
|---|------------|
| 2.3. Algorithme d'Imhof | 27 |
| Chapitre 3. Le test de Hong | 31 |
| 3.1. Description du test | 32 |
| 3.2. Propriétés asymptotiques | 34 |
| Chapitre 4. Simulations | 36 |
| 4.1. Algorithmes de simulations | 36 |
| 4.2. Description de la simulation | 38 |
| 4.3. Résultats de la simulation | 42 |
| 4.3.1. Discussion des résultats de l'étude de niveau | 42 |
| 4.3.2. Discussion des résultats de l'étude de puissance | 56 |
| Chapitre 5. Exemples | 64 |
| 5.1. Premier exemple : Données de méthane | 64 |
| 5.2. Deuxième exemple : Niveau du Nil (NIL) | 72 |
| Conclusion | 80 |
| Annexe | 83 |
| Algorithme des simulations | 83 |
| Code FORTRAN | 86 |
| Bibliographie | 103 |

LISTE DES FIGURES

| | | |
|-------|--|----|
| 4.3.1 | Niveaux empiriques (%) de $\tilde{Q}_{KY}(m, i)$, $i = 0, \dots, m-1$ pour (a) $m = 6$ et (b) $m = 12$, aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(1)$ avec $\phi = 0.8$ et $n = 100$. | 47 |
| 4.3.2 | Niveaux empiriques (%) de $\tilde{Q}_{KY}(m, i)$, $i = 0, \dots, m-1$ pour (a) $m = 18$ et (b) $m = 24$, aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(1)$ avec $\phi = 0.8$ et $n = 100$. | 48 |
| 4.3.3 | Niveaux empiriques (%) de $Q_H(w, T_n)$ aux niveaux nominaux (a) 1%, (b) 5% et (c) 10%. Le modèle est $AR(1)$ avec $\phi = 0.8$ et $n = 100$. | 49 |
| 4.3.4 | Niveaux empiriques (%) de $\tilde{Q}_{KY}(m, i)$, $i = 0, \dots, m-1$ pour (a) $m = 6$ et (b) $m = 12$, aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(2)$ avec $\phi_1 = 0.1$, $\phi_2 = 0.5$ et $n = 100$. | 52 |
| 4.3.5 | Niveaux empiriques (%) de $\tilde{Q}_{KY}(m, i)$, $i = 0, \dots, m-1$ pour (a) $m = 18$ et (b) $m = 24$, aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(2)$ avec $\phi_1 = 0.1$, $\phi_2 = 0.5$ et $n = 100$. | 53 |
| 4.3.6 | Niveaux empiriques (%) de $Q_H(w, T_n)$ aux niveaux nominaux (a) 1%, (b) 5% et (c) 10%. Le modèle est $AR(2)$ avec $\phi_1 = 0.1$, $\phi_2 = 0.5$ et $n = 100$. | 54 |
| 5.1.1 | Série de méthane non différenciée et son ACF. | 65 |
| 5.1.2 | Graphiques des autocorrélations et des autocorrélations partielles de la série de méthane différenciée. | 66 |

| | | |
|-------|--|----|
| 5.1.3 | Données de méthane après une différence régulière et l'ACF de la série résiduelle..... | 67 |
| 5.2.1 | Niveau du Nil non différencié et son ACF..... | 73 |
| 5.2.2 | Graphiques des autocorrélations et des autocorrélations partielles du niveau du Nil différencié..... | 74 |
| 5.2.3 | Niveau du Nil après une différence régulière et l'ACF de la série résiduelle..... | 75 |

LISTE DES TABLEAUX

| | | |
|-------|--|----|
| 4.2.1 | Équations des modèles considérés pour les simulations. | 39 |
| 4.3.1 | Niveaux empiriques (%) du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et du test de Hong $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(1)$ avec $\phi = 0.8$ et $n = 50, 100$ | 45 |
| 4.3.2 | Niveaux empiriques (%) du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et du test de Hong $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(1)$ avec $\phi = 0.8$ et $n = 200, 500$ | 46 |
| 4.3.3 | Niveaux empiriques (%) du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et du test de Hong $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(2)$ avec $\phi_1 = 0.1, \phi_2 = 0.5$ et $n = 50, 100$ | 50 |
| 4.3.4 | Niveaux empiriques (%) du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et du test de Hong $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(2)$ avec $\phi_1 = 0.1, \phi_2 = 0.5$ et $n = 200, 500$ | 51 |

- 4.3.5 Niveaux empiriques (%) du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et du test de Hong $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est un bruit blanc avec $n = 100$ 55
- 4.3.6 Puissances empiriques (%) basées sur les VCE et les VCA (entre parenthèses) de $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et de $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $MA(4)$ avec $\theta_4 = 0.8$ estimé par $AR(1)$, $n=100$ 59
- 4.3.7 Puissances empiriques (%) basées sur les VCE et les VCA (entre parenthèses) de $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et de $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $MA(4)$ avec $\theta_4 = 0.4$ estimé par $AR(1)$, $n=100$ 60
- 4.3.8 Puissances empiriques (%) basées sur les VCE et les VCA (entre parenthèses) de $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et de $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $MA(15)$ avec $\theta_{13} = 0.2$ et $\theta_{14} = 0.2$ $\theta_{15} = 0.5$ et estimé par $AR(1)$, $n=100$ 61
- 4.3.9 Puissances empiriques (%) basées sur les VCE et les VCA (entre parenthèses) de $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et de $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $ARMA(1, 4)$ avec $\phi_1 = 0.8$ et $\theta_4 = 0.8$ estimé par $AR(1)$, $n=100$ 62

| | |
|---|----|
| 4.3.10 Tests les plus puissants pour chaque modèle, $\alpha = 0.05$. BP : Box-Pierce, H(BA) : Hong avec Bartlett, H(DA) : Hong avec Daniell, H(PA) : Hong avec Parzen, H(TR) : Hong avec noyau uniforme tronqué et KY : Koch et Yang. | 63 |
| 5.1.1 Résultats du test de diagnostic du bruit blanc pour la série de méthane. | 68 |
| 5.1.2 Estimation des paramètres du modèle $AR(3)$ ajusté à la série de méthane. | 69 |
| 5.1.3 Résultats du test d'adéquation de Ljung et Box (série de méthane). . | 69 |
| 5.1.4 Résultats du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$ appliqué aux données du méthane. | 70 |
| 5.1.5 Résultats du test de Hong $Q_H(w, T_n)$ appliqué aux données du méthane. | 70 |
| 5.2.1 Résultats du test de diagnostic du bruit blanc pour le niveau du Nil différencié. | 76 |
| 5.2.2 Estimation des paramètres du modèle $AR(4)$ ajusté au niveau du Nil différencié. | 76 |
| 5.2.3 Résultats du test d'adéquation de Ljung et Box (niveau du Nil différencié). | 77 |
| 5.2.4 Résultats du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$ appliqué aux données du niveau du Nil. | 78 |

INTRODUCTION

Lors de toute analyse statistique, il est important de s'assurer de la validité du modèle que l'on choisit pour décrire les observations puisque c'est sur celui-ci que repose l'inférence. Le problème de tester l'adéquation d'un modèle statistique aux données, remonte aux origines de la statistique. Depuis, de nombreux développements sont apparus et récemment de nouvelles stratégies de tests ont commencé à émerger.

Dans le domaine des séries chronologiques, l'ouvrage de Box et Jenkins (1970, 1976) a eu un impact énorme sur l'utilisation des méthodes de séries chronologiques dans plusieurs disciplines scientifiques. Cet ouvrage a popularisé la classe des modèles autorégressifs-moyenne mobile (*ARMA*). La construction de tels modèles à partir de données réelles comporte trois grandes étapes : l'identification d'un modèle potentiel pour décrire le phénomène à l'étude, l'estimation des paramètres du modèle retenu et finalement la validation du modèle estimé par une analyse résiduelle. Avec l'élargissement du champ d'application des séries temporelles, il y a eu émergence de nouveaux créneaux de recherche, ce qui a donné naissance à de nouvelles stratégies de modélisation qui viennent compléter le processus entamé par Box et Jenkins.

Dans ce mémoire, l'attention est portée sur la phase de validation et plus précisément sur la phase des tests basés sur les autocorrélations des résidus. Cette

étape occupe une place importante dans la procédure de modélisation puisqu'elle permet, dans le cas où le modèle identifié et estimé lors des étapes antérieures est accepté, de mieux comprendre le comportement de la série étudiée et par conséquent de mieux contrôler les prévisions.

Parmi les tests usuels pour vérifier l'autocorrélation des résidus, on peut citer celui de Durbin-Watson qui est basé sur l'autocorrélation résiduelle de délai 1. Cependant, dans la modélisation de séries saisonnières il est fréquent que l'autocorrélation de délai 1 soit nulle et qu'il y ait des autocorrélations non nulles à des délais supérieurs à 1. Dans cette situation, le test de Durbin-Watson ne détecte pas la présence de corrélation résiduelle. Pour palier à cette faiblesse, Box et Jenkins (1970, 1976) ont prôné la prise en compte des autocorrélations résiduelles à plusieurs délais, typiquement pour les délais $1, 2, \dots, m$ où $m > 0$ est un entier qui correspond au délai maximum choisi. L'ensemble des m premières autocorrélations est aussi appelé « corrélogramme ». Lors de l'analyse du corrélogramme résiduel, on fait des tests à chaque délai et aussi un « test d'adéquation globale » appelé « tests portemanteaux » dans le jargon des séries chronologiques. L'analyse des autocorrélations résiduelles repose sur l'idée intuitive que si le modèle *ARMA* estimé est adéquat, les résidus devraient se comporter essentiellement comme un bruit blanc et par conséquent, les autocorrélations résiduelles devraient toutes être petites (près de 0).

Dans la pratique, les deux tests portemanteaux les plus utilisés sont ceux de Box et Pierce (1976) et de Ljung et Box (1976). Les statistiques de test pour ceux-ci sont basées sur une somme des carrés des m premières autocorrélations résiduelles, où m est le délai maximal choisi.

D'autres auteurs se sont intéressés aux tests de corrélation entre deux séries. Koch et Yang (1986), ont développé un test d'indépendance de deux séries qui est basé sur les corrélations croisées résiduelles résultant de l'estimation des vrais modèles *ARMA* pour chacune des séries observées. Ces derniers ont montré que leur test englobe celui de Haugh (1976) et qu'il est plus puissant que ce dernier dans plusieurs situations. Hong (1996b) propose une modification du test de Haugh (1976) pour des modèles autorégressifs d'ordre infini qui fournit des tests convergents. Récemment, Duschesne et Roy (2003), ont proposé une version robuste aux valeurs à l'écart du test de Hong. Dans leur ensemble, ces tests s'intéressent à étudier la corrélation sérielle entre deux séries observées.

Dans le premier chapitre de ce mémoire nous abordons avec détail les tests de Box et Pierce et de Ljung et Box. Nous introduisons d'abord la notion de processus stochastique et spécialement la classe des processus *ARMA* qui sont particulièrement utiles pour décrire le comportement des séries chronologiques univariées. Cette présentation suppose que l'on définisse au préalable un certain nombre de notions essentielles à l'analyse des séries chronologiques.

Dans le deuxième chapitre, nous proposons une adaptation du test de Koch et Yang (1986) noté $Q_{KY}(m, i)$, afin de vérifier l'adéquation d'un modèle estimé. En $i = 0$, nous retrouvons la statistique de Box-Pierce. Avec $i > 0$, la statistique $Q_{KY}(m, i)$ prend en compte le fait que des autocorrélations résiduelles consécutives qui sont petites mais de même signe est possiblement une indication que le modèle estimé n'est pas tout à fait adéquat. Nous montrons, sous l'hypothèse nulle que le modèle estimé est adéquat, que la statistique $Q_{KY}(m, i)$ est distribuée asymptotiquement comme une combinaison linéaire de variables aléatoires khi-deux indépendantes. La distribution asymptotique obtenue peut être évaluée

de façon exacte à l'aide de l'algorithme d'Imohf (1961) que nous décrivons brièvement à la section 2.3. Lorsque $i = 0$, nous retrouvons la loi asymptotique khi-deux de la statistique de Box-Pierce. A notre connaissance, c'est la première fois que l'idée de Koch et Yang est exploitée dans le contexte de tests d'adéquation et certains résultats de ce chapitre sont à notre avis originaux.

Dans le troisième chapitre, nous introduisons la statistique spectrale de Hong (1996a) afin de tester l'adéquation d'un modèle autorégressif estimé. Cette dernière est une somme pondérée de toutes les autocorrélations résiduelles existantes. Les poids sont définis par une fonction de noyau et un point de troncature. Avec le noyau uniforme tronqué, nous retrouvons une version normalisée de la statistique de Box-Pierce. La validité asymptotique de la statistique de Hong pour des *ARMA* est établie dans Hong et Lee (2002). Duschesne et Roy (2004) ont établi la validité de la statistique de Hong (1996a) pour des modèles autorégressifs multivariés avec variables exogènes (*VARX*).

Le quatrième chapitre est consacré à la discussion d'une expérience de Monte Carlo pour illustrer la performance de la version modifiée du test de Koch et Yang et de celui de Hong (1996a) appliqués à la série résiduelle d'un processus *ARMA*. La comparaison de ces tests avec celui de Box-Pierce est faite en terme de niveau et de puissance via des simulations informatiques effectuées à l'aide de programmes écrits dans le langage *FORTRAN 77*.

Finalement, à l'aide de données provenant de séries réelles, présentées dans le cinquième chapitre, nous mettrons en évidence l'utilité des tests proposés en comparaison à celui de Box-Pierce et montrerons l'intérêt de leur utilisation dans certains contextes.

Chapitre 1

PRÉLIMINAIRES

Les modèles ARMA (autorégressifs-moyenne mobile) sont souvent utilisés pour la description des séries chronologiques. Ils permettent, en outre, de prévoir la variable étudiée à un instant futur à partir des valeurs antérieures prises par cette variable.

Dans ce chapitre, nous allons introduire les modèles ARMA et présenter leurs principales propriétés.

1.1. PROCESSUS STATIONNAIRES

Le concept de stationnarité joue un rôle déterminant dans l'étude des séries chronologiques. Du point de vue théorique, pour développer des techniques puissantes de prévision, il est important que le processus étudié présente des propriétés qui ne varient pas avec le temps.

Dans cette section, nous rappelons le concept de stationnarité et nous définissons les fonctions d'autocovariance et d'autocorrélation ainsi que leurs propriétés.

Définition 1.1.1. *Un processus $\{X_t : t \in \mathbb{Z}\}$ tel que $E[X_t^2] < \infty$ est dit stationnaire au sens large (SSL) si :*

i) $E(X_t) = \mu_x$ est indépendante de t , $\forall t \in \mathbb{Z}$;

ii) $Cov(X_{t+h}, X_t) = \gamma_h$ est indépendante de t , $\forall t, h \in \mathbb{Z}$.

Dans le cas où la distribution conjointe du vecteur $(X_{1+h}, \dots, X_{n+h})'$ est la même que celle du vecteur $(X_1, \dots, X_n)'$, $\forall n \geq 1$ et $\forall h \in \mathbb{Z}$, on dit que le processus $\{X_t : t \in \mathbb{Z}\}$ est stationnaire au sens strict (SSS).

La fonction $\gamma(\cdot)$ définie par $\gamma_h = Cov(X_{t+h}, X_t)$ est la fonction d'autocovariance du processus $\{X_t, t \in \mathbb{Z}\}$ et γ_h est l'autocovariance de délai h , $h \in \mathbb{Z}$. Cette fonction est paire, c'est-à-dire que $\gamma_{-h} = \gamma_h, \forall h \in \mathbb{Z}$. Elle est aussi définie non négative (d.n.n.), c'est à dire que pour tout entier $n \geq 1$ et pour tout $a_1, a_2, \dots, a_n \in \mathbb{R}$, $\sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma_{i-j} = \mathbf{a}' \Gamma_n \mathbf{a} \geq 0$, où $\mathbf{a} = (a_1, a_2, \dots, a_n)'$ est un vecteur colonne et $\Gamma_n = [\gamma_{i-j}]_{i,j=1}^n$ est la matrice de covariance du vecteur $(X_1, \dots, X_n)'$.

L'autocorrélation de délai h est

$$\rho_h = Corr(X_{t+h}, X_t) = \frac{\gamma_h}{\gamma_0}, \quad h \in \mathbb{Z},$$

et la fonction $\rho(\cdot)$ est la fonction d'autocorrélation du processus $\{X_t, t \in \mathbb{Z}\}$. Elle est aussi paire et définie non négative.

Il faut noter que les fonctions d'autocovariance et d'autocorrélation permettent de mesurer le degré de dépendance linéaire entre les observations de la variable étudiée à des temps différents. Par conséquent, elles ont un rôle déterminant dans la prévision des valeurs futures à partir des valeurs présentes et passées.

Un processus $\{X_t : t \in \mathbb{Z}\}$ est un bruit blanc faible (*BBfaible* $(0, \sigma^2)$) si les variables X_t sont de moyenne nulle, de variance commune σ^2 et si elles sont non corrélées, c'est-à-dire si $Cov(X_{t+h}, X_t) = 0, \forall h \neq 0$.

De même, un processus $\{X_t : t \in \mathbb{Z}\}$ est un bruit blanc fort (*BBfort*) si les X_t sont des variables aléatoires indépendantes et identiquement distribuées (iid) selon la fonction de répartition $F(x) = P(X_t \leq x)$, $\forall x \in \mathbb{R}$, $\forall t \in \mathbb{Z}$. Un bruit blanc fort de moyenne zéro dont la variance existe est un bruit blanc faible.

Nous terminons cette section par la présentation des processus linéaires. Ces derniers offrent un cadre plus général pour l'étude des processus stationnaires. En effet, une grande classe de processus stationnaires au sens large sont, d'après le théorème de la décomposition de Wold (Brockwell et Davis, 2002, p. 76), des processus linéaires causals.

Définition 1.1.2. *Un processus $\{X_t : t \in \mathbb{Z}\}$ est un processus linéaire s'il admet la représentation suivante :*

$$X_t = \sum_{j=-\infty}^{+\infty} \psi_j a_{t-j}, \quad t \in \mathbb{Z} \quad (1.1.1)$$

où $\{a_t\}$ est un *BB faible* $(0, \sigma^2)$ et les ψ_j sont des coefficients réels tels que $\sum_{j=-\infty}^{+\infty} |\psi_j| < \infty$.

En utilisant l'opérateur retard B défini par : $B^m X_t = X_{t-m}$, $m \geq 1$ et $B^0 = 1$, l'opérateur identité, l'équation (1.1.1) peut s'écrire comme suit :

$$X_t = \psi(B) a_t \quad \text{où} \quad \psi(B) = \sum_{j=-\infty}^{+\infty} \psi_j B^j \quad (1.1.2)$$

Le processus 1.1.1 est stationnaire au sens large. Les autocovariances et autocorrélations sont données par

$$\gamma_h = \sigma^2 \sum_{j=-\infty}^{+\infty} \psi_j \psi_{j+h}, \quad h \geq 0, \quad (1.1.3)$$

et

$$\rho_h = \frac{\sum_{j=-\infty}^{+\infty} \psi_j \psi_{j+h}}{\sum_{i=-\infty}^{+\infty} \psi_i^2}, \quad h \geq 1. \quad (1.1.4)$$

En particulier, $\sigma_x^2 = \gamma_0 = \sigma^2 \sum_{j=-\infty}^{+\infty} \psi_j^2$ et nous pouvons écrire

$$\rho_h = \frac{\sigma^2}{\sigma_x^2} \sum_{j=-\infty}^{+\infty} \psi_j \psi_{j+h}, \quad h \geq 1. \quad (1.1.5)$$

Dans le cas où $\psi_j = 0$ pour $j \leq 0$, le processus $\{X_t : t \in \mathbb{Z}\}$ est alors dit causal et par la décomposition de Wold, tout processus stationnaire au sens large qui est purement non déterministe admet une représentation linéaire causale.

1.2. ESTIMATION DES PARAMÈTRES D'UN PROCESSUS STATIONNAIRE

Afin d'ajuster un modèle à une série d'observations, il est nécessaire de passer par l'étape de l'estimation des paramètres du modèle.

Dans cette section, nous décrivons l'estimation de la moyenne, des autocovariances et des autocorrélations d'un processus stationnaire à partir d'une réalisation finie de n observations X_1, \dots, X_n qui est aussi appelée une série chronologique.

1.2.1. Estimation de la moyenne et des autocovariances

Soit X_1, \dots, X_n une série stationnaire. La moyenne μ est en général estimée par $\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$, la moyenne arithmétique des observations.

L'estimateur usuel de la fonction d'autocovariance est donnée par

$$\hat{\gamma}_h = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-h} (X_{t+h} - \bar{X})(X_t - \bar{X}) & \text{si } 0 \leq h \leq n-1, \\ \hat{\gamma}_{-h} & \text{si } 1-n \leq h \leq 0, \\ 0 & \text{si } |h| \geq n. \end{cases} \quad (1.2.1)$$

La fonction $\{\hat{\gamma}_h, h \in \mathbb{Z}\}$ est la fonction d'autocovariance échantillonnale et $\hat{\gamma}_h$ est appelée l'autocovariance échantillonnale de délai h . Cette fonction est aussi paire, définie non négative et $\hat{\gamma}_h$ est un estimateur convergent de γ_h , $1 \leq h \leq m \leq n-1$ où m est fixe par rapport à n .

1.2.2. Estimation des autocorrélations

Comme estimation de ρ_h , on considère la fonction $\hat{\rho}_h$ définie par :

$$\hat{\rho}_h = \begin{cases} \frac{\sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X})(X_t - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2} & \text{si } |h| \leq n-1, \\ 0 & \text{si } |h| \geq n. \end{cases} \quad (1.2.2)$$

La fonction $\{\hat{\rho}_h, h \in \mathbb{Z}\}$ est la fonction d'autocorrélation échantillonnale ou corrélogramme. Elle présente les propriétés suivantes :

- i) $\hat{\rho}_0 = 1$;
- ii) $|\hat{\rho}_h| \leq 1$;
- iii) Elle est une fonction définie non négative.

Dans le cas où $\{X_t\}$ est un processus linéaire satisfaisant certaines hypothèses de régularité et que $\{a_t\}$ est un *BBfort*($0, \sigma^2$), nous avons les résultats suivants :

- i) $\lim_{n \rightarrow \infty} E(\hat{\rho}_h) = \rho_h$, $1 \leq h \leq m \leq n-1$ où m est fixe par rapport à n ;
- ii) pour n fixé, le biais est de l'ordre de $\frac{1}{n}$ et il croît, en général, avec h ;
- iii) $\hat{\rho}_h$ est un estimateur convergent en moyenne quadratique pour ρ_h .

Un résultat qui joue un rôle important dans ce travail est celui donnant la distribution asymptotique d'un vecteur d'autocorrélations échantillonnales. Pour ceci posons $\hat{\rho}_m = (\hat{\rho}_1, \dots, \hat{\rho}_m)'$, le vecteur des autocorrélations échantillonnales et $\rho_m = (\rho_1, \dots, \rho_m)'$, le vecteur correspondant des autocorrélations théoriques.

Théorème 1.2.1. *Si le processus $\{X_t\}$ décrit par (1.1.1) est stationnaire au sens large, que $\{a_t\}$ est un $BBfort(0, \sigma^2)$ et que $\sum_{j=-\infty}^{+\infty} |\psi_j| < \infty$, $\sum_{j=-\infty}^{+\infty} \psi_j^2 |j| < \infty$, alors $\sqrt{n}(\hat{\rho}_m - \rho_m)$ converge en loi vers une multinormale $N_m(\mathbf{0}, \mathbf{W})$ où la matrice de covariance $\mathbf{W} = (w_{ij})_{m \times m}$ dont les éléments w_{ij} sont donnés par la formule de Bartlett :*

$$w_{ij} = \sum_{h=-\infty}^{+\infty} \{ \rho_{h+i} \rho_{h+j} + \rho_{h-i} \rho_h - 2\rho_i \rho_h \rho_{h+j} - 2\rho_j \rho_h \rho_{h+i} + 2\rho_i \rho_j \rho_h^2 \}. \quad (1.2.3)$$

Une démonstration de ce théorème est présentée dans Brockwell et Davis (1991, chap. 7).

Pour n grand, le vecteur des autocorrélations échantillonnales $\hat{\rho}_m$ est asymptotiquement multinormal de vecteur moyenne ρ_m et de matrice de covariance $\frac{1}{n}\mathbf{W}$, que nous abrégeons par $\hat{\rho}_m$ est $AN_m(\rho_m, \frac{1}{n}\mathbf{W})$. Ainsi, $\hat{\rho}_i$ est $AN(\rho_i, \frac{1}{n}w_{ii})$ où w_{ii} est donné par (1.2.3).

$$\text{Dans le cas d'un } BBfort(0, \sigma^2), \text{ nous avons : } w_{ij} = \begin{cases} 0 & , \quad i \neq j, \\ 1 & , \quad i = j. \end{cases}$$

Il en découle que les $\hat{\rho}_i$, $i = 1, \dots, m$, sont asymptotiquement indépendantes et suivent une $N(0, \frac{1}{n})$.

1.3. PROCESSUS ARMA

Une sous-classe importante de processus linéaires est la classe des processus *ARMA* popularisés par Box et Jenkins (1970). Ils sont caractérisés par un nombre fini de paramètres, ce qui est une propriété fort importante car lors de la modélisation d'une série finie, on ne peut estimer qu'un nombre fini de paramètres.

Définition 1.3.1. *Un processus $\{X_t : t \in \mathbb{Z}\}$, tel que $E(X_t^2) < \infty$ et $E(X_t) = 0$, est un processus autorégressif-moyenne mobile d'ordre (p, q) , noté $ARMA(p, q)$ s'il satisfait l'équation suivante :*

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}, t \in \mathbb{Z}, \quad (1.3.1)$$

où $\{a_t\}$ est un bruit blanc faible $(0, \sigma^2)$. Les paramètres ϕ_i et θ_j , $i = 1, \dots, p$, $j = 1, \dots, q$ sont tous des paramètres réels appelés respectivement paramètres autorégressifs et paramètres moyenne mobile.

Le processus $\{X_t\}$ est $ARMA(p, q)$ de moyenne $\mu \neq 0$ si $\{X_t - \mu\}$ est $ARMA(p, q)$ de moyenne 0.

L'équation (1.3.1) peut s'écrire de façon abrégée

$$\Phi(B)X_t = \Theta(B)a_t \quad (1.3.2)$$

où $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ représente l'opérateur autorégressif et $\Theta(B) = \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ l'opérateur moyenne mobile. Un processus $ARMA(p, q)$ est stationnaire causal si et seulement si les racines de $\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p$ sont à l'extérieur du cercle unité.

Si $q = 0$, $\{X_t\}$ est un processus purement autorégressif, noté $AR(p)$.

Si $p = 0$, $\{X_t\}$ est dit processus moyenne mobile, noté $MA(q)$.

La modélisation d'une série chronologique selon l'approche de Box et Jenkins est un processus itératif en trois étapes principales. L'étape d'identification consiste à déterminer, à partir de l'observation des comportements des autocorrélations et des autocorrélations partielles échantillonnables, les ordres p et q du modèle *ARMA*. Vient ensuite l'étape d'estimation, que ce soit par la méthode du maximum de vraisemblance ou par la méthode des moindres carrés, des paramètres autorégressifs et moyenne mobile, de la moyenne et de la variance du processus. La troisième étape est celle relative à la validation du modèle estimé par le biais d'une analyse des résidus.

1.4. LE TEST DE BOX ET PIERCE : DESCRIPTION ET PROPRIÉTÉS

La validation de l'ajustement d'un modèle *ARMA* passe d'abord par une étude des résidus. En effet, le modèle estimé n'est acceptable que si les résidus se comportent essentiellement comme un bruit blanc. Dans cette section, nous présentons d'abord un résultat général sur la distribution asymptotique d'un vecteur d'autocorrélations résiduelles résultant de l'estimation d'un modèle *ARMA*. De ce résultat, nous déduisons la distribution asymptotique du test de Box et Pierce (1970) et de celui de Ljung et Box (1976) qui sont les deux tests les plus fréquemment utilisés pour vérifier l'adéquation d'un modèle *ARMA*.

Soit X_1, \dots, X_n une série générée d'un modèle autorégressif *AR*(p) stationnaire causal :

$$\phi(B)X_t = a_t \tag{1.4.1}$$

Le processus $\{X_t\}$ peut aussi s'écrire également sous la forme :

$$X_t = \psi(B)a_t \text{ où } \psi(B) = \sum_{j=0}^{+\infty} \psi_j B^j.$$

D'où :

$$X_t = \sum_{j=0}^{+\infty} \psi_j a_{t-j} \quad (1.4.2)$$

A partir des observations, les paramètres autorégressifs $\hat{\phi}_i$, $i = 1, \dots, p$, sont estimés par la méthode des moindres carrés conditionnels ou une méthode asymptotiquement équivalente (méthode des moments ou la méthode du maximum de vraisemblance).

A l'instant t , la valeur de X_t est estimée par $\hat{X}_t = \hat{\phi}_1 X_{t-1} + \dots + \hat{\phi}_p X_{t-p}$ et les résidus calculés sont $\hat{a}_t = X_t - \hat{X}_t$, $\forall t = p+1, \dots, n$.

Soit $\dot{\phi} = (\dot{\phi}_1, \dots, \dot{\phi}_p)'$ une valeur quelconque dans l'espace des valeurs admissibles des paramètres. A partir de X_1, \dots, X_n , nous pouvons calculer

$$\dot{a}_t = X_t - \dot{\phi}_1 X_{t-1} - \dots - \dot{\phi}_p X_{t-p}, \quad t = p+1, \dots, n.$$

Lorsque $\dot{\phi} = \phi$, le vecteur des vraies valeurs des paramètres, $\dot{a}_t = a_t$, la vraie erreur à l'instant t et si $\dot{\phi} = \hat{\phi}$, alors

$$\dot{a}_t = \hat{a}_t, \text{ le résidu à l'instant } t.$$

Afin d'obtenir une série de n erreurs $\dot{a}_1, \dots, \dot{a}_n$, il est d'habitude de poser que $\dot{a}_t = 0$, $t = 1, \dots, p$. Une telle hypothèse n'a pas de conséquence sur la distribution asymptotique des autocorrélations échantillonnales ni sur la distribution des estimateurs des paramètres autorégressifs. Les autocorrélations échantillonnales correspondantes sont définies par

$$\hat{\rho}_h(\dot{a}) = \frac{\sum_{t=1}^{n-h} \dot{a}_{t+h} \dot{a}_t}{\sum_{t=1}^n \dot{a}_t^2}, \quad \forall h = 1, 2, \dots, m$$

D'une façon générale, si notre hypothèse de base H_0 est que la série des erreurs a_1, a_2, \dots, a_n est un BB et $\hat{\rho}_h(a)$ est l'autocorrélation échantillonnale de délai h , la statistique de Box et Pierce pour l'hypothèse de bruit blanc est définie par

$$Q_{BP}(m, a) = n \sum_{h=1}^m (\hat{\rho}_h(a))^2. \quad (1.4.3)$$

Par le théorème 1.2.1,

$$\sqrt{n} \hat{\boldsymbol{\rho}}_m(a) \xrightarrow{L} N_m(\mathbf{0}, \mathbf{I}_m).$$

Il s'ensuit que les $\hat{\rho}_h(a)$, $h = 1, \dots, m$ sont asymptotiquement indépendants et $N(0, 1)$ et par conséquent,

$$Q_{BP}(m, a) \xrightarrow{L} \chi_m^2.$$

Pour tester l'hypothèse H_0 , un test basé sur les m premières autocorrélations peut être construit à partir de la statistique $Q_{BP}(m, a) = n \sum_{h=1}^m \hat{\rho}_h(a)^2$. La valeur de $Q_{BP}(m, a)$ dépend des valeurs prises par les $\hat{\rho}_h(a)$. Si H_0 est vraie, alors les $\hat{\rho}_h(a)$ devraient être au voisinage de zéro. Dans ce cas, $Q_{BP}(m, a)$ prend une petite valeur. Dans l'autre cas (si H_0 est fausse), au moins un des $\hat{\rho}_h(a)$ est relativement grand et $Q_{BP}(m, a)$ prendra donc une grande valeur. L'hypothèse H_0 est rejetée lorsque $Q_{BP}(m, a) \geq \chi_{m, 1-\alpha}^2$, où $\chi_{m, p}^2$ désigne le quantile d'ordre p de la loi χ_m^2 .

Si on s'intéresse au modèle AR estimé, ce dernier est adéquat si les résidus $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$ se comportent essentiellement comme un bruit blanc. Pour vérifier ceci, on utilise la statistique de Box et Pierce décrite précédemment. Ainsi, en remplaçant, dans cette statistique, a par \hat{a} , la loi asymptotique est une khi-deux mais les degrés de liberté sont, en général, différents de m . Cependant, avant d'annoncer cette loi asymptotique, il y a lieu de présenter certains résultats préliminaires.

Pour le processus autorégressif $\{X_t\}$ décrit en (1.4.1), posons $\hat{\phi}$ l'estimateur des moindres carrés conditionnels de ϕ . Soit la matrice \mathbf{X} définie comme suit

$$\mathbf{X} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \psi_1 & 1 & \cdot & \cdot \\ \psi_2 & \psi_1 & \cdot & 0 \\ \cdot & \cdot & \cdot & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{m-1} & \psi_{m-2} & \cdots & \psi_{m-p} \end{pmatrix} \quad (1.4.4)$$

Posons également la matrice suivante

$$\mathbf{D} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \quad (1.4.5)$$

qui est une matrice idempotente de rang m .

Dans la suite, nous aurons besoin de l'hypothèse suivante.

Hypothèse 1.4.1. *La série X_1, \dots, X_n est générée par un processus $AR(p)$ stationnaire causal dont les coefficients ψ_j de la représentation 1.4.2 sont tels que*

- i) $\psi_j = O(\frac{1}{\sqrt{n}})$, $j \geq m_n - p$, où $m = m_n$ est un entier qui dépend de n tel que*
- ii) $\frac{m_n}{n} = O(\frac{1}{\sqrt{n}})$.*

Sous cette hypothèse, si n est grand, $m_n \leq \Delta\sqrt{n}$ où Δ est une constante indépendante de n et pour $j \geq m_n - p$, les coefficients ψ_j sont négligeables.

Théorème 1.4.1. *Sous l'hypothèse 1.4.1, nous avons*

$$\hat{\rho}_m(\hat{a}) \text{ est } AN\left(\mathbf{0}, \frac{1}{n}(\mathbf{I}_m - \mathbf{D})\right)$$

et $\mathbf{I}_m - \mathbf{D}$ est aussi une matrice idempotente de rang $m - l$, où l est le nombre de paramètres AR estimés.

Esquisse de démonstration

Si $\hat{\phi}$ est l'estimateur des moindres carrés conditionnels de ϕ , alors $\hat{\phi}$ est la valeur de ϕ qui minimise

$$S(\phi) = \sum_{t=p+1}^n a_t^2 = \sum_{t=p+1}^n (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p})^2.$$

Les équations normales correspondantes sont

$$\frac{\partial S}{\partial \phi_j} = \sum_{t=p+1}^n (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p}) X_{t-j} = 0, \quad j = 1, \dots, p.$$

En remplaçant ϕ_i par $\hat{\phi}_i$, on obtient la condition d'orthogonalité

$$\sum_{t=p+1}^n \hat{a}_t X_{t-j} = 0, \quad j = 1, \dots, p. \quad (1.4.6)$$

En utilisant (1.4.2), il est démontré dans Box et Pierce (1970, p. 1513) que

$$\hat{\rho}_h(\hat{a}) = \hat{\rho}_h(a) + \sum_{j=1}^p (\phi_j - \hat{\phi}_j) \psi_{h-j} + O_p\left(\frac{1}{n}\right), \quad h = 1, \dots, m, \quad (1.4.7)$$

où $O_p(\cdot)$ est la notation habituelle de l'ordre en probabilité et $\psi_i = 0, i < 0$.

Pour représenter le vecteur $\hat{\rho}_m(\hat{a})$ comme une transformation linéaire du vecteur $\hat{\rho}_m(a)$, nous faisons intervenir la matrice \mathbf{X} . L'écriture (1.4.7) devient équivalente à :

$$\hat{\rho}_m(\hat{a}) = \hat{\rho}_m(a) + \mathbf{X}(\phi - \hat{\phi}) + O_p\left(\frac{1}{n}\right) \quad (1.4.8)$$

et d'après la condition d'orthogonalité des résidus (1.4.6), et le fait que $X_t = \sum_{j=0}^{\infty} \hat{\psi}_j \hat{a}_{t-j}$ (Box et Pierce, p. 1512), on a

$$\begin{aligned} 0 &= \sum_{t=1}^n \hat{a}_t \left(\sum_{k=0}^{\infty} \hat{\psi}_k \hat{a}_{t-j-k} \right) \\ &= \sum_{k=0}^{\infty} \hat{\psi}_k \sum_{t=1}^n \hat{a}_t \hat{a}_{t-j-k}. \end{aligned} \quad (1.4.9)$$

En divisant (1.4.9) par $\sum_{t=1}^n \hat{a}_t^2$, on obtient

$$0 = \sum_{k=0}^{\infty} \hat{\psi}_k \hat{\rho}_{j+k}(\hat{a}) \quad (1.4.10)$$

$$\approx \sum_{k=0}^m \hat{\psi}_k \hat{\rho}_{j+k}(\hat{a}) \quad (1.4.11)$$

où $m = m_n$ satisfait l'hypothèse 1.4.1. Pour n grand, l'erreur en approximant (1.4.10) par (1.4.11) est négligeable (voir Box et Pierce, p. 1513).

Sous forme matricielle, (1.4.11) s'écrit

$$\hat{\boldsymbol{\rho}}_m(\hat{a})' \mathbf{X} \approx \mathbf{0} \quad (1.4.12)$$

En multipliant les deux membres de (1.4.8) par \mathbf{D} nous obtenons

$$\mathbf{D} \hat{\boldsymbol{\rho}}_m(\hat{a}) \approx \mathbf{D} \hat{\boldsymbol{\rho}}_m(a) + \mathbf{D}\mathbf{X}(\boldsymbol{\phi} - \hat{\boldsymbol{\phi}}). \quad (1.4.13)$$

Mais

$$\mathbf{D} \hat{\boldsymbol{\rho}}_m(\hat{a}) = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \hat{\boldsymbol{\rho}}_m(\hat{a}) \approx \mathbf{0} \quad (1.4.14)$$

par (1.4.12). Puisque $\mathbf{D}\mathbf{X} = \mathbf{X}$ par (1.4.5) et $\mathbf{X}(\boldsymbol{\phi} - \hat{\boldsymbol{\phi}}) \approx \hat{\boldsymbol{\rho}}_m(\hat{a}) - \hat{\boldsymbol{\rho}}_m(a)$ par (1.4.8), l'équation (1.4.13) devient alors

$$\mathbf{0} \approx \mathbf{D} \hat{\boldsymbol{\rho}}_m(a) + \hat{\boldsymbol{\rho}}_m(\hat{a}) - \hat{\boldsymbol{\rho}}_m(a)$$

et donc

$$\hat{\boldsymbol{\rho}}_m(\hat{a}) \approx (\mathbf{I}_m - \mathbf{D}) \hat{\boldsymbol{\rho}}_m(a). \quad (1.4.15)$$

La matrice \mathbf{D} étant idempotente, la matrice $\mathbf{I}_m - \mathbf{D}$ l'est aussi. Le vecteur des autocorrélations résiduelles étant approximativement une transformation linéaire du vecteur $\hat{\rho}_m(a)$ qui est $AN_m(\mathbf{0}, \frac{1}{n}\mathbf{I}_m)$, il est aussi asymptotiquement multinormal de vecteur moyenne nul et de matrice de covariance $\frac{1}{n}(\mathbf{I}_m - \mathbf{D})$, ce qui complète l'esquisse de démonstration du théorème 1.4.1.

La matrice de covariance des autocorrélations résiduelles $\mathbf{M} = \mathbf{I}_m - \mathbf{D}$ mérite d'être abordée avec plus de détails. La détermination de cette matrice consiste à trouver les coefficients d_{ij} de la matrice \mathbf{D} . Ainsi par (1.1.5), nous pouvons écrire

$$X'X = \begin{pmatrix} \sum \psi_j^2 & \sum \psi_j \psi_{j-1} & \cdots & \sum \psi_j \psi_{j-p+1} \\ \sum \psi_j \psi_{j-1} & \sum \psi_j^2 & \cdots & \sum \psi_j \psi_{j-p+2} \\ \vdots & \vdots & & \vdots \\ \sum \psi_j \psi_{j-p+1} & \sum \psi_j \psi_{j-p+2} & \cdots & \sum \psi_j^2 \end{pmatrix}$$

$$\approx \frac{\sigma_x^2}{\sigma_a^2} \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & & \vdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{pmatrix}.$$

Ainsi, $\sigma_a^2 X'X$ n'est autre que la matrice de covariance du vecteur $(X_1, \dots, X_p)'$. Si on note par c_{ij} l'élément (i, j) de $(X'X)^{-1}$ et si on pose $\epsilon'_j = (\psi_{j-1}, \psi_{j-2}, \dots, \psi_{j-p})$, les éléments d_{ij} de la matrice \mathbf{D} sont donnés par

$$\begin{aligned} d_{ij} &= \epsilon'_i (X'X)^{-1} \epsilon_j \\ &= \sum_{k=1}^p \sum_{l=1}^p \psi_{j-k} c_{kl} \psi_{i-l} \\ &= -n \operatorname{cov}(\hat{\rho}_i, \hat{\rho}_j) \text{ pour } i \neq j. \end{aligned}$$

(voir Box et Pierce, 1970, Équation (2.31)).

Théorème 1.4.2. *Sous l'hypothèse 1.4.1, nous avons*

$$Q_{BP}(m, \hat{a}) \xrightarrow{L} \chi_{m-l}^2.$$

Démonstration

La matrice \mathbf{M} est idempotente et symétrique de rang $m-l$, il existe donc une matrice \mathbf{P} orthogonale telle que

$$\mathbf{P}'\mathbf{M}\mathbf{P} = \text{diag}(\beta_1, \dots, \beta_m) \quad (1.4.16)$$

où les $\beta_i, i = 1, \dots, m$, sont les valeurs propres de la matrice \mathbf{M} et les vecteurs propres associés aux β_i sont les vecteurs colonnes de la matrice \mathbf{P} . De plus, $m-l$ valeurs propres sont égales à 1 et les autres sont 0 et nous pouvons écrire

$$\text{diag}(\beta_1, \dots, \beta_m) = \text{diag}(1, \dots, 1, 0, \dots, 0). \quad (1.4.17)$$

Posons

$$\mathbf{U} = \mathbf{P}' \hat{\boldsymbol{\rho}}_m(a),$$

donc $\hat{\boldsymbol{\rho}}_m(a) = \mathbf{P}\mathbf{U}$ et puisque \mathbf{M} est idempotente, alors

$$\begin{aligned} \hat{\boldsymbol{\rho}}_m'(\hat{a}) \hat{\boldsymbol{\rho}}_m(\hat{a}) &= (\mathbf{M} \hat{\boldsymbol{\rho}}_m(a))' \mathbf{M} \hat{\boldsymbol{\rho}}_m(a) \\ &= \hat{\boldsymbol{\rho}}_m'(a) \mathbf{M}' \mathbf{M} \hat{\boldsymbol{\rho}}_m(a) \\ &= \hat{\boldsymbol{\rho}}_m'(a) \mathbf{M} \hat{\boldsymbol{\rho}}_m(a) \\ &= \mathbf{U}' \mathbf{P}' \mathbf{M} \mathbf{P} \mathbf{U} \\ &= \mathbf{U}' \text{diag}(\beta_1, \dots, \beta_m) \mathbf{U} \\ &= \sum_{i=1}^{m-l} U_i^2 \end{aligned}$$

par (1.4.17).

Puisque le vecteur \mathbf{U} est $AN_m(\mathbf{0}, \frac{1}{n}\mathbf{I}_m)$, le sous-vecteur $(U_1, \dots, U_{m-l})'$ est aussi $AN_{m-l}(\mathbf{0}, \frac{1}{n}\mathbf{I}_{m-l})$. Donc nous avons

$$Q_{BP}(m, \hat{a}) = n \hat{\boldsymbol{\rho}}'_m(\hat{a}) \hat{\boldsymbol{\rho}}_m(\hat{a}) \xrightarrow{L} \chi_{m-l}^2$$

et la démonstration est complétée.

Dans le cas d'un bruit blanc pour lequel on n'estime aucun paramètre, $Q_{BP}(m, a)$ converge vers une loi khi-deux à m degrés de liberté.

Pour tester l'hypothèse de base que le modèle AR estimé est adéquat, on se base sur la statistique $Q_{BP}(m, \hat{a})$ et on rejette H_0 au niveau de signification α si $Q_{BP}(m, \hat{a}) = n \sum_{h=0}^m \hat{\rho}_h(\hat{a})^2 \geq \chi_{m-l, 1-\alpha}^2$ où $\chi_{n,p}^2$ désigne le quantile d'ordre p de la loi χ_m^2 .

Box et Pierce (1970) discutent de la généralisation des théorèmes (1.4.1) et (1.4.2) au cas des modèles *ARMA*. La généralisation de ce résultat aux processus ARMA est possible grâce à la propriété suivante. Si une première série est générée par le processus $\pi(B)X_t = (1 - \pi_1 B - \dots - \pi_{p+q} B^{p+q})X_t = a_t$ et que la deuxième est générée du modèle *ARMA*, $\phi(B)X_t = \theta(B)a_t$, avec les mêmes innovations $\{a_t\}$ et si

$$\pi(B) = \phi(B)\theta(B) \tag{1.4.18}$$

alors les autocorrélations résiduelles des deux séries, lorsque l'estimation est faite par la méthode des moindres carrés, sont très rapprochées. A cet effet, pour n grand, les autocorrélations résiduelles d'un modèle *ARMA*(p, q) correctement estimé vont être distribuées comme celles d'un modèle *AR*($p + q$) satisfaisant la relation (1.4.18). Cette dernière relation lie les paramètres des deux modèles. En particulier, les coefficients ψ_j , $j = 1, \dots, m$ de la matrice \mathbf{X} décrite en (1.4.4)

seront les coefficients de $\psi(B) = [\phi(B)\theta(B)]^{-1}$. Une démonstration plus formelle dans le cas *ARMA* est présentée dans McLeod (1978).

Dans des études subséquentes à l'article de Box et Pierce, il a été remarqué que leur test conduit souvent au non rejet de l'hypothèse nulle alors que des autocorrélations résiduelles à certains délais sont importantes. Des études menées par certains auteurs notamment par Davis, Triggs et Newbold (1977) et Prothero et Wallis (1976), font ressortir que la distribution asymptotique de $Q_{BP}(m, \hat{a})$ peut dévier considérablement de celle d'une χ_{m-l}^2 .

Dans le but d'améliorer l'approximation de la statistique $Q_{BP}(m, \hat{a})$ par une khi-deux, Ljung et Box (1978) ont proposé une modification de cette statistique qui est basée sur la variance exacte des autocorrélations d'un bruit blanc gaussien. Leur statistique notée $Q_{LB}(m, \hat{a})$ est définie par

$$Q_{LB}(m, \hat{a}) = n(n+2) \sum_{h=1}^m (n-h)^{-1} (\hat{\rho}_h(\hat{a}))^2 \quad (1.4.19)$$

Ces auteurs ont montré par des simulations que le test modifié fournit une approximation sensiblement améliorée qui pourrait être adéquate dans la plupart des cas pratiques.

Chapitre 2

LE TEST DE KOCH ET YANG MODIFIÉ

Les tests de Box et Pierce et de Ljung et Box, décrits dans le chapitre précédent sont basés sur le carré des autocorrélations échantillonnales des résidus. Ils permettent de détecter une dépendance, si elle existe, dans la série des résidus. Cependant, dans certaines situations, plusieurs autocorrélations successives sont petites et de même signe, ce qui nous porte à douter de l'adéquation du modèle estimé. Les tests de Box et Pierce et de Ljung et Box ne peuvent pas prendre en compte cette particularité et conduisent la plupart du temps au non rejet du modèle estimé.

Un test asymptotique a été développé par Koch et Yang (1986) pour tester l'indépendance de deux séries chronologiques. Ici, nous reprenons l'idée de ces deux auteurs mais dans le contexte des tests d'ajustement et nous utiliserons ce test pour détecter des patrons particuliers dans des autocorrélations résiduelles successives.

2.1. DESCRIPTION DU TEST

Le test de Koch et Yang modifié est décrit par la statistique suivante :

$$Q_{KY}(m, i) = n \sum_{h=1}^{m-i} \left[\sum_{l=0}^i \hat{\rho}_{h+l}(\hat{a}) \right]^2. \quad (2.1.1)$$

Pour $i = 0$, $Q_{KY}(m, 0) = n \sum_{h=1}^m \hat{\rho}_h^2(\hat{a})$ et nous retrouvons la statistique de Box et Pierce. Pour $i = 1$, nous avons

$$\begin{aligned} Q_{KY}(m, 1) &= n \sum_{h=1}^{m-1} (\hat{\rho}_h(\hat{a}) + \hat{\rho}_{h+1}(\hat{a}))^2 \\ &= n[(\hat{\rho}_1(\hat{a}) + \hat{\rho}_2(\hat{a}))^2 + (\hat{\rho}_2(\hat{a}) + \hat{\rho}_3(\hat{a}))^2 + \cdots + (\hat{\rho}_{m-1}(\hat{a}) + \hat{\rho}_m(\hat{a}))^2]. \end{aligned}$$

et pour $i = 2$, la statistique devient

$$\begin{aligned} Q_{KY}(m, 2) &= n[(\hat{\rho}_1(\hat{a}) + \hat{\rho}_2(\hat{a}) + \hat{\rho}_3(\hat{a}))^2 + (\hat{\rho}_2(\hat{a}) + \hat{\rho}_3(\hat{a}) + \hat{\rho}_4(\hat{a}))^2 + \\ &\quad \cdots + (\hat{\rho}_{m-2}(\hat{a}) + \hat{\rho}_{m-1}(\hat{a}) + \hat{\rho}_m(\hat{a}))^2]. \end{aligned}$$

D'après les deux cas cités précédemment, il est clair que si des autocorrélations successives sont de même signe, cela se reflétera sur la statistique $Q_{KY}(m, i)$. Comme pour le test de Box et Pierce, nous rejetons l'hypothèse nulle pour les grandes valeurs de la statistique.

Dans la section qui suit, nous étudions la loi asymptotique de la statistique de Koch et Yang.

2.2. LOI ASYMPTOTIQUE DE LA STATISTIQUE DE KOCH ET YANG MODIFIÉE

La statistique de Koch et Yang modifiée est aussi une forme quadratique en le vecteur $\hat{\boldsymbol{\rho}}_m(\hat{a})$.

Proposition 2.2.1.

$$Q_{KY}(m, i) = n \hat{\boldsymbol{\rho}}_m'(\hat{a}) \mathbf{A}_i \hat{\boldsymbol{\rho}}_m(\hat{a}) \quad (2.2.1)$$

où

$$\mathbf{A}_i = \mathbf{C}_i' \mathbf{C}_i \quad (2.2.2)$$

est une matrice symétrique de dimension m qui est complètement déterminée par m et i et la matrice \mathbf{C}_i de dimension $(m-i, m)$ est définie par

$$\mathbf{C}_i = \begin{pmatrix} 1 & 1 & \dots & 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & 1 & 1 & 0 & \dots & 0 \\ \vdots & & & & & & & \\ 0 & \dots & \dots & \dots & 0 & 1 & \dots & 1 \end{pmatrix}.$$

Démonstration

Comme les rangées de \mathbf{C}_i sont composées de $(i+1)$ uns et de $(m-i-1)$ zéros, nous voyons immédiatement que

$$\mathbf{C}_i \hat{\boldsymbol{\rho}}_m(\hat{a}) = \begin{pmatrix} \hat{\rho}_1(\hat{a}) + \hat{\rho}_2(\hat{a}) + \dots + \hat{\rho}_{i+1}(\hat{a}) \\ \hat{\rho}_{i+1}(\hat{a}) + \hat{\rho}_{i+2}(\hat{a}) + \dots + \hat{\rho}_{2i+1}(\hat{a}) \\ \vdots \\ \hat{\rho}_{m-i}(\hat{a}) + \dots + \hat{\rho}_{m-1}(\hat{a}) + \hat{\rho}_m(\hat{a}) \end{pmatrix}$$

et donc, par (2.1.1),

$$\begin{aligned} Q_{KY}(m, i) &= n \sum_{h=1}^{m-i} \left[\sum_{l=0}^i \hat{\rho}_{h+l}(\hat{a}) \right]^2 \\ &= n (\mathbf{C}_i \hat{\boldsymbol{\rho}}'_m(\hat{a})) (\mathbf{C}_i \hat{\boldsymbol{\rho}}_m(\hat{a})) \\ &= n \hat{\boldsymbol{\rho}}'_m(\hat{a}) \mathbf{A}_i \hat{\boldsymbol{\rho}}_m(\hat{a}). \end{aligned}$$

Ce qui complète la démonstration.

Rappelons d'abord un résultat sur la distribution d'une forme quadratique.

Théorème 2.2.1. *Soit $\mathbf{X} = (X_1, \dots, X_p)'$ un vecteur aléatoire qui suit une loi multinormale $N_p(\mathbf{0}, \Sigma)$ et soit \mathbf{A} une matrice symétrique de dimension p . Alors, la forme quadratique $\mathbf{X}'\mathbf{A}\mathbf{X}$ suit la même loi de probabilité que $\sum_{i=1}^p \lambda_i Z_i^2$, c'est-à-dire*

$$\mathbf{X}'\mathbf{A}\mathbf{X} \sim \sum_{i=1}^p \lambda_i Z_i^2$$

où les λ_i sont les valeurs propres de $\mathbf{A}\Sigma$ et les Z_i sont des variables aléatoires indépendantes et identiquement distribuées (iid) selon la loi $N(0, 1)$.

Une démonstration de ce théorème se trouve dans Graybill (1976, p. 136). Notons que si $\mathbf{X} \sim N_p(\mathbf{0}, \Sigma)$ et que la matrice de covariance Σ est de plein rang, alors $\mathbf{X}'\Sigma^{-1}\mathbf{X}$ est distribuée selon une loi khi-deux à p degrés de liberté.

Nous aurons aussi besoin du résultat suivant sur la convergence d'une suite transformée de vecteurs aléatoires (Serfling, 1980, section 1.7).

Soient \mathbf{Y}_n , $n = 1, 2, \dots$, une suite de vecteurs aléatoires dans \mathbb{R}^m telle que $\mathbf{Y}_n \xrightarrow{L} \mathbf{Y}$ où \mathbf{Y} est aussi un vecteur aléatoire dans \mathbb{R}^m .

Théorème 2.2.2. *Soit $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$ une fonction continue. Si $\mathbf{Y}_n \xrightarrow{L} \mathbf{Y}$, alors*

$$g(\mathbf{Y}_n) \xrightarrow{L} g(\mathbf{Y}).$$

De ce qui précède, on déduit le résultat suivant.

Théorème 2.2.3. *Sous l'hypothèse 1.4.1, la loi asymptotique de $Q_{KY}(m, i)$ est la même que celle de*

$$\sum_{i=1}^m \lambda_i Z_i^2$$

où les Z_i sont iid selon la loi $N(0, 1)$ et les λ_i sont les valeurs propres de la matrice $\mathbf{M}\mathbf{A}_i$.

Démonstration

Il suffit de poser $\mathbf{Y}_n = \sqrt{n} \hat{\boldsymbol{\rho}}_m(\hat{a})$ et de choisir la fonction $g : \mathbb{R}^m \rightarrow \mathbb{R}$ définie par

$$g(\mathbf{y}) = \mathbf{y}' \mathbf{A}_i \mathbf{y}, \mathbf{y} \in \mathbb{R}^m.$$

La fonction $g(\cdot)$ est continue et par le théorème (1.4.1), $\mathbf{Y}_n \xrightarrow{L} \mathbf{Y}$ où $\mathbf{Y} \sim N_m(\mathbf{0}, \mathbf{M})$. Ainsi, nous avons $Q_{KY}(m, i) = g(\mathbf{Y}_n) \xrightarrow{L} g(\mathbf{Y}) = \mathbf{Y}' \mathbf{A}_i \mathbf{Y}$ et la loi de $\mathbf{Y}' \mathbf{A}_i \mathbf{Y}$ est donnée par le théorème 2.2.1. Ce qui complète la démonstration.

Il est clair que la loi asymptotique de la statistique de Koch et Yang dépend du modèle estimé puisqu'elle fait intervenir la matrice \mathbf{M} . En pratique, nous serons obligés d'utiliser les valeurs propres estimées de $\mathbf{M} \mathbf{A}_i$, ce qui conduit à une valeur critique estimée à partir des observations. Afin de contourner cette difficulté, on peut définir l'analogue de la statistique de Koch et Yang dans notre situation à partir d'une transformation linéaire du vecteur $\hat{\boldsymbol{\rho}}_m(\hat{a})$. Posons

$$\mathbf{U} = \mathbf{P}' \hat{\boldsymbol{\rho}}_m(\hat{a}) \quad (2.2.3)$$

où la matrice \mathbf{P}' est telle que $\mathbf{P}' \mathbf{M} \mathbf{P} = \text{diag}(\beta_1, \beta_2, \dots, \beta_m)$. Ce résultat nous conduit à la variante suivante de la statistique $Q_{KY}(m, i)$:

$$Q_{KY}^*(m, i) = n \sum_{h=1}^{m-i} \left[\sum_{j=0}^i U_{h+j} \right]^2 = n \mathbf{U}' \mathbf{A}_i \mathbf{U} \quad (2.2.4)$$

où le vecteur $\mathbf{U} = (U_1, \dots, U_m)$ est $AN(\mathbf{0}, \frac{1}{n} \mathbf{I})$ et donc, $Q_{KY}^*(m, i)$ est distribuée asymptotiquement comme

$$\sum_{j=1}^m \lambda_j Z_j^2 \quad (2.2.5)$$

où les λ_j sont les valeurs propres de la matrice \mathbf{A}_i .

A première vue, la statistique $Q_{KY}^*(m, i)$ semble plus intéressante d'un point de vue théorique, mais comme l'on remarqué Francq, Roy et Zakoïan (2003), la

matrice \mathbf{M} n'est qu'une approximation de la matrice de covariance asymptotique exacte de $\sqrt{n}\hat{\rho}_m(\hat{a})$ et cette approximation se dégrade rapidement dès que l'on s'écarte même légèrement de l'hypothèse 1.4.1, en particulier, lorsque m prend de petites valeurs. Pour cette raison, nous utiliserons la statistique $Q_{KY}(m, i)$ définie par (2.1.1) dans la suite de ce travail.

2.3. ALGORITHME D'IMHOF

Imhof (1961) a développé un algorithme permettant de calculer la distribution exacte d'une forme quadratique de variables normales qui est souvent la même que celle d'une combinaison linéaire de variables khi-deux indépendantes.

Le contexte est le suivant. Soit \mathbf{X} un vecteur $N_l(\mathbf{0}, \Sigma)$. L'objectif est de calculer la distribution de la forme quadratique $Q = \mathbf{X}'\mathbf{A}\mathbf{X}$ où \mathbf{A} est une matrice symétrique.

Posons $\mathbf{Y} = (\Sigma)^{-\frac{1}{2}}\mathbf{X}$, $\mathbf{Y} \sim N_l(\mathbf{0}, \mathbf{I}_l)$ et si $\mathbf{Y} = (Y_1, \dots, Y_l)'$, alors ses composantes Y_i sont indépendantes et suivent une loi $N(0, 1)$. Nous savons que

$$\begin{aligned} Q &= \mathbf{X}'\mathbf{A}\mathbf{X} \\ &= \mathbf{Y}'(\Sigma)^{\frac{1}{2}}\mathbf{A}(\Sigma)^{\frac{1}{2}}\mathbf{Y} \\ &= \mathbf{Y}'\boldsymbol{\beta}\mathbf{Y} \end{aligned}$$

où $\boldsymbol{\beta} = (\Sigma)^{\frac{1}{2}}\mathbf{A}(\Sigma)^{\frac{1}{2}}$. La matrice $\boldsymbol{\beta}$ étant symétrique, il existe donc une matrice \mathbf{P} orthogonale telle que $\mathbf{P}'\boldsymbol{\beta}\mathbf{P} = \text{diag}(\lambda_1, \dots, \lambda_l)$ où $\lambda_1, \dots, \lambda_l$ sont les valeurs propres de $\boldsymbol{\beta}$ que nous supposons toutes distinctes pour le moment. Posons $\mathbf{Z} = (Z_1, \dots, Z_l)' = \mathbf{P}'\mathbf{Y}$. Alors la forme quadratique Q peut s'écrire

$$\begin{aligned}
Q &= \mathbf{Y}'\mathbf{P}'\text{diag}(\lambda_1, \dots, \lambda_l)\mathbf{P}'\mathbf{Y} \\
&= \mathbf{Z}'\text{diag}(\lambda_1, \dots, \lambda_l)\mathbf{Z} \\
&= \sum_{i=1}^l \lambda_i Z_i^2.
\end{aligned} \tag{2.3.1}$$

Le vecteur \mathbf{Z} est $N_l(0, \mathbf{I}_l)$ et donc les Z_i sont iid $N(0, 1)$.

Dans le cas général où les valeurs propres de la matrice \mathbf{A}_i sont possiblement de multiplicité supérieure ou égale à 1, notons les valeurs propres distinctes par $\lambda_1, \dots, \lambda_m$, de multiplicités respectives, n_1, \dots, n_m telles que, $n_1 + \dots + n_m = l$. La forme quadratique Q est alors équivalente à

$$Q = \sum_{j=1}^m \lambda_j w_j$$

où les w_j sont des variables χ^2 indépendantes de degré de liberté n_j , $j = 1, \dots, m$.

Avant de calculer la fonction de répartition de la variable aléatoire Q , il est utile de rappeler la définition suivante.

Définition 2.3.1. Soit $F(x)$ une fonction de répartition. La fonction caractéristique correspondant à $F(x)$ est définie par :

$$\Phi(t) = \int_{-\infty}^{+\infty} e^{itx} dF(x) \tag{2.3.2}$$

Si $F(x)$ est continue et possède une dérivée $f(x)$ continue, alors (2.3.2) devient

$$\Phi(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx. \tag{2.3.3}$$

Une formule d'inversion décrite par Gil-Peleaz (1951) permet de calculer $F(x)$ à partir de sa fonction caractéristique $\Phi(t)$. Ainsi, on obtient

$$\begin{aligned} P(Q \leq x) &= F(x) \\ &= \frac{1}{2} + \frac{1}{2\pi} \int_0^{+\infty} \frac{e^{itx}\Phi(-t) - e^{itx}\Phi(t)}{it} dt \\ &= \frac{1}{2} - \frac{1}{\pi} \int_0^{+\infty} \frac{1}{t} I(e^{-itx}\Phi(t)) dt. \end{aligned} \quad (2.3.4)$$

Le terme $I(e^{-itx}\Phi(t))$ représente la partie imaginaire du nombre complexe $e^{-itx}\Phi(t)$.

Dans le développement du calcul de la distribution $F(x)$, Imhof utilise le résultat suivant (voir Koerts et Abrahamse, 1969) :

$$F(x) = \frac{1}{2} - \frac{1}{\pi} \int_0^{+\infty} \frac{\sin(\epsilon(u))}{u\delta(u)} d\mu \quad (2.3.5)$$

où,

$$\epsilon(u) = \frac{1}{2} \sum_{j=1}^m [n_j t g^{-1}(\lambda_j u)] - \frac{1}{2} x u$$

et

$$\delta(u) = \prod_{j=1}^m (1 + \lambda_j^2 u^2)^{\frac{1}{4} n_j}.$$

Il montre aussi que

$$\lim_{u \rightarrow 0} \frac{\sin(\epsilon(u))}{u\delta(u)} = \frac{1}{2} \sum_{j=1}^m \beta_j n_j - \frac{1}{2} x. \quad (2.3.6)$$

Il faut signaler que la fonction $u\delta(u)$ croît de façon monotone vers l'infini. Par conséquent, l'intégration dans (2.3.5) sera calculée numériquement sur l'intervalle fini $[0, U]$. Pour $u = 0$, la limite est déduite à partir de (2.3.6).

Deux types d'erreurs sont alors à contrôler : l'erreur de troncature et l'erreur d'intégration. La première notée t_U est donnée par

$$t_U = \frac{1}{\pi} \int_u^{+\infty} \frac{\sin(\epsilon(u))}{u\delta(u)} du.$$

La valeur absolue de t_U possède une borne supérieure T_U égale à

$$T_U^{-1} = \pi K U^K \left[\sum_{j=1}^m m |\beta_j|^{\frac{1}{2}n_j} \right] \quad (2.3.7)$$

où $K = \frac{1}{2} \sum_{j=1}^m n_j$.

Pour que la procédure d'intégration numérique soit performante, Koerts et Abrahamse (1969) suggèrent de spécifier l'erreur de troncature à accepter et de déterminer par la suite la valeur U correspondante.

Le deuxième type d'erreur résulte de l'utilisation de la procédure d'intégration numérique. il est défini comme suit :

$$I = \frac{1}{\pi} \int_0^U \frac{\sin(\epsilon(u))}{u\delta(u)} du \quad (2.3.8)$$

Pour évaluer cette intégrale, la règle de Simpson qui permet d'intégrer une fonction $f(x)$ sur un intervalle fermé $[a, b]$ en divisant cet intervalle en de petites parties égales a été utilisée par Koerts et Abrahamse (1969).

En pratique, ces auteurs suggèrent de retrouver les deux erreurs T_U et I_U selon des procédures itératives jusqu'à ce qu'on se rapproche des valeurs fixées à l'avance.

Chapitre 3

LE TEST DE HONG

Dans le chapitre des préliminaires, nous avons vu les tests de non-corrélation des résidus basés sur la statistique de Box-Pierce et sur celle de Ljung-Box. L'analogie de ces deux tests pour vérifier l'indépendance de deux séries *ARMA* est le test de Haugh (1976). Comme l'ont remarqué Koch et Yang (1986), le dernier test ne permet pas de détecter des patrons dans les corrélations croisées échantillonnelles, comme par exemple, plusieurs corrélations successives de même signe, et par conséquent, conduit souvent au non-rejet de l'hypothèse nulle. Les tests de Box-Pierce et de Ljung-Box présentent la même faiblesse.

Hong (1996a) met l'accent sur un autre inconvénient du test de Box-Pierce qui découle du fait que ce dernier affecte le même poids à toutes les autocorrélations échantillonnelles. Selon Hong, cette affectation n'est pas optimale car pour la majorité des processus stationnaires, les autocorrélations décroissent rapidement vers zéro quand le délai augmente.

Hong propose une famille de tests convergents qui permettent de vérifier si le terme d'erreur est un bruit blanc et utilise pour ceci plusieurs fonctions de distance. Son approche repose sur la comparaison d'un estimateur de la densité spectrale calculé par la méthode du noyau à la densité spectrale sous l'hypothèse

nulle.

La première classe de ces tests, et à laquelle nous nous intéressons, est celle pour laquelle la fonction de distance est la norme quadratique. Le test défini par cette distance fait intervenir une fonction de noyau qui affecte un poids à chaque autocorrélation. En donnant plus de poids aux autocorrélations de petits délais, Hong pense que son test sera plus puissant dans plusieurs situations que ceux de Box-Pierce et de Ljung-Box.

3.1. DESCRIPTION DU TEST

Avant de décrire la statistique proposée par Hong, nous introduisons les fonctions de noyau que nous noterons par $w(\cdot)$ et présentons quelques unes de leurs caractéristiques.

Hypothèse 3.1.1. *la fonction $w : \mathbb{R} \rightarrow [-1, 1]$ est une fonction symétrique continue en zéro, ayant au plus un nombre fini de points de discontinuité et telle que $w(0) = 1$, $\int_{-\infty}^{+\infty} w^2(z) dz < \infty$.*

Il existe plusieurs fonctions de noyau. Les plus fréquemment utilisées en analyse spectrale sont :

$$\text{Noyau uniforme tronqué : } w(z) = \begin{cases} 1 & , \quad |z| \leq 1, \\ 0 & , \quad \text{sinon.} \end{cases}$$

$$\text{Bartlett : } w(z) = \begin{cases} 1 - |z| & , \quad |z| \leq 1, \\ 0 & , \quad \text{sinon.} \end{cases}$$

$$\text{Daniell : } w(z) = \frac{\sin(\pi z)}{\pi z}, \quad z \in \mathbb{R}$$

$$\text{Parzen : } w(z) = \begin{cases} 1 - 6\left(\frac{\pi z}{6}\right)^2 + 6\left|\left(\frac{\pi z}{6}\right)\right|^3 & , \quad |z| \leq \frac{3}{\pi}, \\ 2\left(1 - \left|\frac{\pi z}{6}\right|\right)^3 & , \quad \frac{3}{\pi} \leq |z| \leq \frac{6}{\pi}, \\ 0 & , \quad \text{sinon.} \end{cases}$$

$$\text{Bartlett-Priestley : } w(z) = \frac{9}{5\pi^2 z^2} \left\{ \frac{\sin(\pi\sqrt{\frac{5}{3z}})}{\pi\sqrt{\frac{5}{3z}}} - \cos(\pi\sqrt{\frac{5}{3z}}) \right\}, \quad z \in \mathbb{R}.$$

La statistique proposée par Hong est une somme pondérée des carrés de toutes les autocorrélations résiduelles qui est définie par

$$Q_H(w, T_n) = \frac{n \sum_{h=1}^{n-1} w^2\left(\frac{h}{T_n}\right) \hat{\rho}_h(\hat{a})^2 - C_n(w)}{(2D_n(w))^{\frac{1}{2}}} \quad (3.1.1)$$

où

$$C_n(w) = \sum_{h=1}^{n-1} \left(1 - \frac{h}{n}\right) w^2\left(\frac{h}{T_n}\right)$$

$$D_n(w) = \sum_{h=1}^{n-2} \left(1 - \frac{h}{n}\right) \left(1 - \frac{h+1}{n}\right) w^4\left(\frac{h}{T_n}\right)$$

et w est une fonction de noyau satisfaisant l'hypothèse 3.1.1. Le paramètre T_n est un entier tel que T_n tend vers l'infini mais moins rapidement que n , c'est-à-dire que $\frac{T_n}{n}$ tend vers zéro. Avec des noyaux tels le noyau uniforme tronqué et ceux de Bartlett et de Parzen, T_n est dit « le point de troncature » car pour des délais h plus grands que T_n , le poids $w\left(\frac{h}{T_n}\right)$ est nul.

En prenant le noyau uniforme tronqué et $T_n = m$,

$$Q_H(w, T_n) = \frac{n \sum_{h=1}^m \hat{\rho}_h(\hat{a})^2 - C_n(w)}{(2D_n(w))^{\frac{1}{2}}}$$

qui est une version normalisée de la statistique de Box et Pierce.

Le terme $D_n(w)$ peut être remplacé par $T_n D(w) = T_n \int_0^\infty w^4(z) dz$ sans affecter la distribution asymptotique de $Q_H(w, T_n)$ car $\frac{D_n(w)}{T_n}$ tend vers $D(w)$. De la même

façon, on peut remplacer $C_n(w)$ par $T_n C(w) = T_n \int_0^\infty w^2(z) dz$. Ainsi, une variante de la statistique $Q_H(w, T_n)$ est

$$\tilde{Q}_H(w, T_n) = \frac{n \sum_{h=1}^{n-1} \hat{\rho}_h(\hat{a})^2 w^2(\frac{h}{T_n}) - T_n C(w)}{(2T_n D(w))^{\frac{1}{2}}} \quad (3.1.2)$$

Note

Pour n fini, $Q_H(w, T_n)$ et $\tilde{Q}_H(w, T_n)$ sont différentes mais asymptotiquement, ils ont la même distribution.

Comme pour les tests cités dans les chapitres précédents, nous rejetons l'hypothèse nulle pour les grandes valeurs de $Q_H(w, T_n)$.

3.2. PROPRIÉTÉS ASYMPTOTIQUES

Afin de montrer la convergence en loi de la statistique $Q_H(w, T_n)$, Hong (1996a) a formulé certaines hypothèses.

Hypothèse 3.2.1. *Le processus innovation $\{a_t\}$ du processus $\{X_t\}$ est un bruit blanc fort tel que $E(a_t^2) = \sigma_a^2$ et $E(a_t^4) < \infty$.*

Hypothèse 3.2.2. *L'estimateur $\hat{\phi}$ du vecteur ϕ des paramètres autorégressifs doit être \sqrt{n} -convergent c'est-à-dire $n^{\frac{1}{2}}(\hat{\phi} - \phi) = O_p(1)$.*

L'estimateur des moindres carrés et l'estimateur de vraisemblance maximale de ϕ sont \sqrt{n} -convergents.

Théorème 3.2.1. *Sous les hypothèses 3.1.1, 3.2.1, 3.2.2 et si $T_n \rightarrow \infty$, $\frac{T_n}{n} \rightarrow \infty$, alors $Q_H(w, T_n)$ converge en loi vers une variable aléatoire $N(0, 1)$:*

$$Q_H(w, T_n) \xrightarrow{L} N(0, 1). \quad (3.2.1)$$

Pour une grande classe d'alternatives de corrélation sérielle de forme quelconque des erreurs, Hong (1996a) montre que son test est convergent. Il fait aussi ressortir que le noyau de Daniell est optimal, dans un certain sens, dans une classe de noyaux satisfaisant certaines conditions de régularité. Finalement, la validité du test de Hong pour des processus *ARMA* a été établi tout récemment par Hong et Lee (2003).

Chapitre 4

SIMULATIONS

Dans les chapitres précédents, nous avons présenté différents tests statistiques pour vérifier la non corrélation des résidus d'un processus autorégressif. Dans ce chapitre, nous effectuerons des simulations de Monte Carlo pour étudier le niveau exact et la puissance de ces différents tests. Il est à noter que toute la partie programmation a été faite en FORTRAN 77. Les progiciels S-Plus et Excel ont servi pour les représentations graphiques.

4.1. ALGORITHMES DE SIMULATIONS

Afin de générer une série chronologique de longueur n d'un processus autorégressif avec des erreurs normalement distribuées, nous avons utilisé le sous-programme *G05EGF* de la librairie NAG de FORTRAN. Ce dernier appelle à son tour, le sous-programme *G05EWF* qui permet de générer des termes successifs d'un bruit blanc et d'initier la série à une position stationnaire.

L'estimation des paramètres du vrai modèle par la méthode des moindres carrés est rendue possible grâce au sous-programme *G13AFF*. Les résidus ainsi que les autocorrélations résiduelles, qui serviront de base pour le calcul des statistiques de test sont obtenus à l'aide du sous-programme *G13ABF*.

Considérons le processus autorégressif $\{X_t\}$, tel que décrit par l'équation (1.4.1) et notons $\phi = (\phi_1, \dots, \phi_p)'$ le vecteur des paramètres autorégressifs et $\hat{\phi}$ l'estimateur des moindres carrés de ϕ . Alors, l'autocorrélation résiduelle de délai h correspondante est

$$\hat{\rho}_h(\hat{a}) = \frac{\sum_{t=1}^{n-h} \hat{a}_{t+h} \hat{a}_t}{\sum_{t=1}^n \hat{a}_t^2}, \quad h = 1, \dots, m.$$

La première statistique à laquelle on s'est intéressé est la statistique de Koch et Yang. Rappelons que cette dernière s'écrit de la façon suivante :

$$Q_{KY}(m, i) = n \sum_{h=1}^{m-i} \left[\sum_{j=0}^i \hat{\rho}_{h+j}(\hat{a}) \right]^2 \quad (4.1.1)$$

Pour améliorer l'approximation par la loi asymptotique, nous avons introduit un facteur de correction aux autocorrélations résiduelles et au lieu de travailler avec le vecteur $\hat{\rho}_m(\hat{a})$ directement, nous avons introduit le vecteur $\tilde{\rho}_m(\hat{a}) = (\tilde{\rho}_1(\hat{a}), \dots, \tilde{\rho}_m(\hat{a}))'$ où

$$\tilde{\rho}_h(\hat{a}) = \sqrt{\frac{n}{n-h}} \hat{\rho}_h(\hat{a}), \quad h = 1, \dots, m.$$

La statistique de Koch et Yang modifiée devient

$$\tilde{Q}_{KY}(m, i) = n \sum_{h=1}^{m-i} \left[\sum_{j=0}^i \tilde{\rho}_{h+j}(\hat{a}) \right]^2 \quad (4.1.2)$$

Nous avons commencé par évaluer les coefficients $\psi_j, j = 1, \dots, m-1$, de la matrice \mathbf{X} décrite dans (1.4.4) en résolvant les équations aux différences finies pour un modèle autorégressif. La procédure complète pour le calcul de ces coefficients est décrite dans le sous-programme *covariance* présenté en annexe. Ces coefficients nous ont permis d'établir la matrice \mathbf{D} décrite par (1.4.5) et par conséquent la matrice de covariance \mathbf{M} des autocorrélations échantillonnables selon la relation $\mathbf{M} = \mathbf{I}_m - \mathbf{D}$.

Nous avons calculé la statistique $\tilde{Q}_{KY}(m, i)$ pour différentes valeurs de m et de i .

La deuxième statistique de test est celle de Hong. Cette dernière dépend d'une fonction de noyau et d'un point de troncature. Quatre fonctions de noyau ont été utilisées, celles de Bartlett, Daniell, Parzen et le noyau uniforme tronqué telles que définies à la section 3.1. Les points de troncature ont été choisis de manière à rendre possible la comparaison entre le test de Koch et Yang et celui de Hong. Ainsi, cinq points de troncature ont été choisis, soient 3, 6, 12, 18 et 24. Ceux-ci correspondent aux délais maximum considérés pour le calcul de la statistique de Koch et Yang.

4.2. DESCRIPTION DE LA SIMULATION

Nous voulons étudier le niveau et la puissance des tests proposés. A cette fin, nous examinons, pour un modèle donné, la fréquence empirique de rejet de l'hypothèse de non-corrélation des résidus. Deux modèles différents purement autorégressifs, soit $AR(1)$ et $AR(2)$ et un bruit blanc sont considérés pour l'étude de niveau. Pour l'étude de puissance, nous avons choisi deux modèles $MA(4)$, un modèle $MA(15)$ et un modèle $ARMA(1, 4)$. Les équations de ces modèles sont présentées dans le tableau (4.2.1).

Pour chaque modèle étudié, nous avons généré des séries de longueurs 50, 100, 200 et 500. Les valeurs du point de troncature T_n considérées sont : 3, 6, 12, 18 et 24.

Dans l'étude de niveau, 10 000 réalisations indépendantes de chacun des modèles ont été générées et ce pour chaque taille de série n . Les niveaux empiriques,

TABLEAU 4.2.1. Équations des modèles considérés pour les simulations.

| Modèles | Équations |
|--------------|---|
| Bruit blanc | $X_t = a_t$ |
| $AR(1)$ | $X_t = 0.8X_{t-1} + a_t$ |
| $AR(2)$ | $X_t = 0.1X_{t-1} + 0.5X_{t-2} + a_t$ |
| $MA(4)$ | $X_t = a_t + 0.8a_{t-4}$ |
| $MA(4)$ | $X_t = a_t + 0.4a_{t-4}$ |
| $MA(15)$ | $X_t = a_t + 0.2a_{t-13} + 0.2a_{t-14} + 0.5a_{t-15}$ |
| $ARMA(1, 4)$ | $X_t - 0.8X_{t-1} = at + 0.8a_{t-4}$ |

correspondant aux fréquences de rejet de l'hypothèse nulle, ont été calculés pour chacune des statistiques de test et ce pour les trois différents seuils de signification nominaux ($\alpha = 0.01$, $\alpha = 0.05$ et $\alpha = 0.1$).

Pour un modèle donné, et pour une longueur de série fixée, les étapes suivantes ont été exécutées par le biais du code FORTRAN décrit en annexe, pour chacune des 10 000 réalisations indépendantes :

- 1) Nous avons généré un modèle autorégressif d'ordre p déterminé ($p = 0, 1, 2$).
Le sous-programme de *NAG* utilisé est *G05EGF*.
- 2) Nous avons estimé le vrai modèle par la méthode des moindres carrés à l'aide du sous-programme *G13AFF*.
- 3) Les résidus $\hat{a}_t = X_t - \hat{X}_t$, $t = 1, \dots, n$ ainsi que les autocorrélations résiduelles sont calculés à l'aide du sous-programme *G13ABF*.
- 4) Le sous-programme *covariance* décrit dans le code FORTRAN en annexe, nous a permis de calculer les coefficients ψ_j , $j = 1, \dots, m - 1$ de la matrice **D** et par conséquent d'évaluer la matrice de covariance des autocorrélations résiduelles $\mathbf{M} = \mathbf{I}_m - \mathbf{D}$.

- 5) La statistique de Koch et Yang modifiée $\tilde{Q}_{KY}(m, i)$ définie par 4.1.2 est calculée pour les délais $m = 3, 6, 12, 18$ et 24 et pour chaque m , pour les valeurs de i allant de 0 jusqu'à $m - 1$.
- 6) La statistique de Hong $Q_H(w, T_n)$ est calculée pour chaque fonction de noyau (Bartlett (BA), Daniell (DA), Parzen (PA) et le noyau uniforme tronqué (TR)), et pour chacune des valeurs du point de troncature.
- 7) Pour chaque statistique $\tilde{Q}_{KY}(m, i)$ obtenue, on calcule grâce au sous-programme *G01JDF* de *NAG*

$$P\left[\sum_{i=1}^{n_r} \lambda_i \chi_i^2 < \tilde{Q}_{KY}(m, i)\right]$$

où les λ_i sont les valeurs propres de la matrice \mathbf{MA}_i , n_r est le nombre de valeurs propres non nulles de \mathbf{MA}_i , les χ_i^2 , $i = 1, \dots, n_r$ sont des variables aléatoires indépendantes identiquement distribuées selon une loi khi-deux à 1 degré de liberté. Il faut souligner que pour l'utilisation du sous-programme *G01JDF*, nous devons spécifier la méthode de calcul de la probabilité décrite ci-dessus. La méthode que nous avons choisie est celle qui fait appel à l'algorithme d'Imhof.

- 8) Les probabilités obtenues sont comparées pour chaque test aux seuils nominaux ($\alpha = 0.01$, $\alpha = 0.05$ et $\alpha = 0.1$) et pour chaque réalisation, nous notons si nous rejetons l'hypothèse nulle ou non.
- 9) Pour chaque statistique $Q_H(w, T_n)$ trouvée, nous calculons

$$P[N(0, 1) > Q_H(w, T_n)]$$

Le sous-programme utilisé pour évaluer cette probabilité est *G01EAF*. La probabilité ainsi obtenue est comparée aux seuils nominaux ($\alpha = 0.01$, $\alpha = 0.05$ et $\alpha = 0.1$) et nous notons à chaque fois si nous rejetons l'hypothèse nulle.

- 10) Finalement, avec les 10 000 réalisations indépendantes, nous calculons la proportion de rejet pour chaque statistique de test et pour chaque niveau nominal.

Pour étudier et comparer les propriétés de puissance des tests proposés, la procédure décrite ci-dessus a été appliquée aux modèles $MA(4)$, $MA(15)$ et $ARMA(1, 4)$ que nous avons approximé par un modèle $AR(1)$. Par ailleurs, pour des fins de comparaison, nous avons évalué les valeurs des puissances des tests en considérant les valeurs critiques exactes (VCE) dans un premier lieu et les valeurs critiques asymptotiques (VCA) dans un second lieu. Les valeurs du paramètre autorégressif utilisé pour l'estimation ont été choisies en fonction du modèle de la série générée. Ainsi, pour les modèles moyenne mobile, soit $MA(4)$ et $MA(15)$, la valeur du paramètre autorégressif est nulle. Pour le modèle $ARMA(1, 4)$, la valeur du paramètre autorégressif est 0.8.

Le nombre de rejets X à partir de N réalisations indépendantes est une variable aléatoire qui suit une loi binomiale avec paramètres N et p ($(B(N, p))$) où p est la vraie probabilité d'un rejet. Ainsi, la proportion de rejets $\hat{p} = \frac{X}{N}$ est asymptotiquement $N(p, \frac{p(1-p)}{N})$ que nous abrégeons par $\hat{p} \sim AN(p, \frac{p(1-p)}{N})$.

Dans l'étude de niveau, nous sommes dans une situation de test et nous voulons confronter l'hypothèse $H_0 : p = \alpha$ à la contre-hypothèse $H_1 : p \neq \alpha$. Sous H_0 , $\hat{p} \sim AN(\alpha, \frac{\alpha(1-\alpha)}{N})$ et au niveau de signification 5%, nous rejetons H_0 si

$$|\hat{p} - \alpha| > 1.96 \sqrt{\frac{\alpha(1-\alpha)}{N}}$$

Les limites de signification sont donc $\alpha \pm 1.96 \sqrt{\frac{\alpha(1-\alpha)}{N}}$. Avec $N = 10\,000$ réalisations, l'écart-type $\sqrt{\frac{\alpha(1-\alpha)}{N}}$ prend les valeurs (en %) 0.099% au niveau nominal 1%, 0.218% à 5% et 0.300% à 10%.

Dans l'étude de puissance, nous voulons plutôt un intervalle de confiance, disons de niveau 95%, pour la vraie valeur de la puissance p . La valeur p étant inconnue, l'écart-type $\sqrt{\frac{p(1-p)}{N}}$ est borné par $\frac{1}{2\sqrt{N}}$ qui correspond à l'écart-type

lorsque $p = \frac{1}{2}$. Ainsi, l'intervalle $\hat{p} \pm 1.96 \frac{1}{2\sqrt{N}}$ est un intervalle de confiance pour la vraie puissance p de niveau supérieur ou égal à 95%.

4.3. RÉSULTATS DE LA SIMULATION

4.3.1. Discussion des résultats de l'étude de niveau

Les résultats de l'étude de niveau sont présentés dans les tableaux 4.3.1 à 4.3.5. Le paramètre S correspond au délai maximal m pour la statistique de Koch et Yang et au point de troncature T_n pour la statistique de Hong. Des illustrations graphiques correspondant à ces résultats pour $n = 100$ sont également données dans les figures 4.3.1 à 4.3.6. Les limites de signification à 5% sont illustrées en pointillés sur chaque graphique.

Ces résultats nous permettent de constater que pour les modèles $AR(1)$ et $AR(2)$, la distribution exacte de la statistique modifiée de Koch et Yang $\tilde{Q}_{KY}(m, i)$ est généralement bien approximée par sa distribution asymptotique pour les petites valeurs de i . Par contre lorsque $i = 12, 17$ et 23 , le niveau exact est assez loin du niveau nominal. L'approximation est meilleure quand la taille de la série augmente. Ceci est vrai pour les trois niveaux nominaux considérés.

Par ailleurs, nous constatons que l'approximation est meilleure au niveau 1% et elle se détériore avec l'augmentation de m . En effet, presque tous les niveaux empiriques sont situés à l'intérieur des limites de signification du niveau nominal 1% et ceci pour les différentes tailles échantillonnales. Plus la taille de la série augmente, moins on retrouve des valeurs à l'extérieur des limites de signification.

Au niveaux nominaux 5% et 10%, le niveau du test est moins bien contrôlé quand les valeurs de m et de i sont toutes les deux grandes. Pour ces valeurs, le

test de Koch et Yang sous-rejette l'hypothèse nulle.

Concernant la statistique de Hong ($Q_H(w, T_n)$), les résultats de l'étude de niveau montrent que l'approximation de la distribution exacte de la statistique par une $N(0, 1)$ est bonne pour les niveaux nominaux 5% et 10%.

Au niveau 1%, le test tend à sur-rejeter l'hypothèse nulle. Pour ce niveau nominal, nous constatons que plus la taille de la série augmente plus le niveau empirique s'approche du niveau nominal. Par ailleurs, pour une même taille de série, plus le point de troncature augmente, plus le niveau empirique dépasse le niveau nominal. En d'autres termes, le niveau est mieux contrôlé pour des points de troncature de petites valeurs. Pour des tailles fixes de la série et pour les mêmes valeurs de T_n , le noyau tronqué sur-rejette davantage l'hypothèse nulle comparativement aux noyaux de Bartlett, Daniell et Parzen.

Au niveaux 5%, nous remarquons que le niveau est relativement bien contrôlé. La taille de la série semble ne pas avoir une grande influence sur le niveau puisque nous aboutissons à des résultats relativement similaires pour les différentes tailles des séries ; sauf peut être pour le noyau uniforme tronqué qui donne la meilleure approximation pour des valeurs de T_n plus grandes que 12 et quand la taille de la série est de 500. Pour une même taille de série, le choix du point de troncature a un effet sur le niveau dans la mesure où ce dernier se rapproche du niveau nominal pour les grandes valeurs de T_n . Les fenêtres de Bartlett, Daniell et Parzen conduisent à des résultats semblables.

Au niveau nominal 10%, les résultats sont similaires à ceux obtenus au niveau 5% et les mêmes constatations ont été faites.

Il faut noter que les résultats des modèles $AR(1)$ et $AR(2)$ sont presque similaires pour les deux statistiques du test. Cependant, aux niveaux 5% et 10%, le modèle $AR(2)$ donne de meilleurs résultats que le modèle $AR(1)$ pour le test de Koch et Yang ; on retrouve pour ces deux niveaux moins de valeurs à l'extérieur des limites de signification.

Le modèle bruit blanc a été traité dans l'objectif d'obtenir les valeurs critiques exactes qui serviront pour évaluer les puissances basées sur les valeurs critiques exactes des modèles moyenne mobile considérés dans l'étude de puissance. Pour cette raison, une seule taille de série a été choisie, à savoir $n = 100$, puisque dans l'étude de puissance nous nous sommes limités à $n = 100$.

Les résultats des simulations pour ce modèle, qui sont présentés dans le tableau 4.3.5, montrent que pour le test de Koch et Yang, le niveau est bien contrôlé aux seuils de signification 5% et 10% quand $i = 0$. Pour les valeurs de i supérieures à zéro, le niveau est mal contrôlé. A 1%, le niveau réel est proche du niveau nominal pour $m = 6$, $m = 12$ avec $i < 6$ et $m = 24$ avec i compris entre 2 et 6.

Concernant la statistique de Hong, au niveau nominal 1%, le noyau uniforme tronqué donne une bonne approximation pour les petites valeurs de T_n ($T_n = 3, 6$). Pour $T_n = 12$, le niveau est bien contrôlé pour toutes les fenêtres sauf pour le noyau uniforme tronqué qui sur-rejette l'hypothèse nulle. A $T_n = 18$, c'est le noyau de Bartlett qui donne la meilleure approximation. Au niveau nominal 5%, le niveau est bien contrôlé pour le noyau uniforme tronqué avec les valeurs de T_n plus grandes ou égales à 12 et pour les fenêtres Daniell et Parzen avec $T_n = 24$. Au niveau 10%, le niveau du test de Hong est mal contrôlé.

TABLEAU 4.3.1. Niveaux empiriques (%) du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et du test de Hong $Q_H(w, T_n)$ pour différents noyaux et points de troncation aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(1)$ avec $\phi = 0.8$ et $n = 50, 100$.

| | | Alpha % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----|-----------------|-------------|-----|-----|--------------|-----|-----|-----|-------------|-----|-----|--------------|-----|-----|-----|-------------|-----|-----|--------------|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|-----|-----|-----|-----|
| | | 1 | | | | | | | 5 | | | | | | | 10 | | | | | | | | | | | | | | | | | | | | |
| | | Hong | | | Koch et Yang | | | | Hong | | | Koch et Yang | | | | Hong | | | Koch et Yang | | | | | | | | | | | | | | | | | |
| | | BA DA PA TR | | | valeurs de i | | | | BA DA PA TR | | | valeurs de i | | | | BA DA PA TR | | | valeurs de i | | | | | | | | | | | | | | | | | |
| | | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | | | | | | | | |
| | | n=50 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 1.3 1.4 1.3 1.7 | - | - | - | - | - | - | - | - | - | 2.8 | 2.9 | 3.0 | 3.9 | - | - | - | - | - | 4.4 | 4.6 | 4.5 | 5.9 | - | - | - | - | - | | | | | | | | |
| 6 | 1.5 1.6 1.7 2.3 | 1.5 | 2.0 | 1.9 | - | - | - | - | - | - | 3.4 | 3.5 | 3.7 | 4.5 | 5.4 | 7.5 | 8.4 | - | - | - | 5.1 | 5.4 | 5.4 | 6.9 | 10.2 | 13.2 | 15.0 | - | - | | | | | | | |
| 12 | 1.9 2.2 2.1 2.6 | 1.9 | 2.2 | 1.7 | 1.5 | 1.3 | 1.1 | - | - | - | 4.1 | 4.4 | 4.5 | 5.1 | 5.5 | 6.9 | 6.6 | 6.6 | 5.9 | 5.2 | - | - | 6.1 | 6.4 | 6.6 | 7.9 | 10.1 | 11.7 | 11.4 | 11.8 | 12.0 | 12.0 | - | - | | |
| 18 | 2.2 2.4 2.4 3.0 | 2.3 | 2.5 | 1.8 | 1.0 | 0.9 | 0.8 | 0.7 | 0.4 | - | 4.5 | 4.9 | 5.1 | 6.1 | 6.4 | 6.5 | 5.6 | 4.7 | 4.2 | 3.5 | 2.9 | 1.9 | - | 6.7 | 7.3 | 7.4 | 8.6 | 10.4 | 11.0 | 9.8 | 8.8 | 8.4 | 7.8 | 7.1 | 6.4 | - |
| 24 | 2.4 2.6 2.6 3.3 | 2.6 | 2.9 | 1.9 | 0.9 | 0.5 | 0.4 | 0.3 | 0.3 | 0.1 | 4.9 | 5.2 | 5.3 | 6.6 | 6.9 | 7.0 | 5.4 | 3.9 | 2.9 | 2.3 | 2.0 | 1.1 | 0.5 | 7.2 | 7.7 | 8.1 | 9.5 | 11.1 | 11.0 | 8.9 | 7.4 | 6.0 | 5.4 | 4.8 | 3.9 | 2.8 |
| | | n=100 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 1.2 1.2 1.1 1.6 | - | - | - | - | - | - | - | - | - | 2.5 | 2.6 | 2.6 | 3.7 | - | - | - | - | - | - | - | 3.9 | 4.0 | 4.2 | 5.8 | - | - | - | - | - | - | - | - | | | |
| 6 | 1.3 1.4 1.5 1.9 | 1.3 | 1.7 | 1.6 | - | - | - | - | - | - | 3.0 | 3.2 | 3.3 | 4.4 | 5.2 | 6.6 | 6.8 | - | - | - | - | - | 4.7 | 5.0 | 5.2 | 6.8 | 10.1 | 12.4 | 13.2 | - | - | - | - | - | | |
| 12 | 1.6 1.8 1.8 2.3 | 1.5 | 2.2 | 1.7 | 1.1 | 1.1 | 1.0 | - | - | - | 3.7 | 4.1 | 4.2 | 5.1 | 5.7 | 6.5 | 6.5 | 6.4 | 4.4 | 4.8 | - | - | 5.9 | 6.3 | 6.5 | 8.0 | 10.3 | 11.5 | 11.4 | 11.1 | 11.1 | 10.8 | - | - | | |
| 18 | 1.8 2.0 2.1 2.8 | 2.0 | 2.2 | 1.5 | 0.9 | 0.8 | 1.0 | 0.9 | 0.7 | - | 4.2 | 4.6 | 4.7 | 5.7 | 6.0 | 6.5 | 5.9 | 5.1 | 4.5 | 4.2 | 3.2 | 2.3 | - | 6.6 | 7.3 | 7.2 | 8.6 | 10.1 | 11.1 | 10.5 | 9.6 | 9.3 | 9.0 | 8.7 | 7.0 | - |
| 24 | 2.1 2.3 2.4 2.9 | 2.2 | 2.5 | 1.6 | 0.9 | 0.5 | 0.5 | 0.6 | 0.7 | 0.6 | 4.6 | 5.1 | 5.1 | 6.4 | 6.7 | 6.7 | 5.4 | 4.4 | 3.5 | 3.1 | 2.7 | 2.0 | 1.3 | 7.1 | 7.6 | 7.8 | 9.5 | 11.0 | 11.0 | 9.7 | 8.4 | 7.4 | 7.1 | 6.8 | 5.6 | 4.6 |

TABLEAU 4.3.2. Niveaux empiriques (%) du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et du test de Hong $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(1)$ avec $\phi = 0.8$ et $n = 200, 500$.

| | | Alpha % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-------|-----------------|----------------|-----|-----|-----|-----|-----|-----|-------------|-----|-----|-----|-----|-----|-----|----------------|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|-----|-----|-----|-----|-----|
| | | 1 | | | | | | | 5 | | | | | | | 10 | | | | | | | | | | | | | | | | | | | | |
| Hong | | Koch et Yang | | | | | | | Hong | | | | | | | Koch et Yang | | | | | | | | | | | | | | | | | | | | |
| S | BA DA PA TR | valeurs de i | | | | | | | BA DA PA TR | | | | | | | valeurs de i | | | | | | | | | | | | | | | | | | | | |
| | | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | | | | | | | | |
| n=200 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 1.0 1.0 1.1 1.6 | - | - | - | - | - | - | - | 2.2 | 2.3 | 2.5 | 3.8 | - | - | - | - | - | - | 4.0 | 4.0 | 4.1 | 5.8 | - | - | - | - | - | - | - | - | | | | | | |
| 6 | 1.2 1.3 1.3 1.5 | 0.9 | 1.2 | 1.0 | - | - | - | - | 2.8 | 3.1 | 3.2 | 4.3 | 5.3 | 5.9 | 5.2 | - | - | - | 4.6 | 4.9 | 5.0 | 6.6 | 10.3 | 11.1 | 10.1 | - | - | - | - | - | | | | | | |
| 12 | 1.4 1.5 1.5 1.8 | 1.2 | 1.5 | 1.3 | 1.0 | 0.9 | 0.8 | - | 3.5 | 3.7 | 3.8 | 5.0 | 5.6 | 6.0 | 6.3 | 5.8 | 5.0 | 4.0 | - | 5.7 | 6.0 | 6.1 | 8.0 | 10.3 | 11.4 | 11.1 | 11.1 | 10.4 | 9.5 | - | - | | | | | |
| 18 | 1.4 1.5 1.6 2.2 | 1.6 | 1.6 | 1.4 | 0.8 | 0.9 | 1.1 | 1.1 | 0.8 | - | 3.8 | 4.2 | 4.3 | 5.4 | 5.7 | 5.8 | 5.6 | 5.4 | 5.1 | 4.4 | 3.6 | 2.6 | - | 6.1 | 7.0 | 7.1 | 8.6 | 10.6 | 10.8 | 10.5 | 10.1 | 9.7 | 9.6 | 9.3 | 7.5 | - |
| 24 | 1.6 1.7 1.8 2.5 | 1.8 | 2.0 | 1.5 | 0.9 | 0.8 | 0.7 | 1.1 | 1.3 | 1.2 | 4.2 | 4.5 | 4.5 | 5.7 | 6.0 | 6.1 | 5.5 | 4.9 | 4.4 | 4.2 | 3.1 | 2.7 | 2.3 | 6.8 | 7.6 | 7.5 | 9.1 | 10.7 | 10.6 | 10.0 | 9.4 | 8.8 | 8.4 | 8.3 | 7.0 | 5.6 |
| n=500 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 1.2 1.2 1.2 1.6 | - | - | - | - | - | - | - | 2.4 | 2.4 | 2.6 | 3.7 | - | - | - | - | - | - | 3.8 | 3.8 | 4.0 | 5.7 | - | - | - | - | - | - | - | - | | | | | | |
| 6 | 1.3 1.3 1.4 1.7 | 1.1 | 1.2 | 0.7 | - | - | - | - | 3.0 | 3.2 | 3.3 | 4.3 | 5.3 | 5.6 | 3.9 | - | - | - | 4.6 | 4.8 | 5.0 | 6.9 | 10.6 | 10.1 | 8.0 | - | - | - | - | - | | | | | | |
| 12 | 1.5 1.5 1.5 1.7 | 1.2 | 1.5 | 1.2 | 0.9 | 0.9 | 0.7 | - | 3.5 | 3.7 | 3.7 | 4.9 | 5.5 | 5.8 | 5.6 | 4.9 | 4.1 | 3.0 | - | 5.7 | 6.1 | 6.3 | 7.9 | 10.4 | 11.2 | 10.2 | 9.5 | 8.9 | 7.6 | - | - | | | | | |
| 18 | 1.4 1.6 1.5 1.7 | 1.1 | 1.3 | 1.1 | 0.8 | 0.9 | 0.9 | 1.0 | 0.8 | - | 3.7 | 4.0 | 4.1 | 5.0 | 5.2 | 5.6 | 5.5 | 5.1 | 4.8 | 4.1 | 3.3 | 1.9 | - | 6.2 | 7.0 | 7.1 | 8.5 | 10.7 | 10.8 | 10.5 | 9.8 | 9.8 | 9.7 | 9.2 | 6.4 | - |
| 24 | 1.5 1.6 1.5 1.8 | 1.2 | 1.3 | 1.2 | 0.8 | 0.6 | 0.8 | 1.0 | 1.4 | 1.3 | 4.0 | 4.5 | 4.5 | 4.9 | 5.1 | 5.4 | 4.9 | 4.4 | 4.5 | 4.1 | 3.2 | 3.0 | 2.6 | 6.8 | 7.4 | 7.5 | 8.5 | 10.0 | 10.4 | 9.8 | 9.5 | 9.1 | 9.4 | 9.3 | 7.5 | 5.8 |

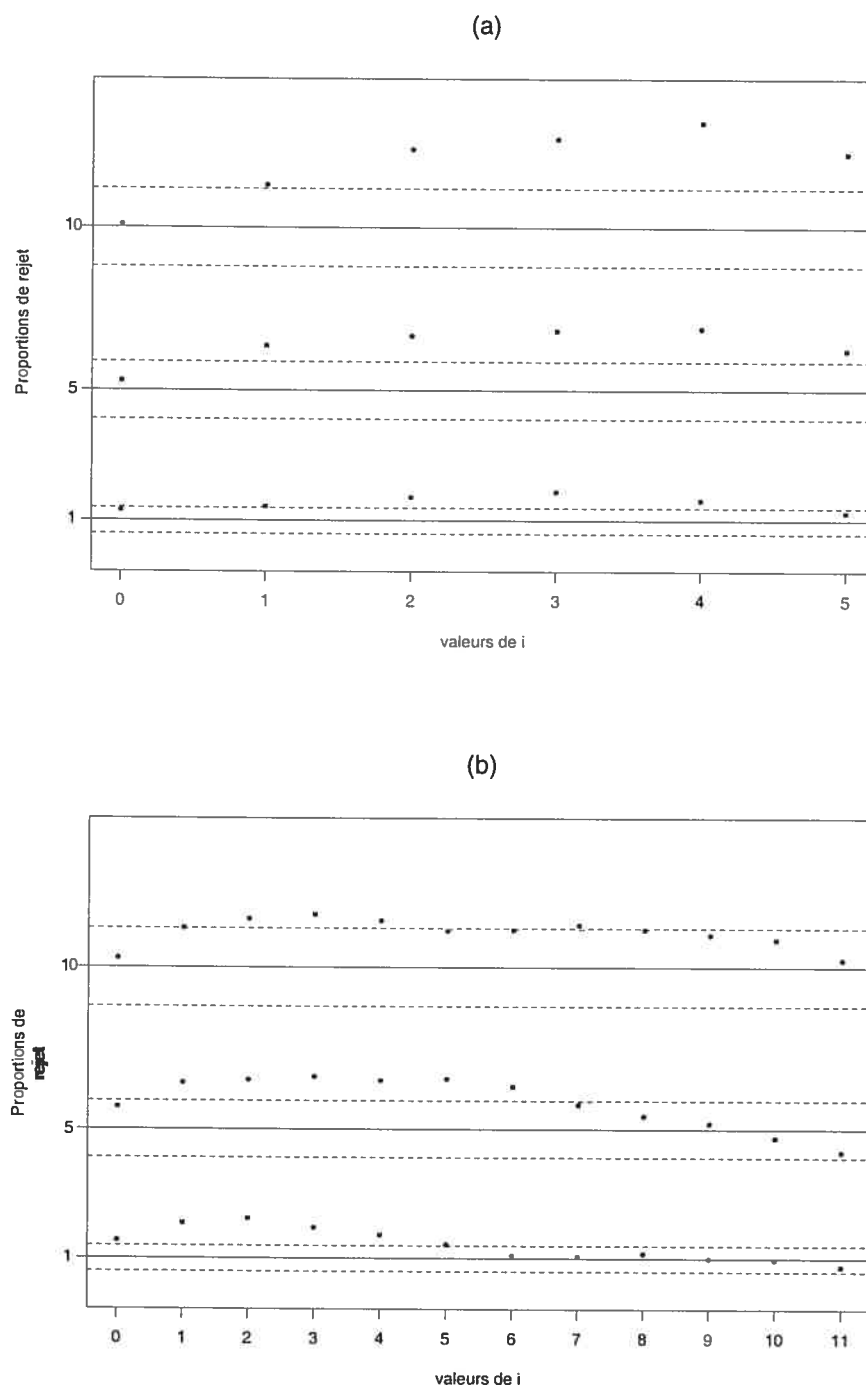


FIGURE 4.3.1. Niveaux empiriques (%) de $\tilde{Q}_{KY}(m, i)$, $i = 0, \dots, m-1$ pour (a) $m = 6$ et (b) $m = 12$, aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(1)$ avec $\phi = 0.8$ et $n = 100$.

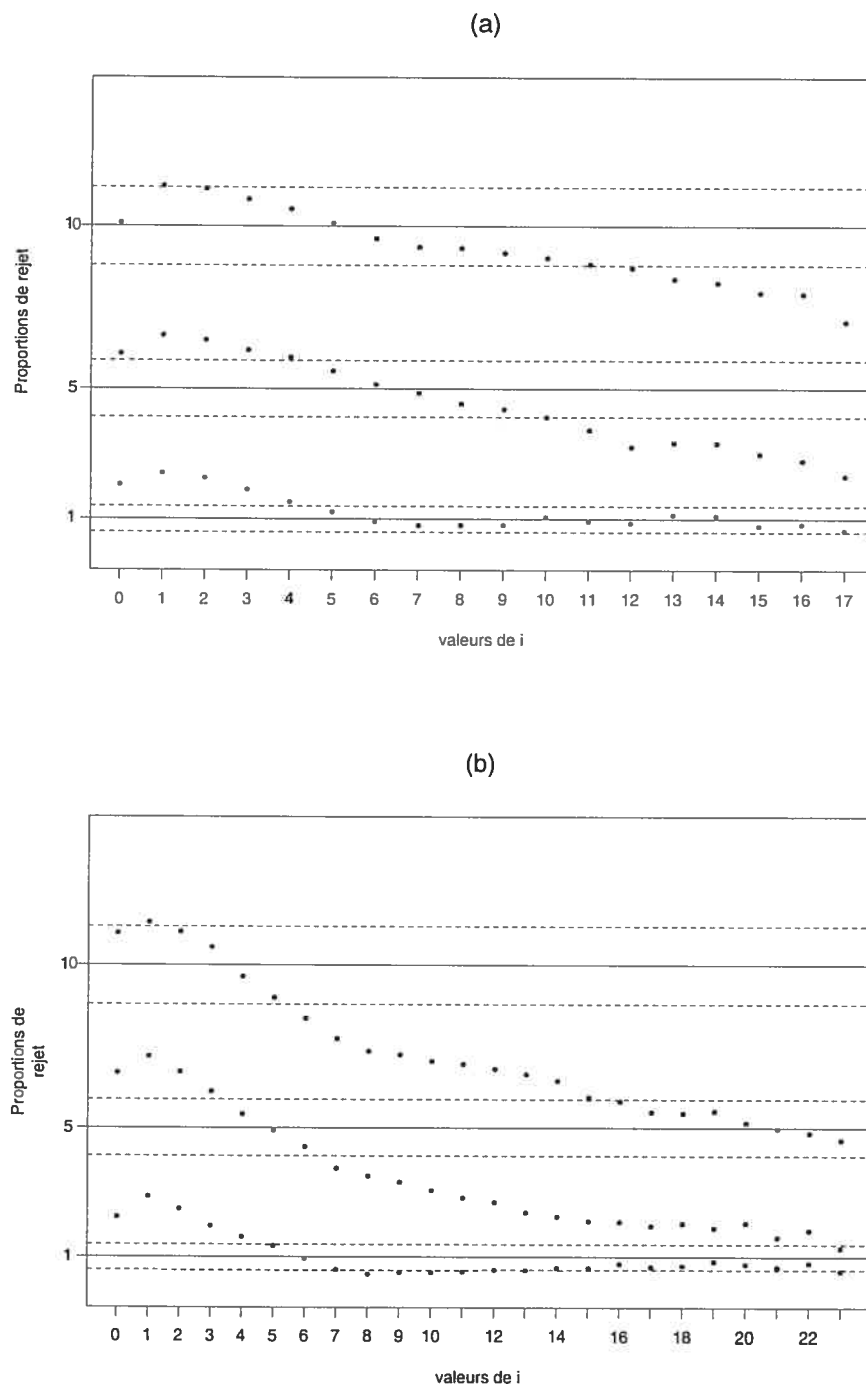


FIGURE 4.3.2. Niveaux empiriques (%) de $\tilde{Q}_{KY}(m, i)$, $i = 0, \dots, m - 1$ pour (a) $m = 18$ et (b) $m = 24$, aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(1)$ avec $\phi = 0.8$ et $n = 100$.

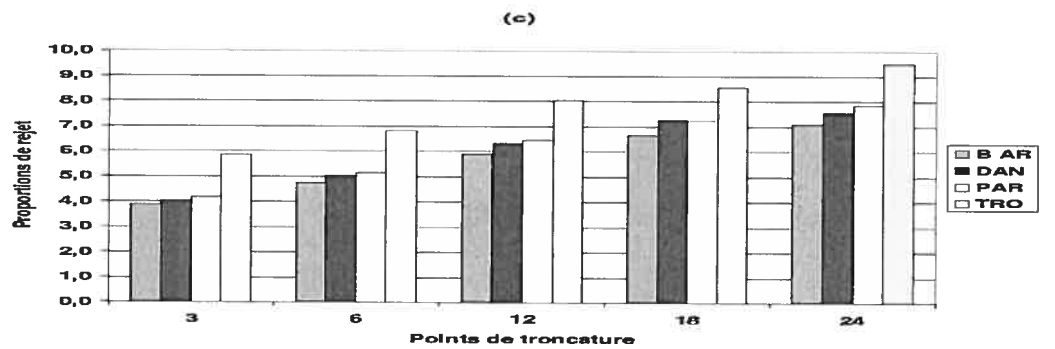
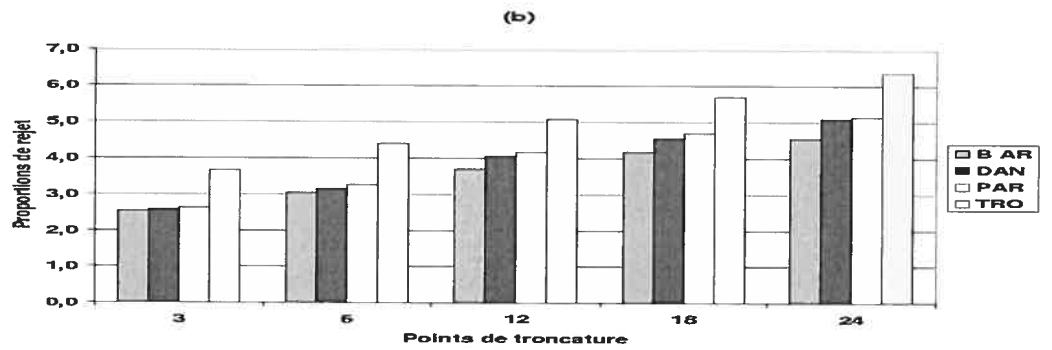
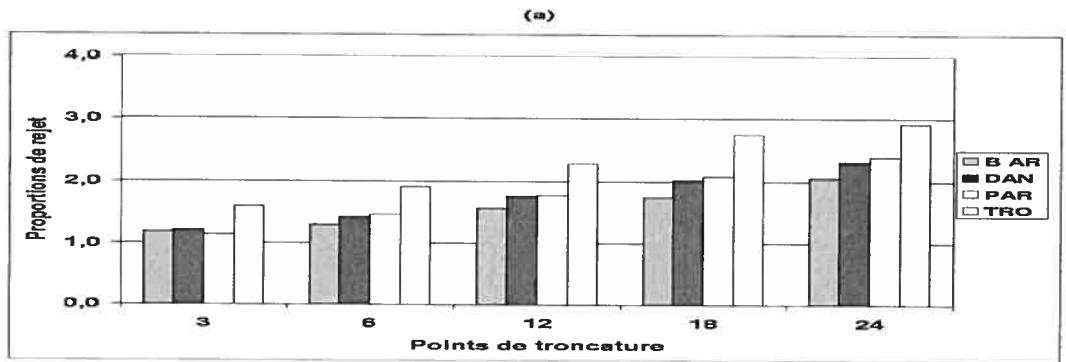


FIGURE 4.3.3. Niveaux empiriques (%) de $Q_H(w, T_n)$ aux niveaux nominaux (a) 1%, (b) 5% et (c) 10%. Le modèle est $AR(1)$ avec $\phi = 0.8$ et $n = 100$.

TABLEAU 4.3.3. Niveaux empiriques (%) du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et du test de Hong $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(2)$ avec $\phi_1 = 0.1$, $\phi_2 = 0.5$ et $n = 50, 100$.

| | | Alpha % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-------|-----------------|--------------|-----|-----|-----|-----|-----|-------------|-----|-----------------|-----------------|-----|-----|--------------|-----|-----|-----|-----|-----------------|-------------|------------------|------|------|------|------|--------------|-----|-----|-----|-----|--|--|--|--|--|--|--|
| | | 1 | | | | | | | | | | | | 5 | | | | | | | | | | | | 10 | | | | | | | | | | | |
| Hong | | Koch et Yang | | | | | | Hong | | | | | | Koch et Yang | | | | | | Hong | | | | | | Koch et Yang | | | | | | | | | | | |
| S | | valeurs de i | | | | | | BA DA PA TR | | | | | | valeurs de i | | | | | | BA DA PA TR | | | | | | valeurs de i | | | | | | | | | | | |
| | | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | | | | | | | | | |
| n=50 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 1.2 1.3 1.2 1.6 | - | - | - | - | - | - | - | - | 3.0 3.1 3.2 4.0 | - | - | - | - | - | - | - | - | 4.5 4.8 6.3 6.5 | - | - | - | - | - | - | - | - | | | | | | | | | | |
| 6 | 1.5 1.8 1.9 2.4 | 1.0 | 1.6 | 2.0 | - | - | - | - | - | 4.2 4.4 4.7 5.2 | 5.0 | 7.6 | 9.2 | - | - | - | - | - | 5.5 6.5 6.8 6.9 | 9.9 | 13.5 | 16.5 | - | - | - | - | - | | | | | | | | | | |
| 12 | 1.8 2.1 2.0 2.5 | 1.5 | 1.7 | 1.3 | 1.3 | 1.3 | 0.7 | - | - | 4.3 4.6 4.9 5.3 | 5.3 | 5.8 | 5.8 | 5.7 | 5.2 | 4.5 | - | - | 6.7 6.9 7.3 7.7 | 9.3 | 10.4 | 10.7 | 11.0 | 10.9 | 10.8 | - | - | | | | | | | | | | |
| 18 | 2.1 2.2 2.4 3.0 | 1.7 | 1.7 | 1.3 | 0.7 | 0.6 | 0.5 | 0.7 | 0.2 | - | 4.7 5.0 5.1 5.4 | 5.7 | 5.6 | 4.7 | 3.9 | 3.3 | 2.4 | 2.0 | 1.4 | - | 6.8 7.5 7.7 8.9 | 9.4 | 9.6 | 8.7 | 7.6 | 7.2 | 6.4 | 5.7 | 4.6 | - | | | | | | | |
| 24 | 2.2 2.4 2.5 3.1 | 2.0 | 2.0 | 1.3 | 0.7 | 0.4 | 0.3 | 0.3 | 0.2 | 0.1 | 5.0 5.1 5.4 5.8 | 5.5 | 5.7 | 4.4 | 3.0 | 2.2 | 1.5 | 1.2 | 0.6 | 0.4 | 7.8 8.0 8.5 10.0 | 9.4 | 9.5 | 7.9 | 6.2 | 5.0 | 4.4 | 3.8 | 2.6 | 2.0 | | | | | | | |
| n=100 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 1.1 1.1 1.0 1.5 | - | - | - | - | - | - | - | - | 2.8 2.9 2.8 3.9 | - | - | - | - | - | - | - | - | 4.2 4.3 4.5 6.1 | - | - | - | - | - | - | - | - | | | | | | | | | | |
| 6 | 1.1 1.3 1.4 1.8 | 0.9 | 1.6 | 1.4 | - | - | - | - | - | 3.2 3.4 3.5 4.7 | 5.0 | 5.7 | 6.9 | - | - | - | - | - | 4.4 5.1 5.3 7.6 | 10.3 | 11.8 | 12.0 | - | - | - | - | - | | | | | | | | | | |
| 12 | 1.5 1.7 1.7 2.2 | 1.0 | 1.4 | 1.2 | 0.9 | 1.2 | 0.7 | - | - | 4.0 4.5 4.8 5.2 | 4.7 | 5.4 | 5.6 | 5.3 | 4.5 | 3.2 | - | - | 6.4 6.8 7.0 7.9 | 9.0 | 10.2 | 10.4 | 10.5 | 10.3 | 9.1 | - | - | - | | | | | | | | | |
| 18 | 1.7 1.8 1.9 2.5 | 1.3 | 1.5 | 1.2 | 0.7 | 0.8 | 0.8 | 0.9 | 0.7 | - | 4.6 4.9 5.3 5.3 | 5.1 | 5.4 | 4.8 | 4.5 | 4.0 | 3.1 | 2.6 | 2.1 | - | 7.2 7.7 7.9 9.0 | 9.5 | 9.9 | 9.5 | 8.8 | 8.7 | 8.5 | 7.8 | 6.0 | - | | | | | | | |
| 24 | 1.8 2.0 2.1 2.7 | 1.7 | 1.8 | 1.3 | 0.7 | 0.4 | 0.4 | 0.5 | 1.0 | 0.6 | 4.8 5.1 5.1 5.4 | 5.5 | 5.7 | 4.9 | 3.8 | 2.9 | 2.5 | 1.8 | 1.9 | 1.2 | 7.5 7.8 8.3 9.8 | 9.6 | 10.1 | 8.7 | 7.8 | 7.0 | 6.6 | 6.2 | 4.5 | 3.4 | | | | | | | |

TABLEAU 4.3.4. Niveaux empiriques (%) du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et du test de Hong $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(2)$ avec $\phi_1 = 0.1$, $\phi_2 = 0.5$ et $n = 200$, 500.

| | | Alpha % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-------|-----------------|-------------|-----|-----|-----|--------------|-----|-----|-----|-------------|-----|-----|-----|--------------|-----|-----|-----|-------------|-----|-----|-----|--------------|-----|------|------|-----|----|----|-----|-----|-----|-----|--|--|--|--|--|
| | | 1 | | | | | | | | | | | | 5 | | | | | | | | | | | | 10 | | | | | | | | | | | |
| | | Hong | | | | Koch et Yang | | | | Hong | | | | Koch et Yang | | | | Hong | | | | Koch et Yang | | | | | | | | | | | | | | | |
| | | BA DA PA TR | | | | valeurs de i | | | | BA DA PA TR | | | | valeurs de i | | | | BA DA PA TR | | | | valeurs de i | | | | | | | | | | | | | | | |
| | | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | | | | | | | | | |
| n=200 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 1.1 1.1 1.2 1.7 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 5.0 | 5.0 | 5.1 | 6.8 | | | | | |
| 6 | 1.3 1.4 1.5 1.6 | 0.8 | 1.3 | 0.7 | - | - | - | - | - | 2.7 | 3.1 | 3.3 | 4.5 | 5.4 | 6.0 | 4.8 | - | - | - | 5.6 | 5.9 | 6.0 | 7.6 | 10.6 | 11.2 | 9.6 | - | - | - | - | - | | | | | | |
| 12 | 1.3 1.4 1.4 1.7 | 1.1 | 1.5 | 1.2 | 1.1 | 0.9 | 0.5 | - | - | 4.5 | 4.7 | 4.8 | 5.0 | 5.3 | 5.8 | 5.5 | 4.9 | 4.0 | 3.1 | - | - | - | - | - | - | - | - | - | - | - | | | | | | | |
| 18 | 1.3 1.4 1.5 2.1 | 1.3 | 1.4 | 1.3 | 0.8 | 1.0 | 1.3 | 1.2 | 0.8 | - | 4.8 | 5.2 | 5.3 | 5.2 | 5.2 | 5.6 | 5.3 | 4.9 | 4.5 | 3.6 | 3.0 | 2.2 | - | - | - | - | - | - | - | - | | | | | | | |
| 24 | 1.5 1.6 1.7 2.4 | 1.4 | 1.6 | 1.2 | 0.8 | 0.5 | 0.7 | 1.0 | 1.4 | 1.0 | 4.8 | 4.9 | 4.9 | 5.3 | 5.4 | 5.6 | 5.0 | 4.2 | 3.9 | 3.0 | 2.6 | 2.9 | 2.0 | - | - | - | - | - | - | - | | | | | | | |
| n=500 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 1.1 1.1 1.1 1.5 | - | - | - | - | - | - | - | - | 3.4 | 3.4 | 3.6 | 4.7 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | | | | | | | |
| 6 | 1.2 1.2 1.3 1.6 | 1.0 | 1.0 | 0.4 | - | - | - | - | - | 4.0 | 4.2 | 4.4 | 4.9 | 5.3 | 5.2 | 3.0 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | | | | | | |
| 12 | 1.4 1.4 1.4 1.6 | 0.9 | 1.2 | 1.0 | 1.1 | 0.9 | 0.7 | - | - | 4.5 | 4.7 | 4.7 | 5.0 | 5.1 | 5.8 | 5.5 | 4.9 | 3.6 | 2.8 | - | - | - | - | - | - | - | - | - | - | - | - | | | | | | |
| 18 | 1.3 1.5 1.4 1.6 | 1.1 | 1.3 | 1.2 | 0.8 | 1.0 | 1.3 | 1.6 | 0.9 | - | 4.7 | 4.9 | 4.9 | 5.2 | 5.4 | 5.5 | 5.2 | 4.9 | 4.8 | 3.3 | 3.1 | 2.3 | - | - | - | - | - | - | - | - | - | | | | | | |
| 24 | 1.4 1.5 1.4 1.7 | 1.0 | 1.3 | 1.2 | 0.9 | 0.7 | 0.9 | 1.2 | 2.0 | 1.5 | 4.8 | 4.9 | 4.9 | 5.1 | 5.3 | 5.5 | 5.1 | 4.8 | 4.5 | 3.7 | 3.1 | 3.6 | 2.8 | - | - | - | - | - | - | - | - | | | | | | |

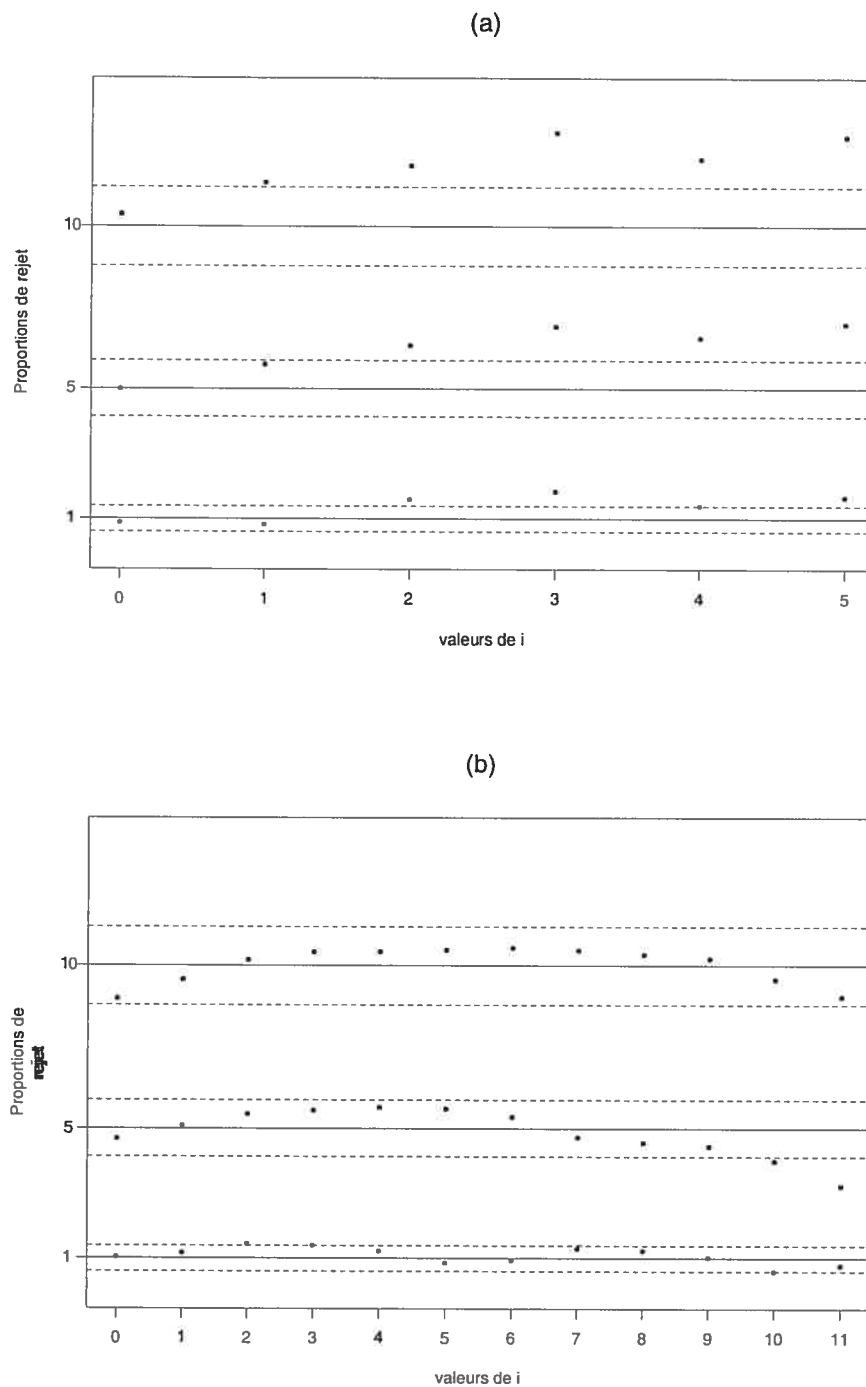


FIGURE 4.3.4. Niveaux empiriques (%) de $\tilde{Q}_{KY}(m, i)$, $i = 0, \dots, m-1$ pour (a) $m = 6$ et (b) $m = 12$, aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(2)$ avec $\phi_1 = 0.1$, $\phi_2 = 0.5$ et $n = 100$.

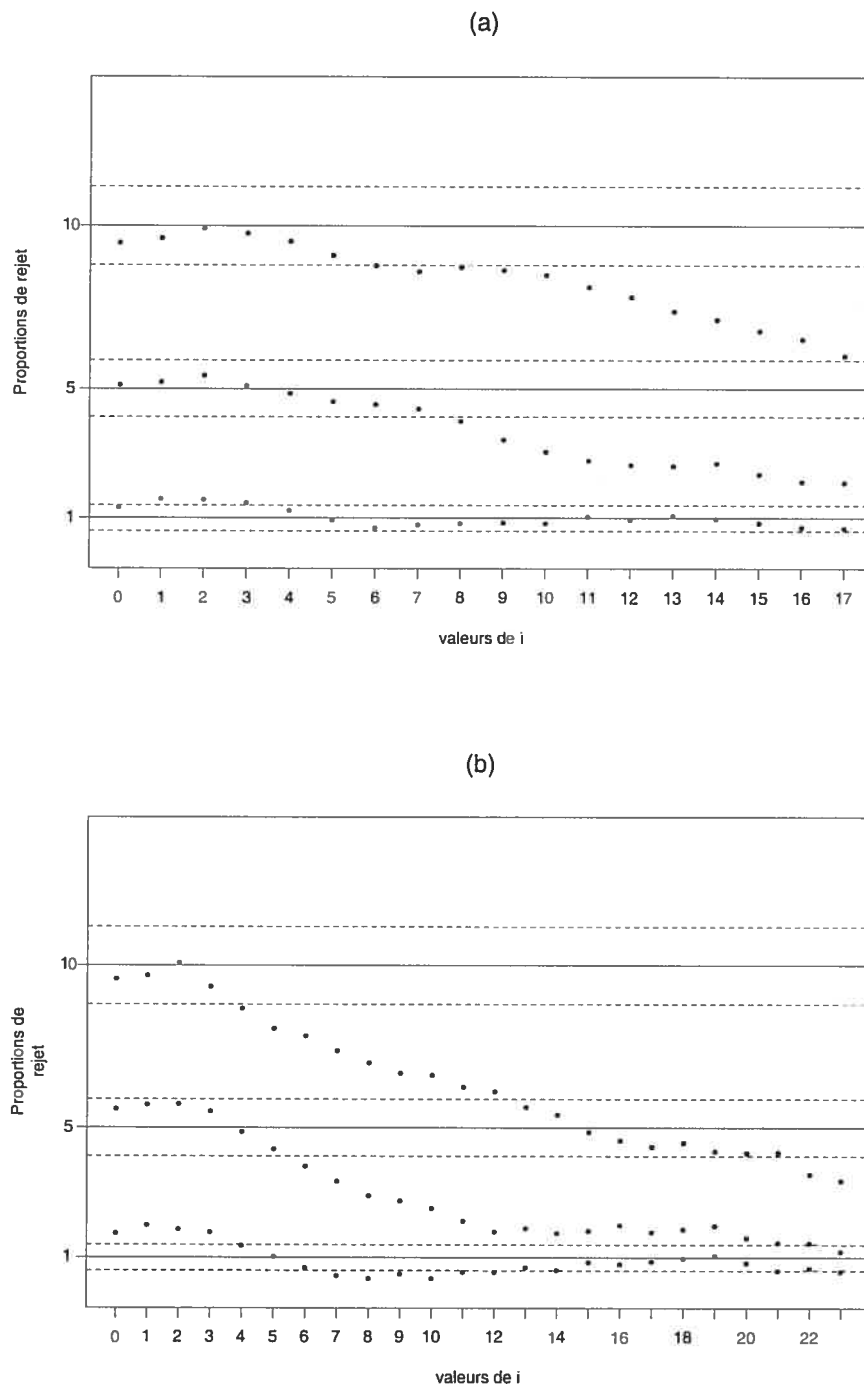


FIGURE 4.3.5. Niveaux empiriques (%) de $\tilde{Q}_{KY}(m, i)$, $i = 0, \dots, m-1$ pour (a) $m = 18$ et (b) $m = 24$, aux niveaux nominaux 1%, 5% et 10%. Le modèle est $AR(2)$ avec $\phi_1 = 0.1$, $\phi_2 = 0.5$ et $n = 100$.

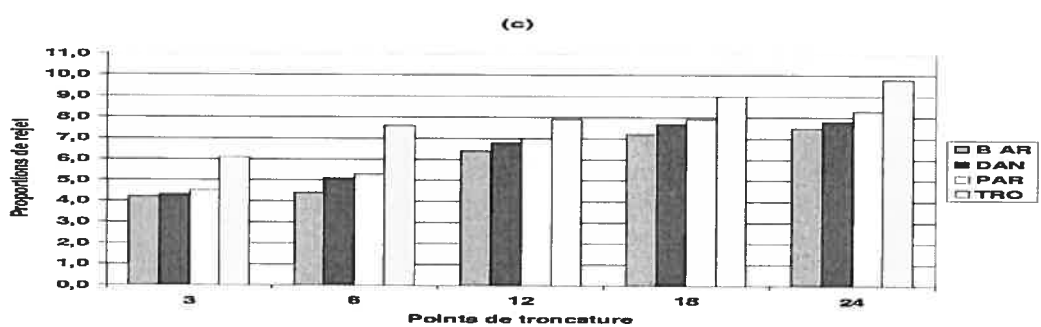
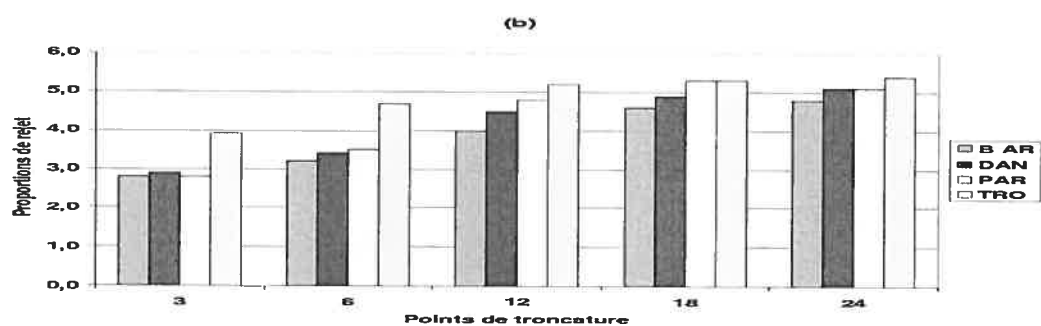
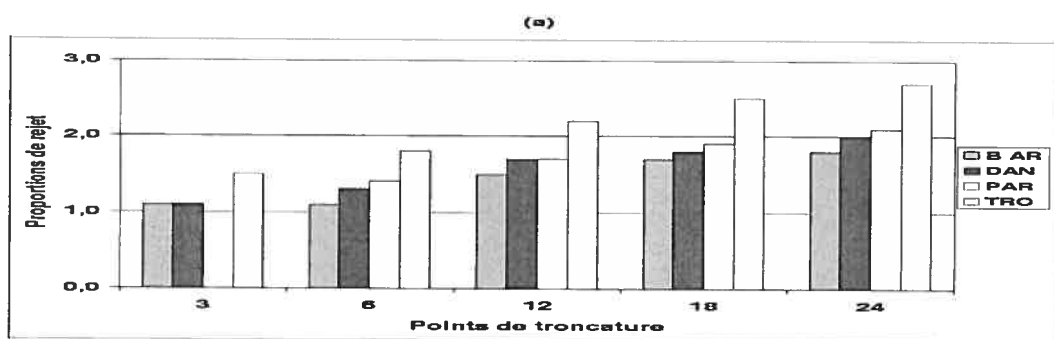


FIGURE 4.3.6. Niveaux empiriques (%) de $Q_H(w, T_n)$ aux niveaux nominaux (a) 1%, (b) 5% et (c) 10%. Le modèle est $AR(2)$ avec $\phi_1 = 0.1$, $\phi_2 = 0.5$ et $n = 100$.

TABLEAU 4.3.5. Niveaux empiriques (%) du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et du test de Hong $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est un bruit blanc avec $n = 100$.

| | | Alpha % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---------------|-----------------|--------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|--------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|--------------|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|----|---|---|---|
| | | 1 | | | | | | | | | | 5 | | | | | | | | | | 10 | | | | | | | | | | | | | | | | | | |
| Hong | | Koch et Yang | | | | | | | | | | Koch et Yang | | | | | | | | | | Koch et Yang | | | | | | | | | | | | | | | | | | |
| S BA DA PA TR | | valeurs de i | | | | | | | | | | valeurs de i | | | | | | | | | | valeurs de i | | | | | | | | | | | | | | | | | | |
| | | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 17 | 23 | | | |
| n=100 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 0.0 0.0 0.1 1.3 | - | - | - | - | - | - | - | - | - | 0.0 | 0.0 | 0.2 | 2.8 | - | - | - | - | - | - | - | 0.1 | 0.1 | 0.3 | 4.6 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | |
| 6 | 0.3 0.5 0.6 1.4 | 1.0 | 0.8 | 1.2 | - | - | - | - | - | - | 0.9 | 1.4 | 1.5 | 3.6 | 4.7 | 3.1 | 2.8 | - | - | - | - | - | 1.6 | 2.4 | 2.8 | 5.9 | 9.2 | 8.4 | 7.4 | - | - | - | - | - | - | - | - | - | | |
| 12 | 0.9 1.2 1.3 2.0 | 1.4 | 0.7 | 0.8 | 2.4 | 2.1 | 1.3 | - | - | - | 2.4 | 3.1 | 3.2 | 4.5 | 5.0 | 3.4 | 2.5 | 2.9 | 2.5 | 2.4 | - | - | 3.9 | 4.9 | 5.0 | 7.2 | 9.5 | 7.7 | 6.8 | 5.0 | 5.4 | 5.6 | - | - | - | - | - | - | | |
| 18 | 1.2 1.6 1.7 2.5 | 2.0 | 0.9 | 0.5 | 1.5 | 2.9 | 3.2 | 2.6 | 0.7 | - | 3.2 | 3.9 | 4.0 | 5.3 | 5.7 | 3.9 | 3.4 | 2.7 | 3.4 | 3.5 | 2.9 | 1.6 | - | 4.9 | 6.0 | 6.2 | 8.1 | 10.0 | 7.7 | 6.7 | 5.5 | 4.2 | 4.1 | 3.2 | 3.2 | - | - | - | - | - |
| 24 | 1.6 2.0 2.1 2.5 | 2.1 | 1.1 | 1.0 | 1.1 | 2.1 | 2.5 | 1.3 | 3.4 | 0.8 | 3.8 | 4.4 | 4.6 | 5.5 | 5.9 | 4.4 | 3.9 | 3.0 | 3.0 | 3.7 | 4.1 | 5.3 | 1.4 | 5.8 | 6.6 | 6.9 | 8.5 | 10.0 | 7.8 | 7.3 | 6.7 | 5.0 | 5.7 | 6.5 | 5.9 | 2.9 | | | | |

4.3.2. Discussion des résultats de l'étude de puissance

La puissance des tests étudiés a été calculée en utilisant les valeurs critiques asymptotiques (VCA) et les valeurs critiques exactes (VCE) que nous avons obtenu à partir de l'étude de niveau. L'obtention des VCE se justifie par ce qui suit :

La représentation autorégressive infinie des processus considérés s'écrit

$$X_t = \sum_{j=1}^{\infty} \pi_j X_{t-j} + a_t ,$$

ou de façon équivalente

$$\Pi(B)X_t = a_t ,$$

où $\Pi(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j$.

Dans le cas d'un $ARMA(p, q)$ stationnaire et inversible

$$\Phi(B)X_t = \Theta(B)a_t$$

$$\Pi(B) = \Theta(B)^{-1}\Phi(B) \tag{4.3.1}$$

De cette dernière relation, on voit que pour les deux modèles $MA(4)$ et pour le modèle $MA(15)$, $\pi_1 = 0$. Par conséquent, sous l'hypothèse que le vrai modèle est $AR(1)$, on devrait avoir $\phi_1 = 0$. Pour cette raison, pour les modèles MA , les VCE proviennent de la simulation exécutée pour l'obtention du tableau 4.3.5 dans laquelle le vrai modèle est un bruit blanc.

Avec le modèle $ARMA(1, 4)$, de (4.3.1), nous avons $\pi_1 = \phi_1$ et puisque $\phi_1 = 0.8$, les VCE proviennent de la simulation effectuée pour l'obtention du tableau 4.3.1.

Les résultats des deux modèles $MA(4)$, du modèle $MA(15)$ et du modèle $ARMA(1, 4)$ sont présentés respectivement dans les tableaux 4.3.6, 4.3.7, 4.3.8 et 4.3.9. Les valeurs entre parenthèses correspondent aux puissances basées sur les

valeurs critiques asymptotiques.

La première remarque à faire en analysant ces données c'est que les puissances basées sur les VCA et celles basées sur les VCE pour tous les modèles étudiés ne sont en écart que lorsque le niveau n'est pas bien contrôlé. Il faut noter également que le modèle $MA(4)$ avec $\theta_4 = 0.8$ et le modèle $ARMA(1, 4)$ donnent des résultats rapprochés. Par ailleurs, les puissances obtenues pour le modèle $MA(4)$ avec $\theta_4 = 0.4$ et le modèle $MA(15)$, sont plus petites.

Pour ce qui est du test de Koch et Yang, nous remarquons que la puissance maximale est atteinte quand $i = 0$, c'est-à-dire quand le test correspond à celui de Box-Pierce. Plus la valeur de i augmente, plus la puissance diminue sauf pour le modèle $MA(15)$ avec $m = 12$. Pour une valeur de i fixée, la puissance du test est maximale quand le délai m est petit et elle décroît, dans la plupart des cas, avec ce dernier. Ceci est vrai pour les 3 niveaux nominaux et pour les valeurs de i qui sont inférieures à 6.

En général, pour des valeurs de i et de m fixées, la puissance maximale est obtenue à 10%, suivie de 5%. Les puissances obtenues à 1% sont plus petites, ce qui n'est pas surprenant.

Concernant la statistique de Hong, nous constatons que la puissance se comporte de la même manière pour les différents noyaux. En effet, à l'exception du noyau uniforme tronqué, dont la puissance est maximale quand T_n est égal à 6, et qui diminue quand le point de troncature augmente, les autres noyaux aboutissent à des puissances qui augmentent avec le point de troncature. Mais quand ce dernier atteint un certain niveau (plus que 18), la puissance tend à se stabiliser. Ce résultat reste vrai pour les trois niveaux nominaux considérés qui donnent des puissances maximales à 10%.

Nous constatons également que les puissances les plus faibles sont enregistrées pour tous les noyaux quand $T_n = 3$. Pour cette valeur, le noyau uniforme tronqué dépasse les autres noyaux pour les deux modèles $MA(4)$ alors que c'est le noyau de Daniell qui est plus puissant pour le modèle $ARMA(1, 4)$.

A $T_n = 6$, c'est toujours le noyau uniforme tronqué qui a la plus forte puissance pour tous les modèles considérés. Pour $T_n = 12$, le noyau Parzen est supérieur aux autres noyaux qui donnent des résultats rapprochés. Pour les valeurs de T_n plus grandes que 12, le noyau uniforme tronqué commence à perdre de sa puissance, ce qui est conforme avec ce que nous avons obtenu pour le test de Koch et Yang, alors que les autres noyaux continuent à donner des puissances rapprochées avec une supériorité du noyau Daniell.

Le tableau 4.3.10 récapitule pour le niveau nominal 5%, les principaux résultats de l'étude de puissance. Il présente, en effet, pour chacun des modèles étudiés et pour chacune des valeurs de S , les tests les plus puissants. Notons qu'à $T_n = 3$, la test de Hong avec le noyau de Daniell est beaucoup plus puissant que les autres noyaux pour le modèle $MA(4)$ avec $\theta = 0.8$ et ceci aux trois niveaux nominaux.

TABLEAU 4.3.6. Puissances empiriques (%) basées sur les VCE et les VCA (entre parenthèses) de $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et de $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $MA(4)$ avec $\theta_4 = 0.8$ estimé par $AR(1)$, $n=100$.

| Alpha % | 1 | | | | | 5 | | | | | 10 | | | | | |
|-----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|---|
| | 3 | 6 | 12 | 18 | 24 | 3 | 6 | 12 | 18 | 24 | 3 | 6 | 12 | 18 | 24 | |
| S | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| KOCH-YANG | 94.8 (88.1) | 82.5 (81.1) | 68.7 (77.8) | 61.4 (72.8) | 88.8 (89.8) | 99.0 (97.0) | 97.8 (96.6) | 92.8 (93.4) | 88.8 (89.8) | 88.8 (89.8) | - | 99.7 (99.3) | 99.5 (98.8) | 97.5 (97.2) | 95.2 (94.8) | |
| | 68.6 (38.1) | 41.5 (31.1) | 27.1 (26.4) | 23.4 (24.8) | 51.5 (47.7) | 88.7 (62.2) | 76.6 (67.7) | 60.1 (53.9) | 51.5 (47.7) | 51.5 (47.7) | - | 94.7 (86.7) | 89.2 (84.6) | 76.3 (70.7) | 68.9 (62.5) | |
| | 28.1 (17.2) | 20.2 (12.6) | 19.1 (11.4) | 17.0 (14.3) | 36.4 (32.1) | 43.3 (26.0) | 45.3 (28.5) | 38.8 (31.5) | 36.4 (32.1) | 36.4 (32.1) | - | 53.7 (44.9) | 58.9 (48.5) | 52.3 (44.4) | 49.2 (42.9) | |
| | - | 14.3 (15.3) | 15.8 (12.1) | 15.4 (11.7) | 33.4 (22.6) | - | 31.1 (18.2) | 33.9 (18.4) | 33.4 (22.6) | 33.4 (22.6) | - | - | 45.1 (25.2) | 46.9 (30.1) | 46.4 (35.2) | |
| | - | 16.6 (12.9) | 13.4 (15.5) | 15.3 (15.1) | 31.4 (19.2) | - | 27.1 (14.4) | 30.0 (17.2) | 31.4 (19.2) | 31.4 (19.2) | - | - | 34.3 (20.0) | 42.2 (20.1) | 44.2 (25.3) | |
| | - | 14.6 (8.8) | 12.0 (15.3) | 14.0 (15.1) | 31.2 (19.9) | - | 22.3 (11.2) | 27.1 (16.4) | 31.2 (19.9) | 31.2 (19.9) | - | - | 27.9 (17.4) | 39.2 (17.9) | 42.8 (25.4) | |
| | - | - | 11.2 (12.4) | 12.2 (10.5) | 29.7 (18.9) | - | - | 24.2 (13.4) | 29.7 (18.9) | 29.7 (18.9) | - | - | - | 33.7 (14.7) | 40.7 (24.7) | |
| | - | - | 09.6 (05.7) | 10.1 (12.6) | 23.8 (17.4) | - | - | 18.2 (08.0) | 23.8 (17.4) | 23.8 (17.4) | - | - | - | 24.1 (12.2) | 33.9 (19.3) | |
| | - | - | - | 08.0 (04.9) | 17.0 (07.2) | - | - | - | 17.0 (07.2) | 17.0 (07.2) | - | - | - | - | 22.9 (10.8) | |
| HONG | 11.9 (00.6) | 46.4 (25.9) | 92.0 (90.6) | 93.4 (94.6) | 91.9 (94.8) | 24.1 (01.7) | 83.7 (44.5) | 99.0 (96.9) | 99.1 (98.5) | 99.1 (98.5) | 34.2 (02.9) | 94.4 (58.3) | 99.8 (98.5) | 99.9 (99.3) | 99.8 (99.4) | |
| DA | 49.8 (02.8) | 47.3 (33.3) | 94.3 (95.0) | 93.8 (96.5) | 91.4 (95.5) | 87.7 (07.5) | 84.0 (53.7) | 99.3 (98.4) | 98.9 (98.7) | 98.9 (98.7) | 97.1 (12.3) | 94.8 (66.9) | 99.9 (99.3) | 99.9 (99.6) | 99.8 (99.5) | |
| PA | 12.1 (02.4) | 74.8 (63.5) | 94.4 (95.7) | 93.0 (96.0) | 89.6 (94.7) | 25.0 (04.9) | 94.6 (81.8) | 99.5 (98.7) | 99.3 (98.9) | 98.6 (98.5) | 34.8 (07.1) | 98.6 (89.4) | 99.9 (99.5) | 99.9 (99.6) | 99.7 (99.4) | |
| TR | 07.5 (08.7) | 94.8 (96.1) | 84.3 (92.0) | 72.7 (86.0) | 64.5 (81.7) | 18.2 (13.5) | 99.0 (98.5) | 98.2 (97.7) | 94.1 (94.5) | 91.2 (92.1) | 26.5 (17.4) | 99.7 (99.2) | 99.6 (99.1) | 98.0 (97.1) | 96.4 (95.3) | |

TABLEAU 4.3.7. Puissances empiriques (%) basées sur les VCE et les VCA (entre parenthèses) de $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et de $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $MA(4)$ avec $\theta_4 = 0.4$ estimé par $AR(1)$, $n=100$.

| Alpha % | 1 | | | | 5 | | | | 10 | | | | | | |
|-----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| | 3 | 6 | 12 | 24 | 3 | 6 | 12 | 24 | 3 | 6 | 12 | 24 | | | |
| S | - | 48.7 | 29.3 | 22.9 | 20.5 | - | 75.2 | 58.4 | 50.4 | 46.2 | - | 84.7 | 73.1 | 64.5 | 60.7 |
| KOCH-YANG | - | (46.5) | (33.9) | (30.6) | (29.0) | - | (73.1) | (57.9) | (51.9) | (44.0) | - | (83.6) | (71.5) | (63.5) | (59.7) |
| 2 | - | 25.3 | 11.5 | 08.6 | 08.4 | - | 47.9 | 31.5 | 25.0 | 22.5 | - | 61.4 | 45.6 | 37.3 | 34.0 |
| 4 | - | (15.9) | (08.0) | (08.5) | (09.1) | - | (33.9) | (24.7) | (21.4) | (20.3) | - | (55.9) | (39.6) | (32.4) | (29.5) |
| 6 | - | 12.1 | 08.2 | 08.6 | 08.5 | - | 23.5 | 22.1 | 21.2 | 20.7 | - | 31.5 | 32.1 | 31.2 | 30.4 |
| 8 | - | (08.2) | (05.5) | (05.1) | (07.5) | - | (13.9) | (13.4) | (16.6) | (17.8) | - | (26.5) | (25.6) | (25.5) | (25.5) |
| 10 | - | - | 07.4 | 08.1 | 08.5 | - | - | 17.2 | 21.0 | 21.3 | - | - | 27.2 | 31.1 | 31.6 |
| 12 | - | - | (08.0) | (07.1) | (06.8) | - | - | (09.5) | (11.0) | (13.8) | - | - | (14.1) | (19.2) | (23.1) |
| 17 | - | - | 09.4 | 07.5 | 09.0 | - | - | 17.2 | 19.1 | 20.5 | - | - | 23.0 | 29.6 | 31.4 |
| 23 | - | - | (07.2) | (10.1) | (10.3) | - | - | (08.5) | (11.5) | (13.2) | - | - | (12.8) | (13.2) | (17.6) |
| BA | - | 08.6 | 07.0 | 18.2 | 18.2 | - | - | 14.9 | 17.7 | 20.8 | - | - | 20.0 | 27.6 | 30.9 |
| DA | - | (05.4) | (10.3) | (10.1) | (10.1) | - | - | (07.5) | (11.1) | (13.9) | - | - | (12.4) | (12.1) | (18.4) |
| PA | - | - | - | 06.9 | 07.1 | - | - | - | 16.4 | 20.3 | - | - | - | 24.2 | 29.9 |
| TR | - | - | (08.7) | (16.1) | (16.1) | - | - | - | (09.5) | (13.0) | - | - | - | (10.3) | (18.5) |
| | - | - | 06.2 | 05.7 | 05.7 | - | - | - | 13.4 | 16.8 | - | - | - | 18.6 | 25.5 |
| | - | - | (03.9) | (09.1) | (09.1) | - | - | - | (05.9) | (13.1) | - | - | - | (09.2) | (14.4) |
| | - | - | - | 04.9 | 04.9 | - | - | - | - | 12.8 | - | - | - | - | 18.3 |
| | - | - | (02.8) | (02.8) | (02.8) | - | - | - | - | (04.8) | - | - | - | - | (07.7) |
| HONG | 07.5 | 15.8 | 41.3 | 42.0 | 39.1 | 17.1 | 39.4 | 69.7 | 70.3 | 67.5 | 26.2 | 58.7 | 82.8 | 82.8 | 81.3 |
| | (00.2) | (08.4) | (38.4) | (44.9) | (45.8) | (00.6) | (15.3) | (55.4) | (61.2) | (61.8) | (01.3) | (21.6) | (64.9) | (70.1) | (70.7) |
| DA | 17.5 | 16.2 | 45.9 | 42.0 | 38.0 | 42.9 | 39.5 | 72.7 | 70.9 | 65.4 | 63.5 | 57.5 | 84.4 | 83.0 | 79.6 |
| | (00.6) | (10.9) | (48.2) | (49.9) | (47.8) | (01.9) | (19.4) | (63.7) | (65.4) | (63.4) | (03.4) | (26.6) | (72.0) | (73.9) | (72.1) |
| PA | 07.3 | 26.4 | 44.8 | 40.8 | 36.2 | 17.4 | 54.2 | 72.8 | 68.8 | 63.8 | 26.5 | 71.9 | 84.2 | 82.2 | 78.2 |
| | (01.0) | (19.1) | (48.3) | (48.7) | (46.6) | (02.4) | (32.3) | (63.9) | (64.4) | (62.2) | (03.8) | (42.0) | (72.8) | (73.0) | (71.2) |
| TR | 04.5 | 48.8 | 30.7 | 24.6 | 21.2 | 12.9 | 75.0 | 59.7 | 51.6 | 48.2 | 20.8 | 84.6 | 74.2 | 65.9 | 62.7 |
| | (05.3) | (54.8) | (41.4) | (37.4) | (35.6) | (09.1) | (69.5) | (57.3) | (52.7) | (50.3) | (12.1) | (77.3) | (67.1) | (61.6) | (59.4) |

TABLEAU 4.3.8. Puissances empiriques (%) basées sur les VCE et les VCA (entre parenthèses) de $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et de $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $MA(15)$ avec $\theta_{13} = 0.2$ et $\theta_{14} = 0.2$ et $\theta_{15} = 0.5$ et estimé par $AR(1)$, $n=100$.

| Alpha % | | 1 | | | | | 5 | | | | | 10 | | | | | |
|---------|-----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| | | 3 | 6 | 12 | 18 | 24 | 3 | 6 | 12 | 18 | 24 | 3 | 6 | 12 | 18 | 24 | |
| S | Koch-Yang | 0 | - | 02.8 | 04.3 | 27.6 | 23.9 | - | 11.1 | 15.4 | 57.9 | 53.4 | - | 19.1 | 24.0 | 72.1 | 67.4 |
| | | 2 | - | (02.8) | (05.7) | (36.0) | (34.4) | - | (10.3) | (15.2) | (59.6) | (55.0) | - | (17.8) | (22.9) | (71.3) | (66.4) |
| | | 4 | - | 03.0 | 07.2 | 25.7 | 22.5 | - | 09.7 | 18.1 | 53.4 | 46.6 | - | 15.9 | 26.8 | 66.5 | 61.3 |
| | | 6 | - | (02.4) | (05.2) | (25.5) | (23.8) | - | (06.4) | (14.7) | (48.4) | (43.6) | - | (14.1) | (23.3) | (61.9) | (56.3) |
| | | 8 | - | 02.6 | 07.4 | 18.0 | 14.9 | - | 07.6 | 18.8 | 38.6 | 34.6 | - | 13.8 | 26.9 | 52.8 | 48.5 |
| | | 10 | - | (02.1) | (05.0) | (10.9) | (13.4) | - | (04.3) | (12.5) | (31.6) | (30.2) | - | (11.0) | (22.0) | (45.2) | (42.0) |
| | | 12 | - | - | 08.5 | 11.9 | 10.9 | - | - | 20.4 | 29.0 | 28.6 | - | - | 29.1 | 41.8 | 41.7 |
| | | 14 | - | - | (08.9) | (09.6) | (08.6) | - | - | (11.5) | (16.6) | (20.7) | - | - | (16.8) | (28.4) | (32.4) |
| | | 16 | - | - | 09.3 | 06.5 | 07.3 | - | - | 22.3 | 20.3 | 21.0 | - | - | 31.4 | 33.8 | 34.8 |
| | | 18 | - | (10.2) | (09.7) | (08.9) | - | - | (11.8) | (11.7) | (13.2) | - | - | - | (18.9) | (15.3) | (20.2) |
| HONG | TR | 0 | - | 09.1 | 05.1 | 05.3 | - | - | 23.0 | 13.6 | 14.6 | - | - | 32.0 | 23.6 | 24.5 | |
| | | 2 | - | (07.6) | (08.0) | (07.6) | - | - | (11.7) | (08.9) | (11.1) | - | - | (20.3) | (10.0) | (15.4) | |
| | | 4 | - | - | 06.7 | 05.2 | - | - | - | 13.0 | 11.9 | - | - | - | 17.7 | 17.1 | |
| | | 6 | - | - | (06.9) | (04.4) | - | - | - | (07.4) | (08.3) | - | - | - | (08.1) | (11.1) | |
| | | 8 | - | - | 07.4 | 05.7 | - | - | - | 13.6 | 11.5 | - | - | - | 17.0 | 15.1 | |
| | | 10 | - | (04.5) | (05.9) | - | - | - | - | (06.5) | (07.8) | - | - | - | (09.6) | (08.7) | |
| | | 12 | - | - | - | 06.1 | - | - | - | - | 10.7 | - | - | - | - | 14.8 | |
| | | 14 | - | - | - | (03.5) | - | - | - | - | (04.8) | - | - | - | - | (07.1) | |
| | | 16 | 05.3 | 03.1 | 03.4 | 04.5 | 07.7 | 13.3 | 11.2 | 11.9 | 15.6 | 23.2 | 20.9 | 19.1 | 21.1 | 26.0 | 37.4 |
| | | 18 | (00.3) | (01.3) | (03.0) | (05.3) | (10.3) | (00.6) | (02.9) | (06.4) | (10.5) | (19.0) | (01.1) | (04.8) | (09.7) | (15.4) | (25.8) |
| PA | 0 | 05.5 | 03.5 | 04.1 | 04.5 | 09.1 | 14.0 | 12.6 | 13.8 | 16.4 | 25.3 | 22.3 | 20.8 | 24.3 | 26.9 | 40.0 | |
| | 2 | (00.3) | (02.1) | (04.6) | (06.7) | (13.6) | (00.7) | (04.4) | (09.4) | (13.0) | (23.7) | (01.2) | (07.1) | (13.4) | (18.3) | (31.6) | |
| | 4 | 04.2 | 03.3 | 03.7 | 06.8 | 11.6 | 12.0 | 10.8 | 12.8 | 20.6 | 31.0 | 19.3 | 19.4 | 22.8 | 33.8 | 47.2 | |
| | 6 | (00.5) | (02.0) | (04.4) | (09.5) | (17.5) | (01.2) | (04.4) | (09.1) | (17.6) | (29.3) | (01.9) | (07.0) | (12.8) | (23.9) | (38.2) | |
| DA | 0 | 02.6 | 02.9 | 04.2 | 25.6 | 21.6 | 10.3 | 11.0 | 15.2 | 54.2 | 51.0 | 18.1 | 19.0 | 23.9 | 69.0 | 65.7 | |
| | 2 | (03.3) | (03.8) | (07.3) | (38.9) | (37.6) | (06.6) | (08.3) | (14.0) | (55.4) | (53.1) | (09.5) | (12.2) | (19.1) | (64.9) | (62.1) | |

TABLEAU 4.3.9. Puissances empiriques (%) basées sur les VCE et les VCA (entre parenthèses) de $\tilde{Q}_{KY}(m, i)$, pour différentes valeurs de m et de i et de $Q_H(w, T_n)$ pour différents noyaux et points de troncature aux niveaux nominaux 1%, 5% et 10%. Le modèle est $ARMA(1, 4)$ avec $\phi_1 = 0.8$ et $\theta_4 = 0.8$ estimé par $AR(1)$, $n=100$.

| Alpha % | 1 | | | | | 5 | | | | | 10 | | | | | |
|-----------|------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | S | 3 | 6 | 12 | 18 | 24 | 3 | 6 | 12 | 18 | 24 | 3 | 6 | 12 | 18 | 24 |
| KOCH-YANG | 0 | - | 91.2 (75.2) | 81.8 (86.5) | 68.4 (78.4) | 57.5 (73.0) | - | 97.9 (95.3) | 96.2 (96.6) | 90.8 (92.6) | 85.1 (88.6) | - | 99.2 (98.7) | 98.7 (98.7) | 96.4 (96.4) | 92.9 (93.8) |
| | 2 | - | 38.9 (38.5) | 08.9 (20.9) | 06.2 (13.7) | 03.8 (11.1) | - | 85.6 (79.2) | 46.0 (59.7) | 28.8 (37.2) | 21.4 (28.6) | - | 93.8 (93.9) | 73.5 (80.1) | 49.9 (55.5) | 38.8 (43.1) |
| | 4 | - | 01.6 (03.0) | 02.5 (04.9) | 04.3 (06.2) | 03.2 (05.7) | - | 09.0 (13.4) | 11.4 (17.0) | 12.5 (15.9) | 12.2 (13.7) | - | 47.9 (26.8) | 23.1 (29.8) | 21.2 (23.6) | 20.2 (21.1) |
| | 6 | - | - | 01.2 (02.4) | 04.6 (04.8) | 04.9 (05.1) | - | - | 06.2 (11.9) | 12.5 (15.1) | 13.3 (13.8) | - | - | 14.9 (22.5) | 21.1 (23.0) | 21.8 (20.2) |
| | 8 | - | - | 03.0 (00.6) | 04.4 (04.1) | 05.4 (03.6) | - | - | 02.09 (04.2) | 12.4 (14.3) | 14.3 (43.3) | - | - | 05.7 (10.3) | 19.7 (22.1) | 21.3 (20.5) |
| | 10 | - | - | 00.6 (01.0) | 04.3 (03.9) | 05.5 (03.8) | - | - | 02.9 (04.9) | 11.4 (13.0) | 14.0 (12.3) | - | - | 06.2 (10.8) | 18.7 (21.4) | 21.5 (19.6) |
| | 12 | - | - | - | 03.0 (03.0) | 05.4 (03.6) | - | - | - | 09.0 (09.6) | 13.4 (10.9) | - | - | - | 15.1 (18.2) | 20.3 (19.0) |
| | 17 | - | - | - | 01.7 (01.9) | 04.3 (03.1) | - | - | - | 06.9 (06.6) | 10.7 (07.2) | - | - | - | 11.9 (14.8) | 13 (15.5) |
| | 23 | - | - | - | - | 02.1 (01.7) | - | - | - | - | 07.1 (04.2) | - | - | - | - | 12.6 (11.2) |
| | HONG | BA | 09.4 (10.1) | 42.9 (48.9) | 85.9 (90.5) | 89.3 (93.8) | 88.6 (94.1) | 25.8 (16.5) | 76.8 (65.7) | 97.2 (96.1) | 98.0 (97.6) | 97.9 (97.7) | 39.6 (22.2) | 89.1 (75.4) | 99.1 (97.8) | 99.4 (98.7) |
| DA | | 17.7 (19.6) | 48.2 (54.4) | 89.8 (94.1) | 90.6 (95.2) | 87.9 (94.3) | 49.2 (33.5) | 79.7 (70.9) | 98.1 (97.5) | 98.3 (98.1) | 97.8 (97.9) | 70.0 (43.7) | 91.5 (79.8) | 99.4 (98.7) | 99.5 (99.1) | 99.3 (98.9) |
| PA | | 11.4 (12.5) | 66.9 (72.5) | 90.2 (94.5) | 89.8 (94.7) | 86.1 (93.6) | 31.0 (20.5) | 90.9 (85.9) | 98.3 (97.8) | 98.1 (98.0) | 97.5 (97.5) | 45.7 (27.5) | 96.1 (91.2) | 99.5 (98.8) | 99.4 (99.1) | 99.1 (98.7) |
| TR | | 14.5 (18.4) | 91.2 (94.4) | 83.0 (91.1) | 71.9 (85.4) | 63.7 (81.3) | 32.8 (28.0) | 97.9 (97.6) | 96.7 (96.7) | 92.5 (93.4) | 88.1 (90.7) | 45.9 (35.1) | 99.2 (98.5) | 98.8 (98.4) | 97.0 (96.3) | 94.9 (94.5) |

TABLEAU 4.3.10. Tests les plus puissants pour chaque modèle, $\alpha = 0.05$. BP : Box-Pierce, H(BA) : Hong avec Bartlett, H(DA) : Hong avec Daniell, H(PA) : Hong avec Parzen, H(TR) : Hong avec noyau uniforme tronqué et KY : Koch et Yang.

| Modèles | Valeurs de S | | | |
|-----------------------|--------------------------------|--------------------------------|--------------------------------|---------------------|
| | 6 | 12 | 18 | 24 |
| $MA(4), \theta = 0.8$ | BP, H(TR) | BP, H(BA), H(DA), H(PA), H(TR) | H(BA), H(DA), H(PA) | H(BA), H(DA), H(PA) |
| $MA(4), \theta = 0.4$ | BP, H(TR) | H(DA), H(PA) | H(BA), H(DA), H(PA) | H(BA), H(DA) |
| $MA(15)$ | BP, H(BA), H(DA), H(PA), H(TR) | KY(12,10) | BP | BP |
| $ARMA(1,4)$ | BP, H(TR) | BP, H(BA), H(DA), H(PA), H(TR) | BP, H(BA), H(DA), H(PA), H(TR) | H(BA), H(DA), H(PA) |

Chapitre 5

EXEMPLES

Dans ce chapitre, nous appliquerons les tests étudiés dans les chapitres précédents à des données réelles. Ainsi, nous considérons deux exemples. Dans le premier, nous utiliserons la série méthane dont les données proviennent de Box et Jenkins (1976, chap.11). Elle représente l'alimentation en méthane d'une fournaise au gaz. Dans le second, les données décrivent l'évolution annuelle du niveau du Nil et proviennent de la dernière édition de la banque de données BD(2002) de Brockwell et Davis (2002). Pour chacun des deux exemples proposés, nous commencerons par la modélisation des données et la validation du modèle selon la méthode de Box et Jenkins en utilisant la procédure *procarima* de *SAS*. Dans une deuxième étape, nous utiliserons le test de Koch et Yang et celui de Hong pour l'analyse des résidus.

5.1. PREMIER EXEMPLE : DONNÉES DE MÉTHANE

Notons par X_t la valeur de la série méthane à l'instant t , $t = 1, \dots, 296$. Le graphe de la série et celui des autocorrélations sont présentés dans la figure 5.1.1. Sur cette figure, nous constatons que les autocorrélations décroissent lentement vers zéro, ce qui permet de douter de la stationnarité de la série. Une différence régulière a été appliquée à la série afin de la stationnariser. D'après la figure 5.1.2, les autocorrélations se présentent sous une forme sinusoidale et les autocorrélations partielles sont importantes aux délais 2 et 3.

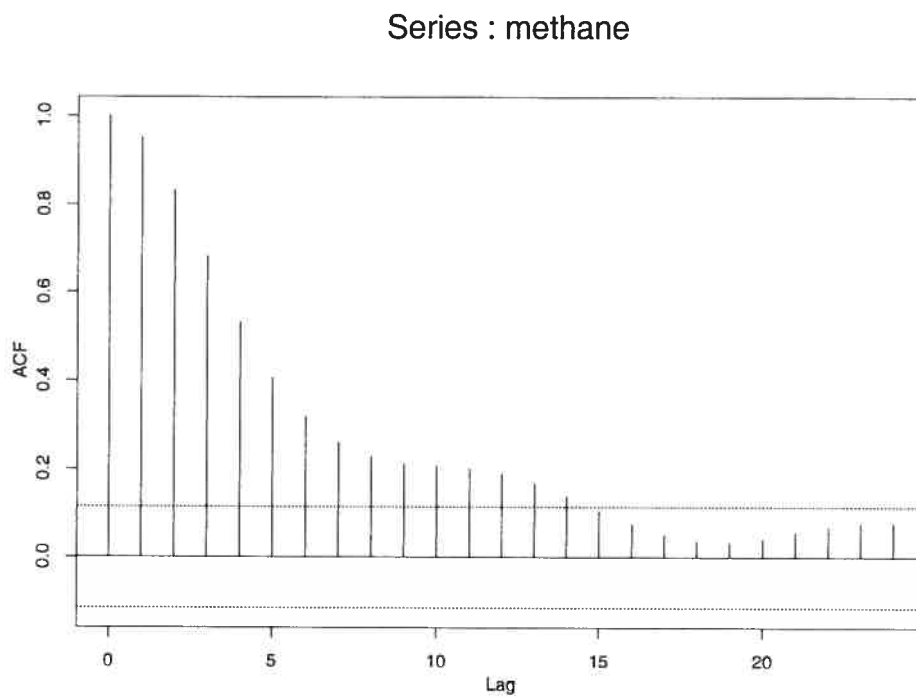
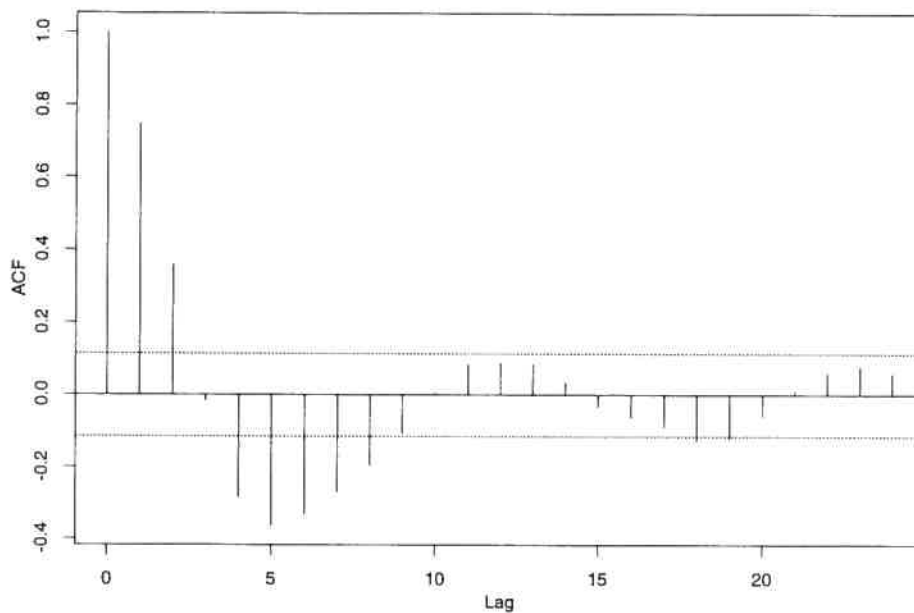


FIGURE 5.1.1. Série de méthane non différenciée et son ACF.

Series : diff(methane)



Series : diff(methane)

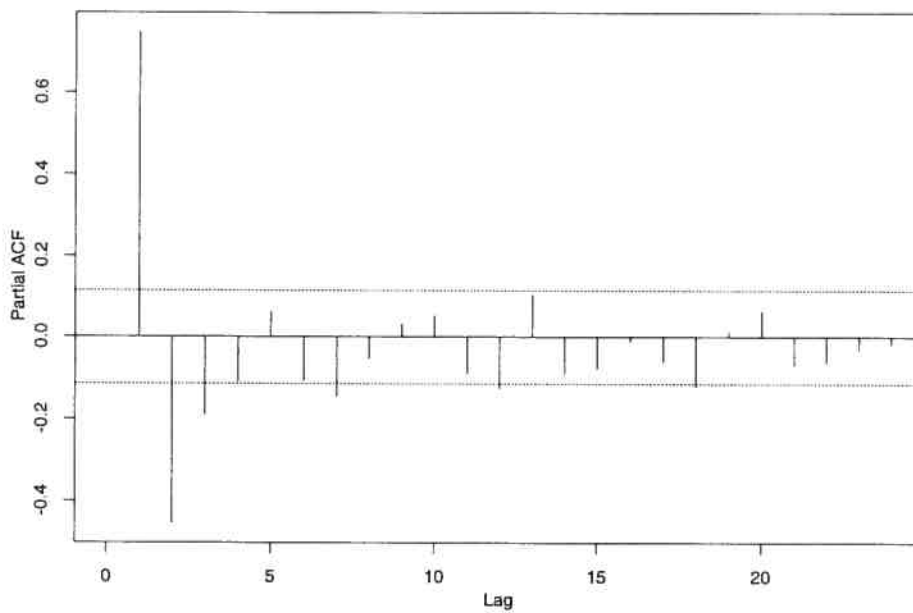
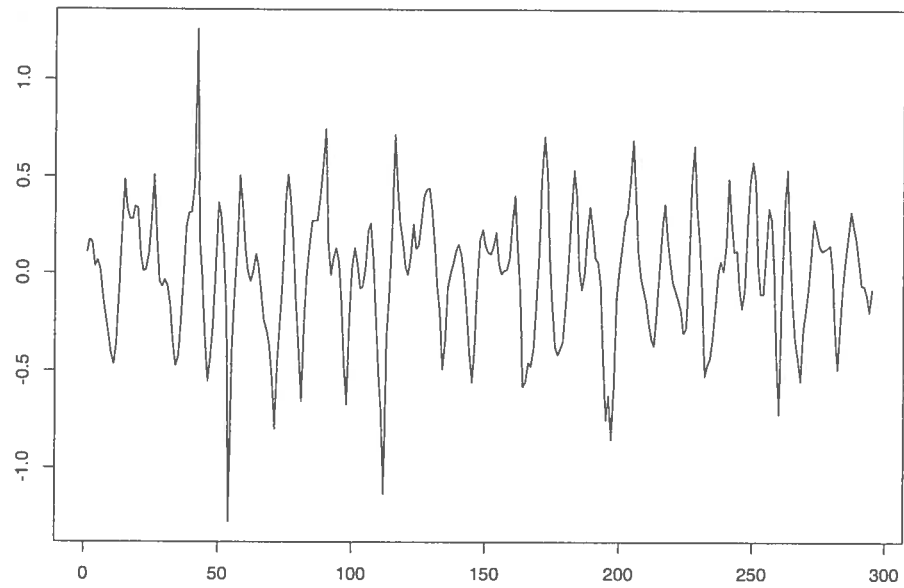


FIGURE 5.1.2. Graphiques des autocorrélations et des autocorrélations partielles de la série de méthane différenciée.

Serie methane avec difference reguliere



Series : serie residuelle methane

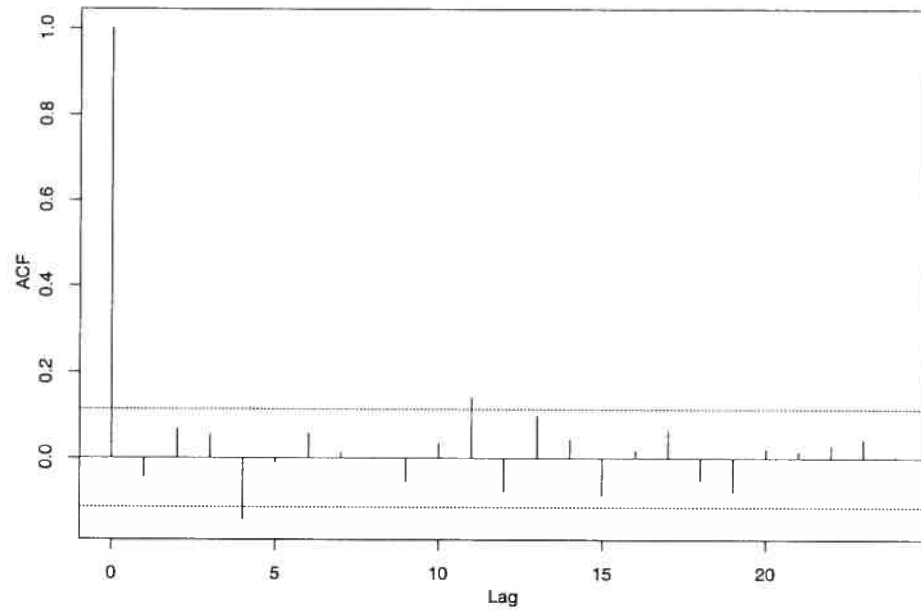


FIGURE 5.1.3. Données de méthane après une différence régulière et l'ACF de la série résiduelle.

Pour rester conforme à la procédure de modélisation proposée par Box et Jenkins (1976), nous avons considéré un modèle $AR(3)$ pour décrire la série originale. Le modèle proposé est alors

$$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + a_t$$

où a_t est un bruit blanc faible $(0, \sigma^2)$, $\hat{\mu} = -0.12$, $\hat{\phi}_1 = 1.97$, $\hat{\phi}_2 = -1.37$, $\hat{\phi}_3 = 0.34$ et $\sigma^2 = 0.0353$.

Afin de modéliser la série par la méthode de Box et Jenkins, le progiciel *SAS* a été utilisé et a donné les résultats suivants :

- 1) Le test de diagnostic du bruit blanc de Ljung et Box (Q_{LB}) donne des niveaux de signification expérimentaux ou valeurs-p tous inférieurs à 0.0001, ce qui veut dire que la série n'est pas un bruit blanc. Les résultats du test sont présentés au tableau 5.1.1.

TABLEAU 5.1.1. Résultats du test de diagnostic du bruit blanc pour la série de méthane.

| Délais | Q_{LB} | Valeur-p |
|--------|----------|----------|
| 6 | 786.35 | < 0.0001 |
| 12 | 874.07 | < 0.0001 |
| 18 | 895.24 | < 0.0001 |
| 24 | 902.52 | < 0.0001 |

- 2) Les paramètres estimés et les valeurs-p correspondantes sont présentés dans le tableau 5.1.2. A l'exception de la moyenne μ , tous les paramètres du modèle sont significativement non nuls à 1%, 5% et 10%.
- 3) Les résultats du test sur les résidus sont présentés dans le tableau 5.1.3.

Les principaux résultats de l'application du test de Koch et Yang et de celui de Hong à la série du méthane sont présentés aux tableaux 5.1.4 et 5.1.5. Dans ces

TABLEAU 5.1.2. Estimation des paramètres du modèle $AR(3)$ ajusté à la série de méthane.

| paramètres | Valeurs | Valeur-p |
|------------|---------|----------|
| ϕ_1 | 1.97 | < 0.0001 |
| ϕ_2 | -1.37 | < 0.0001 |
| ϕ_3 | 0.34 | < 0.0001 |
| μ | -0.12 | 0.2609 |

TABLEAU 5.1.3. Résultats du test d'adéquation de Ljung et Box (série de méthane).

| Délais | Q_{LB} | Valeur-p | Décision à 1% | Décision à 5% | Décision à 10% |
|--------|----------|----------|---------------|---------------|----------------|
| 6 | 10.30 | 0.0162 | non rejet | rejet | rejet |
| 12 | 19.89 | 0.0186 | non rejet | rejet | rejet |
| 18 | 27.92 | 0.0221 | non rejet | rejet | rejet |
| 24 | 31.05 | 0.0729 | non rejet | non rejet | non rejet |

tableaux, nous présentons uniquement les cas pour lesquels un des tests étudiés a enregistré un rejet de l'hypothèse nulle. Ces résultats nous permettent de constater qu'au niveau de signification 1%, le test de Koch et Yang rejette l'hypothèse nulle pour $i = 1$ et $m = 6$. Notons que le test de Ljung et Box ne parvenait pas à détecter cette autocorrélation à 1%.

A 5%, le test de Koch et Yang parvient à détecter une corrélation dans les résidus au délai 6 avec $i = 0$, $i = 1$ et $i = 2$, au délai 12 avec $i = 0$ et $i = 1$ ainsi qu'au délai 18 avec $i = 0$ alors que le test de Hong n'enregistre aucun rejet.

A 10%, le test de Koch et Yang rejette l'hypothèse nulle quand $m = 6$ pour $i = 0$, $i = 1$ et $i = 2$ et quand $m = 12$ pour $i = 0$, $i = 1$, $i = 5$ et $i = 6$. Quand le délai m est plus grand que 12, le test rejette uniquement pour $i = 0$. Concernant le test de Hong, les noyaux de Bartlett, Daniell et Parzen aboutissent au rejet

de l'hypothèse nulle quand le point de troncature est grand ($T_n = 24$). Le noyau uniforme tronqué, conduit au rejet seulement au niveau 10% lorsque $T_n = 12$ et 18.

TABLEAU 5.1.4. Résultats du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$ appliqué aux données du méthane.

| Délais m | valeur de i | valeur-p | Décision à 1% | Décision à 5% | Décision à 10% |
|------------|---------------|----------|---------------|---------------|----------------|
| 6 | 0 | 0.0186 | non rejet | rejet | rejet |
| | 1 | 0.0059 | rejet | rejet | rejet |
| | 2 | 0.0241 | non rejet | rejet | rejet |
| 12 | 0 | 0.0213 | non rejet | rejet | rejet |
| | 1 | 0.0477 | non rejet | rejet | rejet |
| | 5 | 0.0936 | non rejet | non rejet | rejet |
| | 6 | 0.0522 | non rejet | non rejet | rejet |
| 18 | 0 | 0.0241 | non rejet | rejet | rejet |
| 24 | 0 | 0.0796 | non rejet | non rejet | rejet |

TABLEAU 5.1.5. Résultats du test de Hong $Q_H(w, T_n)$ appliqué aux données du méthane.

| T_n | $w(\cdot)$ | Valeur-p | Décision à 1% | Décision à 5% | Décision à 10% |
|-------|------------------|----------|---------------|---------------|----------------|
| 12 | Uniforme tronqué | 0.0643 | non rejet | non rejet | rejet |
| 18 | Uniforme tronqué | 0.0544 | non rejet | non rejet | rejet |
| 24 | Bartlett | 0.0975 | non rejet | non rejet | rejet |
| | Daniell | 0.0901 | non rejet | non rejet | rejet |
| | Parzen | 0.0790 | non rejet | non rejet | rejet |

Nous concluons de cet exemple qu'au niveau 5%, le test de Hong ne rejette pas l'hypothèse nulle alors que le test de Box-Pierce et celui de Koch et Yang rejettent pour plusieurs valeurs de m et de i . D'ailleurs, si nous regardons les autocorrélations résiduelles présentées dans la figure 5.1.3 nous constatons que certaines

autocorrélations sont importantes au délais 4, 11 et 13, ce qui est conforme avec les résultats obtenus avec le test de Koch et Yang.

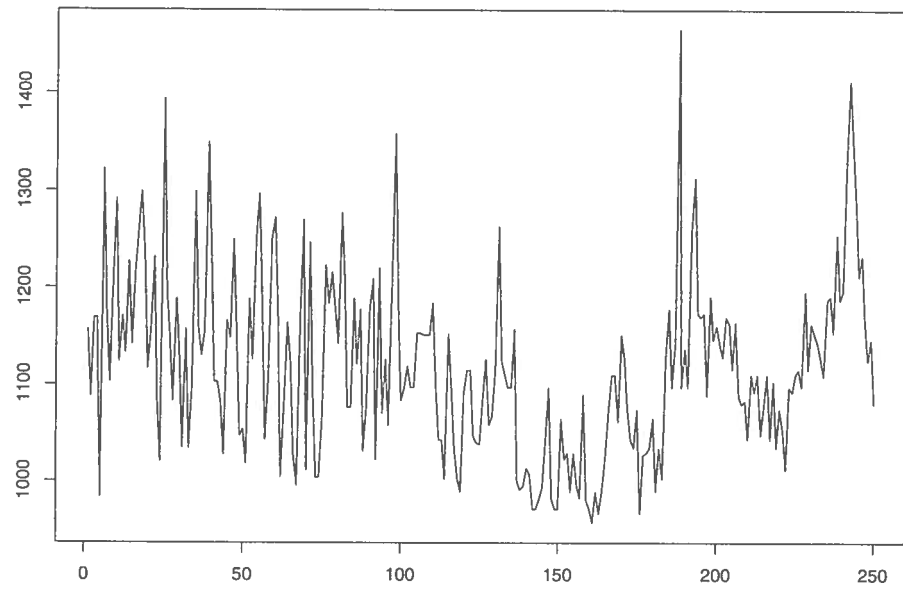
5.2. DEUXIÈME EXEMPLE : NIVEAU DU NIL (NIL)

Soit X_t , le niveau du Nil à l'année t , $t = 1, \dots, 250$. Le niveau du Nil est mesuré au poste Roda près du Caire pour les années 622 – 871. Le graphique des autocorrélations du processus de la figure (5.2.1) montre que la série n'est pas stationnaire puisque les autocorrélations décroissent lentement vers zéro. Une différence d'ordre 1 est appliquée à cette dernière et nous constatons que la série devient stationnaire. Le modèle $AR(4)$ suivant qui minimise le critère AIC est proposé pour décrire la série différenciée :

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \phi_4 Y_{t-4} + a_t$$

où $Y_t = X_t - X_{t-1}$, la série différenciée, a_t un bruit blanc faible $(0, \sigma^2)$, $\hat{\mu} = -0.08$, $\hat{\phi}_1 = -0.58$, $\hat{\phi}_2 = -0.51$, $\hat{\phi}_3 = -0.33$, $\hat{\phi}_4 = -0.15$ et $\sigma^2 = 6393.6$.

Serie NIL



Series : nil

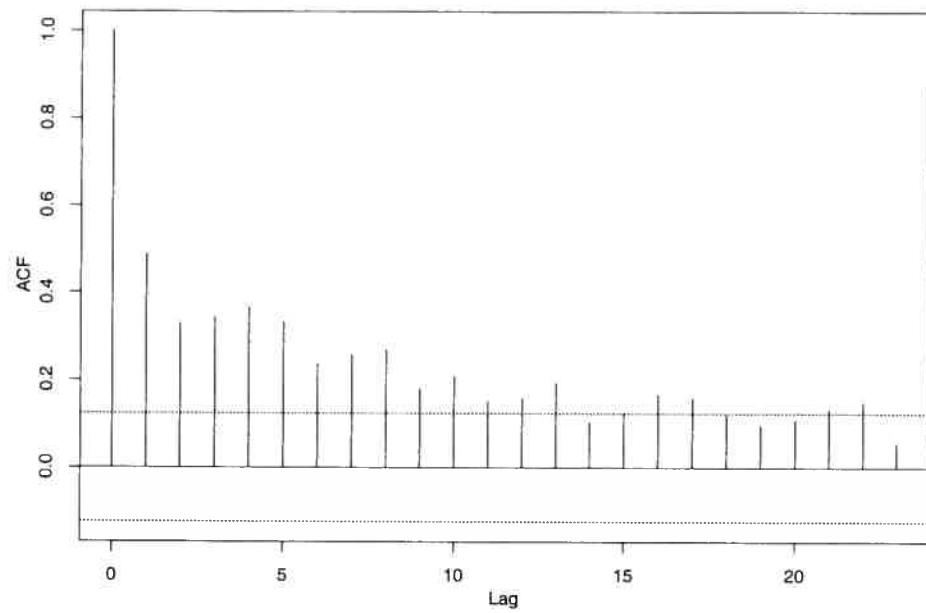
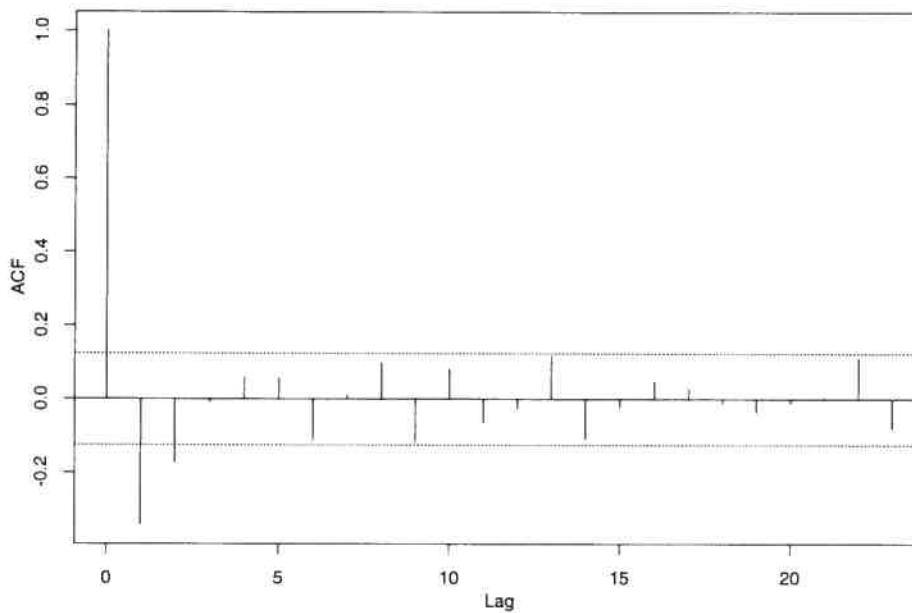


FIGURE 5.2.1. Niveau du Nil non différencié et son ACF.

Series : diff(nil)



Series : diff(nil)

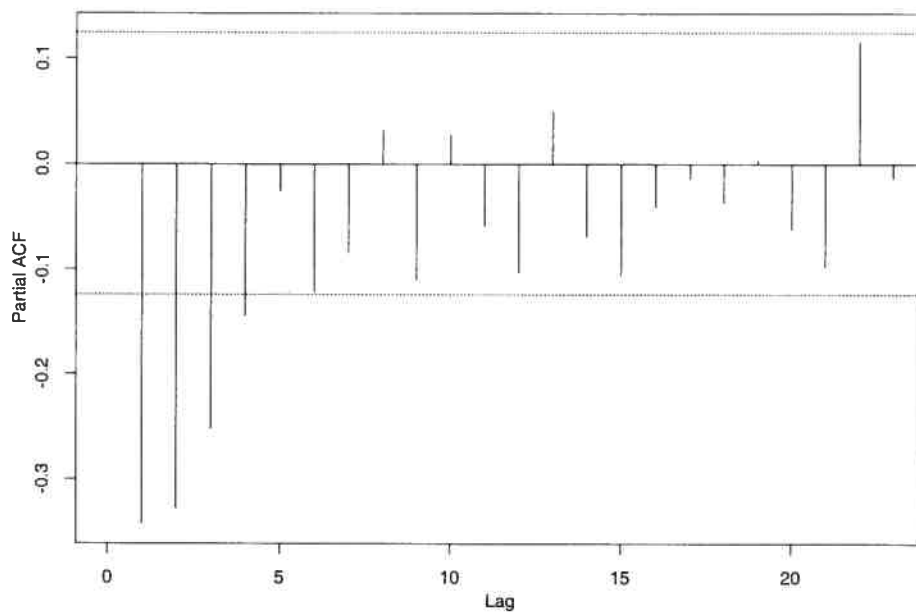
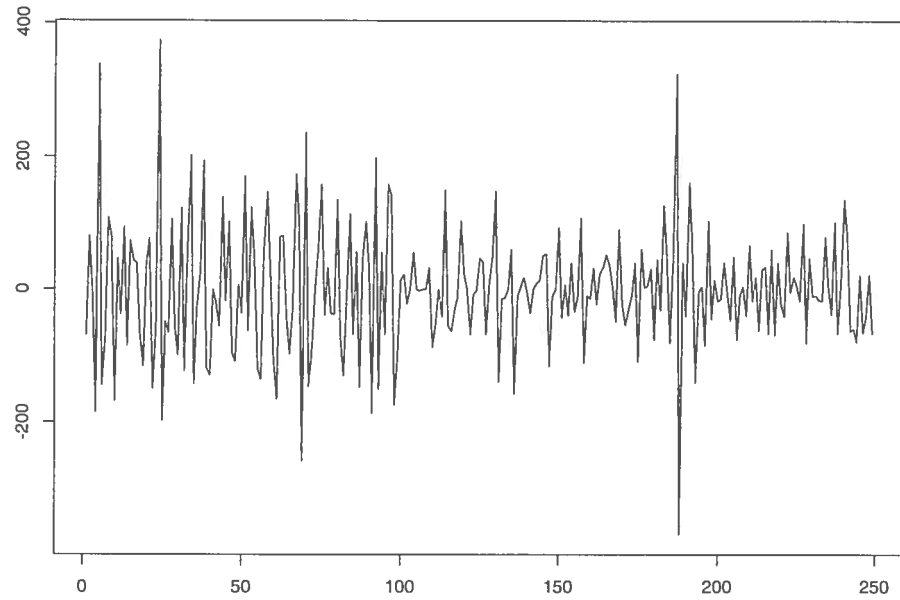


FIGURE 5.2.2. Graphiques des autocorrélations et des autocorrélations partielles du niveau du Nil différencié.

Serie NIL avec difference reguliere



Series : residus Nil

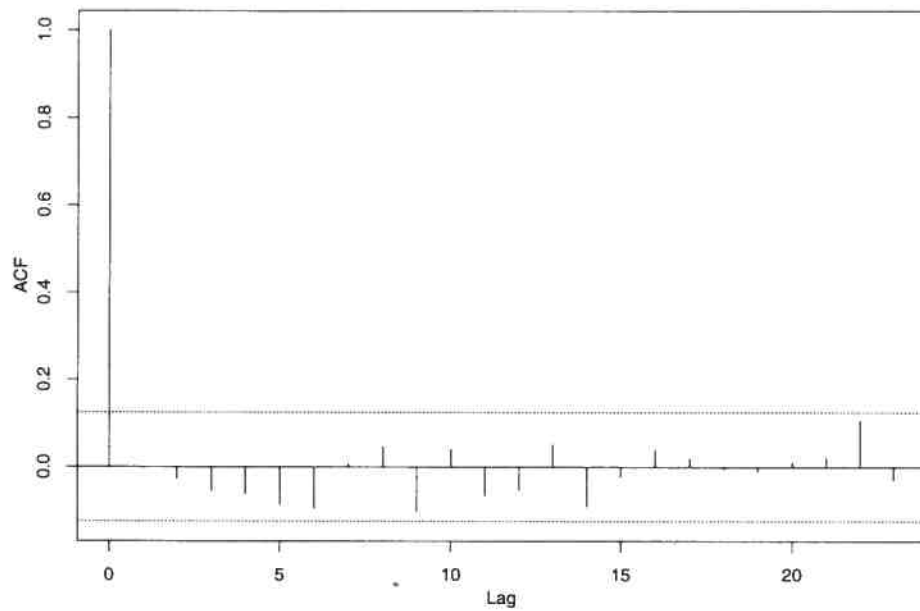


FIGURE 5.2.3. Niveau du Nil après une différence régulière et l'ACF de la série résiduelle.

Les résultats de la procédure *procarima* de *SAS* nous permettent de faire les constatations suivantes :

- (1) Le test de diagnostic du bruit blanc de Ljung et Box donne des valeurs-p qui sont inférieures à 0.0001, donc la série n'est pas un bruit blanc. Les résultats du test sont donnés dans le tableau 5.2.1.

TABLEAU 5.2.1. Résultats du test de diagnostic du bruit blanc pour le niveau du Nil différencié.

| Délais | Q_{LB} | Valeur-p |
|--------|----------|----------|
| 6 | 42.05 | < 0.0001 |
| 12 | 51.16 | < 0.0001 |
| 18 | 58.87 | < 0.0001 |
| 24 | 64.66 | < 0.0001 |

- (2) Les estimateurs des paramètres et les valeurs-p correspondantes sont présentés dans le tableau 5.2.2.

TABLEAU 5.2.2. Estimation des paramètres du modèle $AR(4)$ ajusté au niveau du Nil différencié.

| paramètres | Valeurs | Valeur-p |
|------------|---------|----------|
| ϕ_1 | -0.58 | < 0.001 |
| ϕ_2 | -0.51 | < 0.001 |
| ϕ_3 | -0.33 | < 0.001 |
| ϕ_4 | -0.15 | 0.019 |
| μ | -0.08 | 0.966 |

A 5%, tous les paramètres du modèle, à part la moyenne, sont significativement non nuls.

- (3) Le test de corrélation sérielle donne des valeurs-p toutes supérieures à 5% pour les délais 6, 12, 18 et 24; ce qui veut dire que le modèle choisi est adéquat. Ces résultats sont présentés dans le tableau 5.2.3.

TABLEAU 5.2.3. Résultats du test d'adéquation de Ljung et Box (niveau du Nil différencié).

| Délais | Q_{LB} | Valeur-p | Décision à 1% | Décision à 5% | Décision à 10% |
|--------|----------|----------|---------------|---------------|----------------|
| 6 | 5.95 | 0.0502 | non rejet | non rejet | rejet |
| 12 | 11.47 | 0.1765 | non rejet | non rejet | non rejet |
| 18 | 15.04 | 0.3751 | non rejet | non rejet | nonrejet |
| 24 | 18.99 | 0.5223 | non rejet | non rejet | non rejet |

Les principaux résultats de l'application du test de Koch et Yang et de celui de Hong au niveau du Nil différencié sont présentés au tableau 5.2.4. Comme pour la série traitée dans l'exemple 1, nous présentons uniquement les cas pour lesquels un des tests étudiés a enregistré un rejet de l'hypothèse nulle. Ces résultats nous permettent de constater qu'au niveau de signification 1%, le test de Koch et Yang rejette l'hypothèse nulle avec $m = 12$ et $i = 11$.

A 5%, le test de Koch et Yang détecte la corrélation dans les résidus à $m = 6$ avec toutes les valeurs de i , à $m = 12$ avec $i = 8, 10$ et 11 ainsi qu'à $m = 18$ avec $i = 13, 17$.

A 10%, le test de Koch et Yang rejette l'hypothèse nulle quand $m = 6$ pour toutes les valeurs de i , quand $m = 12$ pour les valeurs de i supérieures à 2 et quand $m = 18$ pour $i = 9, 10, 11, 12, 13, 15$ et 17 .

Concernant le test de Hong, on n'enregistre aucun rejet de l'hypothèse nulle.

Nous pouvons ainsi conclure de cet exemple que le test de Koch et Yang permet de mettre en doute le modèle estimé alors que le test de Hong ne le rejette pas.

TABLEAU 5.2.4. Résultats du test de Koch et Yang $\tilde{Q}_{KY}(m, i)$ appliqué aux données du niveau du Nil.

| Délais m | valeur de i | Valeurs-p | Décision à 1% | Décision à 5% | Décision à 10% |
|------------|---------------|-----------|---------------|---------------|----------------|
| 6 | 0 | 0.0498 | non rejet | rejet | rejet |
| | 1 | 0.0335 | non rejet | rejet | rejet |
| | 2 | 0.0257 | non rejet | rejet | rejet |
| | 3 | 0.0299 | non rejet | rejet | rejet |
| | 4 | 0.0218 | non rejet | rejet | rejet |
| | 5 | 0.0372 | non rejet | rejet | rejet |
| 12 | 3 | 0.0785 | non rejet | non rejet | rejet |
| | 4 | 0.0748 | non rejet | non rejet | rejet |
| | 5 | 0.0638 | non rejet | non rejet | rejet |
| | 6 | 0.0588 | non rejet | non rejet | rejet |
| | 7 | 0.0571 | non rejet | non rejet | rejet |
| | 8 | 0.0331 | non rejet | rejet | rejet |
| | 9 | 0.0549 | non rejet | non rejet | rejet |
| | 10 | 0.0358 | non rejet | rejet | rejet |
| | 11 | 0.0020 | rejet | rejet | rejet |
| 18 | 9 | 0.0955 | non rejet | non rejet | rejet |
| | 10 | 0.0777 | non rejet | non rejet | rejet |
| | 11 | 0.0797 | non rejet | non rejet | rejet |
| | 12 | 0.0510 | non rejet | non rejet | rejet |
| | 13 | 0.0124 | non rejet | rejet | rejet |
| | 15 | 0.0969 | non rejet | non rejet | rejet |
| | 17 | 0.0492 | non rejet | rejet | rejet |

Ces deux exemples font ressortir l'intérêt pratique du test de Koch et Yang. Le test de Hong, par contre, ne rejette que rarement. La question à savoir si ceci est dû au fait que le niveau de ce test est mal contrôlé et que le niveau réel est souvent supérieur au niveau nominal. Ce problème est traité dans Chen et Deo (2003).

CONCLUSION

Dans leur article, Koch et Yang (1986) ont construit un test d'adéquation pour vérifier la dépendance entre deux séries chronologiques. L'objectif premier de notre recherche était d'adapter leur approche au contexte de l'ajustement d'un modèle à une série de données univariée et de comparer sa performance à celle des tests de Hong (1996a), de Box et Pierce (1976) et de Ljung et Box (1976).

Dans un premier temps, nous avons exposé les propriétés asymptotiques sous l'hypothèse nulle des tests de Box et Pierce, de Koch et Yang et de Hong. Alors que la statistique de Box et Pierce converge en loi vers une khi-deux, la statistique de Koch et Yang est distribuée asymptotiquement comme une combinaison linéaire de variables aléatoires indépendantes et distribuées selon des khi-deux et la statistique de Hong est distribuée selon une loi normale.

A l'aide de simulations de Monte Carlo, nous avons étudié de près l'approximation de la vraie distribution des tests de Koch et Yang et de Hong par leurs lois asymptotiques. Nous avons également examiné la puissance de chacun des tests en utilisant les valeurs critiques exactes et les valeurs critiques asymptotiques.

Nous retenons de ces simulations que pour les modèles autorégressifs considérés dans l'étude de niveau, la distribution du test de Koch et Yang est généralement bien approximée par sa distribution asymptotique pour les petites valeurs de i . Au niveau 1%, l'approximation est meilleure mais se détériore avec m . Une amélioration du niveau est généralement observée quand la taille des séries considérées augmente. Toutefois, un sousrejet de l'hypothèse nulle est constaté pour

les grandes valeurs de m et de i .

Concernant la statistique de Hong, l'approximation par la distribution asymptotique est bonne aux niveaux 5% et 10% et pour les grandes valeurs de T_n . Au niveau 1%, le test tend à surrejeter l'hypothèse nulle mais le niveau s'améliore pour les petites valeurs de T_n et quand la taille de la série augmente.

Nous retenons de l'étude de puissance que la puissance des tests étudiés est très dépendante du délai m et de la valeur de i dans le cas du test de Koch et Yang et du point de troncature T_n dans le cas du test de Hong dans presque tous les cas. Le test de Koch et Yang affiche la puissance la plus élevée quand i est nul, c'est-à-dire quand le test est celui de Box-Pierce, et quand m est petit. Cependant, la puissance du test se détériore quand i augmente sauf pour le modèle $MA(15)$ avec $m = 12$.

Le test de Hong donne des puissances qui augmentent avec T_n et qui se stabilisent lorsque ce dernier est assez élevé (plus de 18). Le noyau uniforme tronqué est plus puissant quand $T_n = 6$ alors que le noyau de Daniell est supérieur quand $T_n = 12$. Pour des valeurs de T_n plus grandes, les puissances des quatre noyaux sont rapprochées sauf pour le noyau uniforme tronqué qui perd de sa puissance.

A travers les deux exemples pratiques présentés dans le cinquième chapitre, dans lequel nous avons ajusté deux modèles autorégressifs à deux séries de données réelles, nous avons pu mettre en évidence l'intérêt pratique que présente l'utilisation du test de Koch et Yang pour tester la non corrélation des séries résiduelles dans le cas où celles-ci présentent des autocorrélations successives de même signe. En effet, avec le test de Ljung et Box, le modèle ajusté à la série de méthane n'est pas rejeté au niveau 1% alors qu'avec le test de Koch et Yang, le modèle est rejeté à ce niveau de signification pour certaines valeurs de m et de i . De même, le modèle ajusté à la série du niveau du Nil est rejeté à plusieurs reprises par le test de Koch et Yang aux trois niveaux nominaux alors que le test

de Ljung et Box ne le rejette pas aux niveaux 5% et 10%. Ainsi, nous concluons qu'en introduisant de nouveaux éléments dans le calcul de la statistique de test, le test de Koch et Yang modifié met davantage en évidence la corrélation sérielle des résidus à certains délais et met donc en doute la validité de certains modèles que le test de Ljung et Box ne rejette pas. Comparativement au test de Koch et Yang, notre étude nous a permis de constater que le test de Hong rejette moins souvent le modèle ajusté.

ANNEXE

ALGORITHME DES SIMULATIONS

Nous présentons ici les principales opérations exécutées par le biais du code *Fortran*.

- 1) Lire les paramètres nécessaires à la simulation : n , la taille de la série à générer, m , le délai maximum considéré, pn , le vecteur des points de troncature et $taille$, le nombre de simulations à effectuer.
- 2) Générer une série chronologique de longueur n d'un processus autorégressif avec des erreurs normalement distribuées, à l'aide du sous-programme *G05EGF*. Le paramètre autorégressif est défini selon le modèle considéré.
- 3) Estimer les paramètres du vrai modèle par la méthode des moindres carrés à l'aide du sous-programme *G13AFF*.
- 4) Calculer les résidus ainsi que le vecteur des autocorrélations résiduelles à l'aide du sous-programme *G13ABF*.
- 5) Calculer la statistique de Koch et Yang. Rappelons que cette dernière s'écrit de la façon suivante :

$$Q_{KY}(m, i) = n \sum_{h=1}^{m-i} \left[\sum_{j=0}^i \hat{\rho}_{h+j}(\hat{a}) \right]^2$$

- 6) Calculer les coefficients ψ_j , $j = 1, \dots, m - 1$, et par la suite la matrice de covariance \mathbf{M} des autocorrélations échantillonales.

Les coefficients ψ_j , $j = 1, \dots, m - 1$ sont calculés à partir de la relation

$$\psi_j - \sum_{i=1}^p \phi_i \psi_{j-i} = 0, j = 0, 1, \dots$$

ou $\psi_j = 0$ pour $j < 0$ et les ϕ_k sont les paramètres autorégressifs du modèle considéré.

- 7) Créer la matrice \mathbf{C}'_i puis calculer la matrice $\mathbf{A}_i = \mathbf{C}'_i \mathbf{C}_i$.
- 8) Calculer les valeurs propres de $\mathbf{M} \mathbf{A}_i$ par le sous-programme *F02FAF*.
- 9) Calculer la distribution de la statistique de Koch et Yang par le sous-programme *G01JDF* qui utilise l'algorithme d'Imhof.
- 10) Calculer la statistique de Hong $Q_H(w, T_n)$ pour chaque fonction de noyau (Bartlett, Daniell, Parzen et le noyau uniforme tronqué,) et pour chacune des valeurs du point de troncature.
- 11) Calculer la distribution de la statistique de Hong par le sous programme *G01EAF*.
- 12) Avec 10 000 réalisations indépendantes, calculer la proportion de rejet pour chaque statistique de test et pour chaque niveau nominal.
- 13) Classer les 10 000 statistiques de Koch et Yang obtenues afin de déterminer les valeurs critiques exactes pour chacun des niveaux nominaux. Faire la même opération pour la statistique de Hong.

- 14) Générer des séries d'un modèle $ARMA(p, q)$ (modèles définis dans l'étude de puissance) et estimer un modèle $AR(1)$.
- 15) Calculer les deux statistiques de test et les comparer avec leurs valeurs critiques exactes obtenues dans l'étude de niveau et avec leurs valeurs critiques asymptotiques.
- 16) Calculer les puissances basées sur les valeurs critiques exactes et celles basées sur les valeurs critiques asymptotiques pour chacun des tests.

CODE FORTRAN

Nous présentons seulement un modèle de code pour chacune des simulations effectuées.

I. PROGRAMMES POUR L'ETUDE DE NIVEAU

C Déclaration des paramètres

```

implicit real*8 (a-h,o-z)
integer ifail,ip,iq,nr,n,i,k,nmax,lalmax,choix,ipn,taille
implicit real*8 (a-h,o-z)
parameter (mmax=12,n=200,pnmax=5,nm1=n-1)
parameter (nmax=mmax,lalmax=mmax*mmax)
double precision par(3),c,s,sd(3),cm(1,1),st(1300),res(50000)
&,dd(mmax,mmax),a(12),b(12),r(52),var,w(5000),mean,t
double precision lambda(mmax)
external g05egf,g05cbf,g01eaf,g01jcf,g01jdf,g05zaf,g05fdf
integer mr(7),isf(4),kfc,ai
integer ifail1,m,pnn(pnmax)
double precision vm,vvm,r1(nm1),r2(nm1,1),stat,statis
double precision wt(1000),xmean,s2,s3,s4,xmin,xmax,wsum
double precision mata(mmax,mmax)
double precision koche(taille,mmax),crit(3,mmax)
double precision hong(4,5),hong1(taille,4),hong2(taille,4),
&hong3(taille,4),hong4(taille,4),hong5(taille,4),critere(3,20)
common /r2/r2
common /ai/ai
common /m/m
common /par/par
common /kochpar/lambda,k
common/dd/dd
double precision g01eaf,travail(1000),p
parameter(taille=10000)
integer rej1(0:mmax),rej2(0:mmax),rej3(0:mmax)
integer rejet1(4,pnmax),rejet2(4,pnmax),rejet3(4,pnmax)
double precision propky1(0:mmax),proph1(4,pnmax),p1,proba
double precision propky2(0:mmax),proph2(4,pnmax)
double precision propky3(0:mmax),proph3(4,pnmax)

```

C Initialisation

```

ndon=n

```

```

print *, 'm=?'
read *, m
do ai=0,m-1
  rej1(ai)=0
  rej2(ai)=0
  rej3(ai)=0
enddo
pnn(1)=3
pnn(2)=6
pnn(3)=12
pnn(4)=18
pnn(5)=24
do choix=1,4
  do ipn=1,3
    rejet1(choix,ipn)=0
    rejet2(choix,ipn)=0
    rejet3(choix,ipn)=0
  enddo
enddo
call g05cbf(0)
ip=1
iq=0
l=ip+iq
a(1)=.8d0
b(1)=1.d0
b(2)=.0d0
nr=2*ip+2*iq+4+1
mean=0.0d0
npar=ip+iq
C Debut de la boucle
  do 1000 nsimul=1,taille
C Simulation d'un ARMA(p,q)
  ifail=0
  call g05egf(mean,a,ip,b,iq+1,r,nr,var,ifail)
  if(ifail.ne.0)write(*,*)'ifail=',ifail
  do 2 i=1,n
  ifail=0
  w(i)=g05ewf(r,nr,ifail)
C Estimation
  mr(1)=1
  mr(2)=0
  mr(3)=0

```

```

mr(4)=0
mr(5)=0
mr(6)=0
mr(7)=0
par(1)=0.0d0
par(2)=0.0d0
par(3)=0.0d0
kfc=0
c=0.0d0
ifail=1
nppc=npar+kfc
icm=nppc+1
nst=ip+iq
call g13aff(mr,par,npar,c,kfc,w,n,s,ndf,sd,nppc,cm,icm,st,nst,
&0,500,itc,isf,res,50000,nres,ifail)
C Calcul des autocorrélations du processus
call g13abf(w,n,n-1,vm,vvm,r1,stat,ifail1)
C Calcul des autocorrélations des résidus
call g13abf(res,n,n-1,vm,vvm,r2,stat,ifail1)
C Calcul de la statistique de Koch et Yang
do ai=0,m-1
call covariance(m,ip,iq,dd)
call koch(m,ai,r2,t)
koche(ii,ai+1)=t
call matricema(m,ai,ip,mata,matm,lambda,k)
ifail=0
call g01jdf("I",k,lambda,0,t,p,travail,ifail)
p1=1-p
if (p1.lt.0.01)then
rej1(ai)=rej1(ai)+1
endif
if (p1.lt.0.05)then
rej2(ai)=rej2(ai)+1
endif
if (p1.lt.0.1)then
rej3(ai)=rej3(ai)+1
endif
enddo
C Calcul de la statistique de Hong
do choix=1,4
do ipn=1,5
call pmanteau(r2,n,pnn(ipn),statis,choix)

```

```

        proba = g0leaf ("u",statis,ifail)
        if (proba .lt.0.01)then
        rejet1(choix,ipn)= rejet1(choix,ipn)+1
        endif
        if (proba .lt.0.05)then
        rejet2(choix,ipn)=rejet2(choix,ipn)+1
        endif
        if (proba .lt.0.1)then
        rejet3(choix,ipn)=rejet3(choix,ipn)+1
        endif
        enddo
        hong1(ii,choix)=hong(choix,1)
        hong2(ii,choix)=hong(choix,2)
        hong3(ii,choix)=hong(choix,3)
        hong4(ii,choix)=hong(choix,4)
        hong5(ii,choix)=hong(choix,5)
        enddo
C Fin de la boucle
1000 continue
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C Calcul des proportions de rejet de Ho.      C
C Dans le cas de l'étude de puissances,      C
C ces proportions correspondent aux           C
C puissances asymptotiques des tests étudiés C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
        write(*,*)"m=", m, "n= ", n
        write(*,*)" test de Koch et Yang "
        do ai=0,m-1
        propky1(ai)=(rej1(ai)+0.0d0)/(taille+0.0d0)
        propky2(ai)=(rej2(ai)+0.0d0)/(taille+0.0d0)
        propky3(ai)=(rej3(ai)+0.0d0)/(taille+0.0d0)
        enddo
        Write(*,*) "alpha=0.01, alpha=0.05, alpha=0.1"
        do ai=0,m-1
        write(*,600) ai, propky1(ai), propky2(ai),propky3(ai)
        enddo
        write(*,*)" test de Hong "
        do choix=1,4
        do ipn=1,5
        proph1(choix,ipn)=(rejet1(choix,ipn)+0.0d0)/(taille+0.0d0)
        proph2(choix,ipn)=(rejet2(choix,ipn)+0.0d0)/(taille+0.0d0)
        proph3(choix,ipn)=(rejet3(choix,ipn)+0.0d0)/(taille+0.0d0)

```

```

        enddo
        enddo
        Write(*,*) "alpha=0.01"
        do choix=1,4
        write(*,600)(proph1(choix,ipn),ipn=1,5)
        enddo
        Write(*,*) "alpha=0.05"
        do choix=1,4
        write(*,600)(proph2(choix,ipn),ipn=1,5)
        enddo
        Write(*,*) "alpha=0.1"
        do choix=1,4
        write(*,600)(proph3(choix,ipn),ipn=1,5)
        enddo
600 format(12f8.4)
        stop
        end
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C Calcul des valeurs critiques exactes C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
        do j=1,m
        call m01caf(koche(1,j),1,taille,'A',ifail)
        enddo
        do i=1,m
        crit(1,i)=koche(taille*(1-0.01),i)
        crit(2,i)=koche(taille*(1-0.05),i)
        crit(3,i)=koche(taille*(1-0.1),i)
        enddo
        do i=1,m
        write(6,*)(crit(j,i),j=1,3)
        enddo
        do j=1,4
        call m01caf(hong1(1,j),1,taille,'A',ifail)
        call m01caf(hong2(1,j),1,taille,'A',ifail)
        call m01caf(hong3(1,j),1,taille,'A',ifail)
        call m01caf(hong4(1,j),1,taille,'A',ifail)
        call m01caf(hong5(1,j),1,taille,'A',ifail)
        enddo
        do i=1,4
        critere(1,i)=hong1(taille*(1-0.01),i)
        critere(2,i)=hong1(taille*(1-0.05),i)
        critere(3,i)=hong1(taille*(1-0.1),i)

```



```

        enddo
        do i=5,8
            critere(1,i)=hong2(taille*(1-0.01),i-4)
            critere(2,i)=hong2(taille*(1-0.05),i-4)
            critere(3,i)=hong2(taille*(1-0.1),i-4)
        enddo
        do i=9,12
            critere(1,i)=hong3(taille*(1-0.01),i-8)
            critere(2,i)=hong3(taille*(1-0.05),i-8)
            critere(3,i)=hong3(taille*(1-0.1),i-8)
        enddo
        do i=13,16
            critere(1,i)=hong4(taille*(1-0.01),i-12)
            critere(2,i)=hong4(taille*(1-0.05),i-12)
            critere(3,i)=hong4(taille*(1-0.1),i-12)
            write(*,*)(critere(j,i),j=1,3)
        enddo
        do i=17,20
            critere(1,i)=hong5(taille*(1-0.01),i-16)
            critere(2,i)=hong5(taille*(1-0.05),i-16)
            critere(3,i)=hong5(taille*(1-0.1),i-16)
            write(*,*)(critere(j,i),j=1,3)
        enddo
        do i=1,20
            write(2,*)(critere(j,i),j =1,3)
        enddo
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C          Sous-programme pour le calcul          C
C    de la statistique de Koch et Yang          C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
        subroutine koch(m,ai,r2,t)
        implicit real*8 (a-h,o-z)
        integer nin, nout,ai
        parameter (mmax=12, lw=64*mmax, n=200, nm1=n-1)
        Double precision r2(nm1,1),q,q1,t1,t,r22(nm1,1),correc(nm1)
        parameter (nin=5,nout=6)
C Transformation des rho_h(â):
        do i=1,m
            correc(i)= dsqrt(n*1.0d0/((n+0.0d0)-i))
        enddo
        do i=1,m
            r22(i,1)= correc(i) * r2(i,1)

```

```

        enddo
C calcul final de Koch et Yang
        t1=0.0d0
        do 28 h=1,m-ai
            q=0.0d0
            do 29 j=0,ai
                q=q+r22(h+j,1)
29         continue
            q1=q**2
            t1=t1+q1
            t=t1*n
28         continue
        return
        end
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C Sous-programme pour le calcul des matrices M et Ai      C
C et des valeurs propres de la matrice produit M*Ai      C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
        subroutine matricema(m,ai,ip,mata,matm,lambda,k)
            implicit real*8 (a-h,o-z)
            parameter (mmax=12, lw=64*mmax, n=200, nm1=n-1)
            integer nin, nout,ai,j,i,a,h,l,ifail
            parameter (nin=5,nout=6)
            integer lda, lwork
            character uplo
            double precision matc(mmax,mmax),matct(mmax,mmax)
            &,w2(mmax),lambda(mmax),work(lw), mata(mmax,mmax),mat(mmax,mmax)
            &,wi(mmax),dd(mmax,mmax)
            external f02faf, x04caf,f01ckf,f02eaf
            common /j/j
            common/par/par
            common/dd/dd
            double precision par(3)
            double precision matm(mmax,mmax),matd(mmax,mmax)
            double precision matam(mmax,mmax)
            lda=mmax
            lwork=64*m
            uplo='l'
C Matrice matd(m,m)
600    format(12f8.4)
        do 7 i=1,m
            do 8 j=1,m

```

```

        matd(i,j)=(par(1)**(i+j-2))*(1-(par(1)**2))
C      matd(i,j)=dd(i,j)
      8      continue
      7      continue
C Matrice matm(m,m)
      do 10 i =1,m
      do 11 j=1,m
        matm(i,j)=(0-matd(i,j))
        matm(i,i)=(1-matd(i,i))
11      continue
10      continue
C      matrice A
      if (ai.le.m-1)then
      do 31 l=1,m-ai
      do 32 j=1,m
        matc(l,j)=0
32      continue
31      continue
      do 33 l=1,m-ai
      do 34 j= l,ai+1
        matc(l,j)=1
34      continue
33      continue
      do 36 i=1,m
      do 37 j=1,m-ai
        matct(i,j)=matc(j,i)
37      continue
36      continue
      do 50 i=1,m
      do 51 j=1,m
        mata(i,j)=0.0d0
      do 52 il=1,m-ai
52      mata(i,j)=mata(i,j)+matct(i,il)*matc(il,j)
51      continue
50      continue
      endif
      call mult(matm,m,m,mata,m,matam)
C Valeurs propres de la matrice m*a
      do 40 i=1,m
      do 41 j=1,i
        mat(i,j)=matam(i,j)
41      continue

```

```

40    continue
      ifail = 0
      call f02faf('vectors',uplo,m,mat,lda,w2,work,lwork,ifail)
      ifail = 0
C Valeurs propres non nulles de la matrice m*a
      k=0
      do i=1,m
        if(abs(w2(i)).ge..0000001)then
          k=k+1
          lambda(k)= w2(i)
        endif
      enddo
      return
    end

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C Sous-programme pour le calcul de la matrice      C
C de covariance des autocorrélations résiduelles C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      subroutine covariance(m,ip,iq,dd)
      implicit real*8 (a-h,o-z)
      parameter (mmax=12,ipmax=12)
      double precision psi(-mmax:mmax),tat,par(3),matx(mmax,ipmax),
&matxt(ipmax,mmax),matxtx(ipmax,ipmax),xxtinv(ipmax,ipmax),
&d(mmax,ipmax),dd(mmax,mmax)
      integer j,k,m
      common/par/par
      external f01adf
      psi(0)=1.
      psi(1)=par(1)
      do j =2,(m-1)
        tat = 0.
        do k = 1,ip
          tat = tat + par(k) * psi(j - k)
        enddo
        psi(j)= tat
      enddo
      do i=1,m
        do j=1,ip
          matx(i,j)=0
        enddo
      enddo
      do i=1,m

```

```

do j=1,ip
  if (i.ge.j) then
    matx(i,j)=psi(i-j)
  endif
enddo
enddo
do i=1,ip
  do j=1,m
    matxt(i,j)=matx(j,i)
  enddo
enddo
call mult(matxt,ip,m,matx,ip,matxtx)
call f0ladf (ip,matxtx,ipmax,ifail)
do i=1,ip
  do j=1,i
    xxtinv(i,j)=matxtx(i+1,j)
  enddo
enddo
do i=1,ip
  do j=i,ip
    xxtinv(i,j)= xxtinv(j,i)
  enddo
enddo
call mult(matx,m,ip,xxtinv,ip,d)
call mult(d,m,ip,matxt,m,dd)
return
end

```

```

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C Sous-programme pour le produit de deux matrices C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

```

```

subroutine mult(a,n,m,b,k,c)
parameter (mmax=12,ipmax=12)
double precision a(ipmax,mmax),b(mmax,ipmax),c(ipmax,ipmax)
implicit real*8 (a-h,o-z)
do1i=1,n
do2j=1,k
c(i,j)=0.0d0
do3il=1,m
3 c(i,j)=c(i,j)+a(i,il)*b(il,j)
2 continue
1 continue
return

```

```

end
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C Sous-programme pour le calcul de la statistique de Hong C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
  subroutine pmanteau(r2,n,pn,statis,choix)
  implicit double precision (a-g,o-z)
  integer n, choix,pn, tronque
  double precision r2(n-1,1),qrn,statis, cnk, dnk
  double precision bartlett, parzen, daniell
c   choix = 1  bartlett
c   choix = 2  daniell
c   choix = 3  parzen
c   choix = 4  tronque
  if (choix .eq. 1) then
    cnk = 0.0d0
    do i=1,n-1
      cnk = cnk + (1.0d0 - i/(n+0.0d0))*(bartlett(i/(pn+0.0d0))**2)
    enddo
    dnk = 0.0d0
    do i=1,n-2
      dnk = dnk + (1.0d0-i/(n+0.0d0))*
+ (1.0d0-(i+1.0d0)/(n+0.0d0))*
+ (bartlett(i/(pn+0.0d0))**4)
    enddo
    qrn = 0.0d0
    do i=1,n-1
      QRn = QRn +(n*(bartlett(i/(pn+0.0d0))**2 * r2(i,1)**2))/
+ dsqrt(2.0d0* dnk)
    enddo
    statis = QRn - cnk/dsqrt(2.0d0* dnk)
  endif
  if (choix .eq. 2) then
    cnk = 0.0d0
    do i=1,n-1
      cnk = cnk + (1.0d0 - i/(n+0.0d0))*(daniell(i/(pn+0.0d0))**2)
    enddo
    dnk = 0.0d0
    do i=1,n-2
      dnk = dnk + (1.0d0-i/(n+0.0d0))*
+ (1.0d0-(i+1.0d0)/(n+0.0d0))*
+ (daniell(i/(pn+0.0d0))**4)
    enddo

```

```

qrn = 0.0d0
do i=1,n-1
qrn = qrn +(n*(daniell(i/(pn+0.0d0))**2 * r2(i,1)**2))/
+ dsqrt(2.0d0* dnk)
enddo
statis = qrn - cnk/dsqrt(2.0d0* dnk)
endif
if (choix .eq. 3) then
cnk = 0.0d0
do i=1,n-1
cnk = cnk + (1.0d0 - i/(n+0.0d0))*(parzen(i/(pn+0.0d0))**2)
enddo
dnk = 0.0d0
do i=1,n-2
dnk = dnk + (1.0d0-i/(n+0.0d0))*
+ ( 1.0d0-(i+1.0d0)/(n+0.0d0))*
+ (parzen(i/(pn+0.0d0))**4)
enddo
qrn = 0.0d0
do i=1,n-1
qrn = qrn +(n*(parzen(i/(pn+0.0d0))**2 * r2(i,1)**2))/
+ dsqrt(2.0d0* dnk)
enddo
statis = qrn - cnk/dsqrt(2.0d0* dnk)
endif
if (choix .eq. 4) then
cnk = 0.0d0
do i=1,n-1
cnk = cnk + (1.0d0 - i/(n+0.0d0))*(tronque(i/(pn+0.0d0))**2)
enddo
dnk = 0.0d0
do i=1,n-2
dnk = dnk + (1.0d0-i/(n+0.0d0))*
+ (1.0d0-(i+1.0d0)/(n+0.0d0))*
+ (tronque(i/(pn+0.0d0))**4)
enddo
qrn = 0.0d0
do i=1,n-1
qrn = qrn +(n*(tronque(i/(pn+0.0d0))**2 * r2(i,1)**2))/
+ dsqrt(2.0d0* dnk)
enddo
statis = qrn - cnk/dsqrt(2.0d0* dnk)

```

```

        endif
        return
    end
CCCCCCCCCCCCCCCCCCCC
C Fonction Bartlett C
CCCCCCCCCCCCCCCCCCCC
    double precision function bartlett(i)
    implicit double precision (a-g,o-z)
    double precision i
    if (dabs(i+0.d0).le.1.0d0) then
        bartlett=1.0d0-dabs(i+0.0d0)
    else bartlett=0.0d0
    endif
    return
end
CCCCCCCCCCCCCCCCCCCC
C Fonction Daniell C
CCCCCCCCCCCCCCCCCCCC
    double precision function daniell(i)
    implicit double precision (a-g,o-z)
    double precision i,pi
    parameter(pi=3.141593)
    if (i .ne. 0) then
        daniell=dsin(pi*i)/(pi*i)
    else
        daniell=1
    endif
    return
end
CCCCCCCCCCCCCCCCCCCC
C Fonction Parzen C
CCCCCCCCCCCCCCCCCCCC
    double precision function parzen(i)
    implicit double precision (a-g,o-z)
    double precision i,pi
    parameter (pi=3.141593)
    if (dabs(i).le.3/pi)then
        parzen=1-6*((pi*i)/6)**2+6*(dabs((pi*i)/6))**3
    endif
    if (dabs(i).ge.3/pi .and. dabs(i).le.6/pi )then
        parzen=2*(1-dabs((i*pi)/6))**3
    endif

```



```
        if (dabs(i).gt.6/pi)then
            parzen=0.0d0
        endif
        return
    end

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C Fonction noyau uniforme tronque C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
    integer function tronque(i)
    implicit double precision (a-g,o-z)
    double precision i
    if (dabs(i).le.1.0d0) then
        tronque=1
    else
        tronque=0
    endif
    return
end
```

II. PROGRAMMES POUR L'ETUDE DE PUISSANCE

Pour la génération, l'estimation, le calcul des statistiques de test et les valeurs critiques exactes, le programme décrit ci-dessus a été adapté pour chacun des modèles utilisés pour générer les séries et pour chacun des modèles estimés. Pour éviter la répétition, nous présentons seulement la portion du programme qui a servi pour le calcul des puissances en faisant appel aux valeurs critiques exactes calculées à l'aide du premier programme.

```

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C Portion du programme pour      C
C le calcul des puissances avec C
C valeurs critiques exactes      C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

C Lecture des valeurs critiques exactes
  open (unit=1,file="fort.12",status="old")
  do i=1,m
  READ (1,*)(crit(i,j),j=1,3)
  enddo
  open (unit=2,file="fort.2",status="old")
  do i=1,20
  READ (2,*)(critere(i,j),j=1,3)
  enddo

C Clacul des statistiques et tests
C Koch et Yang
  do ai=0,m-1
  call koch(m,ai,r2,t)
  if (t .gt. crit(ai+1,1)) then
  rejet11=1
  else if (t .lt. crit(ai+1,1))then
  rejet11=0
  endif
  if (t .gt.crit(ai+1,2))then
  rejet12=1
  else if ( t .lt.crit(ai+1,2)) then
  rejet12=0
  endif
  if ( t .gt.crit(ai+1,3))then
  rejet13=1
  else if ( t .lt.crit(ai+1,3))then

```

```

rejet13=0
endif
vectr6(ai)=rejet11
sumvectr6(ai)=sumvectr6(ai)+vectr6(ai)
vectr4(ai)=rejet12
sumvectr4(ai)=sumvectr4(ai)+vectr4(ai)
vectr3(ai)=rejet13
sumvectr3(ai)=sumvectr3(ai)+vectr3(ai)
enddo

```

C Hong

```

do choix=1,4
  compt=0
  do ipn=1,5
    call pmanteau(r2,n,pnn(ipn),statis,choix)
    if (statis.gt. critere(choix+compt,1)) then
      rejet21=1
    else if (statis.lt. critere(choix+compt,1))then
      rejet21=0
    endif
    if (statis.gt.critere(choix+compt,2))then
      rejet22=1
    else if (statis.lt.critere(choix+compt,2)) then
      rejet22=0
    endif
    if (statis .gt.critere(choix+compt,3))then
      rejet23=1
    else if (statis .lt.critere(choix+compt,3))then
      rejet23=0
    endif
    element6(choix,ipn)=rejet21
    element2(choix,ipn)=rejet22
    element3(choix,ipn)=rejet23
    somme6(choix,ipn)=somme6(choix,ipn)+element6(choix,ipn)
    somme2(choix,ipn)=somme2(choix,ipn)+element2(choix,ipn)
    somme3(choix,ipn)=somme3(choix,ipn)+element3(choix,ipn)
    compt=compt+4
  enddo
enddo

```

C Proportion de rejet de Ho: valeurs critiques exactes

```

write(*,*)"Puissances:MA(15) estime par AR(1)"
write(*,*)"proportion de rejet : valeurs critiques exactes"
write(*,*)"m = ", m, "n = ", n, " taille = ", taille

```

```
write(*,*)" test de Koch "  
do ai=0,m-1  
  probrejet1(ai)=(sumvectr6(ai)+0.0d0)/(taille+0.0d0)  
  probrejet2(ai)=(sumvectr4(ai)+0.0d0)/(taille+0.0d0)  
  probrejet3(ai)=(sumvectr3(ai)+0.0d0)/(taille+0.0d0)  
enddo  
Write(*,*) "alpha=0.01 alpha=0.05 alpha=0.1"  
do ai=0,m-1  
  write(*,600) ai, probrejet1(ai),probrejet2(ai),probrejet3(ai)  
enddo  
write(*,*)" test de Hong "  
do choix=1,4  
  do ipn=1,5  
    probelem1(choix,ipn)=(somme6(choix,ipn)+0.0d0)/(taille+0.0d0)  
    probelem2(choix,ipn)=(somme2(choix,ipn)+0.0d0)/(taille+0.0d0)  
    probelem3(choix,ipn)=(somme3(choix,ipn)+0.0d0)/(taille+0.0d0)  
  enddo  
enddo  
Write(*,*) "alpha=0.01"  
do choix=1,4  
  write(*,600)(probelem1(choix,ipn),ipn=1,5)  
enddo  
Write(*,*) "alpha=0.05"  
do choix=1,4  
  write(*,600)(probelem2(choix,ipn),ipn=1,5)  
enddo  
Write(*,*) "alpha=0.1"  
do choix=1,4  
  write(*,600)(probelem3(choix,ipn),ipn=1,5)  
enddo  
600 format(12f8.3)  
500 format(1x,12f7.4)  
stop  
end
```

Bibliographie

Box, G. E. P. et Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65, 1509-26.

Brockwell, P.J. et Davis, R.A (1991). *Time Series : Theory and Methods*, 2e édition, Springer-Verlag, New York.

Brockwell, P.J. et Davis, R.A (2002). *Introduction to Time Series and Forecasting*, 2e édition, Springer, New York.

Chen, W.W. et Deo, R.S. (2003). Power transformation to induce normality and their applications. *Journal of the Royal Statistical Society, series B*, à paraître.

Davis, N., Triggs, C. M. et Newbold, P. (1977). Significance levels of the Box-Pierce portmanteau statistic in finite samples. *Biometrika*, 64, 517-22.

Duchesne, P. et Roy, R. (2003). Robust tests for independence of two time series. *Statistica Sinica*, 13, 827-852.

Duchesne, P. et Roy, R. (2004). On consistent testing for serial correlation of unknown form in vector time series models. *Journal of Multivariate Analysis*, à paraître.

Franco, C., Roy, R. et Zakoïan, J-M (2003). Goodness-of-fit tests for ARMA models with uncorrelated errors. Rapport de recherche CRM-2925, Université de Montréal, juillet 2003.

Graybill, F.A. (1976). *Theory and Application of the Linear Model*, Duxbury Press, North Scituate, MA.

Haugh, L. D. (1976). Checking the Independence of Two Covariance Stationary Time Series : A Univariate Residual Cross-Correlation Approach. *Journal of the American Statistical Association*, 71, 378-385.

Hong, Y. (1996a). Consistent testing for serial correlation of unknown form. *Econometrica*, 64, 837-864.

Hong, Y. (1996b). Testing for independence between two covariance stationary time series. *Biometrika*, 83, 615-625.

Hong, Y. et Lee, Y.J. (2003). Consistent testing for serial correlation of unknown form under general conditional heteroscedasticity. Rapport de recherche, Department of economics and department of statistical science, Cornell University, avril 2003.

Imhof, J., P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48, 419-426.

Koerts, J. et Abrahamse, A. P. J. (1969). *On the Theory and Application of the General Linear Model*, Rotterdam University Press.

Koch, P. D. et Yang, S. (1986). A method for testing the independence of two time series that accounts for a potential pattern in the cross-correlation function. *Journal of the American Statistical Association*, 81, 533-544.

Ljung, G. M. et Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika* 65, 297-303.

McLeod, A., I. (1978). On the distribution of residual autocorrelations in Box-Jenkins models. *Journal of the Royal Statistical Society, series B.* 40, 296-302.

Prothero, D. L. et Wallis, K. F. (1976). Modelling macroeconomic time series (with discussion). *Journal of the Royal Statistical Society, series A*, 139, 468-500.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*, John Wiley and Sons, New York.

Shumway R. et S. Stoffer D. (2000). *Time Series Analysis and its Applications*, Springer, New York.

