

Université de Montréal

Sur l'étude de la transformation des tests
portemanteaux pour séries chronologiques
multivariées

par

Jennifer Poulin

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 2006



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

**Sur l'étude de la transformation des tests
portemanteaux pour séries chronologiques
multivariées**

présenté par

Jennifer Poulin

a été évalué par un jury composé des personnes suivantes :

Yves Lepage

(président-rapporteur)

Pierre Duchesne

(directeur de recherche)

David Haziza

(membre du jury)

Mémoire accepté le:

SOMMAIRE

Dans le contexte des séries chronologiques multivariées, les tests portemanteaux sont couramment utilisés pour valider l'ajustement de modèles. Dans le cas des séries chronologiques stationnaires et partiellement non-stationnaires, Duchesne et Roy (2004) et Duchesne (2005a) ont proposé des statistiques de test visant à vérifier l'hypothèse de non-corrélation des erreurs. Leurs tests sont basés sur la comparaison de la densité spectrale des erreurs sous l'hypothèse nulle de non-corrélation avec un estimateur de la densité spectrale par la méthode du noyau, utilisant une métrique quadratique. Malgré la distribution asymptotique normale des tests, la distribution empirique des statistiques de test se révèle être très asymétrique vers la droite en échantillons finis, causant fréquemment une distorsion appréciable dans les niveaux des tests.

Le but du présent travail consiste à proposer une transformation des statistiques de test de Duchesne et Roy (2004) et Duchesne (2005a) afin de corriger en bonne partie le problème d'asymétrie survenant dans la distribution exacte des tests pour une taille d'échantillon finie. Pour y arriver, nous abordons le problème en considérant le choix d'une transformation appropriée de type puissance, en nous basant sur les travaux effectués par Chen et Deo (2004a). De plus, suite à une étude approfondie des statistiques de test, nous obtenons de meilleures expressions de la moyenne et de la variance de la mesure de distance intervenant dans la définition des statistiques de test. Ceci permet d'améliorer considérablement les propriétés en échantillons finis des tests étudiés dans un bon nombre de situations.

Une étude de simulation Monte Carlo est menée afin de comparer les statistiques de test originales de Duchesne et Roy (2004) et Duchesne (2005a) avec

leurs versions transformées. Plus précisément, nous effectuons une étude de niveau et de puissance. Concernant l'étude de niveau, les taux de rejet de l'hypothèse nulle lorsque cette dernière est vraie sont rapportés. Pour l'étude de puissance, plusieurs alternatives sont considérées. Les puissances calculées sont déterminées selon les points critiques issus de la théorie asymptotique et selon les points critiques empiriques (ces derniers sont déterminés à l'aide de l'étude de niveau). De plus, nous comparons les tests originaux et transformés au test portemanteau de Hosking (1980).

De l'étude de simulation, plusieurs conclusions s'imposent. Nous pouvons affirmer que les statistiques de test transformées proposées présentent beaucoup d'avantages comparativement à leurs versions originales. En effet, les niveaux empiriques des tests transformés sont de manière générale beaucoup plus près des niveaux nominaux, ce qui est grandement souhaitable dans les applications. De plus, les puissances empiriques sont assez près des puissances asymptotiques pour les tests transformés. Conséquemment, étant donné qu'en pratique les valeurs critiques asymptotiques sont utilisées, les statistiques de test transformées se comparent fort avantageusement aux statistiques de test originales.

Mots clés : Cointégration, estimation de la densité spectrale par la méthode du noyau, séries chronologiques multivariées, tests portemanteaux, transformation puissance, densité spectrale.

SUMMARY

In multivariate time series, portmanteau test statistics are frequently used to check the adequacy of fitted models. For stationary and partially non-stationary vector time series, Duchesne and Roy (2004) and Duchesne (2005a) have proposed test statistics designed to test the null hypothesis of uncorrelated error terms. Their tests are based on the comparison of the spectral density of the errors under the null hypothesis to a kernel-based spectral density estimator, using a quadratic distance measure. Both tests have been shown to be asymptotically normal, but being heavily right skewed in finite samples often causes a distortion in the levels of the tests.

In this Master's thesis, we propose a power transformation of Duchesne and Roy (2004) and Duchesne's (2005a) test statistics to improve the normal approximation in finite samples. Our work is based on the simple method proposed by Chen and Deo (2004a) for obtaining a power transformation of the test statistics. In addition, we provide better finite sample expressions for the mean and variance of the distance measure intervening in the test statistics. This leads to improvements of the finite sample properties of the test statistics under study in many situations.

A Monte Carlo simulation is conducted in order to compare Duchesne and Roy (2004) and Duchesne's (2005a) test statistics with their transformed versions. Specifically, a level and power study is carried out. For the level analysis, the rejection rates of the null hypothesis are calculated when the error terms are uncorrelated. The empirical powers are then calculated using asymptotic critical values and empirical critical values (which are determined from the level study).

We also compare the original and transformed test statistics to Hosking's (1980) portmanteau test statistic.

The comparative study described above confirms that the transformed test statistics have superior empirical properties to their original versions. Indeed, the empirical levels of the transformed tests are much closer to the nominal levels, which is highly desirable in practical applications. In addition, the empirical powers are quite close to the asymptotic powers. Therefore, since in practice asymptotic critical values are used, these transformed test statistics enhance their original versions.

Keywords : Cointegration, kernel-based spectral density estimation, multiple time series, portmanteau test statistics, power transformation, spectral density.

TABLE DES MATIÈRES

Sommaire	iii
Summary	v
Liste des figures	viii
Liste des tableaux	ix
Remerciements	1
Introduction	2
Chapitre 1. Préliminaires	7
1.1. Séries chronologiques multivariées.....	7
1.2. Processus stationnaires.....	8
1.2.1. Notions de base.....	8
1.2.2. Estimation de la fonction d'autocovariance.....	11
1.2.3. Estimation d'un processus autorégressif multivarié.....	12
1.3. Modèles non-stationnaires.....	14
1.3.1. Notions de base.....	15
1.3.2. Estimation d'un processus autorégressif multivarié partiellement non-stationnaire.....	19
1.3.2.1. Estimation par la méthode du maximum de vraisemblance.	19
1.3.2.2. Estimation à deux étapes.....	21
1.4. Densité spectrale.....	22
1.4.1. Notions de base.....	22

1.4.2. Estimation de la densité spectrale.....	24
Chapitre 2. L'article.....	27
Conclusion.....	65
Annexe A. Simulations complémentaires.....	A-i
A.1. Description des expériences.....	A-i
A.2. Discussion des résultats empiriques.....	A-iii
A.2.1. Étude de niveau.....	A-iv
A.2.2. Étude de puissance.....	A-v
Annexe B. Démonstrations des principaux résultats.....	B-i
Bibliographie.....	i

LISTE DES FIGURES

- A.1 Représentation graphique des noyaux utilisés pour les statistiques de test transformée T_n^* définie par (2.0.16) et originale T_n donnée par (2.0.6). A-iii

LISTE DES TABLEAUX

2.1	Model coefficients of the data generating processes in the level analysis.	42
2.2	Additional model coefficients of the data generating processes in the power analysis.	45
2.3	Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP_1	53
2.4	Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP_2	54
2.5	Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP_3	55
2.6	Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6),	

- for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP₄. 56
- 2.7 Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP₅. 57
- 2.8 Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP₆. 58
- 2.9 Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP₁. 59
- 2.10 Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for

- the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP_2 60
- 2.11 Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP_3 61
- 2.12 Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP_4 62
- 2.13 Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP_5 63
- 2.14 Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for

- the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP₆. 64
- A.1 Valeurs des paramètres des modèles générés selon les DGP₇ et DGP₈. A-ii
- A.2 Valeurs de la puissance β définie par (2.0.15) et intervenant dans la statistique de test transformée $T_n^*(k; p_n)$ définie par (2.0.16) A-iv
- A.3 Niveaux empiriques (en pourcentage) de la statistique de test transformée T_n^* définie par (2.0.16), de la statistique de test originale T_n donnée par (2.0.6), pour différents noyaux et différents paramètres de lissage ou points de troncature p_n , et la statistique de test portemanteau de Hosking's (1980) (2.0.5) utilisant $M = p_n$, pour le processus VAR(1) de dimension $d = 3$ généré par le DGP₇. A-vii
- A.4 Niveaux empiriques (en pourcentage) de la statistique de test transformée T_n^* définie par (2.0.16), de la statistique de test originale T_n donnée par (2.0.6), pour différents noyaux et différents paramètres de lissage ou points de troncature p_n , et la statistique de test portemanteau de Hosking's (1980) (2.0.5) utilisant $M = p_n$, pour le processus VAR(3) de dimension $d = 2$ généré par le DGP₈. A-viii
- A.5 Puissances empiriques basées sur les valeurs critiques asymptotiques et empiriques (entre parenthèses) de la statistique de test transformée T_n^* définie par (2.0.16), de la statistique de test originale T_n donnée par (2.0.6), pour différents noyaux et différents paramètres de lissage ou points de troncature p_n , et de la statistique de test portemanteau de Hosking's (1980) (2.0.5) utilisant $M = p_n$, pour le processus VARMA(1,1) de dimension $d = 3$ généré par le DGP₇. A-ix
- A.6 Puissances empiriques basées sur les valeurs critiques asymptotiques et empiriques (entre parenthèses) de la statistique de test transformée T_n^* définie par (2.0.16), de la statistique de test originale T_n donnée par (2.0.6), pour différents noyaux et différents paramètres de lissage

ou points de troncature p_n , et de la statistique de test portemanteau de Hosking's (1980) (2.0.5) utilisant $M = p_n$, pour le processus VARMA(3,1) de dimension $d = 2$ généré par le DGP₈. A-x

REMERCIEMENTS

Dans un premier temps, je tiens à remercier mon directeur de recherche Pierre Duchesne. Tout au long de la réalisation de ce travail de recherche, il a su me donner des conseils judicieux. Il m'a toujours soutenu par sa grande disponibilité et par nos rapports cordiaux entretenus. Je tiens également à remercier son aide financière qui m'a permis d'aller jusqu'au bout de ce travail.

Dans un second temps, je veux remercier ma famille, mon père Normand, ma mère Lisa et mon frère Steven, pour leurs encouragements et leur support moral et financier tout au long de mes études. Pour ma grand-mère Ida : « Nonna, ti ringrazio per tutto l'aiuto e l'amore che mi hai dato in tutti questi anni ».

Je tiens également à remercier mes amis de longue date pour leurs encouragements et mes confrères de classe qui ont rendu cette expérience plus enrichissante. Finalement, je tiens à remercier une personne qui m'est très chère pour m'avoir généreusement offert son aide.

INTRODUCTION

Depuis quelques années, un grand nombre de chercheurs ont manifesté un intérêt croissant pour des méthodes de validation de l'ajustement des modèles de séries chronologiques multivariées. Ces derniers se sont concentrés sur des modèles de séries chronologiques stationnaires et partiellement non-stationnaires. Une approche possible visant à valider l'ajustement de modèles consiste à étudier le comportement des autocovariances résiduelles. Certains exemples de recherches effectuées dans le contexte des modèles stationnaires sont Hosking (1980, 1981a, 1981b), Li et McLeod (1981) et Duchesne et Roy (2004), pour ne citer que ces travaux. Dans leurs ouvrages, ils ont étudié les propriétés théoriques des autocovariances résiduelles ainsi que la distribution asymptotique des statistiques de test portemanteaux. Ce n'est que récemment que les propriétés asymptotiques des autocovariances résiduelles et les procédures de tests visant à vérifier l'indépendance des erreurs des modèles ajustés ont été développées pour les modèles partiellement non-stationnaires. En effet, Brüggemann, Lütkepohl et Saikkonen (2006) ont étudié les propriétés des autocovariances résiduelles pour des modèles partiellement non-stationnaires pouvant s'exprimer selon la représentation dite à correction d'erreur. De plus, ils ont obtenu des statistiques de test qui permettent de détecter la présence d'autocorrélation résiduelle. Duchesne (2005a) a aussi adapté les statistiques de test proposées par Duchesne et Roy (2004) dans le contexte des séries chronologiques partiellement non-stationnaires.

Les statistiques de test de Duchesne et Roy (2004) et de Duchesne (2005a) ont été développées dans le domaine des fréquences pour les séries chronologiques multivariées. Les tests sont basés sur la comparaison entre la densité spectrale

des erreurs sous l'hypothèse nulle d'indépendance et d'un estimateur non paramétrique de la densité spectrale calculé selon la méthode du noyau, à l'aide d'une mesure de distance, comme la métrique L_2 . De plus, les tests spectraux sont exprimés comme une somme pondérée d'une fonction des autocovariances résiduelles pour les $n - 1$ premiers délais de la série chronologique de longueur n . Les poids intervenant dans les tests sont déterminés par deux éléments : le premier est une fonction de noyau et le second est un point de troncature ou un paramètre de lissage, tout dépendant de la nature du noyau. Il est aussi démontré, autant pour les modèles stationnaires que pour les modèles partiellement non-stationnaires, que les statistiques de test spectraux de Duchesne et Roy (2004) et Duchesne (2005a) convergent vers une distribution normale sous l'hypothèse nulle d'adéquation.

Malgré la convergence normale des statistiques de test proposées par Duchesne et Roy (2004) et Duchesne (2005a), la distribution empirique de ces dernières est très asymétrique vers la droite en échantillons finis, ce qui cause une distorsion dans les niveaux des tests. Chen et Deo (2004a) ont remarqué que des variables aléatoires exprimées comme une combinaison linéaire non-négative de variables aléatoires positives peuvent être asymétriques à la droite en échantillons finis. Ils ont également remarqué que malgré une certaine standardisation, ces statistiques demeurent asymétriques. Pour améliorer l'approximation normale de ces variables aléatoires standardisées, ils ont proposé une procédure simple qui corrige en bonne partie le problème d'asymétrie. La procédure consiste à déterminer une transformation de type puissance de la statistique. Pour illustrer leur approche, ils ont appliqué leur méthode de transformation aux statistiques de test de Hong (1996). Ces statistiques de test, servant à vérifier l'existence d'autocorrélation sérielle pour des séries chronologiques univariées stationnaires, sont en fait un cas particulier des statistiques de test de Duchesne et Roy (2004) lorsque la dimension du processus est unitaire. Comme dans le cas des statistiques de test de Duchesne et Roy (2004), la distribution exacte des statistiques de test de Hong (1996) est très asymétrique pour une taille d'échantillon finie. Chen et Deo (2004a) ont remarqué qu'une transformation de type puissance n'était pas suffisante en soi. En effet, certains ajustements supplémentaires de la moyenne

et de la variance de la mesure de distance intervenant dans la définition des statistiques de test contribuent à améliorer grandement les propriétés des niveaux empiriques des statistiques de test. Ces ajustements dans la façon de standardiser les statistiques de test reposent sur la distribution asymptotique des autocorrélations résiduelles. De plus, suite aux études empiriques menées par Chen et Deo (2004a), ils ont découvert, pour certaines hypothèses alternatives, que les versions corrigées des statistiques de test affichent un gain en puissance empirique relativement appréciable en échantillons finis.

Dans ce travail, nous déterminons la transformation appropriée afin de corriger en bonne partie le problème d'asymétrie des tests de Duchesne et Roy (2004) et Duchesne (2005a) en échantillons finis. Nous adaptons la méthode de Chen et Deo (2004a) dans notre contexte. Également, nous démontrons que la transformation de type puissance des statistiques de test est invariante par rapport au nombre de variables définissant le processus multivarié. De plus, nous proposons à l'aide des distributions des autocovariances résiduelles obtenues par Li et McLeod (1981) dans le cas des séries chronologiques stationnaires, et par Brüggemann *et al.* (2006) dans le cas de séries chronologiques partiellement non-stationnaires, de meilleures approximations de la moyenne et de la variance de la mesure de distance intervenant dans la définition des statistiques de test.

Dans une étude empirique, nous comparons les propriétés en échantillons finis des statistiques de test de Duchesne et Roy (2004) et Duchesne (2005a) avec leurs versions corrigées obtenues par la méthode de transformation de Chen et Deo. Étant donné que Brüggemann *et al.* (2006) démontrent que dans le cas de séries chronologiques partiellement non-stationnaires le test de Hosking (1980) est approprié, nous comparons les tests spectraux à celui-ci. En effet, Brüggemann *et al.* (2006) notent l'utilité du test de Hosking (1980) dans certaines circonstances, particulièrement lorsque les autocorrélations résiduelles de délais modérés ou grands sont significatives. Cependant, ils remarquent que les degrés de liberté de ce test doivent être ajustés afin de tenir compte du rang de cointégration. À des fins pratiques, nous comparons et évaluons empiriquement à l'aide de simulations Monte Carlo les statistiques de test de Duchesne et Roy (2004) et Duchesne

(2005a) avec leurs versions corrigées et avec le test modifié de Hosking (1980), en terme de niveau et de puissance exacts.

Le présent mémoire est organisé comme suit. Au premier chapitre, nous présentons quelques notions de base sur les séries chronologiques multivariées. Pour être plus précis, nous introduisons les concepts de séries chronologiques stationnaires et partiellement non-stationnaires pour des modèles autorégressifs moyennes mobiles multivariés (VARMA). Également, nous introduisons les procédures d'estimation pour les processus autorégressifs multivariés (VAR). De plus, une revue sur l'analyse spectrale est présentée pour les séries chronologiques multivariées.

Au second chapitre, nous présentons une copie intégrale d'un article qui est actuellement soumis dans un journal scientifique avec comité de lecture. Le premier auteur de l'article est également l'auteur de ce mémoire. Ce chapitre introduit les séries chronologiques stationnaires et la forme à correction d'erreur des séries chronologiques partiellement non-stationnaires. Pour ces deux types de séries chronologiques, les matrices d'autocovariances résiduelles sont définies et les propriétés asymptotiques des autocovariances résiduelles sont énoncées. Les procédures de tests proposées par Hosking (1980), Duchesne et Roy (2004) et Duchesne (2005a) sont également présentées. De plus, l'article explique la méthode de transformation de Chen et Deo (2004a), et les versions corrigées des statistiques de test de Duchesne et Roy (2004) et Duchesne (2005a) sont obtenues. Les propriétés en échantillons finis des statistiques de test transformées de Duchesne et Roy (2004) et Duchesne (2005a) sont étudiées à l'aide de simulations Monte Carlo. Pour terminer, ces statistiques de test sont comparées, en terme de niveau et de puissance, à leurs versions originales et au test modifié de Hosking (1980). Finalement, quelques remarques concluantes sont présentées.

Nous concluons en discutant les contributions originales de l'article. Des études de simulation complémentaires sont fournies en annexes afin de compléter l'étude des propriétés en échantillons finis des statistiques de test transformées de Duchesne et Roy (2004) et Duchesne (2005a). Des détails supplémentaires sur les principaux résultats obtenus sont également fournis. Les modèles additionnels

considérés sont un modèle autorégressif d'ordre un stationnaire de dimension 3 et un modèle autorégressif d'ordre 3 partiellement non-stationnaire de dimension 2.

Chapitre 1

PRÉLIMINAIRES

Dans ce chapitre, nous présentons une brève revue des notions sur les séries chronologiques multivariées qui touchent principalement le présent travail. Tout d'abord, une série chronologique univariée est un ensemble d'observations recueillies d'un phénomène quelconque pour un intervalle de temps donné. Lorsque plusieurs séries chronologiques sont observées simultanément dans le temps, nous parlons de séries chronologiques multivariées. Certains exemples d'applications sont les cotations de plusieurs indices boursiers, l'imagerie du cerveau et les fréquences de répliques sismiques de tremblements de terre. Voir par exemple Shumway et Stoffer (2006) ou Tsay (2005) pour ce genre d'applications.

1.1. SÉRIES CHRONOLOGIQUES MULTIVARIÉES

Une série chronologique multivariée est définie comme une réalisation finie d'un processus stochastique multivarié $\mathbf{Y} = \{\mathbf{Y}_t, t \in \mathbb{Z}\}$, où le vecteur aléatoire $\mathbf{Y}_t = (Y_t(1), \dots, Y_t(d))^\top$ est de dimension d et $Y_t(i)$ représente la $i^{\text{ième}}$ composante de \mathbf{Y}_t . Pour une telle série multivariée avec $E\{Y_t^2(i)\} < \infty$ pour tout t et tout i , les propriétés du deuxième ordre sont spécifiées par le vecteur des moyennes,

$$\boldsymbol{\mu}_t = E(\mathbf{Y}_t) = (\mu_t(1), \dots, \mu_t(d))^\top, \quad (1.1.1)$$

et les matrices de covariance,

$$\boldsymbol{\Gamma}(t, t-h) = E\{(\mathbf{Y}_t - \boldsymbol{\mu}_t)(\mathbf{Y}_{t-h} - \boldsymbol{\mu}_{t-h})^\top\} = [\gamma_{ij}(t, t-h)]_{i,j=1}^d. \quad (1.1.2)$$

Les covariances croisées $\gamma_{ij}(t, t - h)$ donnent non seulement une indication de la dépendance entre les observations d'une même série (lorsque $i = j$), mais aussi entre les observations de différentes séries (lorsque $i \neq j$). Lorsque les propriétés du deuxième ordre sont invariantes du temps, nous parlons de stationnarité (au sens large). Sinon, le processus $\{\mathbf{Y}_t\}$ n'est pas stationnaire.

1.2. PROCESSUS STATIONNAIRES

Lors de l'analyse des séries chronologiques multivariées, le concept de processus stationnaire joue un rôle fort important. Les théories développées permettent la modélisation et l'ajustement de modèles ainsi que la prévision d'observations. Dans cette section, quelques notions de base sont abordées, et nous décrivons l'estimation des paramètres d'un modèle de série chronologique lorsque ce dernier peut être représenté par une structure dite autorégressive.

1.2.1. Notions de base

Définition 1.2.1. *Un processus $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ de dimension d , tel que $E\{Y_t^2(i)\} < \infty$ pour tout t et tout i , est stationnaire au sens large si les premiers et deuxièmes moments sont invariants du temps. Le vecteur des moyennes et la matrice de covariance de délai h sont respectivement :*

- (i) $\boldsymbol{\mu} = E(\mathbf{Y}_t) = (\mu(1), \dots, \mu(d))^\top$ et
- (ii) $\Gamma(h) = E\{(\mathbf{Y}_t - \boldsymbol{\mu})(\mathbf{Y}_{t-h} - \boldsymbol{\mu})^\top\} = [\gamma_{ij}(h)]_{i,j=1}^d$.

Dans le cas où le processus $\{\mathbf{Y}_t\}$ est stationnaire avec une matrice de covariance $\Gamma(\cdot)$, la série univariée $\{Y_t(i)\}$ est stationnaire pour tout i avec fonction d'autocovariance $\gamma_{ii}(\cdot)$. La fonction $\gamma_{ij}(\cdot)$, $i \neq j$, est appelée la fonction de covariance croisée des deux séries $\{Y_t(i)\}$ et $\{Y_t(j)\}$.

Voici quelques propriétés sur la matrice de covariance $\Gamma(\cdot)$:

- (i) $\gamma_{ij}(h) = \gamma_{ji}(-h)$;
- (ii) $\Gamma(h) = \Gamma^\top(-h)$;
- (iii) $|\gamma_{ij}(h)| \leq [\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2}$, $i, j = 1, \dots, d$;
- (iv) $\Gamma(\cdot)$ est définie semi-positive, c'est-à-dire $\sum_{k=1}^m \sum_{l=1}^m \boldsymbol{\alpha}_k^\top \Gamma(k-l) \boldsymbol{\alpha}_l \geq 0$ pour tout $m \in \mathbb{N}$ et pour tout vecteur réel $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m \in \mathbb{R}^d$.

La matrice de corrélation $\mathbf{R}(\cdot)$ du processus $\{\mathbf{Y}_t\}$ est définie par

$$\mathbf{R}(h) = \mathbf{D}^{-1/2} \mathbf{\Gamma}(h) \mathbf{D}^{-1/2} = [\rho_{ij}(h)]_{i,j=1}^d = [\gamma_{ij}(h) / [\gamma_{ii}(0) \gamma_{jj}(0)]^{1/2}]_{i,j=1}^d \quad (1.2.1)$$

où $\mathbf{D} = \text{diag}[\gamma_{11}(0), \dots, \gamma_{dd}(0)]$ est une matrice diagonale dont le $i^{\text{ème}}$ élément de la diagonale est la variance de la $i^{\text{ème}}$ composante du processus $\{Y_t(i)\}$ et $\rho_{ii}(\cdot)$ est la fonction d'autocorrélation de la $i^{\text{ème}}$ composante du processus $\{Y_t(i)\}$.

La matrice de corrélation $\mathbf{R}(\cdot)$ possède les mêmes propriétés que la matrice de covariance $\mathbf{\Gamma}(\cdot)$, à l'exception de la propriété additionnelle suivante :

$$(v) \quad \rho_{ii}(0) = 1, \quad \forall i = 1, \dots, d.$$

Nous allons maintenant introduire dans ce qui suit certaines classes de séries chronologiques stationnaires : dans un premier temps, nous allons définir le bruit blanc multivarié, qui est possiblement la classe la plus simple, et par la suite le processus autorégressif moyenne mobile multivarié (VARMA). La classe des modèles VARMA est sans contredit fort importante dans les applications car de nombreuses séries chronologiques réelles peuvent être modélisées par ce type de modèles.

Définition 1.2.2. *Le processus $\{\mathbf{a}_t, t \in \mathbb{Z}\}$ de dimension d est un bruit blanc faible si et seulement si $\{\mathbf{a}_t\}$ est stationnaire avec un vecteur de moyenne $\mathbf{0}$ et une matrice de covariance*

$$\mathbf{\Gamma}(h) = \begin{cases} \mathbf{\Sigma}, & \text{si } h = 0; \\ \mathbf{0}, & \text{sinon,} \end{cases}$$

où $\mathbf{0}$ désigne la matrice nulle de dimension $d \times d$. Si les vecteurs aléatoires du processus $\{\mathbf{a}_t, t \in \mathbb{Z}\}$ sont indépendents et identiquement distribués (iid), alors $\{\mathbf{a}_t\}$ est appelé un bruit blanc fort.

Lorsque nous parlons d'un bruit blanc sans spécifier s'il est fort ou faible, alors il est sous-entendu qu'il s'agit d'un bruit blanc faible.

Définition 1.2.3. *Un processus stationnaire $\{\mathbf{Y}_t\}$ autorégressif moyenne mobile de dimension d et d'ordres p et q , noté VARMA(p, q), est défini par :*

$$\Phi(B) \mathbf{Y}_t = \theta_0 + \Theta(B) \mathbf{a}_t, \quad t \in \mathbb{Z}, \quad (1.2.2)$$

où :

- B est l'opérateur retard et entraîne un retard temporel, c'est-à-dire $B\mathbf{Y}_t = \mathbf{Y}_{t-1}$;
- $\Phi(B) = \mathbf{I}_d - \sum_{j=1}^p \Phi_j B^j$ de dimension $d \times d$ est l'opérateur autorégressif multivarié tel que $\Phi_p \neq \mathbf{0}$;
- \mathbf{I}_d est la matrice identité de dimension $d \times d$;
- $\Theta(B) = \mathbf{I}_d + \sum_{j=1}^q \Theta_j B^j$ de dimension $d \times d$ est l'opérateur moyenne mobile multivarié tel que $\Theta_q \neq \mathbf{0}$;
- θ_0 de dimension $d \times 1$ est un terme constant.

Lorsque le terme d'erreur $\{\mathbf{a}_t\}$ est un bruit blanc fort, nous parlons souvent de VARMA fort. Sinon, dans le cas où $\{\mathbf{a}_t\}$ consiste seulement d'une suite de vecteurs aléatoires non-corrélés, nous parlons plutôt de VARMA faible. Nous remarquons qu'un processus VARMA avec un terme d'erreur $\{\mathbf{a}_t\}$ gaussien est nécessairement un VARMA fort.

Lorsque $p = 0$, le processus $\{\mathbf{Y}_t\}$ est un processus moyenne mobile multivarié d'ordre q , noté VMA(q). Dans ce cas, l'opérateur autorégressif multivarié est la matrice identité. Similairement, lorsque $q = 0$, le processus $\{\mathbf{Y}_t\}$ est un processus autorégressif multivarié d'ordre p , noté VAR(p); dans un tel cas, l'opérateur moyenne mobile est la matrice identité.

Les critères de causalité et d'inversibilité sont des notions importantes reliées aux processus VARMA(p, q). Effectivement, un processus VARMA(p, q) causal peut être exprimé comme une combinaison linéaire infinie d'un bruit blanc. De plus, un processus VARMA(p, q) pouvant s'exprimer comme une combinaison linéaire infinie des valeurs passées est dit inversible. Les théorèmes suivants traitent respectivement des critères de causalité et d'inversibilité.

Théorème 1.2.1. *Le processus VARMA(p, q) représenté à l'équation (1.2.2) est causal si*

$$\det\{\Phi(z)\} \neq 0, \quad \text{pour tout } z \in \mathbb{C} \text{ tel que } |z| \leq 1. \quad (1.2.3)$$

Dans ce cas, le processus VARMA(p, q) admet la représentation suivante :

$$\mathbf{Y}_t = \sum_{j=0}^{\infty} \Psi_j \mathbf{a}_{t-j} = \Psi(B)\mathbf{a}_t, \quad (1.2.4)$$

et les matrices Ψ_j sont uniquement déterminées par

$$\Psi(B) = \sum_{j=0}^{\infty} \Psi_j B^j = \Phi^{-1}(B)\Theta(B). \quad (1.2.5)$$

Une preuve de ce théorème se trouve dans Brockwell et Davis (1991, section 11.3). Il est à noter que si un processus VARMA(p, q) admet une représentation causale comme en (1.2.4), il est forcément stationnaire.

Théorème 1.2.2. *Le processus VARMA(p, q) représenté à l'équation (1.2.2) est inversible si*

$$\det\{\Theta(z)\} \neq 0, \quad \text{pour tout } z \in \mathbb{C} \text{ tel que } |z| \leq 1. \quad (1.2.6)$$

Dans ce cas, le processus VARMA(p, q) admet la représentation suivante :

$$\mathbf{a}_t = \sum_{j=0}^{\infty} \Pi_j \mathbf{Y}_{t-j} = \Pi(B)\mathbf{Y}_t, \quad (1.2.7)$$

et les matrices Π_j sont uniquement déterminées par

$$\Pi(B) = \sum_{j=0}^{\infty} \Pi_j B^j = \Theta^{-1}(B)\Phi(B). \quad (1.2.8)$$

L'intérêt d'avoir des représentations inversibles peut être expliqué intuitivement lorsque nous désirons calculer des prévisions. En effet, dans une telle situation, un prédicteur de \mathbf{Y}_{n+1} , disons, peut être obtenu en exprimant \mathbf{Y}_{n+1} en fonction de $\mathbf{Y}_n, \mathbf{Y}_{n-1}, \dots, \mathbf{Y}_1$. Évidemment, dans les applications générales, il faut souvent adopter une stratégie de troncation de (1.2.7), puisque nous disposons rarement de tout l'historique du processus. Plus techniquement, le théorème 1.2.2 portant sur l'inversibilité d'un processus multivarié est quelque peu analogue au théorème 1.2.1. En fait, les arguments principaux de la preuve sont plutôt similaires. Pour plus de détails, voir Brockwell et Davis (1991).

1.2.2. Estimation de la fonction d'autocovariance

En pratique, nous disposons que d'une réalisation de la série chronologique $\{\mathbf{Y}_t, t = 1, \dots, n\}$, et la plupart du temps sa fonction d'autocovariance est inconnue. Cependant, lorsque la série chronologique est présumée issue d'un processus

stationnaire, il est possible de développer des versions échantillonnelles de la fonction d'autocovariance théorique. Posons

$$\hat{\Gamma}_{\mathbf{Y}}(h) = \begin{cases} n^{-1} \sum_{t=h+1}^n (\mathbf{Y}_t - \bar{\mathbf{Y}})(\mathbf{Y}_{t-h} - \bar{\mathbf{Y}})^{\top}, & h = 0, 1, \dots, n-1; \\ \hat{\Gamma}_{\mathbf{Y}}^{\top}(-h), & h = -1, -2, \dots, -n+1, \end{cases}$$

où $\bar{\mathbf{Y}} = n^{-1} \sum_{t=1}^n \mathbf{Y}_t$. La fonction $\{\hat{\Gamma}_{\mathbf{Y}}(h), |h| < n\}$ est appelée la fonction d'autocovariance échantillonnale. Elle a été étudiée par plusieurs auteurs. Par exemple, Chitturi (1974, 1976), Hosking (1980), Li et McLeod (1981) sont probablement les premiers auteurs à avoir étudié la distribution asymptotique des autocovariances échantillonnelles pour des processus VAR(p) et également dans le cas plus général des processus VARMA(p, q). Voir aussi Fuller (1996) et Brockwell et Davis (1991).

1.2.3. Estimation d'un processus autorégressif multivarié

Nous allons maintenant discuter d'une technique particulière afin d'estimer une série chronologique $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ de dimension d et de longueur n générée par un processus VAR(p) stationnaire. La raison principale pour laquelle nous présentons en détails cette technique d'estimation est que cette dernière a été utilisée dans nos études de simulation. Plus précisément, le processus est défini par :

$$\mathbf{Y}_t = \mathbf{c} + \Phi_1 \mathbf{Y}_{t-1} + \dots + \Phi_p \mathbf{Y}_{t-p} + \mathbf{a}_t, \quad (1.2.9)$$

où $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_d)^{\top}$ est un vecteur de termes constants de dimension $d \times 1$, les matrices $\Phi_i, i = 1, \dots, p$ sont les coefficients autorégressifs de dimension $d \times d$ et $\{\mathbf{a}_t\}$ est présumé un bruit blanc fort avec une matrice de covariance Σ_a non singulière de dimension $d \times d$.

L'estimation par la méthode des moindres carrés généralisés estimés (MCGE) multivariée suppose d'abord que p observations pré-échantillonnées sont disponibles pour chaque composante de \mathbf{Y}_t . Introduisons les notations suivantes avec

leurs dimensions respectives entre parenthèses :

$$\mathbf{Y} = [\mathbf{Y}_1, \dots, \mathbf{Y}_n], \quad (d \times n); \quad (1.2.10)$$

$$\mathbf{y} = \text{vec}(\mathbf{Y}), \quad (dn \times 1); \quad (1.2.11)$$

$$\mathbf{B} = [\mathbf{c}, \Phi_1, \dots, \Phi_p], \quad (d \times (dp + 1)); \quad (1.2.12)$$

$$\boldsymbol{\beta} = \text{vec}(\mathbf{B}), \quad ((d^2p + d) \times 1); \quad (1.2.13)$$

$$\mathbf{Z}_t = [1, \mathbf{Y}_t^\top, \dots, \mathbf{Y}_{t-p+1}^\top]^\top, \quad ((dp + 1) \times 1); \quad (1.2.14)$$

$$\mathbf{Z} = [\mathbf{Z}_0, \dots, \mathbf{Z}_{n-1}], \quad ((dp + 1) \times n); \quad (1.2.15)$$

$$\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n], \quad (d \times n); \quad (1.2.16)$$

$$\mathbf{a} = \text{vec}(\mathbf{A}), \quad (dn \times 1), \quad (1.2.17)$$

où l'opérateur $\text{vec}(\cdot)$ désigne l'empilement des colonnes \mathbf{q}_i de dimension $m \times 1$ d'une matrice $\mathbf{Q} = (\mathbf{q}_1, \dots, \mathbf{q}_n)$ de dimension $m \times n$, c'est-à-dire $\text{vec}(\mathbf{Q}) = (\mathbf{q}_1^\top, \dots, \mathbf{q}_n^\top)^\top$.

La technique présentée dans ce qui suit est générale dans la mesure où nous pouvons présumer des contraintes linéaires dans les paramètres. Ainsi, nous supposons que les vecteurs $\boldsymbol{\beta}$ et $\boldsymbol{\gamma}$ sont reliés par une matrice \mathbf{R} , qui est une matrice connue de contraintes linéaires sur les paramètres du processus. Un type d'exemple important dans les applications concerne les contraintes linéaires correspondant aux paramètres présumés nuls du modèle. Ainsi, il arrive fréquemment dans les applications que les praticiens adoptent une stratégie en deux temps. Par exemple, dans une première étape, un modèle sans contrainte est ajusté. Après un examen des coefficients significatifs, ceux qui sont déclarés non significatifs sont fixés à 0 lors d'une seconde étape d'estimation. Ceci peut être effectué avec un choix judicieux de la matrice \mathbf{R} . En général, le choix de la matrice \mathbf{R} permet de tenir compte de diverses contraintes linéaires sur les paramètres composant le vecteur $\boldsymbol{\beta}$.

Ainsi, supposons les contraintes linéaires sur les paramètres en posant la relation $\boldsymbol{\beta} = \mathbf{R}\boldsymbol{\gamma}$. Le processus VAR(p) défini à l'équation (1.2.9) peut maintenant s'écrire sous la forme compacte suivante : $\mathbf{Y} = \mathbf{BZ} + \mathbf{A}$. En vectorisant cette représentation et en utilisant la relation $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^\top \otimes \mathbf{A})\text{vec}(\mathbf{B})$, où $\mathbf{A}, \mathbf{B}, \mathbf{C}$

sont des matrices de dimensions appropriées et ' \otimes ' désigne le produit de Kronecker, nous obtenons la forme suivante du processus : $\mathbf{y} = (\mathbf{Z}^\top \otimes \mathbf{I}_d)\boldsymbol{\beta} + \mathbf{a}$, où \mathbf{y} , \mathbf{Z} , $\boldsymbol{\beta}$, et \mathbf{a} sont définis par (1.2.11), (1.2.15), (1.2.13) et (1.2.17), respectivement.

La première étape d'estimation consiste à trouver l'estimateur de Σ_a et $\boldsymbol{\gamma}$. Plus précisément, l'estimateur de la matrice de covariance des erreurs Σ_a est obtenu à partir de l'estimateur par moindres carrés (MC) de $\boldsymbol{\gamma}$, soit :

$$\boldsymbol{\gamma}^* = [\mathbf{R}^\top (\mathbf{Z}\mathbf{Z}^\top \otimes \mathbf{I}_d)\mathbf{R}]^{-1}\mathbf{R}^\top (\mathbf{Z}^\top \otimes \mathbf{I}_k)\mathbf{y}.$$

En se basant sur l'estimateur $\boldsymbol{\gamma}^*$, nous pouvons déterminer les résidus $\mathbf{a}^* = \mathbf{y} - (\mathbf{Z}^\top \otimes \mathbf{I}_d)\mathbf{R}\boldsymbol{\gamma}^*$, et donc l'estimateur de la matrice de covariance résiduelle $\Sigma_a^* = n^{-1}\mathbf{A}^*\mathbf{A}^{*\top}$, où $\text{vec}(\mathbf{A}^*) = \mathbf{a}^*$.

La deuxième étape de cette technique d'estimation consiste à obtenir l'estimateur MCGE de $\boldsymbol{\gamma}$ à partir des résidus obtenus par MC. L'estimateur résultant de $\boldsymbol{\gamma}$ est alors noté $\hat{\boldsymbol{\gamma}}$ et il est obtenu comme suit :

$$\hat{\boldsymbol{\gamma}} = [\mathbf{R}^\top (\mathbf{Z}\mathbf{Z}^\top \otimes \Sigma_a^{*-1})\mathbf{R}]^{-1}\mathbf{R}^\top (\mathbf{Z}^\top \otimes \Sigma_a^{*-1})\mathbf{y}. \quad (1.2.18)$$

Il peut être démontré sous certaines conditions que l'estimateur MCGE $\hat{\boldsymbol{\gamma}}$ de $\boldsymbol{\gamma}$ est un estimateur convergent de $\boldsymbol{\gamma}$, c'est-à-dire que $\hat{\boldsymbol{\gamma}} \xrightarrow{p} \boldsymbol{\gamma}$. Parmi les conditions permettant d'obtenir le résultat annoncé, il faut que les quatrièmes moments du bruit blanc fort $\{\mathbf{a}_t\}$ existent, que la matrice \mathbf{R} soit de rang plein et que $\hat{\Sigma}_a \xrightarrow{p} \Sigma_a$. Pour plus de détails entourant cette technique d'estimation et les résultats asymptotiques concernant l'estimation de $\boldsymbol{\gamma}$, Lütkepohl (1993, chapitre 5) peut être consulté.

1.3. MODÈLES NON-STATIONNAIRES

Un grand nombre de séries chronologiques observées affichent des caractéristiques de non-stationnarité. Un exemple simple survient lorsqu'un terme de tendance est présent. Par exemple, une tendance linéaire ou plus généralement polynomiale pourrait décrire l'évolution de la moyenne en fonction du temps. Dans un tel cas, la moyenne est fonction de l'indice temporel, causant ainsi une forme possible de non-stationnarité. Une autre forme de non-stationnarité survient lorsque

des racines unitaires sont présentes dans l'opérateur autorégressif d'un modèle VARMA. Nous parlons dans un tel cas de processus cointégrés ou encore partiellement non-stationnaires. C'est cette dernière forme de non-stationnarité qui va retenir principalement notre attention dans ce qui suit. Une caractéristique intéressante des processus cointégrés est qu'ils sont non-stationnaires, mais certaines combinaisons linéaires sont stationnaires. Il est donc important de développer et de fournir des notions afin de pouvoir bien analyser ce type de processus. Quelques notions de base ainsi que l'estimation de processus non-stationnaires sont introduites dans les prochaines sections.

1.3.1. Notions de base

Comme mentionné précédemment, un processus multivarié $\{\mathbf{Y}_t\}$ affiche de la non-stationnarité si son vecteur de moyennes défini à l'équation (1.1.1) et/ou ses matrices de covariance définies à l'équation (1.1.2) sont fonction du temps. Comme dans le cas de processus stationnaires, une classe importante de séries chronologiques non-stationnaires est le processus autorégressif moyenne mobile.

Définition 1.3.1. *Un processus non-stationnaire $\{\mathbf{Y}_t\}$ est VARMA(p, q) s'il est défini par :*

$$\Phi(B)\mathbf{Y}_t = \boldsymbol{\theta}_0 + \Theta(B)\mathbf{a}_t \quad t \in \mathbb{Z}, \quad (1.3.1)$$

où B est l'opérateur retard, $\Phi(B)$ et $\Theta(B)$ sont les opérateurs autorégressif et moyenne mobile multivariés tels que $\Phi_p \neq \mathbf{0}$ et $\Theta_q \neq \mathbf{0}$ et $\{\mathbf{a}_t\}$ est un bruit blanc fort.

Il est à noter que dans la définition 1.2.3 les équations de récurrence fournies à l'équation (1.2.2) étaient définies pour un processus $\{\mathbf{Y}_t\}$ présumé stationnaire. Ici, la définition 1.3.1 est généralisée dans la situation où $\{\mathbf{Y}_t\}$ n'est pas forcément stationnaire. Par exemple, $\det\{\Phi(z)\}$ pourrait avoir des racines sur le disque unité.

Les processus VAR multivariés non-stationnaires peuvent être définis de la même façon que dans le cas des processus stationnaires. De plus, la notion d'inversibilité d'un processus s'applique également aux processus VARMA(p, q) non-stationnaires. Un moyen pour vérifier si un processus n'est pas stationnaire est

d'étudier la nature des racines du déterminant de l'opérateur autorégressif, c'est-à-dire étudier les racines de $\det\{\Phi(z)\}$.

En pratique, une procédure souvent adaptée pour transformer un processus $\{\mathbf{Y}_t\}$ non-stationnaire en un processus stationnaire consiste à différencier chaque composante $\{Y_t(i)\}$, $i = 1, \dots, d$, de $\{\mathbf{Y}_t\}$ un nombre k_i de fois afin de les rendre stationnaires. Dans ce cas, chaque composante du processus définit un processus autorégressif moyenne mobile intégré (souvent noté ARIMA), et le processus multivarié est présumé admettre la représentation suivante :

$$\Phi_1(B)\mathbf{D}(B)\mathbf{Y}_t = \Theta(B)\mathbf{a}_t, \quad (1.3.2)$$

où $\mathbf{D}(B) = \text{diag}[(1-B)^{k_1}, \dots, (1-B)^{k_d}]$ est un opérateur de différence, k_1, \dots, k_d sont des entiers non-négatifs et les racines de $\det\{\Phi_1(z)\} = 0$ sont plus grandes que un en module. Le processus résultant $\{\mathbf{W}_t = \mathbf{D}(B)\mathbf{Y}_t\}$ est alors un processus VARMA(p, q) stationnaire.

Cette procédure de différenciation d'un processus peut engendrer certaines complications lors de l'ajustement de modèles. En effet, la sur-différenciation d'au moins une des composantes de $\{\mathbf{Y}_t\}$ peut engendrer une représentation non-inversible du processus (Box et Tiao, 1977) et donc causer un biais dans l'estimation des paramètres du modèle (Hillmer et Tiao, 1979). Une alternative à la différenciation de chaque composante de \mathbf{Y}_t est d'analyser les modèles en incorporant leur structure associée aux racines. Cette procédure sera décrite un peu plus loin dans cette section. De plus, il est à noter qu'en pratique il est souvent préférable de modéliser directement le processus observé $\{\mathbf{Y}_t\}$, plutôt que de formuler des modèles pour des processus transformés. Typiquement, les conclusions tirées pour $\{\mathbf{W}_t\}$ doivent être adaptées pour le processus original $\{\mathbf{Y}_t\}$.

Nous introduisons maintenant une classe de modèles permettant de décrire des séries chronologiques affichant une certaine forme de non-stationnarité. Ces modèles permettent de décrire des processus dits cointégrés ou partiellement non-stationnaires. Puisque des racines unitaires sont présentes sur le disque unité de l'opérateur autorégressif, il est possible d'exploiter cette caractéristique en réécrivant ce processus sous la forme dite à correction d'erreur. Plusieurs hypothèses

et définitions reliées aux processus partiellement non-stationnaires se retrouvent dans Cédras (1997), Pham, Roy et Cédras (2000) et Pham, Roy et Cédras (2003).

Définition 1.3.2. *Un processus $\{\mathbf{Y}_t\}$ VARMA(p, q) de dimension d est partiellement non-stationnaire si $\det\{\Phi(z)\} = 0$ pour $d_0 < d$ racines unitaires, et que toutes les autres racines sont plus grandes que un en module.*

Hypothèse 1.3.1. *Soit le processus VARMA(p, q) partiellement non-stationnaire de dimension d . Le déterminant de $\Phi(z)$ admet une racine unitaire de multiplicité $d_0 \leq d$, le rang de $\Phi(1)$ est $r = d - d_0$ et $\Theta(1)$ est non singulière.*

Dans l'analyse d'un processus partiellement non-stationnaire, afin d'incorporer la structure associée aux racines unitaires, il est possible d'exprimer le processus sous sa forme dite à correction d'erreur. Cette forme est décrite dans la définition suivante.

Définition 1.3.3. *Soit un processus autorégressif multivarié VAR(p) partiellement non-stationnaire satisfaisant (1.3.1) avec le terme constant $\theta_0 = \mathbf{0}$. Le processus peut être représenté sous la forme à correction d'erreur comme suit :*

$$\Phi^*(B)\mathbf{W}_t = \mathbf{C}\mathbf{Y}_{t-1} + \mathbf{a}_t, \quad (1.3.3)$$

où $\mathbf{W}_t = (1 - B)\mathbf{Y}_t$, $\mathbf{C} = -\Phi(1)$, $\Phi^*(B) = \mathbf{I}_d - \sum_{j=1}^{p-1} \Phi_j^* B^j$, $\Phi_j^* = \sum_{k=j+1}^p \Phi_k$ et $\Phi(1) = \mathbf{I}_d - \sum_{j=1}^p \Phi_j$.

La forme à correction d'erreur d'un processus multivarié $\{\mathbf{Y}_t\}$ est pratique étant donné que la non-stationnarité de l'opérateur autorégressif multivarié est concentrée dans le comportement de la matrice \mathbf{C} .

Généralement, les processus non-stationnaires peuvent être transformés afin que la transformation s'exprime comme la somme d'une composante non-stationnaire et d'une composante stationnaire. Avant de présenter ce type de transformation, nous allons introduire la notion de racines unitaires dans le contexte de processus univariés menant aux processus intégrés, ainsi que le concept de cointégration propre aux processus multivariés.

Définition 1.3.4. *Un processus univarié $\{X_t\}$ ayant k racines unitaires dans l'opérateur autorégressif univarié est appelé un processus intégré d'ordre k , noté $\{X_t\} \sim I(k)$.*

Si un processus univarié $\{X_t\}$ est $I(k)$, alors le processus peut être rendu stationnaire en le différenciant k fois, c'est-à-dire $\{(1 - B)^k X_t\}$ est stationnaire.

Définition 1.3.5. Les composantes du processus $\{\mathbf{Y}_t\}$ de dimension d sont cointégrés d'ordre (k, l) , noté $\mathbf{Y}_t \sim CI(k, l)$, si $\mathbf{Y}_t(i) \sim I(k), \forall i \in [1, d]$ et qu'il existe une combinaison linéaire $\mathbf{c}^\top \mathbf{Y}_t$ telle que $\mathbf{c}^\top \mathbf{Y}_t \sim I(k - l)$, $l > 0$, et $\mathbf{c} = (c_1, \dots, c_d)^\top \neq \mathbf{0}$ est nommé un vecteur de cointégration.

Il est à noter qu'un vecteur de cointégration n'est pas unique. La proposition suivante présente le type de transformation décrit précédemment.

Proposition 1.3.1. Soit $\{\mathbf{Y}_t\}$ un processus ARMA(p, q) partiellement non-stationnaire satisfaisant aux conditions énoncées dans l'hypothèse 1.3.1. Alors $\{\mathbf{Y}_t\}$ peut s'exprimer par la combinaison linéaire suivante :

$$\mathbf{Y}_t = \mathbf{P}_1 \mathbf{Z}_{1t} + \mathbf{P}_2 \mathbf{Z}_{2t}, \quad (1.3.4)$$

où \mathbf{P}_1 est une matrice de rang plein de dimension $d \times d_0$ telle que $\Phi(1)\mathbf{P}_1 = \mathbf{0}$ et \mathbf{P}_2 est une matrice de dimension $d \times r$ telle que $\mathbf{P} = (\mathbf{P}_1, \mathbf{P}_2)$ est non singulière. De plus, le processus $\{\mathbf{Z}_{1t}\}$ de dimension d_0 est à accroissements stationnaires, c'est-à-dire que $\{(1 - B)\mathbf{Z}_{1t}\}$ est stationnaire, et le processus $\{\mathbf{Z}_{2t}\}$ de dimension r est stationnaire.

Une preuve de la proposition précédente se retrouve dans Pham, Roy et Cédras (2000). Ils définissent la matrice $\mathbf{Q} = \mathbf{P}^{-1} = [\mathbf{Q}_1^\top, \mathbf{Q}_2^\top]^\top$, où \mathbf{Q}_1 est une matrice de dimension $d_0 \times d$ et \mathbf{Q}_2 est une matrice de dimension $r \times d$. Ils notent que $\mathbf{P}_1 \mathbf{Q}_1 + \mathbf{P}_2 \mathbf{Q}_2 = \mathbf{I}_d$, et que $\{\mathbf{Z}_{1t} = \mathbf{Q}_1 \mathbf{Y}_t\}$ est non-stationnaire tel que $\{(1 - B)\mathbf{Z}_{1t}\}$ est stationnaire et que $\{\mathbf{Z}_{2t} = \mathbf{Q}_2 \mathbf{Y}_t\}$ définit un processus stationnaire.

D'après la proposition précédente, il s'en suit que $\{\mathbf{W}_t = (1 - B)\mathbf{Y}_t\}$ est stationnaire. De plus, puisque $\{\mathbf{Z}_{2t} = \mathbf{Q}_2 \mathbf{Y}_t\}$ est stationnaire, alors le processus $\{\mathbf{Y}_t\}$ est cointégré d'ordre $(1, 1)$. De plus, les vecteurs de cointégration sont donnés par les r lignes linéairement indépendantes de la matrice \mathbf{Q}_2 . Puisqu'il existe r combinaisons linéaires indépendantes des composantes de $\{\mathbf{Y}_t\}$ étant stationnaires, alors le processus $\{\mathbf{Y}_t\}$ est dit être cointégré de rang r .

1.3.2. Estimation d'un processus autorégressif multivarié partiellement non-stationnaire

Soit le processus $\{\mathbf{Y}_t\}$ autorégressif de degré p partiellement non-stationnaire ayant la représentation décrite à l'équation (1.3.3). Étant donné que la matrice \mathbf{C} n'est pas de rang plein ($\text{rang}(\mathbf{C}) = r$), il est possible d'exprimer \mathbf{C} comme $\mathbf{C} = \mathbf{A}\mathbf{B}$, où \mathbf{A} et \mathbf{B} sont des matrices de rang plein de dimensions $d \times r$ et $r \times d$ respectivement. La matrice \mathbf{B} est normalisée afin que $\mathbf{B} = [\mathbf{I}_r, \mathbf{B}_0]$, où \mathbf{B}_0 est une matrice $r \times (d - r)$ de paramètres inconnus (voir Ahn et Reinsel, 1990). L'estimation de la forme à correction d'erreur d'un processus multivarié $\{\mathbf{Y}_t\}$ est une alternative à la différenciation composante par composante. Elle permet une meilleure compréhension de la nature de la non-stationnarité parmi les composantes d'un processus (Reinsel 1993, section 6.3.3).

Nous allons discuter de deux méthodes d'estimation d'un processus partiellement non-stationnaire. La première technique est la méthode du maximum de vraisemblance, qui peut être mise en pratique en développant un algorithme de type Newton-Raphson. Une seconde technique, fort séduisante dans les applications compte tenu de sa facilité d'implantation, repose sur une estimation en seulement deux étapes. C'est cette dernière méthode que nous avons implanté dans nos études empiriques. Elle est plus simple numériquement que la technique du maximum de vraisemblance et il a été démontré théoriquement qu'elle fournit des estimateurs asymptotiquement efficaces. Voir Ahn et Reinsel (1990).

1.3.2.1. Estimation par la méthode du maximum de vraisemblance

Soit un processus $\{\mathbf{Y}_t\}$ partiellement non-stationnaire tel que défini à l'équation (1.3.3) et la matrice \mathbf{C} comme étant le produit des matrices \mathbf{A} et \mathbf{B} décrites précédemment. Posons le vecteur de paramètres inconnus $\boldsymbol{\beta} = (\boldsymbol{\beta}_0^\top, \boldsymbol{\alpha}^\top)^\top$ de dimension $b \times 1$, où $\boldsymbol{\beta}_0 = \text{vec}(\mathbf{B}_0^\top)$, $\boldsymbol{\alpha} = \text{vec}\{(\mathbf{A}, \boldsymbol{\Phi}_1^*, \dots, \boldsymbol{\Phi}_{p-1}^*)^\top\}$, $b = r(d - r) + rd + d^2(p - 1)$ représente le nombre de paramètres inconnus du modèle et r est le rang de la matrice \mathbf{C} . L'estimateur du maximum de vraisemblance de $\boldsymbol{\beta}$ basé sur les n premières observations du processus $\{\mathbf{Y}_t\}$ minimise la fonction

de vraisemblance suivante :

$$l(\boldsymbol{\beta}, \boldsymbol{\Sigma}_a) = -\frac{n}{2} \log |\boldsymbol{\Sigma}_a| - \frac{1}{2} \sum_{t=1}^n \mathbf{a}_t^\top \boldsymbol{\Sigma}_a^{-1} \mathbf{a}_t. \quad (1.3.5)$$

L'estimateur du maximum de vraisemblance de $\boldsymbol{\beta}$ est obtenu par une méthode itérative semblable à celle de l'algorithme de Newton-Raphson. L'estimateur de $\boldsymbol{\beta}$ à la $(i+1)$ ^{ème} itération est donné par la relation suivante :

$$\hat{\boldsymbol{\beta}}^{(i+1)} = \hat{\boldsymbol{\beta}}^{(i)} + \left\{ \sum_{t=1}^n \mathbf{U}_t^* \boldsymbol{\Sigma}_a^{-1} \mathbf{U}_t^{*\top} \right\}^{-1} \left\{ \sum_{t=1}^n \mathbf{U}_t^* \boldsymbol{\Sigma}_a^{-1} \mathbf{a}_t \right\} \hat{\boldsymbol{\beta}}^{(i)}, \quad (1.3.6)$$

où $\hat{\boldsymbol{\beta}}^{(i)}$ désigne l'estimateur à la i ^{ème} itération, $\mathbf{U}_t^* = [(\mathbf{A}^\top \otimes \mathbf{J}^\top \mathbf{Y}_{t-1})^\top, \mathbf{I}_d \otimes \tilde{\mathbf{U}}_{t-1}^\top]^\top$ est une matrice de dimension $b \times d$, $\tilde{\mathbf{U}}_{t-1} = [(\mathbf{B}\mathbf{Y}_{t-1})^\top, \mathbf{W}_{t-1}^\top, \dots, \mathbf{W}_{t-p+1}^\top]^\top$ et $\mathbf{J}^\top = [\mathbf{0}, \mathbf{I}_{d-r}]$ est une matrice de dimension $(d-r) \times d$.

Le point de départ des itérations définies à l'équation (1.3.6) doit être un estimateur convergent de $\boldsymbol{\beta}$. Ahn et Reinsel (1990) proposent un estimateur initial de $\boldsymbol{\beta}$ basé sur les estimateurs à rang plein obtenus par la méthode des moindres carrés. Pour obtenir les estimateurs par moindres carrés, il faut d'abord supposer que $p+1$ observations pré-échantillonnées soient disponibles pour chaque composante de \mathbf{Y}_t . Posons $\boldsymbol{\Upsilon} = (\mathbf{C}, \boldsymbol{\Phi}_1^*, \dots, \boldsymbol{\Phi}_{p-1}^*)$ une matrice de paramètres de dimension $d \times dp$ et $\mathbf{H}_t = (\mathbf{Y}_{t-1}^\top, \mathbf{W}_{t-1}^\top, \dots, \mathbf{W}_{t-p+1}^\top)^\top$ un vecteur de dimension $dp \times 1$. Le modèle à correction d'erreur défini à l'équation (1.3.3) peut être réécrit de la forme $\mathbf{W}_t = \boldsymbol{\Upsilon} \mathbf{H}_t + \mathbf{a}_t$. L'estimation par moindres carrés $\hat{\boldsymbol{\Upsilon}}$ de $\boldsymbol{\Upsilon}$ est donné par :

$$\hat{\boldsymbol{\Upsilon}} = \left(\sum \mathbf{W}_t \mathbf{H}_t^\top \right) \left(\sum \mathbf{H}_t \mathbf{H}_t^\top \right)^{-1}. \quad (1.3.7)$$

Les résidus sont donnés par $\hat{\mathbf{a}}_t = \mathbf{W}_t - \hat{\boldsymbol{\Upsilon}} \mathbf{H}_t$ et la matrice de covariance associée aux résidus est $\hat{\boldsymbol{\Sigma}}_a$.

La prochaine étape consiste à décomposer l'estimateur des moindres carrés de \mathbf{C} en $\hat{\mathbf{C}} = (\hat{\mathbf{C}}_1, \hat{\mathbf{C}}_2)$ et à poser $\hat{\mathbf{A}} = \hat{\mathbf{C}}_1$, où $\hat{\mathbf{A}}$, qui est un estimateur convergent de \mathbf{A} , est une matrice $d \times r$. Ahn et Reinsel (1990) ont démontré qu'un estimateur convergent de \mathbf{B}_0 est donné par $\tilde{\mathbf{B}}_0 = (\hat{\mathbf{A}}^\top \hat{\boldsymbol{\Sigma}}_a^{-1} \hat{\mathbf{A}})^{-1} (\hat{\mathbf{A}}^\top \hat{\boldsymbol{\Sigma}}_a^{-1} \hat{\mathbf{C}}_2)$, une matrice de dimension $d \times (d-r)$. L'estimateur initial de $\boldsymbol{\beta}$ résultant est donné par $\hat{\boldsymbol{\beta}}^{(0)} = (\tilde{\boldsymbol{\beta}}_0^\top, \hat{\boldsymbol{\alpha}}^\top)^\top$, où $\tilde{\boldsymbol{\beta}}_0 = \text{vec}(\tilde{\mathbf{B}}_0^\top)$ et $\hat{\boldsymbol{\alpha}} = \text{vec}\{(\hat{\mathbf{A}}, \hat{\boldsymbol{\Phi}}_1^*, \dots, \hat{\boldsymbol{\Phi}}_{p-1}^*)^\top\}$. Les estimateurs

par moindres carrés $\hat{\Phi}_i^*$, $i = 1, \dots, p-1$, tels que définis à l'équation (1.3.7) sont des estimateurs convergents de Φ_i^* , $i = 1, \dots, p-1$.

Cette méthode d'estimation peut poser des inconvénients lors des études de simulation Monte Carlo. En effet, la convergence de l'algorithme va survenir lorsque les valeurs initiales sont adéquatement choisies. Même en utilisant la procédure citée plus haut, il peut arriver que l'algorithme soit sujet à des problèmes de convergence. Ceci peut être perçu comme un désavantage à bien des égards, particulièrement lors de longues études de simulation où la convergence doit être assurée à chaque itération, et ce dans un temps raisonnable. La prochaine section discute d'une technique alternative d'estimation.

1.3.2.2. Estimation à deux étapes

Nous discutons maintenant de la deuxième méthode d'estimation d'un processus partiellement non-stationnaire, soit la méthode d'estimation à deux étapes. La première étape de cette procédure consiste à estimer les paramètres du modèle à correction d'erreur défini à l'équation (1.3.3) par une technique telle que la méthode des moindres carrés décrite à la section 1.3.2.1.

La deuxième étape de cette méthode d'estimation est l'obtention des estimateurs à rang réduit. D'abord, un estimateur convergent de \mathbf{B} est donné par $\tilde{\mathbf{B}} = (\mathbf{I}_r, \tilde{\mathbf{B}}_0)$, où $\tilde{\mathbf{B}}_0$ est obtenu de la même façon que décrite à la section 1.3.2.1. Les estimateurs $\tilde{\mathbf{A}}$, $\tilde{\Phi}_i^*$, $i = 1, \dots, p-1$ de \mathbf{A} , Φ_i^* , $i = 1, \dots, p-1$ sont alors obtenus en effectuant la régression des moindres carrés de \mathbf{W}_t sur $\tilde{\mathbf{B}}\mathbf{Y}_{t-1}, \mathbf{W}_{t-1}, \dots, \mathbf{W}_{t-p+1}$, où $\tilde{\mathbf{B}}$ est présumé fixé. Il a été démontré par Ahn et Reinsel (1990) que les estimateurs résultants sont asymptotiquement efficaces.

Cette méthode d'estimation est fort simple à programmer et elle garantit des estimateurs convergents des paramètres du modèle à correction d'erreur définis à l'équation (1.3.3). Il est fort vraisemblable que cette technique d'estimation soit la plus populaire dans les applications puisqu'elle nécessite principalement des routines capables de calculer des estimateurs par moindres carrés.

1.4. DENSITÉ SPECTRALE

Dans les sections précédentes, les séries chronologiques ont été analysées dans le domaine du temps. Une approche alternative est l'analyse des séries chronologiques dans le domaine des fréquences. Cette approche s'avère utile puisque des résultats généraux sur les processus stochastiques montrent qu'un processus stationnaire peut se décomposer approximativement comme une certaine somme aléatoire de sinus et cosinus, présentant des périodicités qui peuvent être expliquées par l'étude des fréquences.

1.4.1. Notions de base

Il peut être démontré que la fonction d'autocovariance d'un processus stationnaire $\{\mathbf{Y}_t\}$ peut s'exprimer à l'aide d'une fonction particulière, la fonction de distribution spectrale. Cette fonction présente essentiellement les caractéristiques d'une fonction de répartition d'une variable aléatoire. Lorsque cette fonction est absolument continue, alors selon des résultats d'analyse mathématique, il peut être démontré qu'une dérivée de cette fonction existe, qui est appelée la densité spectrale. Une condition suffisante pour la détermination de cette fonction est que la fonction d'autocovariance soit sommable. La fonction de densité spectrale est donnée à la définition suivante.

Définition 1.4.1. *La matrice de densité spectrale du processus multivarié $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ ayant la matrice de covariance $\Gamma(h)$ est définie par :*

$$\mathbf{f}(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \Gamma(h) e^{-ih\omega}, \quad \omega \in [-\pi, \pi], \quad (1.4.1)$$

où $e^{i\omega} = \cos(\omega) + i \sin(\omega)$ et $i = \sqrt{-1}$.

La matrice d'autocovariance de délai h est la transformée de Fourier de $\mathbf{f}(\omega)$ et admet la représentation suivante :

$$\Gamma(h) = \int_{-\pi}^{\pi} e^{ih\omega} \mathbf{f}(\omega) d\omega, \quad h \in \mathbb{Z}.$$

Le (k, l) ^{ième} élément de la matrice de covariance $\Gamma(h)$ est donné par $\gamma_{kl}(h) = \int_{-\pi}^{\pi} e^{ih\omega} f_{kl}(\omega) d\omega$.

Dans le cas d'une série chronologique univariée, la densité spectrale est non-négative et réelle. De manière générale, dans le cas multivarié, la fonction de densité spectrale définie par (1.4.1) est à valeurs complexes. Ainsi, il peut être montré que l'élément (k, k) de $\mathbf{f}(\omega)$, noté $f_{kk}(\omega)$, est non-négative et réelle, comme dans le cas univarié, pour $k = 1, \dots, d$. En fait, $f_{kk}(\omega)$ est la densité spectrale de $\{Y_t(k)\}$, et elle est parfois nommée la fonction de densité autospectrale de la $k^{\text{ième}}$ composante de $\{\mathbf{Y}_t\}$. L'élément (k, l) de la fonction $\mathbf{f}(\omega)$, $k \neq l$, est la fonction de densité spectrale croisée de $\{Y_t(k)\}$ et $\{Y_t(l)\}$; elle est représentée par

$$f_{kl}(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_{kl}(h) e^{-ih\omega}, \quad \omega \in [-\pi, \pi]. \quad (1.4.2)$$

Cette fonction est généralement à valeurs complexes et nous notons que $\gamma_{kl}(h) \neq \gamma_{kl}(-h)$ pour $k \neq l$ (en fait, rappelons $\gamma_{kl}(h) = \gamma_{lk}(-h)$, en vertu des propriétés de base de la fonction d'autocovariance multivariée). Les fonctions de densité autospectrales et spectrales croisées sont continues et périodiques en ω . De plus, il peut être montré que $\mathbf{f}(\omega)$ est hermitienne, ce qui implique que $\mathbf{f}^*(\omega) = \mathbf{f}(-\omega)^\top$, où \mathbf{A}^* représente la transposée du conjugué complexe (c'est-à-dire que $\mathbf{A}^* = \bar{\mathbf{A}}^\top$). De plus, cette fonction est définie semi-positive. Voir Fuller (1996) qui démontre ces propriétés.

La densité spectrale est un outil permettant de décrire la dépendance en terme des fréquences. Nous remarquons cependant qu'elle s'exprime en fonction des autocovariances théoriques pour tous les délais. Ainsi, en pratique, la densité spectrale doit être estimée, puisque nous disposons d'une série chronologique $\{\mathbf{Y}_t, t = 1, \dots, n\}$. Une technique générale repose sur l'estimation non paramétrique à l'aide de la méthode du noyau, que nous abordons dans la prochaine section.

1.4.2. Estimation de la densité spectrale

Plusieurs techniques existent afin d'estimer la densité spectrale multivariée d'un processus stationnaire. Une première technique est le périodogramme matriciel. Soit

$$d_{Y^{(k)}}(\omega) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n Y_t(k) e^{-i\omega t}, \quad k = 1, \dots, d, \quad (1.4.3)$$

la transformée de Fourier discrète de $\{Y_t(k)\}$. Nous posons alors le périodogramme matriciel $\mathbf{I}_n(\omega) = (I_{n,kl}(\omega))$, où $I_{n,kl}(\omega) = d_{Y^{(k)}}(\omega) d_{Y^{(l)}}^*(\omega)$. En fait, si $k = l$, $I_{n,kk}(\omega)$ représente le périodogramme usuel de la série univariée $\{Y_t(k), t = 1, \dots, n\}$. Pour $k \neq l$, $I_{n,kl}(\omega)$ est appelé le périodogramme croisé entre $\{Y_t(k)\}$ et $\{Y_t(l)\}$. Un intérêt du périodogramme matriciel est que ce dernier peut être interprété comme la version échantillonnale de la densité spectrale. En fait, il est possible de montrer qu'aux fréquences de Fourier :

$$\mathbf{I}_n(\omega) = \frac{1}{2\pi} \sum_{s=-n+1}^{n-1} \hat{\Gamma}(s) e^{-i\omega s}, \quad (1.4.4)$$

où $\{\hat{\Gamma}(s), |s| < n\}$ est la fonction d'autocovariance échantillonnale. Nous notons que les fréquences de Fourier sont de la forme $2\pi j/n$. Le périodogramme matriciel possède des propriétés intéressantes : par exemple, il peut être montré qu'il est un estimateur asymptotiquement sans biais de la densité spectrale. Cependant, il souffre également des mêmes problèmes que le périodogramme usuel pour une série chronologique univariée. Ainsi, ce n'est pas un estimateur convergent de la densité spectrale. Une raison de cet état de fait peut être expliqué intuitivement. Nous notons que dans (1.4.4), chaque autocovariance résiduelle a la même importance dans le périodogramme. Autrement formulé, chaque autocovariance résiduelle se voit attribuer le même poids. Or, $\hat{\Gamma}(s)$ avec s proche de $n - 1$ est beaucoup plus imprécise que pour un s plus petit, la raison étant que $\hat{\Gamma}(n - 1)$ est calculé avec beaucoup moins d'observations que $\hat{\Gamma}(1)$, à titre d'exemple. Ainsi, globalement, il peut être montré que la variance de $\mathbf{I}_n(\omega)$ ne tend pas vers 0 et ainsi $\mathbf{I}_n(\omega)$ n'est pas un estimateur convergent. Voir aussi Brillinger (1981).

Une alternative au périodogramme est un estimateur non paramétrique de la densité spectrale $\mathbf{f}(\lambda)$ du processus $\{\mathbf{Y}_t\}$. Cette technique repose sur le choix d'un noyau satisfaisant l'hypothèse suivante :

Hypothèse 1.4.1. *Le noyau $k : \mathbb{R} \rightarrow [-1, 1]$ est une fonction symétrique, continue en zéro, ayant au plus un nombre fini de points de discontinuités tel que $k(0) = 1$ et $\int_{-\infty}^{\infty} k^2(z) dz < \infty$.*

Typiquement, les poids associés aux petits délais sont plus importants, de sorte que $k(j/p_n)$ pour un petit délai j prend une plus grande valeur que pour un grand délai j , pour un paramètre p_n donné (et discuté plus bas). Rappelons qu'au point $z = 0$ la fonction $k(z)$ doit satisfaire $k(0) = 1$, suggérant qu'un poids unitaire devrait être maximum. De plus, la condition $\int_{-\infty}^{\infty} k^2(z) dz < \infty$ exige une décroissance en fonction de z , de sorte que $k(z) \rightarrow 0$ lorsque $z \rightarrow \infty$. Il est à noter qu'en analyse spectrale, la grande majorité des noyaux utilisés en pratique satisfont l'hypothèse 1.4.1.

L'estimateur de la densité spectrale calculé avec la méthode du noyau est défini par l'estimateur général suivant :

$$\hat{\mathbf{f}}(\omega) = \frac{1}{2\pi} \sum_{j=-n+1}^{n+1} k(j/p_n) \hat{\Gamma}(j) e^{-ij\omega}, \quad (1.4.5)$$

où $k(\cdot)$ est un noyau ou un générateur de fenêtre. Le paramètre p_n correspond à un point de troncature lorsque le noyau est de support compact ou un paramètre de lissage lorsque le support du noyau n'est pas borné. De plus, le point de troncature satisfait les conditions suivantes : $p_n \rightarrow \infty$ et $p_n/n \rightarrow 0$.

Les noyaux les plus utilisés sont ceux attribuant un poids plus élevé aux petits délais et un poids moins élevé aux plus grands délais. Une exception digne de mention, compte tenu de son importance dans les applications, concerne le noyau uniforme tronqué, dont la définition est $k_T(z) = \mathbb{I}[|z| \leq 1]$ où $\mathbb{I}(A)$ est la fonction indicatrice de l'ensemble A . Ce noyau attribue le même poids à tous les délais considérés et amène essentiellement au périodogramme matriciel, d'où l'importance de ce noyau en pratique, même si d'autres noyaux sont supérieurs sous certaines conditions. D'autres noyaux sont le noyau de Bartlett, Daniell ou

encore le noyau de Parzen. Pour des définitions précises de ces noyaux, voir par exemple Priestley (1981).

Les propriétés asymptotiques des estimateurs non paramétriques de la densité spectrale multivariée ont été étudiées par plusieurs auteurs, notamment Hannan (1970) et Fuller (1996), à titre d'exemples. En fait, il peut être montré que la variance de l'estimateur (1.4.5) est d'ordre p_n/n , qui est supposée par hypothèse tendre vers 0. De plus, le biais asymptotique de (1.4.5) peut être démontré être d'ordre p_n^{-q} , où q est l'indice caractéristique du noyau (par exemple $q = 1$ lorsque le noyau est celui de Bartlett, et $q = 2$ pour les noyaux de Daniell et Parzen). Voir Hannan (1970, p. 286). Ainsi, puisque $p_n \rightarrow \infty$ tel que $p_n/n \rightarrow 0$, alors l'erreur quadratique moyenne tend vers 0 et l'estimateur non paramétrique de la densité spectrale est convergent pour $f(\omega)$.

Comme complément d'information, mentionnons qu'en choisissant le point de troncature ou paramètre de lissage tel que $p_n \propto n^{1/(2q+1)}$, l'erreur quadratique moyenne peut être rendue la plus petite possible et il devient alors possible d'exprimer $p_n = cn^{1/(2q+1)}$. Un meilleur choix de c peut alors être déterminé, mais ces considérations techniques sortent quelque peu des objectifs du présent travail.

En résumé, nous avons discuté de considérations générales entourant les processus stationnaires et partiellement non-stationnaires, et nous avons fait un bref survol des techniques spectrales. Dans l'article qui suit, nous abordons la question du diagnostic des modèles de séries chronologiques multivariées en utilisant des tests portemanteaux calculés selon une approche spectrale.

Chapitre 2

L'ARTICLE

Le présent chapitre est une copie intégrale de l'article intitulé "On the Power Transformation of Kernel-Based Tests for Serial Correlation in Vector Time Series : Some Finite Sample Results" soumis dans une revue scientifique avec un comité de lecture en juin 2006. Le premier auteur est également l'auteur du présent mémoire.

On the Power Transformation of Kernel-Based Tests for Serial Correlation in Vector Time Series : Some Finite Sample Results*

June 14, 2006

Jennifer POULIN
Université de Montréal

and

Pierre DUCHESNE
Université de Montréal

ABSTRACT

Portmanteau test statistics represent useful diagnostic tools for checking the adequacy of multivariate time series models. For stationary and partially non-stationary vector time series models, Duchesne and Roy (2004) and Duchesne (2005a) have proposed kernel-based test statistics, obtained by comparing the spectral density of the errors under the null hypothesis of non-correlation with a kernel-based spectral density estimator; these test statistics are asymptotically standard normal under the null hypothesis of non-correlation in the error term of the model. Following the method of Chen and Deo (2004a), we determine an appropriate power transformation to improve the normal approximation in small samples. Additional corrections for the mean and variance of the distance measures intervening in these test statistics are obtained. In a Monte Carlo study, comparisons are made under various alternatives between : the new corrected test statistics, the original spectral test statistics and the classical Hosking portmanteau test statistic.

Keywords and phrases : Cointegration ; kernel-based spectral density estimation ; multiple time series ; portmanteau test statistics ; power transformation ; spectral density.

Mathematics subject classification codes (1991) : primary 62M10 ; secondary 62M15.
JEL classification : C12, C32, C52.

*Abbreviated title : "On the power transformation of kernel-based tests".

Corresponding author : Pierre Duchesne, Université de Montréal; Département de mathématiques et de statistique; C.P. 6128 Succursale Centre-Ville; Montréal, Québec H3C 3J7; Canada. Tel : (514) 343-7267; Fax : (514) 343-5700; e-mail :



1. INTRODUCTION

In recent years, there has been an increasing interest in checking for residual autocorrelation in a fitted model from stationary or partially non-stationary multivariate time series models. Test procedures based on the residual autocovariances and/or autocorrelations have been found useful for examining the adequacy of a particular model. In stationary vector time series, the theoretical properties of residual autocovariances and autocorrelations as well as the asymptotic distributions of portmanteau test statistics have been studied in Chitturi (1974, 1976), Hosking (1980, 1981a, 1981b), Li and McLeod (1981), Ahn (1988), Paparoditis (2005), Duchesne and Roy (2004) and Duchesne (2005b), amongst others. For example, Hosking's (1980) multivariate portmanteau test statistic is advocated in popular textbooks : see for example Li (2004) and Lütkepohl (1993, 2005). Recent studies involving economic and financial multivariate time series find the presence of unit roots. See, e.g., Ahn and Reinsel (1990), Reinsel and Ahn (1992) and Yap and Reinsel (1995). However, the asymptotic properties of residual autocorrelations and test procedures designed for partially non-stationary time series models have been developed only recently. Brüggemann, Lütkepohl and Saikkonen (2006) have studied the properties of residual autocorrelations of vector error correction models and they have derived test statistics for residual autocorrelation. Other test procedures are proposed by Duchesne (2005a), who has extended the test procedures of Duchesne and Roy (2004) for partially non-stationary time series models.

The Duchesne and Roy (2004) and Duchesne (2005a) test procedures rely on a spectral approach. Their test statistics are based on a distance measure, such as the L_2 norm, between a nonparametric kernel-based spectral density estimator and the spectral density under the null hypothesis of non-correlation. These spectral test statistics can be expressed as a weighted sum of a function of all the residual autocovariance matrices, for lags $j = 1, \dots, n - 1$. Interestingly, these test statistics converge toward a standard normal distribution, and this asymptotic distribution is justified when the underlying data generating process (DGP) is stationary or partially non-stationary. The weights intervening in the

spectral tests are determined using a kernel function $k(\cdot)$, and a truncation point or a smoothing parameter, denoted p_n . The generalized Hosking test statistic is obtained when the truncated uniform kernel is adopted.

As noted by Chen and Deo (2004a), random variables which are non-negative linear combinations of positive independent random variables can display heavily right-skewed finite sample distribution, even though they might be normally distributed asymptotically, after an appropriate standardization. In view of this, Chen and Deo (2004a) proposed a simple method for determining an appropriate power transformation to ameliorate the normal approximation in small samples. They used their transformation for improving the finite sample approximation of the distribution of Hong's (1996) test statistic for serial correlation, which is designed for stationary univariate time series models. Hong's test statistic can be viewed as a special case of Duchesne and Roy's (2004) test statistic when the dimension of the variable of interest is one. Chen and Deo (2004a) found that test statistics have markedly better empirical level properties when they are based on a power transformation with additional mean and variance corrections, which rely on the asymptotic distribution of the autocorrelations. From their empirical studies, they found that, for several alternatives, finite sample evidence suggests that the corrected versions can display an overall increase in empirical power.

The first aim of this paper is to determine what we call the Chen-Deo power transformation, in order to improve the finite sample approximation to asymptotic normality for the spectral test statistics proposed by Duchesne and Roy (2004) and Duchesne (2005a). We demonstrate that the power transformation is independent of the dimension of the variable of interest and that it improves normality. Using the asymptotic distributions of the residual autocovariances established by Li and McLeod (1981) in the stationary case and those derived in Brüggemann *et al.* (2006) for partially non-stationary time series, better approximations for the mean and variance of the distance measures intervening in the test procedures of Duchesne and Roy (2004) and Duchesne (2005a) are derived.

In pursuit of the paper's second objective, we conduct finite-sample investigations of portmanteau test statistics designed for vector time series. Several

comparative studies of portmanteau test statistics have been carried out for univariate time series. See, e.g., the empirical studies described in Kheoh and McLeod (1992) and Kwan, Sim and We (2005). In the multivariate case, few empirical studies seem to be available, to the best of our knowledge. For stationary DGPs, we use simulation studies to compare Hosking's test statistic, the spectral test statistics introduced in Duchesne and Roy (2004), and the new corrected versions. When the DGP is partially non-stationary, Brüggemann *et al.* (2006) have shown that the Hosking portmanteau test statistic is appropriate in several practical circumstances, particularly when moderate and large residual autocorrelations are being tested. However, the degrees of freedom of this portmanteau test statistic need to be adjusted for the cointegration rank. Note that Duchesne's (2005a) spectral test statistics were not included in the empirical experiments of Brüggemann *et al.* (2006). Since it may be of practical interest to empirically evaluate these test statistics, with respect to exact levels and exact empirical powers, they are included in our study and compared via a small Monte Carlo experiment.

The paper is organized as follows. In Section 2, some preliminaries are given. In Section 2.1, the stationary vector autoregressive moving-average (VARMA) model is introduced; the sample residual autocovariance matrices are defined; and the test procedures are given. The vector error correction form (VECM) model is presented in Section 2.2. Furthermore, the asymptotic distribution of the residual autocovariances matrices are stated, and, using the results obtained by Ahn (1988), the approximate distribution of the multivariate portmanteau test statistics is provided. In Section 3, the corrected spectral test statistics are proposed. Simulation experiments are conducted in Section 4. Finally, in Section 5, some concluding remarks are presented.

2. PRELIMINARIES

Let $\mathbf{Y} = \{\mathbf{Y}_t, t \in \mathbb{Z}\}$ be a multivariate stochastic process, where $\mathbf{Y}_t = (Y_t(1), \dots, Y_t(d))^T$ corresponds to a d -dimensional random vector. In Section 2.1, we suppose that \mathbf{Y} can be represented by a second-order stationary VARMA model, while in Section 2.2 the stochastic process \mathbf{Y} is assumed to be generated by

a DGP which can be written in a VECM form.

2.1 Case where \mathbf{Y} is a stationary stochastic process

Consider the following multivariate autoregressive moving-average model VAR-MA(p, q), defined by :

$$\Phi(B)\mathbf{Y}_t = \boldsymbol{\theta}_0 + \Theta(B)\mathbf{a}_t, \quad (2.0.1)$$

where $\Phi(B) = \mathbf{I}_d - \sum_{j=1}^p \Phi_j B^j$, $\Phi_p \neq \mathbf{0}$, \mathbf{I}_d being the $d \times d$ identity matrix, $\boldsymbol{\theta}_0$ denoting the constant term, $\Theta(B) = \sum_{j=0}^q \Theta_j B^j$, $\Theta_0 = \mathbf{I}_d$, $\Theta_q \neq \mathbf{0}$ and B corresponding to the usual backward shift operator. The stochastic process $\mathbf{a} = \{\mathbf{a}_t, t \in \mathbb{Z}\}$ intervening in (2.0.1) is strong white noise, that is $\mathbf{a}_t = (a_t(1), \dots, a_t(d))^\top$, $t \in \mathbb{Z}$, are identically and independently distributed (iid) random vectors with mean zero and regular covariance matrix $\Sigma_{\mathbf{a}}$. We assume that \mathbf{Y} is stationary and invertible, that is all roots of $\det\{\Phi(z)\}$ and $\det\{\Theta(z)\}$ lie outside the unit disk, respectively, where z represents a complex number, that is $z \in \mathbb{C}$, and $\det(\mathbf{A})$ stands for the determinant of the matrix \mathbf{A} .

Let $\mathbf{e} = \{\mathbf{e}_t, t \in \mathbb{Z}\}$ be an arbitrary second order stationary process whose mean is $\mathbf{0}$, where $\mathbf{e}_t = (e_t(1), \dots, e_t(d))^\top$. The $d \times d$ autocovariance matrix at lag j is defined by :

$$\Gamma_{\mathbf{e}}(j) = E(\mathbf{e}_t \mathbf{e}_{t-j}^\top), \quad j \in \mathbb{Z}.$$

Given $\mathbf{e}_1, \dots, \mathbf{e}_n$ a finite realization of length n of the stochastic process \mathbf{e} , the sample autocovariances at lag j , $0 \leq |j| \leq n-1$, are represented by :

$$\mathbf{C}_{\mathbf{e}}(j) = \begin{cases} n^{-1} \sum_{t=j+1}^n \mathbf{e}_t \mathbf{e}_{t-j}^\top, & j = 0, 1, \dots, n-1, \\ \mathbf{C}_{\mathbf{e}}^\top(-j), & j = -1, \dots, -n+1. \end{cases} \quad (2.0.2)$$

Once a particular VARMA model has been estimated, the residual time series $\hat{\mathbf{a}}_t$, $t = 1, \dots, n$ can be determined. The residual autocovariance at lag j is obtained from (2.0.2) with \mathbf{e}_t replaced naturally by $\hat{\mathbf{a}}_t$ and it is denoted $\mathbf{C}_{\hat{\mathbf{a}}}(j)$, $|j| \leq n-1$. Let $\mathbf{c}_{\hat{\mathbf{a}}}(j) = \text{vec}\{\mathbf{C}_{\hat{\mathbf{a}}}(j)\}$ be a $d^2 \times 1$ random vector, $|j| \leq n-1$, and $\mathbf{c}_{\hat{\mathbf{a}}} = (\mathbf{c}_{\hat{\mathbf{a}}}^\top(1), \dots, \mathbf{c}_{\hat{\mathbf{a}}}^\top(M))^\top$ be a $d^2 M \times 1$ vector, where $\text{vec}(\mathbf{A})$ stands for the usual matrix operator that transforms the matrix \mathbf{A} into a vector by stacking its columns and $M < n$ denotes a fixed constant with respect to n . The asymptotic distribution of the residual autocovariances and autocorrelations in

vector autoregressive (VAR) and VARMA models have been studied by many authors, see Hosking (1980) and Li and McLeod (1981), amongst others. For a textbook exposition, see Lütkepohl (1993). Let ' $\xrightarrow{\mathcal{L}}$ ' be the symbol for convergence in distribution and let ' \otimes ' be the usual Kronecker product (see Harville (1997) and Lütkepohl (1993) for a review of the properties of this operator). Let $\{\Phi(B)\}^{-1} = \sum_{j \geq 0} \Psi_j B^j$, where $\Psi_0 = \mathbf{I}_d$ and $\Psi_j = \mathbf{0}$, $j < 0$. The asymptotic distribution of the vector of autocovariances in VAR models is given by :

$$n^{1/2} \mathbf{c}_{\hat{\mathbf{a}}} \xrightarrow{\mathcal{L}} N_{d^2 M}(\mathbf{0}, \mathbf{I}_M \otimes \Sigma_{\mathbf{a}} \otimes \Sigma_{\mathbf{a}} - \bar{\mathbf{G}}\{\Gamma_{\mathbf{Y}^{\#}}^{-1}(0) \otimes \Sigma_{\mathbf{a}}\} \bar{\mathbf{G}}^{\top}), \quad (2.0.3)$$

where $\bar{\mathbf{G}} = \mathbf{G}^{\# \top} \otimes \mathbf{I}_d$, and $\mathbf{G}^{\#}$ corresponds to the following $dp \times dM$ matrix :

$$\mathbf{G}^{\#} = \begin{pmatrix} \Sigma_{\mathbf{a}} & \Psi_1 \Sigma_{\mathbf{a}} & \cdots & \Psi_{M-1} \Sigma_{\mathbf{a}} \\ \mathbf{0} & \Sigma_{\mathbf{a}} & \cdots & \Psi_{M-2} \Sigma_{\mathbf{a}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \Psi_{M-p} \Sigma_{\mathbf{a}} \end{pmatrix}, \quad (2.0.4)$$

and $\Gamma_{\mathbf{Y}^{\#}}(0)$ represents the covariance matrix of $\mathbf{Y}_t^{\#} = (\mathbf{Y}_t^{\top}, \dots, \mathbf{Y}_{t-p+1}^{\top})^{\top}$. See Lütkepohl (1993, Proposition 4.5, pp. 146-147) for a simple proof of this result. When the DGP is generated by a VARMA model, the asymptotic normality of the residual autocovariances still holds, but with a different asymptotic matrix of variances and covariances. Explicit expressions of this asymptotic matrix can be found in Hosking (1980) and Li and McLeod (1981), amongst others.

The asymptotic distribution of the autocovariance matrices allows us to develop test procedures at a particular lag j , say, since $n \mathbf{c}_{\hat{\mathbf{a}}}^{\top}(j) \hat{\Omega}_{jj}^{-1} \mathbf{c}_{\hat{\mathbf{a}}}(j) \xrightarrow{\mathcal{L}} \chi_{d^2}^2$, where $\hat{\Omega}_{jj}$ represents a consistent estimator of the asymptotic variance Ω_{jj} of $n^{1/2} \mathbf{c}_{\hat{\mathbf{a}}}(j)$, that is $\lim_{n \rightarrow \infty} n \text{cov}(\mathbf{c}_{\hat{\mathbf{a}}}(j), \mathbf{c}_{\hat{\mathbf{a}}}(k)) = \Omega_{jk}$, where $j, k \in \{1, \dots, M\}$. It is certainly sound practice in real applications to investigate the remaining dependence at individual lags. See for example Duchesne (2005b) for these kinds of applications in VAR models with exogenous variables. In order to check the adequacy of VAR and VARMA models, the so-called multivariate portmanteau test statistic originally proposed by Hosking (1980) represents a popular procedure :

$$P_M = n^2 \sum_{j=1}^M (n-j)^{-1} \text{tr} \{ \mathbf{C}_{\hat{\mathbf{a}}}^{\top}(j) \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \mathbf{C}_{\hat{\mathbf{a}}}(j) \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \}, \quad (2.0.5)$$

which can be used for testing the joint significance of $\mathbf{c}_{\hat{\mathbf{a}}}(j)$, $j = 1, \dots, M$. The asymptotic distribution of Hosking's (1980) test statistic is approximately chi-square under the null hypothesis of model adequacy and the critical points are taken from the $\chi_{d^2(M-p-q)}^2$ distribution. The correction factor $n/(n-j)$ improves the finite sample properties. The Hosking test statistic, other test procedures, and alternative finite sample modifications are discussed in Li and McLeod (1981), Kheoh and McLeod (1992) and Lütkepohl (1993), amongst others.

Duchesne and Roy (2004) and Duchesne (2005a) have proposed test statistics using a spectral approach for checking the adequacy of stationary and cointegrated time series models. The test statistics rely fundamentally on kernel-based spectral density estimators of the residual time series. The kernel-based test statistics are defined by :

$$T_n(k; p_n) = \frac{n \sum_{j=1}^{n-1} k^2(j/p_n) \text{tr}\{\mathbf{C}_{\hat{\mathbf{a}}}^{\top}(j)\mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0)\mathbf{C}_{\hat{\mathbf{a}}}(j)\mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0)\} - M_n(k)}{\{2V_n(k)\}^{1/2}}, \quad (2.0.6)$$

where $M_n(k)$ and $V_n(k)$ are given by :

$$M_n(k) = d^2 \sum_{j=1}^{n-1} (1 - j/n) k^2(j/p_n), \quad (2.0.7)$$

$$V_n(k) = d^2 \sum_{j=1}^{n-2} (1 - j/n)(1 - (j+1)/n) k^4(j/p_n). \quad (2.0.8)$$

In the test statistic $T_n \equiv T_n(k; p_n)$ defined by (2.0.6), the expression

$$\mathcal{Q}_n = n \sum_{j=1}^{n-1} k^2(j/p_n) \text{tr}\{\mathbf{C}_{\hat{\mathbf{a}}}^{\top}(j)\mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0)\mathbf{C}_{\hat{\mathbf{a}}}(j)\mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0)\} \quad (2.0.9)$$

measures the distance between a spectral density estimator calculated using the data and the spectral density under the null hypothesis of adequacy, whereas the quantities $M_n(k)$ and $2V_n(k)$ represent essentially the mean and variance of this distance measure. Consequently, we can interpret these expressions as standardization factors, in order to achieve asymptotic normality. The kernel-based test statistic T_n depends on a kernel $k(\cdot)$ and a truncation point or a smoothing parameter p_n . When the truncated uniform kernel $k_{TR}(z) = I(\{|z| \leq 1\})$ is adopted, where $I(A)$ denotes the indicator function of the set A , this gives a generalized Hosking test statistic. For stationary and cointegrated time series, by including

$p_n \rightarrow \infty$ such that $p_n/n \rightarrow 0$, we find that under certain conditions $T_n(k; p_n)$ converges toward an asymptotic normal distribution under the null hypothesis of adequacy, that is $T_n(k; p_n) \xrightarrow{\mathcal{L}} N(0, 1)$. See Duchesne and Roy (2004) and Duchesne (2005a) for more details.

2.2 Case where \mathbf{Y} is a cointegrated stochastic process

In this set-up we suppose that \mathbf{Y}_t , $t = 1, \dots, n$ is an integrated time series of order one, with $r < d$ cointegration relations. These random vectors are assumed to be generated by a DGP which can be written in the following VECM form :

$$\Phi^*(B)\mathbf{W}_t = \boldsymbol{\nu} + \mathbf{A}\{\mathbf{B}^\top \mathbf{Y}_{t-1} - (t-1)\boldsymbol{\tau}\} + \mathbf{u}_t, \quad (2.0.10)$$

where $\mathbf{W}_t = (1-B)\mathbf{Y}_t$, $\Phi^*(B) = \mathbf{I}_d - \sum_{j=1}^{p-1} \Phi_j^* B^j$, \mathbf{A} and \mathbf{B} are full-rank matrices of dimension $d \times r$, $\boldsymbol{\nu}$ is a $d \times 1$ vector of constants, $\boldsymbol{\tau}$ represents a r -dimensional vector of trend slope parameters and Φ_j^* , $j = 1, \dots, p-1$ denote $d \times d$ shortrun parameter matrices. The error process $\mathbf{u} = \{\mathbf{u}_t\}$ is assumed to be strong white noise with mean zero and regular covariance matrix $\Sigma_{\mathbf{u}}$. According to Granger's representation theorem (see, e.g., Johansen (1995)), it follows that the stochastic process \mathbf{Y} can be expressed as :

$$\mathbf{Y}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t + \mathbf{X}_t, \quad (2.0.11)$$

where the vectors $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ can be expressed in terms of the parameters in (2.0.10). Let \mathbf{A}_\perp be an orthogonal complement of the $n \times m$ matrix \mathbf{A} , where \mathbf{A} is of full column rank ($n > m$). The process $\mathbf{X} = \{\mathbf{X}_t\}$ satisfies :

$$\mathbf{X}_t = \boldsymbol{\Upsilon} \sum_{j=1}^t \mathbf{u}_j + \Pi(B)\mathbf{u}_t + \mathbf{P}_{\mathbf{B}_\perp} \mathbf{X}_0,$$

where the matrix $\boldsymbol{\Upsilon}$ is given by $\boldsymbol{\Upsilon} = \mathbf{B}_\perp \{\mathbf{A}_\perp^\top (\mathbf{I}_d - \sum_{j=1}^{p-1} \Phi_j^*) \mathbf{B}_\perp\}^{-1} \mathbf{A}_\perp^\top$, the vector \mathbf{X}_0 contains the initial values, the projection matrix $\mathbf{P}_{\mathbf{B}_\perp}$ is defined by $\mathbf{P}_{\mathbf{B}_\perp} = \mathbf{B}_\perp (\mathbf{B}_\perp^\top \mathbf{B}_\perp)^{-1} \mathbf{B}_\perp^\top$, and $\Pi(z)$ corresponds to a geometrically convergent power series in the variable z .

Consider the residual autocovariances $\mathbf{C}_{\hat{\mathbf{u}}}(j)$, $|j| \leq n-1$ based on the residuals of the VECM model :

$$\hat{\mathbf{u}}_t = \hat{\Phi}^*(B)\mathbf{W}_t - \hat{\nu} - \hat{\mathbf{A}}\{\hat{\mathbf{B}}^\top \mathbf{Y}_{t-1} - (t-1)\hat{\boldsymbol{\tau}}\}, \quad j = 1, \dots, n.$$

The parameters and residuals in the VECM model are computed using the reduced-rank (RR) regression method or other asymptotically equivalent estimation procedures (see Ahn and Reinsel (1990) or Yap and Reinsel (1995) for alternative estimation techniques, amongst others). Brüggeman *et al.* (2006) have established, under general conditions, the asymptotic normality of $\mathbf{c}_{\hat{\mathbf{u}}} = (\mathbf{c}_{\hat{\mathbf{u}}}^\top(1), \dots, \mathbf{c}_{\hat{\mathbf{u}}}^\top(M))^\top$. More precisely, suppose that the process satisfies the relation (2.0.10) and that the parameters are estimated by RR regression (or some other asymptotically equivalent estimation method), and assume that the estimators satisfy the following rates of convergence : $\hat{\mathbf{B}} - \mathbf{B} = \mathbf{O}_P(n^{-1})$, $\hat{\boldsymbol{\tau}} - \hat{\mathbf{B}}^\top \boldsymbol{\mu}_1 = \mathbf{O}_P(n^{-3/2})$. Let $\mathbf{U}_t = (\mathbf{u}_{t-1}^\top, \dots, \mathbf{u}_{t-M}^\top)^\top$ be a $Md \times 1$ vector and $\mathbf{Z}_t = ((\mathbf{B}^\top \mathbf{X}_{t-1})^\top, \mathbf{V}_{t-1}^\top, \dots, \mathbf{V}_{t-p+1}^\top)^\top$ be a $\{d(p-1) + r\} \times 1$ vector, where the random vector \mathbf{V}_t is defined as $\mathbf{V}_t = (1 - B)\mathbf{X}_t$. Introduce the following matrices :

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathbf{ZU}} &= \text{plim } n^{-1} \sum_{t=1}^n \mathbf{Z}_t \mathbf{U}_t^\top = \boldsymbol{\Sigma}_{\mathbf{UZ}}^\top, \\ \boldsymbol{\Sigma}_{\mathbf{ZZ}} &= \text{plim } n^{-1} \sum_{t=1}^n \mathbf{Z}_t \mathbf{Z}_t^\top. \end{aligned}$$

Then the asymptotic distribution of $\mathbf{c}_{\hat{\mathbf{u}}}$ is multivariate normal, that is :

$$n^{1/2} \mathbf{c}_{\hat{\mathbf{u}}} \xrightarrow{\mathcal{L}} N_{d^2 M}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{UU} \cdot \mathbf{Z}} \otimes \boldsymbol{\Sigma}_{\mathbf{u}}), \quad (2.0.12)$$

where $\boldsymbol{\Sigma}_{\mathbf{UU} \cdot \mathbf{Z}} = \boldsymbol{\Sigma}_{\mathbf{UU}} - \boldsymbol{\Sigma}_{\mathbf{UZ}} \boldsymbol{\Sigma}_{\mathbf{ZZ}}^{-1} \boldsymbol{\Sigma}_{\mathbf{ZU}}$, and note that $\boldsymbol{\Sigma}_{\mathbf{UU}} = \mathbf{I}_M \otimes \boldsymbol{\Sigma}_{\mathbf{u}}$ is a non-singular matrix.

Brüggeman *et al.* (2006) discuss how to use the relation (2.0.11) in order to transform the original model from (2.0.10) to the model :

$$\Phi^*(B)\mathbf{V}_t = \boldsymbol{\nu}^{(0)} + \mathbf{A}\{\mathbf{B}^\top \mathbf{X}_{t-1} - (t-1)\boldsymbol{\tau}^{(0)}\} + \mathbf{u}_t,$$

where $\boldsymbol{\nu}^{(0)} = \boldsymbol{\nu} + \mathbf{A}\mathbf{B}^\top \boldsymbol{\mu}_{0o} - \Phi^*(1)\boldsymbol{\mu}_{1o}$ and $\boldsymbol{\tau}^{(0)} = \boldsymbol{\tau} - \mathbf{B}^\top \boldsymbol{\mu}_{1o}$. Here $\boldsymbol{\mu}_{0o}$ and $\boldsymbol{\mu}_{1o}$ correspond to the true values in (2.0.11). This transformation is particularly useful for deriving the asymptotic distribution of $\mathbf{c}_{\hat{\mathbf{u}}}$.

Using the auxiliary model

$$\Phi^*(B)\mathbf{V}_t = \mathbf{A}\mathbf{B}^\top \mathbf{X}_{t-1} + \mathbf{u}_t,$$

and using results obtained by Ahn (1988), Brüggeman *et al.* (2006) demonstrate that the approximate distribution of the portmanteau test P_M defined by (2.0.5) is $\chi_{Md^2-d^2(p-1)-dr}^2$. Consequently, the degrees of freedom of the chi-square distribution are different than those of the chi-square distribution of the multivariate portmanteau test statistic for diagnosing a stationary VAR(p) model, where the degrees of freedom are $d^2(M - p)$. See for example Lütkepohl (1993, 2005).

The spectral approach of Duchesne (2005a) may also be regarded as a test procedure for testing serial correlation in (2.0.10). The kernel-based test statistic T_n defined by (2.0.6) admits a standard normal distribution under the null hypothesis of adequacy, and, interestingly, there is no need to adjust the critical points for the cointegration rank as in the multivariate portmanteau test statistic P_M . From a theoretical point of view, the method of proof to establish the asymptotic distribution of T_n for cointegrated time series differs considerably from the one developed in the stationary case, but, for practical purposes, the same test statistic can be used for stationary or cointegrated time series models. See Duchesne (2005a) for additional details.

3. NEW CORRECTED TEST STATISTICS

Chen and Deo (2004a) note that positive test statistics which can be expressed as non-negative linear combinations of positive independent random variables can have heavily right-skewed finite sample distributions, even though the asymptotic distribution of such procedures might be normally distributed asymptotically, after some kind of standardization. One important example of a test statistic in this category is the generalized portmanteau test statistic for serial correlation proposed by Hong (1996). In fact, in Monte Carlo studies, Chen and Deo (2004b) found that the asymptotic theory for establishing the standard normal distribution of the Hong test statistic did not hold well in finite samples : the distribution of the Hong test was found to be heavily right-skewed. Chen and Deo (2004a) proposed a simple transformation to ameliorate the finite sample distribution to

asymptotic normality. More precisely, a power transformation, with additional sample mean and variance corrections, was found to provide markedly better finite sample properties, with empirical levels much closer to the nominal levels. Furthermore, under certain alternatives, the corrected test statistics exhibited better power than the Box-Ljung-Pierce portmanteau test.

The test statistic T_n represents a generalization of Hong's test for vector time series and it admits essentially the same form as that of the Hong test : the test statistic is motivated by the distance measure Q_n given by (2.0.9), which is positive with a probability equal to one. Empirical studies reveal that the finite sample distribution can be heavily right-skewed with empirical levels relatively far from the nominal levels, as in the univariate case. Furthermore, the multivariate nature of the problem seems to create additional distortions. In general, it seems that the exact finite sample distribution of T_n and the asymptotic normal distribution are only close for rather large sample sizes. For stationary time series, Duchesne and Roy (2004) showed that :

$$\sum_{j=1}^{n-1} k^2(j/p_n) [\text{tr} \{ \mathbf{C}_a^{-1}(0) \mathbf{C}_a(j) \mathbf{C}_a^{-1}(0) \mathbf{C}_a^{\top}(j) \} - \text{tr} \{ \mathbf{C}_{\hat{a}}^{-1}(0) \mathbf{C}_{\hat{a}}(j) \mathbf{C}_{\hat{a}}^{-1}(0) \mathbf{C}_{\hat{a}}^{\top}(j) \}] = o_p(\sqrt{p_n}/n). \quad (2.0.13)$$

A similar result holds for partially non-stationary time series. See Duchesne (2005a, p. 582). Furthermore, for a given j it is well-known that :

$$n^{1/2} \text{vec} \{ \mathbf{C}_a(j) \} \xrightarrow{\mathcal{L}} N_{d^2}(\mathbf{0}, \boldsymbol{\Sigma}_a \otimes \boldsymbol{\Sigma}_a).$$

See for example Lütkepohl (1993, Proposition 4.4, p. 140). Using the relation $\text{tr}(\mathbf{A}^{\top} \mathbf{B} \mathbf{C} \mathbf{D}^{\top}) = \{ \text{vec}(\mathbf{A}) \}^{\top} (\mathbf{D} \otimes \mathbf{B}) \{ \text{vec}(\mathbf{C}) \}$ (see, e.g., Harville (1997, p. 342)), it follows that :

$$\text{tr} \{ \mathbf{C}_a^{-1}(0) \mathbf{C}_a(j) \mathbf{C}_a^{-1}(0) \mathbf{C}_a^{\top}(j) \} = \mathbf{c}_a^{\top}(j) \{ \mathbf{C}_a^{-1}(0) \otimes \mathbf{C}_a^{-1}(0) \} \mathbf{c}_a(j).$$

When \mathbf{a} corresponds to an iid stochastic process, classical results on quadratic forms give :

$$n \text{tr} \{ \mathbf{C}_a^{-1}(0) \mathbf{C}_a(j) \mathbf{C}_a^{-1}(0) \mathbf{C}_a^{\top}(j) \} \xrightarrow{\mathcal{L}} \chi_{d^2}^2.$$

Consequently, the expression \mathcal{Q}_n given by (2.0.9) intervening in the Duchesne and Roy (2004) test statistic can be approximately expressed as :

$$\mathcal{S}_n = \sum_{j=1}^{n-1} a_{jn} \mathcal{X}_j, \quad (2.0.14)$$

where $\{\mathcal{X}_j\}$ represents a sequence of iid $\chi_{d^2}^2$ random variables and a_{jn} , $j = 1, \dots, n-1$, denote the non-negative weights, that is $a_{jn} = k^2(j/p_n) \geq 0$, $j = 1, \dots, n-1$. Following the method of Chen and Deo (2004a), we now determine the power transformation of the Duchesne and Roy (2004) test statistic. To this end, consider the random variable \mathcal{S}_n defined by (2.0.14). The mathematical expectation and variance of the scaled random variable $p_n^{-1}\mathcal{S}_n$ are given by :

$$\begin{aligned} E(p_n^{-1}\mathcal{S}_n) &= \mu p_n^{-1} \sum_{j=1}^{n-1} a_{jn} \equiv \mu_{\mathcal{S}}, \\ \text{var}(p_n^{-1}\mathcal{S}_n) &= \sigma^2 p_n^{-2} \sum_{j=1}^{n-1} a_{jn}^2 \equiv \sigma_{\mathcal{S}}^2, \end{aligned}$$

where the first two moments of the chi-square distribution $\chi_{d^2}^2$ are $\mu = d^2$ and $\sigma^2 = 2d^2$. Chen and Deo (2004a) have shown that :

$$\frac{\mathcal{S}_n^\beta - \{\mu_{\mathcal{S}}^\beta + \frac{1}{2}\beta(\beta-1)\mu_{\mathcal{S}}^{\beta-2}\sigma_{\mathcal{S}}^2\}}{\beta\mu_{\mathcal{S}}^{\beta-1}\sigma_{\mathcal{S}}} \xrightarrow{\mathcal{L}} N(0, 1),$$

where

$$\beta = 1 - \frac{\mu E(\mathcal{X}_1 - \mu)^3 (\sum_{j=1}^{n-1} a_{jn}) (\sum_{j=1}^{n-1} a_{jn}^3)}{3\sigma^4 (\sum_{j=1}^{n-1} a_{jn}^2)^2}.$$

Given that $E(\mathcal{X}_1^3) = d^2(d^2 + 2)(d^2 + 4)$, this analysis suggests that T_n will be approximately normal if the following power transformation is adopted :

$$\beta = 1 - \frac{2 (\sum_{j=1}^{n-1} k_j^2) (\sum_{j=1}^{n-1} k_j^6)}{3 (\sum_{j=1}^{n-1} k_j^4)^2}, \quad (2.0.15)$$

where we note $k_j = k(j/p_n)$, $j = 1, \dots, n-1$. Interestingly, the power transformation β is invariant with respect to the dimension of the d -dimensional vector \mathbf{Y}_t . In the univariate case (that is, when $d = 1$), Chen and Deo (2004a) note that the power transformation provides an excellent approximation of the normal distribution for the Hong test statistic. However, empirical studies have suggested other kinds of problems, since the empirical levels of the test statistic still did not

match the nominal levels, under the null hypothesis. A careful investigation explained this discrepancy, as the expressions $M_n(k)$ and $2V_n(k)$, which correspond essentially to the mean and variance of the distance measure in the Hong test statistic, are rather inaccurate formulas in finite samples. Following an argumentation similar to that of Chen and Deo (2004a), but adapted to the multivariate case, we first remark that :

$$n \sum_{j=1}^{n-1} k^2(j/p_n) \text{tr} \{ \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \mathbf{C}_{\hat{\mathbf{a}}}(j) \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \mathbf{C}_{\hat{\mathbf{a}}}^{\top}(j) \} =$$

$$n \sum_{j=1}^{n-1} k^2(j/p_n) \mathbf{c}_{\hat{\mathbf{a}}}^{\top}(j) \{ \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \} \mathbf{c}_{\hat{\mathbf{a}}}(j) = n \mathbf{c}_{\hat{\mathbf{a}}}^{\top} \mathbf{K}^* \mathbf{c}_{\hat{\mathbf{a}}},$$

where $\mathbf{K}^* = \mathbf{K}^2 \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0)$, with $\mathbf{K} = \text{diag}(k_1, \dots, k_{n-1})$, where the matrix $\text{diag}(\mathbf{A})$ corresponds to the $n \times n$ diagonal matrix composed of the diagonal of the $n \times n$ matrix \mathbf{A} .

Intuitively, $\mathbf{c}_{\hat{\mathbf{a}}}$ can be thought of as a normal random vector with an expectation equal to $\mathbf{0}$ and with variance $n^{-1}\mathbf{\Omega}$, where $\mathbf{\Omega} = \mathbf{I}_{n-1} \otimes \mathbf{\Sigma}_{\mathbf{a}} \otimes \mathbf{\Sigma}_{\mathbf{a}} - \bar{\mathbf{G}}\{\Gamma_{\mathbf{Y}^{\#}}^{-1}(0) \otimes \mathbf{\Sigma}_{\mathbf{a}}\}\bar{\mathbf{G}}^{\top}$ in the stationary case, and, in the cointegrated case, $\mathbf{\Omega}$ represents the asymptotic variance given in expression (2.0.12). Using well-known results on quadratic forms (see, e.g., Searle (1971, pp. 56-57)), the mathematical expectation and the variance of $\mathbf{X}^{\top} \mathbf{A} \mathbf{X}$, where $\mathbf{X} \sim N_p(\mathbf{0}, \mathbf{\Sigma})$ is a p -dimensional normal random vector, are given by $E(\mathbf{X}^{\top} \mathbf{A} \mathbf{X}) = \text{tr}(\mathbf{A} \mathbf{\Sigma})$ and $\text{var}(\mathbf{X}^{\top} \mathbf{A} \mathbf{X}) = 2\text{tr}(\mathbf{A} \mathbf{\Sigma} \mathbf{A} \mathbf{\Sigma})$, respectively. Let $\hat{\mathbf{\Omega}}$ be a consistent estimator of $\mathbf{\Omega}$. Consequently, the first and second moments of $n \mathbf{c}_{\hat{\mathbf{a}}}^{\top} \mathbf{K}^* \mathbf{c}_{\hat{\mathbf{a}}}$ can be estimated as :

$$E(\mathcal{S}_n) = M_n^*(k),$$

$$\text{var}(\mathcal{S}_n) = 2V_n^*(k),$$

where $M_n^*(k) = \text{tr}(\mathbf{K}^* \hat{\mathbf{\Omega}})$ and $V_n^*(k) = \text{tr}(\mathbf{K}^* \hat{\mathbf{\Omega}} \mathbf{K}^* \hat{\mathbf{\Omega}})$. This analysis suggests a new testing procedure, which is the corrected test statistic defined by :

$$T_n^*(k; p_n) = \frac{\mathcal{Q}_n^{\beta} - [\{M_n^*(k)\}^{\beta} + \frac{1}{2}\beta(\beta-1)\{M_n^*(k)\}^{\beta-2}\{2V_n^*(k)\}]}{\beta\{M_n^*(k)\}^{\beta-1}\{2V_n^*(k)\}^{1/2}}, \quad (2.0.16)$$

where \mathcal{Q}_n and β are given by expressions (2.0.9) and (2.0.15), respectively. If the matrix $\hat{\mathbf{\Omega}}$ is replaced by $\mathbf{I}_{n-1} \otimes \mathbf{C}_{\hat{\mathbf{a}}}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}(0)$ and if $\beta = 1$, the relations

$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC} \otimes \mathbf{BD})$ and $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr}(\mathbf{A})\text{tr}(\mathbf{B})$ (see, e.g., Harville (1997, pp. 337-338)) allow us to easily show that we retrieve essentially the Duchesne and Roy (2004) test statistic. However, the power transformation and the mean and variance corrections are expected to improve the finite sample properties, particularly under the null hypothesis. In the next section, empirical comparisons are made between : the new corrected test statistic T_n^* , the untransformed test statistic T_n , and the classical portmanteau test statistic P_M , with respect to their exact levels and also their empirical powers.

4. SIMULATION RESULTS

Using the Chen-Deo power transformation, we have proposed new test statistics $T_n^* \equiv T_n^*(k; p_n)$ defined by (2.0.16) that should offer interesting finite sample properties, particularly under the null hypothesis of adequacy. From a practical point of view, it appears rather essential to conduct some Monte Carlo experiments. It seems quite relevant to investigate the exact levels and powers for reasonable time series lengths, to see whether the empirical levels of the corrected kernel-based test statistics display notable improvements compared to the original untransformed proposals.

To partially answer these questions, we use Monte Carlo simulations to compare the transformed and untransformed kernel-based test statistics. Since Hosking's (1980) multivariate portmanteau test statistic defined by (2.0.5) represents a popular procedure it is also included in our study.

4.1 Level analysis

For the given bivariate VAR(1) models described below, we first perform a level study. To this end, we examine the frequencies of rejection of the null hypothesis when it is in fact true, using the kernel-based test statistics and the multivariate portmanteau test statistic. The nominal levels 1, 5 and 10% are adopted in the level analysis. Small and moderate time series lengths are considered : $n = 50, 100$ and 200 . Three kernels are included in the study : the truncated uniform kernel (TR), which delivers the generalized Hosking test statistic; the

Bartlett (BAR) kernel; and the Daniell (DAN) kernel. Based on theoretical and empirical grounds, Duchesne and Roy (2004) and Duchesne (2005a) have found that the Bartlett and Daniell kernels display high power and that they may be recommended for practical purposes. The precise definitions of these kernels can be found in Table 1 of Duchesne and Roy (2004). See also Priestley (1981) for more details on the properties of these kernels.

For our investigations of the exact levels of the test statistics, we consider six DGPs, denoted DGP_l , $l \in \{1, 2, 3, 4, 5, 6\}$. They are all defined by the following VAR(1) model :

$$DGP_l : \mathbf{Y}_t = \mathbf{c}_l + \Phi_l \mathbf{Y}_{t-1} + \mathbf{a}_t, \quad t = 1, \dots, \quad (2.0.17)$$

where $\{\mathbf{a}_t\}$ denotes an iid Gaussian stochastic process such that $\mathbf{a}_t \sim N_2(\mathbf{0}, \Sigma_l)$. The definitions of the parameter values specifying DGP_l , $l \in \{1, 2, 3, 4, 5, 6\}$, are given in Table 2.1.

TAB. 2.1. Model coefficients of the data generating processes in the level analysis.

$$\begin{array}{l} \Phi_1 = \begin{pmatrix} -0.5 & 0.5 \\ -1.4 & -0.2 \end{pmatrix}, \quad \Phi_2 = \begin{pmatrix} -0.015 & 0.036 \\ 0.659 & 0.100 \end{pmatrix}, \quad \Phi_3 = \begin{pmatrix} 1.2 & -0.5 \\ 0.6 & 0.3 \end{pmatrix}, \\ \Phi_4 = \begin{pmatrix} 0.8 & 0.3 \\ 0.1 & 0.6 \end{pmatrix}, \quad \Phi_5 = \begin{pmatrix} 0.4 & 0.2 \\ -0.2 & 0.8 \end{pmatrix}, \quad \Phi_6 = \begin{pmatrix} 0.60 & 1.00 \\ 0.12 & 0.70 \end{pmatrix}, \\ \nu_1 = \begin{pmatrix} 3.0 \\ 2.0 \end{pmatrix}, \quad \nu_2 = \begin{pmatrix} 0.0 \\ 0.0 \end{pmatrix}, \quad \mathbf{V}_1 = \begin{pmatrix} 1.00 & 0.75 \\ 0.75 & 1.00 \end{pmatrix}, \\ \mathbf{V}_2 = \begin{pmatrix} 25.0 & 5.4 \\ 5.4 & 9.0 \end{pmatrix}. \end{array}$$

In the experiments DGP_1 - DGP_5 the values of the parameters are : $\mathbf{c}_l = \nu_1$, $\Sigma_l = \mathbf{V}_1$, $l \in \{1, 2, 3, 4, 5\}$. Under DGP_6 : $\mathbf{c}_6 = \nu_2$, $\Sigma_6 = \mathbf{V}_2$. Stationary VAR(1) models are generated under stochastic processes DGP_1 - DGP_5 . The process in experiment DGP_6 corresponds to a cointegrated VAR(1) process of rank $r = 1$. The autoregressive matrices Φ_l , $l = 1, \dots, 6$ of our VAR(1) models are taken from previous simulation studies and data analysis; their eigenvalues present different

and realistic characteristics. The matrix Φ_1 (with eigenvalues $-0.35 \pm 0.82i$, $i = \sqrt{-1}$) comes from the simulation study of Duchesne and Roy (2004) and Φ_3 (with eigenvalues $0.75 \pm 0.31i$) is inspired from an example studied in Reinsel (1993, p. 71). Since the roots of $\det(\lambda I_2 - \Phi_l) = 0$ are complex numbers, $l \in \{1, 3\}$, the correlations of these VAR(1) processes exhibit damped sinusoidal behavior. The VAR(1) model based on Φ_2 (with real eigenvalues 0.21, -0.12) corresponds essentially to the retained model of the Dow Jones and All Ordinaries Indices data analyzed in Brockwell and Davis (1996). The matrices Φ_4 and Φ_5 (with real eigenvalues 0.9, 0.5 and 0.6, 0.6, respectively) have been considered by Wei (1990). Finally, the cointegrated VAR(1) model in DGP₆ has been analyzed in the simulation example of Ahn and Reinsel (1990, Section 8). See also Duchesne (2005a). The eigenvalues of Φ_6 are 1.0 and 0.3.

In the level study, 1000 independent realizations were generated from DGP _{l} , $l \in \{1, 2, 3, 4, 5, 6\}$, for each value of n , and the computations were made in the following way :

1. The Gaussian white noises $\{\mathbf{a}_t\}$ were generated independently using the built-in generator provided in S-PLUS.
2. Using the initial value \mathbf{Y}_0 corresponding to the expected value of \mathbf{Y}_t , a total of $N = 2n + 1$ values of \mathbf{Y}_t , $t = 1, \dots, N$ were obtained. Only the last n observations were retained, providing a time series of length n and the first $n + 1$ values were discarded ; this practice minimizes the effect of the initial value \mathbf{Y}_0 .
3. For each realization, the true DGP was estimated by classical least-squares for experiments DGP₁-DGP₅. For DGP₆, the two-step approach described in Ahn and Reinsel (1990) has been used : in the first step, least-squares estimators are calculated and, in the second step, RR estimators are determined. Once the models were adjusted, the residual time series were obtained.
4. With each residual time series, the untransformed test statistic T_n defined by (2.0.6) and the transformed test statistic T_n^* given by (2.0.16) were computed for BAR, DAN, and TR kernels. These test statistics rely on

p_n . Thus we considered low values of p_n , often associated with powerful test statistics, and also rates in accordance with previous studies (see, e.g., Hong (1996), Duchesne and Roy (2004) and Duchesne (2005a)). This gives for $n = 50 : p_n = 2, 3, 4, 8, 10$, for $n = 100 : p_n = 2, 3, 5, 9, 12$ and for $n = 200 : p_n = 2, 3, 5, 10, 15$. Hosking's (1980) portmanteau test statistic has also been considered. For stationary DGP_1 - DGP_5 , the usual critical points have been calculated, that is they were taken from the chi-square distribution $\chi_{d^2(M-p)}^2$ with $d = 2$, $p = 1$ and $M = p_n$. For the cointegrated DGP_6 , they were computed from the chi-square distribution $\chi_{Md^2-d^2(p-1)-dr}^2$, with $d = 2$, $p = 1$, $r = 1$ and $M = p_n$.

5. Finally, for each time series of length n , for each spectral kernel $k(\cdot)$, for each value of p_n and $M = p_n$, and for each nominal level, we obtained from the 1000 replications the empirical frequencies of rejection of the null hypothesis of adequacy. Based on 1000 replications, acceptable results should be in the intervals $[0.4, 1.6]$, $[3.6, 6.4]$, and $[8.1, 11.9]$, at the 1, 5 and 10% nominal levels, respectively.

4.2 Power analysis

The power analysis was conducted in a similar manner. The same DGPs were considered, but the stochastic process $\{\mathbf{a}_t\}$ was generated by a vector moving-average process of order one, noted VMA(1), rather than a strong white noise; consequently the resulting model is a VARMA(1,1) model. More precisely, we considered :

$$\mathbf{a}_t = \mathbf{b}_t - \Theta_l \mathbf{b}_{t-1}, \quad (2.0.18)$$

where $\{\mathbf{b}_t\}$ is an iid Gaussian stochastic process such that $\mathbf{b}_t \sim N_2(\mathbf{0}, \Sigma_l)$. The same set of parameters \mathbf{c}_l , Φ_l and Σ_l defining DGP_l was used. The additional values of the parameters corresponding to the moving-average part of the VARMA model are given in Table 2.2.

For DGP_1 , DGP_3 and DGP_4 , we set $\Theta_l = \Xi_1$, $l \in \{1, 3, 4\}$; for DGP_2 and DGP_5 we consider $\Theta_l = \Xi_2$, $l \in \{2, 5\}$, and finally for DGP_6 we let $\Theta_6 = \Xi_3$.

TAB. 2.2. Additional model coefficients of the data generating processes in the power analysis.

$$\Xi_1 = \begin{pmatrix} 0.18 & 0.04 \\ 0.00 & 0.02 \end{pmatrix}, \quad \Xi_2 = \begin{pmatrix} 0.63 & 0.14 \\ 0.00 & 0.07 \end{pmatrix}, \quad \Xi_3 = \begin{pmatrix} 0.24 & 0.08 \\ 0.00 & 0.04 \end{pmatrix}.$$

We computed the empirical powers using the asymptotical critical values and the exact (empirical) critical values obtained from the level analysis. The use of empirical critical values offers the advantage of comparing all the test statistics on an equal basis, since they are size-adjusted. For practical applications, a satisfactory test procedure should offer comparable power results using asymptotic and empirical critical values. The results are presented for $n = 100, 200$ and for the nominal levels 5 and 10%. The empirical powers based on the empirical values are given in parentheses.

4.3 Discussion of the empirical results

We begin our discussion by analyzing the results from the level study. They are presented in Tables 2.3-2.8. In general, when the transformed and original kernel-based test statistics T_n^* and T_n are compared for a given nominal level, kernel, smoothing parameter p_n and sample size, the Chen-Deo power transformation has improved substantially the normal approximation for all nominal levels.

It seems that the power transformation provided reasonable empirical levels for all the sample sizes under investigation. Pronounced right-skewed finite sample distributions of the test statistic T_n appeared for some DGPs but even in these cases the power transformation improved the normal approximation considerably. Some improvement was observed as a function of n , particularly at the 10% nominal level and for small and moderate values of the smoothing parameter p_n . The kernels BAR and DAN displayed rather similar characteristics, with a slight advantage for the BAR kernel; the test statistic T_n^* based on the kernel TR offered a different behavior and was somewhat inferior.

For all the DGPs considered, the test statistic T_n displayed sometimes rather severe underrejection. These findings are in agreement with those obtained by

Chen and Deo (2004a) for diagnosing ARMA models in the univariate case. This is particularly true under DPG_2 and DPG_5 , where very large underrejection occurred for the test statistic T_n , even at the 10% nominal level. In other simulation studies involving dynamic models, reasonable empirical levels often occur at this nominal level (see, e.g, Hong (1996), Duchesne and Roy (2004)). Interestingly, under DGP_2 and for a given sample size, all the empirical levels of the test statistic T_n based on the BAR kernel range from 0.0 to 1.0% at the 5% nominal level. By comparison, all the empirical levels of T_n^* fall in the interval $[3.6, 6.4]$ and are consequently very satisfactory. At the 10% nominal level, the empirical levels of the test statistic T_n^* either belong to the interval $[8.1, 11.9]$, or show empirical levels very close to this interval. In general, some deterioration of the normal approximation appeared for large p_n , but the empirical results of T_n^* were always preferable to those of the test statistic T_n , at least in our experiments.

Under scenarios DPG_1 , DPG_3 and DPG_4 , Hosking's (1980) multivariate portmanteau test statistic offered severe overrejection for small values of $M = p_n$. In all these scenarios, at least one eigenvalue of the matrix Φ_l , $l \in \{1, 3, 4\}$, is large and close to one in module. These findings seem compatible with the results of the level study of Kheoh and McLeod (1992, p. 103) : in AR(1) models with large autoregressive parameter ($\phi = 0.9$), the Box-Ljung-Pierce test statistic offered some overrejection in their empirical experiments, even for large M . However, for moderate and large values of M , the empirical levels of Hosking's (1980) test statistic were close to the nominal levels, at least in our empirical experiments. For the cointegrated process DGP_6 , the critical points adjusted for the cointegrated rank determined by Brüggeman *et al.* (2006) provided an excellent approximation of the asymptotic chi-square distribution. This suggests that the critical points of Hosking's (1980) portmanteau test statistic are different for stationary and cointegrated models and that the critical values of Brüggeman *et al.* (2006) display reasonable precision for small and moderate values of M .

Compared to the multivariate Hosking portmanteau test statistic, the transformed test statistic T_n^* offered reasonable empirical levels in most situations, and when the empirical levels differed significantly from the nominal levels for these

test statistics, T_n^* was generally more conservative than P_M . As noted by Kheoh and McLeod (1992), conservative test statistics are often preferable, because, if the hypothesis of model adequacy is rejected at the 5% nominal level, say, then the correct significance level of a conservative test is less than 5%. By comparison, the exact significance level may be worse than 5% for a liberal test. Consequently, a liberal test may suggest the inclusion of additional parameters which could lead to poor forecasting performances. See also the discussion of Kheoh and McLeod (1992, Section 3) for more details.

We now turn to the study of the powers, which are presented in Tables 2.9-2.14. In general, for the transformed test statistics T_n^* , the BAR and DAN kernels behaved very similarly, and the truncated uniform kernel and Hosking's portmanteau test statistic were inferior. For these alternatives, weighting low order lags more heavily than high order lags delivers higher power, as expected from previous studies (see, e.g., Hong (1996), Duchesne and Roy (2004) and Duchesne (2005a)). Interestingly, for the transformed test statistic T_n^* , the results based on the empirical and asymptotic quantiles are seen to be reasonably comparable for all nominal levels, which is highly desirable in practical applications. Larger differences occurred for the original test statistic T_n , reflecting the fact that the nominal levels were less well controlled under the null hypothesis of adequacy.

Based on the asymptotic critical values, T_n^* appears more powerful than T_n . This seems to agree with the empirical findings of Chen and Deo (2004a). However, using the empirical critical values (where all the test statistics are size adjusted), it appears that T_n^* and T_n display essentially the same power under DGP_1 , DGP_4 , DGP_5 , and DGP_6 . Under DGP_2 , the untransformed test statistic T_n reaches a higher power than T_n^* , for a given kernel and value of the smoothing parameter p_n , particularly when p_n is small and moderate. Smaller differences in power occurred for $p_n = 12$ ($p_n = 15$) when $n = 100$ ($n = 200$). Under DGP_3 , the transformed test statistic T_n^* offered more power than T_n , and the power gains can be significant, particularly for small p_n and small n .

Summarizing, the results of our limited Monte Carlo study suggests that the new transformed test statistic T_n^* displays very satisfactory empirical levels, which

come very close to the nominal levels. The level study demonstrates that the empirical levels of the spectral test statistics are as satisfactory as those of Hosking's (1980) classical multivariate portmanteau test statistic. However, our empirical findings show that much higher power is attained by the flexible non-uniform weighting : in several practical applications, weighting low order lags more heavily may give more powerful test statistics than the weighting scheme offered by the truncated uniform kernel or by Hosking's (1980) test statistic. Interestingly, the power study suggests that the power transformation does not have any serious effect on the power performance of the spectral test statistics based on non-uniform kernels, since the empirical powers of T_n and T_n^* seem very comparable under several scenarios. Overall, for the alternatives considered, stationary or partially non-stationary time series models, the Chen-Deo power transformation delivered a test statistic which compared very favorably with the original test statistic T_n . In general, the empirical levels of T_n^* appeared preferable to those of T_n and T_n^* delivered high power, at least in our simulation experiments.

5. CONCLUDING REMARKS

In this paper, we have proposed corrected test statistics for checking the hypothesis of serial correlation in stationary and cointegrated time series models, relying on a power transformation. For stationary and partially non-stationary time series, the finite sample distribution of the kernel-based test statistics of Duchesne and Roy (2004) and Duchesne (2005a) can be heavily right skewed ; as a result, the exact empirical levels can be relatively far from the nominal levels. Using the Chen and Deo (2004a) method, corrected versions of the spectral test statistics have been proposed. Interestingly, the appropriate power transformation in the multivariate case appears to be the same as the one obtained for univariate time series : the transformation is invariant to the dimension of the time series. Additional finite sample mean and variance adjustments have been proposed, using the asymptotic distributions of the sample autocovariances for stationary and cointegrated time series models. Monte Carlo studies have been undertaken. We found that the new corrected versions have generally better empirical

levels than the untransformed test statistics and the gains can be substantial; these empirical properties are highly desirable in real applications. In terms of empirical powers, the corrected test statistics exhibited high power. In several circumstances, the empirical power appears similar to that of the untransformed test statistic. These empirical findings demonstrate that to transform Duchesne and Roy's (2004) test statistics using the Chen and Deo (2004a) method does not seriously affect the power performance of the spectral test statistics based on non-uniform kernels. In fact, for some alternatives, the corrected versions reached a power as high as and even higher than that of the original untransformed test statistics. Overall, the corrected test statistics should prove to be useful for diagnosing stationary and cointegrated time series models.

ACKNOWLEDGEMENTS

The second author acknowledges grants from National Science and Engineering Research Council of Canada and Fonds québécois de la recherche sur la nature et les technologies du Québec (Canada).

REFERENCES

- Ahn, S. K. and Reinsel, G. C. (1988), 'Nested reduced-rank autoregressive models for multiple time series', *Journal of the American Statistical Association* **83**, 849–856.
- Ahn, S. K. and Reinsel, G. C. (1990), 'Estimation for partially nonstationary multivariate autoregressive models', *Journal of the American Statistical Association* **85**, 813–823.
- Brockwell, P. J. and Davis, R. A. (1996), *Introduction to Time Series and Forecasting*, Springer-Verlag: New York.
- Brüggemann, R., Lütkepohl, H. and Saikkonen, P. (2006), 'Residual autocorrelation testing for vector error correction models', *Journal of Econometrics*, **134**, 579–604.
- Chen, W. W. and Deo, R. S. (2004a), 'Power transformations to induce normality and their applications', *Journal of the Royal Statistical Society, Ser. B* **66**, 117–130.
- Chen, W. W. and Deo, R. S. (2004b), 'A generalized portmanteau goodness-of-fit test for time series models', *Econometric Theory* **20**, 382–416.
- Chitturi, R. V. (1974), 'Distribution of residual autocorrelations in multiple autoregressive schemes', *Journal of the American Statistical Association* **69**, 928–934.
- Chitturi, R. V. (1976), 'Distribution of multivariate white noise autocorrelations', *Journal of the American Statistical Association* **71**, 223–226.
- Duchesne, P. (2005a), 'Testing for serial correlation of unknown form in cointegrated time series models', *Annals of the Institute of Statistical Mathematics* **57**, 575–595.
- Duchesne, P. (2005b), 'On the asymptotic distribution of residual autocovariances in VARX models with applications', *Test* **14**, 449–473.
- Duchesne, P. and Roy, R. (2004), 'On consistent testing for serial correlation of unknown form in vector time series models', *Journal of Multivariate Analysis* **89**, 148–180.
- Harville, D. A. (1997), *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag: Berlin.
- Hong, Y. (1996), 'Consistent testing for serial correlation of unknown form', *Econometrica* **64**, 837–864.

- Hosking, J. (1980), 'The multivariate portmanteau statistic', *Journal of the American Statistical Association* **75**, 602–608.
- Hosking, J. (1981a), 'Equivalent forms of the multivariate portmanteau statistic', *Journal of the Royal Statistical Society, Ser. B* **43**, 261–262, (Corr: V51/1989, p. 303).
- Hosking, J. (1981b), 'Lagrange-multiplier tests of multivariate time series models', *Journal of the Royal Statistical Society, Ser. B* **43**, 219–230.
- Johansen, S. (1995), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press: Oxford.
- Kheoh, T. S. and McLeod, A. I. (1992), 'Comparison of two modified portmanteau tests for model adequacy', *Computational Statistics & Data Analysis* **14**, 99–106.
- Kwan, A. C. C., Sim, A.-B. and Wu, Y. (2005), 'A comparative study of the finite-sample performance of some portmanteau test for randomness of a time series', *Computational Statistics & Data Analysis* **48**, 391–413.
- Li, W. K. (2004), *Diagnostic Checks in Time Series*, Chapman & Hall/CRC: New York.
- Li, W. K. and McLeod, A. I. (1981), 'Distribution of the residual autocorrelations in multivariate ARMA time series models', *Journal of the Royal Statistical Society, Ser. B* **43**, 231–239.
- Lütkepohl, H. (1993), *Introduction to Multiple Time Series Analysis*, Second Edition, Springer-Verlag: Berlin.
- Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer: Berlin.
- Paparoditis, E. (2005), 'Testing the fit of a vector autoregressive moving average model', *Journal of Time Series Analysis* **26**, 543–568.
- Priestley, M. B. (1981), *Spectral Analysis and Time Series*, Vol. 1: *Univariate Series*, Academic Press: New York.
- Reinsel, G. C. (1993), *Elements of Multivariate Time Series Analysis*, Springer-Verlag: New York.
- Reinsel, G. C. and Ahn, S. K. (1992), 'Vector AR models with unit roots and reduced rank structure: estimation, likelihood ratio test, and forecasting', *Journal of Time Series Analysis* **13**, 353–375.
- Searle, S. R. (1971), *Linear Models*, Wiley: New York.

Yap, S. F. and Reinsel, G. C. (1995), 'Estimation and testing for unit roots in a partially nonstationary vector autoregressive moving average model', *Journal of the American Statistical Association* **90**, 253–267.

Wei, W. W.-S. (1990), *Time Serie Analysis, Univariate and Multivariate Methods*, Addison-Wesley Publishing Company: Redwood City.

TAB. 2.4. Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP_2 .

		$\alpha = 0.01$							$\alpha = 0.05$							$\alpha = 0.10$												
n		T_n			T_n^*				T_n			T_n^*				T_n			T_n^*									
=50	p_n	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR
	2	0.0	0.0	0.2	1.2	1.1	0.7	1.1	0.0	0.0	0.6	5.3	4.4	5.3	6.4	0.0	0.0	1.2	11.3	8.8	10.5	12.0						
	3	0.0	0.0	0.4	0.6	0.6	1.1	1.5	0.1	0.1	1.5	5.0	4.7	5.7	6.8	0.2	0.3	2.7	9.6	8.1	9.5	11.5						
	4	0.1	0.1	0.6	0.5	0.5	0.8	1.3	0.2	0.3	2.1	5.4	4.6	4.2	6.6	0.4	0.3	3.8	9.7	8.8	9.1	11.4						
	8	0.2	0.4	0.6	0.6	0.7	0.4	1.2	0.6	0.9	2.4	4.7	3.8	2.1	5.1	1.1	2.2	3.8	9.2	7.8	4.2	10.0						
	10	0.3	0.4	0.7	0.6	0.5	0.3	1.1	0.7	1.3	2.7	4.3	3.0	1.4	4.9	1.9	2.6	4.7	8.4	6.2	3.3	9.5						
n		T_n			T_n^*				T_n			T_n^*				T_n			T_n^*									
=100	p_n	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR
	2	0.0	0.0	0.2	1.2	1.1	0.9	1.0	0.1	0.1	0.6	5.1	4.5	5.0	5.6	0.2	0.2	1.1	11.4	9.5	10.1	10.6						
	3	0.0	0.0	0.5	0.7	0.5	0.9	1.2	0.2	0.2	1.3	4.4	4.3	4.1	4.4	0.3	0.3	1.8	11.0	9.8	9.4	10.1						
	5	0.2	0.1	0.4	0.6	0.5	0.8	1.2	0.3	0.3	1.6	4.7	4.5	3.9	4.7	0.4	0.5	2.6	9.4	8.5	7.0	9.8						
	9	0.2	0.1	0.4	0.8	0.7	0.4	0.6	0.6	0.9	1.9	3.7	3.2	2.0	3.6	1.5	1.8	3.2	7.8	6.6	5.4	8.9						
	12	0.0	0.3	0.4	0.9	0.8	0.2	0.8	1.0	1.4	1.9	3.6	2.7	1.5	4.1	1.8	1.8	4.0	6.6	5.5	4.2	8.1						
n		T_n			T_n^*				T_n			T_n^*				T_n			T_n^*									
=200	p_n	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR	BARD	DAN	TR
	2	0.0	0.0	0.0	0.9	1.3	1.1	1.1	0.0	0.0	0.6	6.0	5.6	5.1	5.3	0.0	0.0	1.2	10.0	9.3	11.6	12.0						
	3	0.0	0.0	0.1	0.9	0.9	0.7	0.7	0.0	0.0	0.7	5.7	5.4	4.3	4.6	0.3	0.1	1.5	10.0	9.2	9.1	9.7						
	5	0.0	0.0	0.4	0.8	0.7	0.6	0.6	0.0	0.1	1.2	5.5	5.3	3.3	3.8	0.4	0.4	2.6	9.9	9.1	7.3	8.4						
	10	0.0	0.0	0.5	0.3	0.3	0.5	0.9	0.1	0.4	1.7	4.4	3.9	2.7	3.6	1.3	2.1	3.3	8.7	6.9	6.0	8.1						
	15	0.0	0.3	0.9	0.6	0.6	0.5	0.9	0.6	1.2	2.0	3.7	3.2	2.0	3.0	1.8	2.4	3.4	7.4	6.2	4.8	7.7						

TAB. 2.5. Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP₃.

		$\alpha = 0.01$									$\alpha = 0.05$									$\alpha = 0.10$								
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M	T_n			T_n^*			P_M						
=50	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR							
	2	0.1	0.2	0.5	1.0	0.7	1.1	2.5	1.2	1.1	1.9	5.4	4.7	5.5	13.1	2.5	2.6	2.7	10.7	10.3	10.5	22.4						
	3	0.1	0.1	0.9	0.4	0.2	1.2	2.0	1.3	1.3	2.0	5.3	5.0	5.3	9.1	3.0	3.2	3.7	10.5	9.5	10.7	16.7						
	4	0.3	0.3	1.1	0.8	0.6	1.5	2.5	1.3	1.6	3.0	5.1	4.5	5.8	8.8	3.0	2.8	5.2	10.2	9.2	11.2	15.4						
	8	0.6	0.8	1.3	0.9	0.7	0.7	1.7	2.2	2.3	3.5	5.7	4.7	3.7	7.4	3.4	4.1	6.4	10.8	9.6	7.4	13.2						
	10	0.7	0.6	1.4	0.7	0.6	0.6	1.9	2.5	3.0	3.2	5.4	4.3	2.5	7.0	4.5	4.6	6.1	10.5	8.6	4.9	11.9						
=100	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR							
	2	0.4	0.4	0.4	1.1	1.1	1.0	2.4	1.4	1.4	1.4	5.2	4.5	5.2	11.4	1.9	1.8	2.8	10.9	10.7	9.8	22.4						
	3	0.3	0.3	0.4	1.1	1.0	0.8	1.6	1.4	1.4	1.6	4.7	4.2	5.1	8.1	1.8	1.7	2.7	9.6	8.9	10.4	16.5						
	5	0.5	0.5	1.1	0.7	0.6	1.2	1.7	1.0	0.9	2.6	4.9	4.3	4.6	6.4	2.1	2.6	4.2	9.6	8.9	10.0	13.6						
	9	0.6	0.6	0.7	0.8	0.8	0.6	1.1	1.1	1.9	2.7	4.9	3.7	3.7	6.1	3.0	3.1	5.4	9.0	8.3	8.0	11.3						
	12	0.7	0.7	0.6	0.9	1.0	0.5	1.2	2.0	2.1	2.8	4.0	3.6	2.4	5.5	3.4	3.7	5.4	9.7	7.7	6.0	11.0						
=200	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR							
	2	0.3	0.3	0.7	0.5	0.4	1.0	2.0	0.6	0.7	1.2	4.1	3.7	3.8	8.8	1.4	1.3	2.1	9.3	8.7	7.5	19.4						
	3	0.4	0.5	0.3	0.5	0.5	1.0	1.3	0.6	0.6	1.3	3.3	3.4	4.0	5.8	1.4	1.3	2.3	8.7	8.7	8.1	13.2						
	5	0.5	0.5	0.4	0.7	0.7	0.8	1.0	0.8	0.9	1.3	3.4	3.1	3.4	4.1	1.7	1.5	2.4	7.2	7.5	8.1	10.1						
	10	0.4	0.3	0.8	0.5	0.7	0.9	1.4	0.9	1.0	2.3	3.1	2.6	3.2	4.2	1.9	2.2	3.7	8.0	7.1	6.5	8.3						
	15	0.4	0.5	0.9	0.6	0.6	0.6	1.3	1.0	1.7	2.0	3.2	2.4	2.2	3.9	1.9	2.8	4.1	7.4	6.0	4.8	7.7						

TAB. 2.6. Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP_4 .

		$\alpha = 0.01$						$\alpha = 0.05$						$\alpha = 0.10$								
n		T_n		T_n^*		P_M		T_n		T_n^*		P_M		T_n		T_n^*		P_M				
=50	p_n	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	
	2	0.3	0.3	0.4	1.3	1.2	0.5	2.6	1.0	0.9	1.5	6.6	5.4	5.5	13.4	2.1	1.9	2.6	11.6	10.0	12.0	23.9
	3	0.4	0.4	0.8	0.9	0.8	0.9	1.8	0.8	0.8	1.9	5.6	5.1	5.1	8.8	1.8	2.1	4.1	11.8	10.7	10.7	17.2
	4	0.4	0.4	0.9	0.8	0.6	1.1	2.0	1.0	1.2	2.7	4.8	4.4	5.1	8.1	1.9	1.9	4.3	11.6	10.4	10.7	15.3
	8	0.5	0.5	1.6	0.7	0.7	0.8	1.8	1.7	2.2	3.2	5.1	4.7	3.2	6.4	3.4	3.9	5.6	10.2	8.1	6.9	11.7
	10	0.5	0.7	1.6	0.7	0.6	0.6	2.3	2.3	2.5	3.7	5.1	4.2	2.9	6.3	3.9	4.5	5.8	9.9	7.4	5.2	11.9
=100	p_n	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	
	2	0.6	0.5	0.9	1.4	1.3	1.2	3.0	1.3	1.3	2.2	6.0	5.5	6.5	12.9	1.9	1.9	3.0	11.2	10.7	10.9	22.5
	3	0.6	0.5	0.8	1.1	0.9	1.4	1.7	1.3	1.3	1.8	5.7	5.2	5.2	9.7	2.3	2.4	3.5	11.7	11.4	10.7	18.4
	5	0.6	0.6	0.8	0.9	0.8	1.3	2.1	1.2	1.5	2.5	6.2	5.9	4.5	6.3	2.7	2.5	4.2	11.2	10.5	9.3	13.7
	9	0.8	0.9	1.0	1.3	1.2	0.9	1.5	1.5	2.2	2.6	5.1	4.3	3.2	5.0	2.8	3.2	4.4	9.9	8.8	6.9	10.3
	12	1.0	1.2	0.9	1.4	1.3	0.7	1.3	1.8	1.8	3.1	4.0	3.3	3.0	6.1	3.4	3.3	6.2	9.1	7.7	6.6	10.3
=200	p_n	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	
	2	0.4	0.3	0.5	1.2	1.1	1.1	2.8	1.2	1.1	1.6	4.4	4.6	5.2	11.2	1.8	1.8	3.3	9.9	9.7	8.4	20.3
	3	0.3	0.4	0.6	1.2	1.1	0.7	1.6	1.5	1.3	1.6	3.9	3.6	4.0	7.1	1.9	1.9	3.2	9.4	8.6	8.0	14.1
	5	0.5	0.4	0.6	0.9	0.8	0.8	1.1	1.2	1.2	2.0	4.1	4.2	4.1	5.6	2.0	2.1	3.6	9.1	8.6	8.0	11.0
	10	0.5	0.4	0.9	0.6	0.6	0.9	1.1	1.2	1.7	2.4	4.3	3.9	3.6	4.8	2.7	3.0	4.4	8.2	7.1	7.1	9.3
	15	0.5	0.6	1.1	0.9	0.9	1.0	1.3	1.7	2.1	2.2	3.8	3.5	2.6	4.0	2.7	3.4	4.1	8.0	6.3	6.0	8.9

TAB. 2.7. Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP_5 .

		$\alpha = 0.01$								$\alpha = 0.05$								$\alpha = 0.10$																																
n		T_n				T_n^*				P_M				T_n				T_n^*				P_M				T_n				T_n^*				P_M																
=50	p_n	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR									
	2	0.0	0.0	0.0	1.5	1.2	0.2	0.7	0.0	0.0	0.2	5.8	4.8	4.4	5.8	0.1	0.1	0.7	12.2	9.6	10.0	13.7	0.0	0.0	2.3	10.1	8.8	10.1	11.7	0.2	0.2	3.3	10.2	8.5	9.0	11.9	1.2	1.6	3.6	9.1	7.2	4.6	11.5	1.4	2.0	3.7	8.7	6.6	3.0	10.0
	3	0.0	0.0	0.5	0.6	0.4	0.7	1.1	0.0	0.0	1.1	3.7	3.7	4.5	5.5	0.0	0.0	2.3	10.1	8.8	10.1	11.7	0.0	0.0	0.5	0.6	0.4	0.7	1.1	0.0	0.0	1.1	3.7	3.7	4.5	5.5	0.0	0.0	2.3	10.1	8.8	10.1	11.7							
	4	0.0	0.0	0.6	0.2	0.2	1.0	1.6	0.0	0.1	1.8	3.7	3.3	4.5	5.9	0.2	0.2	3.3	10.2	8.5	9.0	11.9	0.0	0.0	0.6	0.2	0.2	1.0	1.6	0.0	0.1	1.8	3.7	3.3	4.5	5.9	0.2	0.2	3.3	10.2	8.5	9.0	11.9							
	8	0.1	0.2	0.9	0.7	0.6	0.7	1.2	0.6	1.1	1.5	4.2	3.5	1.6	4.9	1.2	1.6	3.6	9.1	7.2	4.6	11.5	0.1	0.2	0.9	0.7	0.6	0.7	1.2	0.6	1.1	1.5	4.2	3.5	1.6	4.9	1.2	1.6	3.6	9.1	7.2	4.6	11.5							
	10	0.3	0.4	0.8	0.8	0.6	0.4	1.2	0.9	1.3	1.8	3.9	2.7	1.6	4.7	1.4	2.0	3.7	8.7	6.6	3.0	10.0	0.3	0.4	0.8	0.8	0.6	0.4	1.2	0.9	1.3	1.8	3.9	2.7	1.6	4.7	1.4	2.0	3.7	8.7	6.6	3.0	10.0							
=100	p_n	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR									
	2	0.0	0.0	0.1	0.9	1.0	1.2	1.2	0.0	0.0	0.8	6.0	4.7	5.1	6.4	0.2	0.2	1.3	12.1	10.4	9.5	12.7	0.0	0.0	0.1	0.9	1.0	1.2	1.2	0.0	0.0	0.8	6.0	4.7	5.1	6.4	0.2	0.2	1.3	12.1	10.4	9.5	12.7							
	3	0.0	0.0	0.3	0.9	0.9	1.0	1.0	0.1	0.1	1.0	4.4	4.2	3.6	4.8	0.2	0.3	1.7	10.6	9.6	10.3	11.5	0.0	0.0	0.3	0.9	0.9	1.0	1.0	0.1	0.1	1.0	4.4	4.2	3.6	4.8	0.2	0.3	1.7	10.6	9.6	10.3	11.5							
	5	0.0	0.1	0.5	0.8	0.6	0.6	0.8	0.4	0.4	1.3	4.2	3.8	3.4	4.3	0.8	1.0	2.9	9.2	8.0	7.9	9.7	0.0	0.1	0.5	0.8	0.6	0.6	0.8	0.4	0.4	1.3	4.2	3.8	3.4	4.3	0.8	1.0	2.9	9.2	8.0	7.9	9.7							
	9	0.3	0.3	0.7	0.7	0.6	0.6	1.1	0.6	0.9	1.6	3.2	2.9	2.5	4.8	1.5	1.8	4.1	8.8	7.3	5.9	9.3	0.3	0.3	0.7	0.7	0.6	0.6	1.1	0.6	0.9	1.6	3.2	2.9	2.5	4.8	1.5	1.8	4.1	8.8	7.3	5.9	9.3							
	12	0.3	0.3	0.6	0.7	0.6	0.5	0.8	1.0	1.1	2.1	2.8	2.3	2.0	4.2	1.6	1.8	4.2	8.2	6.2	4.4	9.2	0.3	0.3	0.6	0.7	0.6	0.5	0.8	1.0	1.1	2.1	2.8	2.3	2.0	4.2	1.6	1.8	4.2	8.2	6.2	4.4	9.2							
=200	p_n	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR	BARDAN	TR									
	2	0.0	0.0	0.1	0.4	0.4	1.0	1.2	0.0	0.0	0.8	4.5	4.1	3.6	4.6	0.0	0.1	1.3	8.9	7.5	8.0	10.3	0.0	0.0	0.1	0.4	0.4	1.0	1.2	0.0	0.0	0.8	4.5	4.1	3.6	4.6	0.0	0.1	1.3	8.9	7.5	8.0	10.3							
	3	0.0	0.0	0.1	0.6	0.5	0.5	0.5	0.0	0.0	0.5	3.9	3.8	3.5	3.7	0.3	0.2	1.2	7.7	7.1	7.6	8.8	0.0	0.0	0.1	0.6	0.5	0.5	0.5	0.0	0.0	0.5	3.9	3.8	3.5	3.7	0.3	0.2	1.2	7.7	7.1	7.6	8.8							
	5	0.0	0.0	0.4	0.8	0.6	0.6	0.7	0.1	0.1	1.2	3.2	3.0	3.3	3.7	0.6	0.7	2.1	7.5	7.2	7.3	8.3	0.0	0.0	0.4	0.8	0.6	0.6	0.7	0.1	0.1	1.2	3.2	3.0	3.3	3.7	0.6	0.7	2.1	7.5	7.2	7.3	8.3							
	10	0.1	0.2	0.8	0.5	0.6	1.0	1.4	0.5	0.6	1.8	3.0	2.5	3.2	4.3	1.2	1.4	4.2	7.1	6.8	7.0	8.1	0.1	0.2	0.8	0.5	0.6	1.0	1.4	0.5	0.6	1.8	3.0	2.5	3.2	4.3	1.2	1.4	4.2	7.1	6.8	7.0	8.1							
	15	0.4	0.4	1.0	0.6	0.5	0.7	1.1	0.7	1.1	2.0	2.9	2.9	2.1	3.7	1.4	2.1	4.0	7.1	5.6	5.2	7.4	0.4	0.4	1.0	0.6	0.5	0.7	1.1	0.7	1.1	2.0	2.9	2.9	2.1	3.7	1.4	2.1	4.0	7.1	5.6	5.2	7.4							

TAB. 2.8. Empirical levels (in percentage) of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VAR(1) process defined by (2.0.17), using the parameters values corresponding to DGP_6 .

		$\alpha = 0.01$									$\alpha = 0.05$									$\alpha = 0.10$																								
n		T_n			T_n^*			P_M			T_n			T_n^*			P_M			T_n			T_n^*			P_M																		
=50	p_n	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR				
	2	0.9	0.8	0.5	1.1	0.9	0.5	0.5	1.5	1.8	2.3	4.8	4.0	4.7	5.4	3.0	2.9	4.7	8.6	8.3	8.9	10.6																						
	3	0.8	0.7	0.8	1.0	0.9	0.7	0.8	1.4	1.6	2.3	4.1	4.0	3.5	4.5	2.5	2.3	4.5	8.8	8.0	8.5	10.0																						
	4	0.5	0.5	1.0	0.9	0.6	0.6	1.1	1.4	1.4	3.8	4.0	3.2	4.2	5.2	2.2	2.5	5.5	9.2	8.4	7.3	10.3																						
	8	0.8	0.9	1.7	0.8	0.5	0.6	1.4	2.0	2.8	3.9	3.5	3.2	3.1	5.6	3.9	4.4	7.3	8.2	6.6	5.2	9.9																						
	10	0.9	0.9	1.8	0.7	0.6	0.4	1.5	2.5	3.3	4.5	3.7	3.2	2.0	5.5	4.7	5.8	7.3	7.8	6.4	4.2	10.0																						
=100	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR			
	2	0.6	0.6	0.8	1.3	1.3	0.9	1.1	1.7	1.7	2.4	5.0	4.9	4.8	5.3	2.8	3.1	4.3	9.1	8.5	10.1	10.8																						
	3	0.8	0.8	0.8	1.0	0.9	0.7	0.8	1.8	1.8	2.6	4.8	4.7	4.7	5.2	3.1	3.1	5.2	9.3	8.9	9.3	10.1																						
	5	0.5	0.6	1.0	0.9	0.7	0.8	0.9	1.6	1.9	3.4	5.0	4.5	4.0	4.9	3.5	3.5	5.8	9.1	8.5	7.5	9.2																						
	9	0.8	0.9	0.9	1.0	0.7	0.6	0.9	2.0	2.3	3.3	4.6	3.5	2.7	4.2	4.3	5.1	6.1	9.0	7.9	6.2	10.1																						
	12	0.8	0.8	1.0	0.8	0.5	0.5	0.8	2.1	2.5	4.2	3.7	2.7	2.1	5.2	4.2	5.2	7.4	7.6	6.4	5.6	9.6																						
=200	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR
	2	0.4	0.4	0.9	0.7	0.7	0.9	0.9	1.1	1.0	2.1	3.6	3.3	4.2	4.2	1.4	1.6	3.4	7.6	7.0	8.2	8.5																						
	3	0.5	0.6	0.9	0.6	0.6	0.9	0.9	1.0	1.0	2.0	3.5	3.4	3.4	3.8	1.7	1.6	3.8	8.1	7.9	7.6	7.9																						
	5	0.6	0.5	0.9	0.8	0.7	0.9	0.9	1.5	1.5	2.5	3.5	3.9	3.2	3.5	2.0	2.5	4.4	8.3	7.8	6.7	7.1																						
	10	0.7	0.6	1.2	0.7	0.5	1.1	1.2	1.8	2.1	3.5	3.8	3.3	3.5	4.4	3.6	3.9	5.8	7.3	6.2	6.4	8.3																						
	15	0.6	0.7	1.1	0.6	0.5	0.3	1.0	2.5	2.9	2.9	3.6	3.3	2.4	4.0	3.8	4.2	6.1	6.7	5.5	5.1	8.3																						

TAB. 2.9. Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP_1 .

		$\alpha = 0.05$							$\alpha = 0.10$						
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
=100	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	25.6	25.4	16.7	32.9	31.0	23.6	51.8	33.7	33.7	23.6	47.7	46.5	37.4	68.2
		(33.4)	(34.2)	(20.4)	(29.7)	(31.1)	(20.0)	(20.3)	(47.0)	(46.4)	(33.5)	(48.1)	(47.2)	(34.9)	(33.5)
	3	24.5	24.2	13.4	31.7	30.9	21.2	35.7	33.0	32.9	20.5	45.1	43.8	33.5	50.5
		(32.2)	(32.0)	(18.7)	(31.6)	(31.9)	(18.2)	(18.1)	(42.9)	(43.7)	(30.9)	(43.6)	(43.7)	(31.8)	(30.4)
	5	21.0	19.3	10.1	28.9	27.1	14.9	23.5	29.1	28.0	16.0	42.0	39.5	25.9	35.6
		(25.6)	(24.6)	(15.6)	(26.2)	(26.3)	(16.3)	(15.0)	(41.3)	(39.8)	(28.4)	(39.8)	(37.6)	(28.5)	(27.9)
	9	16.3	15.1	9.5	23.5	19.9	10.5	16.1	23.5	21.7	15.0	37.0	30.8	18.3	25.1
		(24.0)	(21.8)	(13.0)	(25.3)	(21.4)	(13.1)	(12.9)	(36.2)	(34.5)	(21.9)	(37.5)	(33.7)	(21.6)	(21.2)
	12	15.2	13.0	10.5	20.0	15.8	9.7	14.5	21.7	18.7	14.3	31.6	24.4	14.5	22.6
		(23.6)	(19.5)	(14.2)	(22.3)	(18.9)	(14.5)	(14.2)	(33.5)	(30.1)	(21.8)	(34.1)	(31.4)	(22.0)	(21.3)
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
=200	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	59.8	59.8	44.8	67.3	66.4	56.1	80.3	68.4	67.8	54.4	80.1	79.2	68.8	89.5
		(64.4)	(65.0)	(53.2)	(65.3)	(66.5)	(53.4)	(53.2)	(79.6)	(79.9)	(67.3)	(81.3)	(81.1)	(68.8)	(67.1)
	3	58.5	58.3	34.7	67.2	66.5	47.4	63.5	67.2	67.3	45.5	79.0	78.3	61.8	76.7
		(65.5)	(66.7)	(49.5)	(65.1)	(66.5)	(51.7)	(49.5)	(80.2)	(80.9)	(62.5)	(79.7)	(80.9)	(62.6)	(62.5)
	5	51.5	49.2	29.7	63.4	60.5	39.5	47.2	61.8	60.5	39.3	75.7	73.9	51.9	59.7
		(64.0)	(62.5)	(45.3)	(63.8)	(60.9)	(46.1)	(44.5)	(76.1)	(73.9)	(55.1)	(76.3)	(74.8)	(55.5)	(54.7)
	10	42.6	38.7	22.5	53.5	47.7	27.2	32.6	52.1	48.2	31.9	66.2	60.8	41.2	45.9
		(55.7)	(52.6)	(32.6)	(56.3)	(53.5)	(33.4)	(31.3)	(68.5)	(64.3)	(46.0)	(68.5)	(64.1)	(46.7)	(45.6)
	15	37.0	32.4	20.5	46.9	38.7	20.6	28.3	46.1	41.8	29.5	59.7	51.5	31.9	38.3
		(53.2)	(47.1)	(30.4)	(52.9)	(48.4)	(31.3)	(29.8)	(62.6)	(56.6)	(40.7)	(62.8)	(57.3)	(40.9)	(39.8)

TAB. 2.10. Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP₂.

		$\alpha = 0.05$							$\alpha = 0.10$						
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
=100	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	9.7 (55.3)	10.4 (57.9)	11.5 (36.1)	37.2 (37.2)	37.4 (38.8)	26.6 (26.6)	38.4 (35.8)	14.5 (64.6)	15.2 (70.0)	16.3 (53.1)	50.1 (47.0)	50.4 (51.6)	41.1 (40.9)	54.0 (52.9)
	3	11.7 (53.6)	11.6 (53.6)	15.3 (43.5)	36.9 (38.7)	35.9 (38.5)	30.4 (34.3)	41.2 (43.5)	15.7 (63.2)	15.4 (63.5)	23.1 (56.5)	50.4 (48.9)	48.7 (50.1)	45.2 (46.9)	56.9 (56.0)
	5	12.6 (49.0)	12.6 (47.6)	12.7 (29.4)	35.8 (36.7)	33.3 (35.4)	18.4 (24.8)	27.2 (27.9)	17.4 (64.3)	18.1 (63.2)	19.1 (40.6)	49.5 (52.2)	46.4 (51.5)	31.2 (35.6)	40.0 (40.2)
	9	13.0 (43.4)	13.4 (40.8)	10.4 (22.0)	29.3 (34.6)	24.7 (32.9)	10.3 (18.3)	17.8 (20.8)	21.1 (60.7)	21.4 (57.7)	16.6 (33.1)	43.0 (49.5)	37.4 (48.2)	19.1 (29.5)	30.2 (32.3)
	12	12.8 (40.3)	13.0 (34.9)	9.9 (17.7)	25.7 (33.0)	19.2 (28.3)	8.6 (15.4)	14.2 (16.8)	20.7 (56.7)	21.0 (52.1)	14.7 (30.3)	36.8 (47.0)	30.6 (43.5)	12.9 (27.8)	26.5 (29.4)
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
=200	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	36.3 (83.4)	38.1 (88.6)	42.6 (75.2)	71.7 (67.8)	75.3 (73.6)	65.9 (65.7)	75.5 (75.2)	43.5 (91.1)	45.6 (94.4)	50.7 (84.4)	82.0 (82.1)	84.6 (86.2)	78.4 (74.2)	86.9 (84.2)
	3	40.5 (83.2)	40.8 (83.2)	54.3 (83.0)	72.9 (71.2)	72.8 (70.9)	74.3 (77.8)	81.5 (82.8)	47.4 (91.0)	47.2 (90.3)	66.2 (89.0)	83.5 (83.5)	81.9 (83.1)	84.7 (85.3)	88.6 (89.0)
	5	48.0 (84.3)	48.9 (83.4)	42.5 (68.3)	77.6 (76.2)	76.0 (74.7)	56.2 (63.0)	64.3 (68.2)	57.4 (93.6)	57.9 (93.8)	55.3 (79.8)	86.9 (86.9)	85.0 (86.0)	70.6 (75.6)	77.1 (79.8)
	10	51.2 (80.8)	49.6 (78.9)	31.2 (49.3)	71.7 (74.5)	66.1 (72.3)	34.2 (44.3)	42.1 (47.0)	60.3 (89.6)	60.2 (87.6)	41.3 (62.5)	82.6 (84.5)	78.8 (82.4)	48.0 (58.2)	57.4 (61.7)
	15	46.1 (75.1)	43.2 (69.0)	26.0 (41.7)	62.0 (67.1)	52.3 (63.3)	24.7 (38.6)	35.6 (40.3)	57.5 (85.1)	54.8 (81.3)	37.8 (55.8)	75.8 (80.7)	65.2 (76.2)	37.7 (52.6)	48.8 (54.5)

TAB. 2.11. Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP_3 .

		$\alpha = 0.05$							$\alpha = 0.10$						
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
=100	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	15.6	15.3	11.0	37.3	35.9	23.9	38.6	22.1	22.5	15.8	51.5	50.7	36.1	54.9
		(33.7)	(33.6)	(23.4)	(37.2)	(37.4)	(23.9)	(23.5)	(45.7)	(45.7)	(35.2)	(49.8)	(50.0)	(36.2)	(35.0)
	3	15.4	15.4	10.3	34.0	33.0	20.0	26.8	20.0	20.4	15.3	48.7	47.1	33.3	40.5
		(33.8)	(33.8)	(20.4)	(35.7)	(34.5)	(19.8)	(20.2)	(44.1)	(44.6)	(29.5)	(49.0)	(49.0)	(32.7)	(29.5)
	5	12.6	12.4	8.5	29.7	27.9	15.2	18.2	17.8	17.3	12.7	42.7	39.6	25.7	28.4
		(29.7)	(29.0)	(15.9)	(30.0)	(29.3)	(16.5)	(15.7)	(42.2)	(40.4)	(25.1)	(44.1)	(42.9)	(25.7)	(24.5)
	9	11.1	10.7	8.8	23.0	17.4	11.1	14.9	15.3	14.3	13.7	36.6	30.7	18.2	22.6
		(21.7)	(20.1)	(11.8)	(23.3)	(20.2)	(13.0)	(12.4)	(36.3)	(34.1)	(20.2)	(39.1)	(33.9)	(21.0)	(19.6)
	12	10.7	9.2	8.8	18.8	14.2	8.4	13.7	14.1	14.0	13.4	32.7	25.5	14.8	21.4
		(19.6)	(17.1)	(12.9)	(21.3)	(20.0)	(12.7)	(12.7)	(34.2)	(29.4)	(20.5)	(32.9)	(29.1)	(21.1)	(19.7)
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
=200	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	44.3	45.0	34.5	72.7	71.3	54.6	69.0	55.1	55.7	43.9	83.7	83.2	67.1	81.0
		(72.3)	(72.1)	(58.3)	(75.0)	(74.1)	(61.2)	(58.2)	(82.2)	(82.1)	(71.7)	(85.6)	(85.0)	(73.0)	(71.6)
	3	44.1	44.3	30.4	69.5	68.3	47.8	57.0	53.5	54.1	39.6	80.8	80.4	62.8	69.2
		(72.3)	(71.9)	(50.7)	(75.5)	(74.9)	(53.7)	(50.5)	(81.4)	(81.5)	(65.2)	(84.3)	(83.5)	(67.7)	(65.1)
	5	40.5	39.3	27.4	62.8	60.6	43.3	44.4	50.2	49.2	36.9	75.8	72.7	55.7	57.8
		(68.0)	(66.0)	(46.8)	(69.7)	(67.4)	(49.1)	(46.5)	(78.3)	(77.6)	(57.9)	(80.3)	(78.6)	(59.9)	(57.7)
	10	36.0	33.8	21.9	55.3	47.6	28.3	31.0	44.3	41.8	30.7	67.9	61.8	42.2	45.3
		(59.6)	(54.0)	(34.1)	(62.0)	(56.5)	(35.8)	(33.6)	(70.6)	(66.3)	(49.4)	(71.0)	(67.9)	(50.9)	(48.8)
	15	32.7	29.3	18.8	47.8	40.5	20.9	26.4	41.5	38.6	28.2	61.6	53.2	32.9	38.4
		(52.4)	(48.5)	(32.6)	(54.4)	(51.1)	(33.6)	(32.3)	(64.5)	(60.2)	(44.0)	(65.9)	(61.4)	(45.9)	(42.9)

TAB. 2.12. Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP_4 .

		$\alpha = 0.05$						$\alpha = 0.10$							
n		T_n			T_n^*				T_n			T_n^*			P_M
=100	p_n	BAR	DAN	TR	BAR	DAN	TR	P_M	BAR	DAN	TR	BAR	DAN	TR	P_M
	2	19.8	20.0	18.0	50.9	49.8	34.2	47.3	26.7	26.5	25.3	64.7	62.8	47.0	63.6
		(43.3)	(43.0)	(28.9)	(48.3)	(48.8)	(31.8)	(28.7)	(59.6)	(59.3)	(41.8)	(62.6)	(62.0)	(44.9)	(41.6)
	3	21.0	20.9	15.9	48.2	46.7	29.2	36.3	27.1	26.8	22.0	61.2	59.7	43.0	51.0
		(44.0)	(44.2)	(27.1)	(46.0)	(46.2)	(28.8)	(27.0)	(54.3)	(54.6)	(37.3)	(59.3)	(58.4)	(40.7)	(37.3)
	5	20.5	20.1	11.6	41.4	37.0	19.9	22.7	26.3	26.2	17.2	56.3	53.2	34.5	37.5
		(35.5)	(33.0)	(19.6)	(37.6)	(34.9)	(20.8)	(19.5)	(49.0)	(47.6)	(32.0)	(52.1)	(51.4)	(36.7)	(31.6)
	9	16.0	15.2	10.6	33.1	27.8	14.0	17.3	24.0	23.4	15.8	47.2	40.1	24.0	28.5
		(30.5)	(27.8)	(18.0)	(31.3)	(30.3)	(20.7)	(17.4)	(42.3)	(38.2)	(28.9)	(47.4)	(42.7)	(30.0)	(28.3)
	12	14.8	13.3	10.9	28.5	21.5	11.0	15.6	23.2	20.8	16.1	42.0	33.8	18.4	23.7
		(28.6)	(25.6)	(14.5)	(30.0)	(28.5)	(15.4)	(13.8)	(40.2)	(37.6)	(24.1)	(43.8)	(39.7)	(25.2)	(23.5)
n		T_n			T_n^*				T_n			T_n^*			P_M
=200	p_n	BAR	DAN	TR	BAR	DAN	TR	P_M	BAR	DAN	TR	BAR	DAN	TR	P_M
	2	60.3	60.1	53.0	85.4	85.4	70.7	81.0	69.9	69.9	60.4	91.7	91.6	80.3	89.7
		(84.1)	(85.1)	(69.4)	(86.1)	(87.4)	(70.4)	(69.2)	(90.8)	(90.6)	(80.1)	(91.7)	(91.9)	(81.7)	(80.1)
	3	62.1	61.5	45.3	83.3	83.2	62.9	70.3	69.4	69.8	55.4	89.6	89.2	74.4	80.7
		(84.9)	(84.7)	(64.5)	(85.6)	(85.4)	(67.3)	(64.4)	(89.4)	(88.5)	(75.1)	(90.2)	(90.2)	(77.9)	(75.0)
	5	60.0	58.3	39.2	78.2	75.9	52.0	55.2	67.7	66.0	48.3	86.8	84.5	64.3	66.9
		(78.1)	(75.6)	(54.1)	(79.9)	(77.5)	(56.7)	(53.5)	(86.5)	(84.3)	(65.9)	(88.1)	(86.5)	(68.0)	(65.2)
	10	49.9	46.4	27.7	68.0	61.1	35.4	38.4	59.7	54.6	37.1	78.2	71.9	50.8	51.9
		(68.1)	(65.0)	(40.2)	(71.5)	(68.3)	(43.2)	(39.1)	(79.5)	(73.7)	(53.7)	(81.0)	(76.0)	(56.6)	(53.5)
	15	45.2	40.8	22.3	59.0	49.9	25.3	30.9	53.2	48.7	32.3	72.0	62.9	38.6	42.5
		(63.2)	(55.8)	(33.7)	(65.3)	(57.8)	(36.1)	(32.7)	(72.5)	(66.5)	(46.1)	(73.8)	(68.6)	(48.7)	(44.9)

TAB. 2.13. Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP_5 .

		$\alpha = 0.05$							$\alpha = 0.10$						
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
=100	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	2.3	2.4	6.6	26.1	26.0	28.3	36.2	3.2	4.0	12.1	37.5	37.2	45.5	52.8
		(26.2)	(28.6)	(30.7)	(24.2)	(26.6)	(28.1)	(30.9)	(36.9)	(39.5)	(47.4)	(34.2)	(37.0)	(46.7)	(47.2)
	3	2.4	2.5	11.3	25.9	26.6	25.7	30.8	4.3	5.0	16.1	42.3	42.8	42.5	45.1
		(29.9)	(30.6)	(31.3)	(28.1)	(29.9)	(30.1)	(31.3)	(44.8)	(46.0)	(42.3)	(41.7)	(43.1)	(41.9)	(42.0)
	5	5.3	6.8	19.9	29.4	28.3	29.1	34.0	9.4	10.8	25.5	47.1	45.3	44.2	47.9
		(35.7)	(34.6)	(36.4)	(34.8)	(33.8)	(37.9)	(35.6)	(51.4)	(52.6)	(49.3)	(49.0)	(51.0)	(47.8)	(48.4)
	9	12.2	14.9	18.2	31.8	29.8	22.3	29.5	19.3	21.4	27.9	49.1	43.9	34.7	43.2
		(40.8)	(41.4)	(31.9)	(39.9)	(39.1)	(31.9)	(31.7)	(53.8)	(52.3)	(45.9)	(52.7)	(50.9)	(46.7)	(45.2)
	12	14.9	17.7	16.2	32.4	28.6	15.2	23.3	22.0	26.0	24.1	46.2	41.8	24.7	35.7
		(40.3)	(40.0)	(26.9)	(38.7)	(38.2)	(27.1)	(25.9)	(52.0)	(52.0)	(38.1)	(53.1)	(51.8)	(39.0)	(37.0)
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
=200	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	11.5	13.8	31.8	53.5	56.3	67.4	72.8	16.5	19.6	42.9	65.8	68.2	80.7	86.0
		(57.7)	(61.4)	(73.9)	(55.9)	(60.7)	(73.1)	(74.0)	(69.3)	(73.5)	(85.6)	(68.8)	(72.8)	(85.4)	(85.6)
	3	17.2	19.0	39.6	61.7	66.3	62.4	65.3	23.9	25.7	48.6	77.9	81.6	76.0	77.7
		(70.9)	(77.2)	(71.4)	(68.7)	(73.0)	(71.9)	(71.2)	(83.5)	(86.6)	(80.3)	(83.1)	(85.6)	(79.4)	(80.2)
	5	29.7	33.0	58.7	71.0	70.9	74.1	75.7	39.5	43.0	69.0	83.5	83.7	84.1	85.7
		(77.7)	(78.6)	(80.2)	(76.6)	(78.0)	(80.6)	(79.8)	(89.5)	(90.2)	(87.0)	(89.7)	(88.7)	(86.9)	(86.8)
	10	53.9	58.1	59.1	79.8	78.0	66.1	70.6	65.6	69.4	69.4	88.3	86.7	79.5	82.0
		(86.1)	(86.1)	(74.3)	(85.3)	(85.8)	(73.9)	(72.8)	(91.7)	(91.4)	(86.0)	(91.5)	(91.0)	(86.3)	(85.7)
	15	60.6	63.3	49.1	80.4	76.3	50.6	58.4	70.9	72.9	61.4	88.7	86.1	64.8	71.7
		(86.1)	(85.3)	(65.1)	(85.8)	(85.0)	(64.6)	(63.7)	(91.2)	(90.2)	(76.7)	(91.1)	(90.3)	(76.2)	(75.8)

TAB. 2.14. Empirical powers based on the asymptotic and empirical (in parentheses) critical values of the transformed test statistic T_n^* defined by (2.0.16), the untransformed test statistic T_n given by (2.0.6), for different kernels and different smoothing parameters or truncation values p_n , and of Hosking's (1980) portmanteau test statistic (2.0.5) using $M = p_n$, for the VARMA(1,1) process defined by (2.0.17) and (2.0.18), using the parameters values corresponding to DGP_6 .

		$\alpha = 0.05$							$\alpha = 0.10$						
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
=100	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	42.0	42.3	28.0	61.8	60.8	39.4	40.8	49.4	49.5	37.3	72.9	70.8	54.0	55.2
		(61.1)	(61.2)	(40.0)	(61.9)	(61.3)	(40.8)	(39.8)	(73.1)	(72.8)	(53.9)	(73.6)	(73.4)	(53.8)	(53.9)
	3	40.3	39.8	23.6	58.6	57.7	31.2	32.3	48.3	48.2	33.1	69.6	68.9	44.2	45.9
		(60.3)	(60.9)	(32.7)	(60.6)	(61.4)	(32.6)	(32.1)	(70.6)	(70.6)	(46.7)	(70.3)	(71.0)	(46.3)	(45.9)
	5	34.7	33.3	18.2	50.1	45.5	19.9	21.8	44.1	41.9	25.5	62.9	58.8	32.1	35.0
		(48.9)	(48.7)	(22.6)	(50.9)	(49.3)	(22.7)	(22.0)	(64.4)	(61.6)	(36.7)	(64.9)	(62.4)	(36.8)	(36.7)
	9	27.4	24.6	14.3	37.1	29.3	12.6	16.6	36.7	33.8	22.6	50.9	42.9	22.2	26.6
		(40.2)	(33.5)	(19.4)	(40.3)	(34.1)	(19.4)	(19.6)	(53.8)	(48.1)	(28.1)	(53.8)	(48.1)	(27.5)	(26.6)
	12	24.8	20.9	14.1	30.4	21.4	10.1	16.1	33.7	30.1	21.1	43.9	34.0	17.5	24.9
		(36.0)	(30.1)	(16.9)	(36.5)	(29.9)	(16.3)	(15.6)	(48.0)	(42.3)	(27.5)	(48.5)	(42.7)	(27.1)	(26.6)
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
=200	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	78.1	78.3	67.3	91.1	89.8	77.6	77.8	84.3	84.0	74.7	95.3	94.4	86.6	86.6
		(92.4)	(93.1)	(81.8)	(93.1)	(93.0)	(81.7)	(81.6)	(95.7)	(96.0)	(88.3)	(96.0)	(96.0)	(88.2)	(88.2)
	3	77.0	77.4	59.1	89.3	89.3	66.1	66.9	83.9	83.9	67.5	94.4	94.0	77.2	77.8
		(91.4)	(91.2)	(71.3)	(91.5)	(91.4)	(71.2)	(70.9)	(94.9)	(95.3)	(81.5)	(95.1)	(95.3)	(81.6)	(81.0)
	5	73.6	72.4	47.3	84.8	83.0	51.8	53.6	79.8	79.1	57.3	91.3	89.9	65.3	66.4
		(87.3)	(85.8)	(59.9)	(87.4)	(85.6)	(60.1)	(59.5)	(92.1)	(91.3)	(72.3)	(92.1)	(91.3)	(72.5)	(72.0)
	10	64.2	58.9	34.0	73.3	65.6	33.9	37.2	72.4	68.1	46.0	83.4	76.9	47.7	51.4
		(78.4)	(72.7)	(41.2)	(78.4)	(73.0)	(41.2)	(41.1)	(86.5)	(82.3)	(56.4)	(86.7)	(82.3)	(56.1)	(55.4)
	15	57.2	49.4	29.7	63.5	51.9	25.7	31.3	66.1	60.1	39.8	75.2	64.5	38.2	44.2
		(71.2)	(63.1)	(38.2)	(71.6)	(63.0)	(37.9)	(36.8)	(80.2)	(72.0)	(50.6)	(80.4)	(72.1)	(50.7)	(48.8)

CONCLUSION

Dans ce mémoire, nous nous sommes intéressés aux statistiques de test corrigées de Duchesne et Roy (2004) et Duchesne (2005a). Ces tests corrigés visent à vérifier l'hypothèse de corrélation sérielle pour des séries chronologiques stationnaires et partiellement non-stationnaires. Tout d'abord, nous avons effectué une transformation de type puissance des statistiques de test de Duchesne et Roy (2004) et Duchesne (2005a) à l'aide de la méthode proposée par Chen et Deo (2004a). Par la suite, nous avons déterminé des expressions plus précises de la moyenne et de la variance de la variable aléatoire Q_n intervenant dans la définition des statistiques de test. Ceci a conduit à une amélioration considérable des propriétés des tests originaux en échantillons finis.

Dans un premier temps, nous avons obtenu la puissance β intervenant dans la transformation des statistiques de test de Duchesne et Roy (2004) et Duchesne (2005a). Nous avons montré qu'elle est indépendante de la dimension de la série chronologique. De plus, les distributions asymptotiques des autocovariances résiduelles, obtenues par Li et McLeod (1981) dans le cas de séries chronologiques stationnaires, et par Brüggemann *et al.* (2006) dans le cas de séries chronologiques partiellement non-stationnaires, nous ont permis de calculer de nouvelles expressions pour la moyenne et pour la variance de la mesure de distance Q_n . Nous avons démontré que la statistique de test corrigée $T_n^*(k; p_n)$ converge vers une distribution normale autant pour les processus stationnaires que partiellement non-stationnaires.

Dans un second temps, nous avons effectué une étude comparative en terme de niveau et de puissance via des simulations Monte Carlo. Nous avons comparé les

statistiques de test transformées à leurs versions originales et au test portemanteau de Hosking (1980) dans le contexte des séries chronologiques stationnaires et partiellement non-stationnaires. Nous avons vu que les tests transformés corrigent considérablement le problème d'asymétrie des tests de Duchesne et Roy (2004) et Duchesne (2005a). Nous avons également constaté que les tests corrigés performant mieux lorsqu'un petit point de troncature ou un paramètre de lissage est considéré. Nous avons remarqué que le test de Hosking (1980) donne d'excellents résultats lorsqu'un grand nombre de délais sont inclus dans la statistique de test. De manière générale, les statistiques de test transformées donnent de meilleurs résultats de niveau que leurs versions non-transformées. De plus, les puissances empiriques obtenues pour les statistiques de test transformées sont similaires, et même meilleures dans certains cas, aux puissances empiriques obtenues par les statistiques de test originales.

En conclusion, les statistiques de test transformées proposées dans l'article présentent beaucoup d'avantages face à leurs versions non-transformées. En effet, les niveaux empiriques des tests transformés sont de manière générale beaucoup plus près des niveaux nominaux, ce qui est grandement souhaitable dans les applications. De plus, les puissances empiriques sont assez près des puissances asymptotiques pour les tests transformés. Conséquemment, étant donné qu'en pratique les valeurs critiques asymptotiques sont utilisées, les statistiques de test transformées se comparent fort avantageusement aux statistiques de test originales.

Annexe A

SIMULATIONS COMPLÉMENTAIRES

Dans l'article, six processus générateurs de données (DGP) ont été considérés. Afin de compléter l'étude des propriétés en échantillons finis de la statistique de test corrigée $T_n^*(k; p_n)$ donnée par (2.0.16), quelques expériences Monte Carlo additionnelles sont effectuées. Les études de niveau et de puissance sont conduites de la même façon que décrites dans l'article.

A.1. DESCRIPTION DES EXPÉRIENCES

Dans l'étude de simulation complémentaire, deux nouveaux modèles sont considérés. Le premier modèle, généré selon le DGP₇ défini plus bas, correspond à un processus VAR(1) stationnaire de dimension $d = 3$. Le second modèle correspondant au DGP₈, correspond à un processus VAR(3) partiellement non-stationnaire de dimension $d = 2$. Des réalisations de $\{\mathbf{Y}_t, t = 1, \dots, n\}$ selon les DGP₇ et DGP₈ sont obtenues à partir des processus suivants :

$$\text{DGP}_7 : \mathbf{Y}_t = \mathbf{c} + \mathbf{A}_1 \mathbf{Y}_{t-1} + \mathbf{a}_{t7}, \quad t = 1, \dots, n, \quad (\text{A.1.1})$$

$$\text{DGP}_8 : \mathbf{Y}_t = \mathbf{A}_1 \mathbf{Y}_{t-1} + \mathbf{A}_2 \mathbf{Y}_{t-2} + \mathbf{A}_3 \mathbf{Y}_{t-3} + \mathbf{a}_{t8}, \quad t = 1, \dots, n, \quad (\text{A.1.2})$$

où n est la longueur de la série. Pour l'étude de niveau, le terme d'erreur $\{\mathbf{a}_{tl}\}$ est défini par $\{\mathbf{a}_{tl} = \mathbf{u}_{tl}\}$ et pour l'étude de puissance, $\{\mathbf{a}_{tl}\}$ est défini par le processus VMA(1) $\{\mathbf{a}_{tl} = \mathbf{u}_{tl} - \mathbf{\Xi}_l \mathbf{u}_{t-1,l}\}$, où $\{\mathbf{u}_{tl}\}$ est un processus stochastique gaussien iid tel que $\mathbf{u}_{tl} \sim N_d(\mathbf{0}, \mathbf{\Sigma}_l)$, $l \in \{7, 8\}$. Les valeurs des paramètres décrits dans

(A.1.1) et (A.1.2) sont présentées dans le tableau A.1. Pour le modèle stationnaire, les valeurs des paramètres \mathbf{c} , $\mathbf{\Lambda}_1$ et $\mathbf{\Sigma}_7$ proviennent de Lütkepohl (1993, p. 120) et la matrice $\mathbf{\Xi}_7$ est inspirée de Duchesne et Roy (2004). Les racines du déterminant de l'opérateur autorégressif $\Lambda(z) = \mathbf{I}_3 - \mathbf{\Lambda}_1 z$, $z \in \mathbb{C}$, sont 2.00, 2.15 et -15.49 . Pour le modèle partiellement non-stationnaire, les paramètres autorégressifs $\mathbf{A}_1, \mathbf{A}_2$, et \mathbf{A}_3 proviennent de Reinsel (1993, p. 186). Les racines du déterminant de l'opérateur autorégressif $\mathbf{A}(z) = \mathbf{I}_2 - \mathbf{A}_1 z - \mathbf{A}_2 z^2 - \mathbf{A}_3 z^3$, $z \in \mathbb{C}$, sont 1.00, $1.18 \pm 0.51i$ et $0.55 \pm 1.44i$. Les valeurs des matrices $\mathbf{\Sigma}_8$ et $\mathbf{\Xi}_8$ sont inspirées de Duchesne (2005a).

TAB. A.1. Valeurs des paramètres des modèles générés selon les DGP₇ et DGP₈.

$$\begin{array}{l}
 \mathbf{\Lambda}_1 = \begin{pmatrix} 0.5 & 0.0 & 0.0 \\ 0.1 & 0.1 & 0.3 \\ 0.0 & 0.2 & 0.3 \end{pmatrix}, \quad \mathbf{A}_1 = \begin{pmatrix} 1.492 & 0.065 \\ 1.296 & 1.404 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} -0.999 & 0.000 \\ -1.979 & -0.582 \end{pmatrix}, \\
 \mathbf{A}_3 = \begin{pmatrix} 0.441 & 0 \\ 0.863 & 0 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 0.01 \\ 0.02 \\ 0.00 \end{pmatrix}, \quad \mathbf{\Sigma}_7 = \begin{pmatrix} 2.25 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.50 \\ 0.00 & 0.50 & 0.74 \end{pmatrix}, \\
 \mathbf{\Sigma}_8 = \begin{pmatrix} 25.0 & 5.4 \\ 5.4 & 9.0 \end{pmatrix}, \quad \mathbf{\Xi}_7 = \begin{pmatrix} 0.288 & 0.000 & 0.000 \\ 0.144 & 0.648 & 1.440 \\ 0.000 & 0.000 & 0.072 \end{pmatrix}, \quad \mathbf{\Xi}_8 = \begin{pmatrix} 0.96 & 0.32 \\ 0.00 & 0.16 \end{pmatrix}.
 \end{array}$$

L'étude de simulation complémentaire est effectuée de la même façon que l'étude de simulation décrite dans l'article. Les trois statistiques de test utilisées sont la statistique de test non-transformée $T_n(k; p_n)$ définie par (2.0.6), sa version corrigée $T_n^*(k; p_n)$ définie par (2.0.16) et la statistique de test portemanteau multivariée P_M définie par (2.0.5). Les noyaux considérés pour les tests reposant sur des noyaux sont les noyaux Bartlett (BAR), Daniell (DAN) et uniforme tronqué (TR). La représentation graphique de ces trois noyaux se retrouve à la figure A.1. Les points de troncature considérés sont $p_n = 2, 3, [\log(n)], [3.5n^{0.2}], [3n^{0.3}]$, où $[a]$ désigne le nombre a arrondi à l'unité près. Nous posons $M = p_n$ pour le test portemanteau de Hosking.

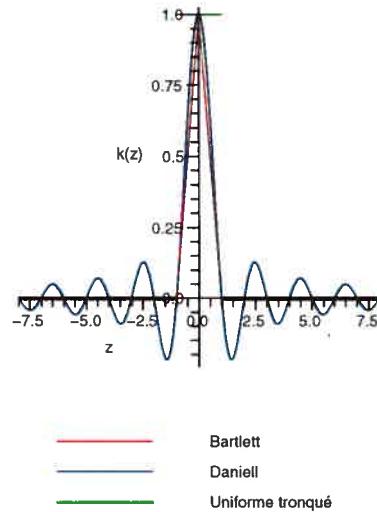


FIG. A.1. Représentation graphique des noyaux utilisés pour les statistiques de test transformée T_n^* définie par (2.0.16) et originale T_n donnée par (2.0.6).

Dans un premier temps, une étude de niveau est effectuée pour les DGP₇ et DGP₈ de longueur $n = 50, 100$ et 200 . Pour chaque longueur de série chronologique et pour chaque noyau, les valeurs de la transformation puissance β définie par (2.0.15) et intervenant dans les statistiques de test transformées sont reportées dans le tableau A.1. Les niveaux considérés sont les niveaux de 1, 5 et 10%. Les fréquences empiriques de rejet de l'hypothèse nulle sont calculées à l'aide des valeurs critiques asymptotiques. Pour chaque longueur de série, 1000 réalisations indépendantes sont générées à partir des DGP₇ et DGP₈. Les calculs des tests et des niveaux empiriques sont effectués de la même façon que décrits dans l'article.

Dans un second temps, l'étude de puissance est conduite de façon similaire à l'étude de niveau. Elle est caractérisée par la génération d'erreurs corrélées. Les puissances empiriques sont calculées à l'aide des valeurs critiques asymptotiques et des valeurs critiques exactes qui sont obtenues dans l'étude de niveau précédente. Les tests sont de niveau 5 et 10 % et des séries de longueur $n = 100$ et 200 sont également considérés.

TAB. A.2. Valeurs de la puissance β définie par (2.0.15) et intervenant dans la statistique de test transformée $T_n^*(k; p_n)$ définie par (2.0.16)

$n = 50$	p_n				
	2	3	4	8	10
BAR	0.3333	0.2503	0.2284	0.2111	0.2093
DAN	0.2065	0.1413	0.1455	0.1712	0.1780
TR	0.3333	0.3333	0.3333	0.3333	0.3333
$n = 100$	p_n				
	2	3	5	9	12
BAR	0.3333	0.2503	0.2196	0.2100	0.2084
DAN	0.2032	0.1373	0.1483	0.1661	0.1733
TR	0.3333	0.3333	0.3333	0.3333	0.3333
$n = 200$	p_n				
	2	3	5	10	15
BAR	0.3333	0.2503	0.2196	0.2093	0.2076
DAN	0.2016	0.1353	0.1456	0.1639	0.1715
TR	0.3333	0.3333	0.3333	0.3333	0.3333

A.2. DISCUSSION DES RÉSULTATS EMPIRIQUES

Dans cette section, nous allons discuter des résultats de l'étude de niveau et des résultats de l'étude de puissance pour les processus générés selon les DGP₇ et DGP₈.

A.2.1. Étude de niveau

Les résultats liés à l'étude de niveau pour les DGP₇ et DGP₈ se retrouvent dans les tableaux A.3 et A.4. Nous remarquons que pour le DGP₈, les niveaux de la statistique de test de Hosking (1980) ne sont pas donnés pour $p_n = 2$. En effet, ce test n'est valide que pour des degrés de liberté positifs, c'est-à-dire, lorsque $p_n > 2$.

D'abord, nous remarquons que pour tous les niveaux des tests (1%, 5% et 10%) et pour les deux modèles, un sous-rejet important de la statistique de test non-transformée $T_n(k; p_n)$ est survenu. De plus, les niveaux empiriques pour le

modèle partiellement non-stationnaire sont pratiquement nuls. Ceci est dû à la forte asymétrie vers la droite de la statistique de test. Pour les deux modèles, la transformation de la statistique de test $T_n(k; p_n)$ améliore nettement les résultats de niveau. Nous remarquons de manière générale qu'il existe une détérioration des niveaux lorsque le point de troncature augmente, mais pour un point de troncature petit ou modéré, les résultats de niveau sont très satisfaisants. De plus, malgré que les niveaux sont faibles pour le point de troncature $p_n = 3n^{0.3}$, ils ne contredisent pas les résultats obtenus dans l'étude de simulation de l'article. D'ailleurs, dans la majorité des cas, le noyau Bartlett donne de meilleurs résultats empiriques que les noyaux Daniell ou encore uniforme tronqué. Nous remarquons que la statistique de test de Hosking (1980) performe très bien lorsqu'un grand nombre de délais est considéré, surtout lorsque les données sont générées par le DGP₈. Par contre, lorsqu'un petit nombre de délais est considéré dans le calcul du test de Hosking (1980), nous constatons pour les deux modèles un grand sur-rejet du test, comme anticipé.

A.2.2. Étude de puissance

Les résultats reliés à l'étude de puissance se retrouvent dans les tableaux A.5 et A.6. De même que dans l'étude de niveau, les puissances asymptotiques de la statistique de test de Hosking (1980) du DGP₈ sont données seulement lorsque $p_n > 2$.

Pour le DGP₇, nous remarquons que pour tous les tests considérés dans l'étude de simulation, la puissance empirique tend à croître pour un point de troncature petit à modéré et décroît par la suite. Ceci va quelque peu à l'encontre de l'intuition suggérant que la puissance maximale serait donnée par un petit p_n quand les erreurs sont générées par un processus VMA(1). Le comportement contre-intuitif des résultats de puissance peut être expliqué par le fait que si nous générons un processus VARMA(1,1) de dimension $d = 3$ et que nous ajustons un processus VAR(1), alors des problèmes surviennent lors de l'étape de l'estimation du modèle. En effet, l'estimateur du paramètre autorégressif Λ_1 devient un estimateur

biaisé, ce qui a pour effet de perturber la dépendance dans les résidus. En conséquence, les résidus ne sont pas tout à fait expliqués par un modèle VMA(1). Dans ce cas, pour obtenir une puissance maximale pour les tests, nous devons choisir un point de troncature modérément petit. Nous remarquons également pour le DGP_7 que les puissances empiriques pour le test transformé sont moins puissantes que le test spectral original. Les écarts entre les puissances empiriques des deux tests spectraux tendent cependant à diminuer en fonction de p_n . Nous observons que les puissances empiriques sont beaucoup plus près des puissances asymptotiques pour le test transformé, surtout lorsque des points de troncature petits à modérés sont considérés.

Pour le DGP_8 , nous remarquons les effets positifs du test transformé. Effectivement, à partir des puissances asymptotiques, nous remarquons que le test non-transformé ne détecte pas la présence d'une dépendance dans les résidus quand les points de troncature sont petits. Nous remarquons également que si les tests sont ramenés au même niveau, alors nous constatons pour le test transformé un gain de puissance. Concernant le test de Hosking (1980), celui-ci est généralement plus puissant lorsqu'un point de troncature modéré est considéré. De plus, l'écart entre les puissances asymptotiques et empiriques est négligeable lorsqu'un grand nombre de délais est considéré dans le calcul du test.

TAB. A.3. Niveaux empiriques (en pourcentage) de la statistique de test transformée T_n^* définie par (2.0.16), de la statistique de test originale T_n donnée par (2.0.6), pour différents noyaux et différents paramètres de lissage ou points de troncature p_n , et la statistique de test portemanteau de Hosking's (1980) (2.0.5) utilisant $M = p_n$, pour le processus VAR(1) de dimension $d = 3$ généré par le DGP₇.

		$\alpha = 0.01$						$\alpha = 0.05$						$\alpha = 0.10$																	
$n = 50$		T_n			T_n^*			P_M			T_n			T_n^*			P_M			T_n			T_n^*			P_M					
	p_n	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR
	2	0.0	0.0	0.1	0.6	0.7	0.5	4.6	0.2	0.2	0.5	5.3	4.2	3.8	18.1	0.3	0.4	1.3	11.6	10.0	8.8	31.2									
	3	0.1	0.1	0.3	0.6	0.6	0.7	2.8	0.2	0.2	1.2	4.4	3.7	3.4	12.1	0.3	0.3	1.8	10.4	8.9	7.0	23.4									
	4	0.1	0.1	0.4	0.6	0.6	0.5	2.6	0.1	0.1	1.4	4.3	3.6	2.8	10.9	0.5	0.5	2.6	8.5	6.5	6.4	20.3									
	8	0.1	0.1	0.5	0.5	0.5	0.1	2.0	0.5	1.1	2.0	2.6	1.9	1.2	8.4	1.5	1.6	4.4	6.6	4.6	3.3	15.2									
	10	0.1	0.2	0.6	0.5	0.2	0.2	2.0	0.8	1.1	2.5	2.6	1.7	0.6	8.5	1.7	1.9	5.3	5.8	3.8	1.8	15.5									
$n = 100$		T_n			T_n^*			P_M			T_n			T_n^*			P_M			T_n			T_n^*			P_M					
	p_n	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR
	2	0.0	0.1	0.2	1.0	1.1	1.1	4.1	0.4	0.4	0.5	5.0	4.4	4.1	17.5	0.6	0.6	1.5	11.2	9.0	9.8	28.2									
	3	0.1	0.0	0.1	0.8	0.9	0.8	2.7	0.6	0.6	1.0	4.3	3.7	3.8	12.9	0.7	0.6	1.7	10.1	8.5	8.7	22.7									
	5	0.1	0.3	0.0	0.8	0.8	0.3	1.7	0.6	0.6	1.1	3.2	2.9	3.1	8.1	0.7	0.7	2.3	9.3	7.7	6.7	14.8									
	9	0.0	0.0	0.6	0.6	0.3	0.5	1.5	0.6	0.6	1.8	3.2	2.4	2.0	5.9	0.9	1.4	3.1	7.4	5.5	3.3	12.1									
	12	0.1	0.0	0.8	0.4	0.3	0.4	1.6	0.7	0.9	1.9	2.4	1.7	1.5	5.4	1.5	1.8	3.8	6.2	3.8	2.9	12.0									
$n = 200$		T_n			T_n^*			P_M			T_n			T_n^*			P_M			T_n			T_n^*			P_M					
	p_n	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR
	2	0.0	0.0	0.1	0.5	0.6	0.8	3.5	0.0	0.0	0.6	4.5	3.9	3.7	15.5	0.0	0.1	1.2	9.3	9.8	8.8	26.3									
	3	0.0	0.0	0.1	0.5	0.5	0.8	1.9	0.0	0.0	0.8	4.8	4.0	3.3	11.1	0.0	0.1	1.2	8.8	8.2	8.3	20.5									
	5	0.0	0.0	0.3	0.4	0.5	0.5	1.9	0.2	0.3	0.8	3.1	3.2	3.5	9.4	0.3	0.5	2.5	7.6	6.7	8.8	18.1									
	10	0.1	0.1	0.3	0.7	0.7	0.5	1.3	0.6	0.8	1.5	3.2	2.5	2.5	6.8	1.4	1.5	3.8	7.2	6.2	5.9	13.7									
	15	0.2	0.2	0.5	0.7	0.5	0.3	1.1	0.9	1.3	1.9	2.9	2.2	1.6	5.8	1.6	2.1	4.1	6.3	5.1	4.0	11.8									

TAB. A.4. Niveaux empiriques (en pourcentage) de la statistique de test transformée T_n^* définie par (2.0.16), de la statistique de test originale T_n donnée par (2.0.6), pour différents noyaux et différents paramètres de lissage ou points de troncature p_n , et la statistique de test portemanteau de Hosking's (1980) (2.0.5) utilisant $M = p_n$, pour le processus VAR(3) de dimension $d = 2$ généré par le DGP₈.

		$\alpha = 0.01$						$\alpha = 0.05$						$\alpha = 0.10$																			
$n = 50$	p_n	T_n			T_n^*			P_M			T_n			T_n^*			P_M			T_n			T_n^*			P_M							
		BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR					
	2	0.0	0.0	0.0	0.8	0.7	1.2	0.0	0.0	0.1	4.0	2.6	4.4	0.0	0.0	0.2	9.4	7.0	8.8														
	3	0.0	0.0	0.0	1.0	0.8	0.7	9.8	0.0	0.0	0.0	3.3	2.8	4.2	34.8	0.0	0.0	0.0	8.5	7.3	8.5	54.5											
	4	0.0	0.0	0.0	0.9	0.8	0.5	2.9	0.0	0.0	0.0	3.6	2.7	3.4	15.5	0.0	0.0	0.2	8.0	6.8	8.2	25.7											
	8	0.0	0.0	0.0	0.5	0.3	0.1	1.3	0.0	0.0	0.1	3.4	2.5	2.5	5.6	0.0	0.0	0.4	7.3	6.5	5.6	11.5											
	10	0.0	0.0	0.1	0.3	0.1	0.2	0.8	0.0	0.0	0.2	3.4	2.4	1.6	6.1	0.0	0.0	0.6	7.8	5.9	4.4	11.0											
$n = 100$	p_n	T_n			T_n^*			P_M			T_n			T_n^*			P_M			T_n			T_n^*			P_M							
		BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR		
	2	0.0	0.0	0.0	0.6	0.4	0.7		0.0	0.0	0.0	4.8	4.2	3.6		0.0	0.0	0.0	9.6	7.9	8.7												
	3	0.0	0.0	0.0	0.3	0.1	0.6	7.0	0.0	0.0	0.0	4.8	3.9	4.0	30.8	0.0	0.0	0.1	9.5	8.1	8.5	50.1											
	5	0.0	0.0	0.1	0.3	0.3	0.7	1.4	0.0	0.0	0.1	3.6	3.3	3.7	7.1	0.0	0.0	0.3	9.0	7.7	7.9	14.9											
	9	0.0	0.0	0.1	0.6	0.6	0.9	1.5	0.0	0.0	0.4	3.5	2.7	3.8	5.4	0.0	0.1	1.2	7.5	6.0	7.4	11.0											
	12	0.0	0.0	0.1	0.6	0.5	0.3	1.2	0.1	0.1	0.3	3.0	2.8	3.3	4.7	0.1	0.1	1.2	7.0	5.8	5.2	9.4											
$n = 200$	p_n	T_n			T_n^*			P_M			T_n			T_n^*			P_M			T_n			T_n^*			P_M							
		BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR	BAR	DAN	TR		
	2	0.0	0.0	0.0	1.0	0.9	0.9		0.0	0.0	0.1	4.8	4.0	4.2		0.0	0.0	0.1	8.6	7.9	7.1												
	3	0.0	0.0	0.0	0.7	0.7	1.1	5.2	0.0	0.0	0.1	4.3	3.9	3.2	24.2	0.0	0.0	0.1	7.9	7.2	7.5	44.4											
	5	0.0	0.0	0.0	0.6	1.0	0.6	1.3	0.1	0.1	0.0	3.9	3.6	4.2	6.5	0.1	0.1	0.2	7.0	6.5	8.4	12.2											
	10	0.0	0.0	0.1	0.6	0.4	0.7	0.9	0.1	0.1	0.2	3.2	3.2	2.9	4.1	0.1	0.1	0.7	6.0	6.1	7.5	9.8											
	15	0.0	0.0	0.0	0.4	0.4	0.4	0.7	0.0	0.0	0.4	3.7	3.0	3.1	4.4	0.1	0.1	0.9	6.2	5.0	6.0	9.3											

TAB. A.5. Puissances empiriques basées sur les valeurs critiques asymptotiques et empiriques (entre parenthèses) de la statistique de test transformée T_n^* définie par (2.0.16), de la statistique de test originale T_n donnée par (2.0.6), pour différents noyaux et différents paramètres de lissage ou points de troncature p_n , et de la statistique de test portemanteau de Hosking's (1980) (2.0.5) utilisant $M = p_n$, pour le processus VARMA(1,1) de dimension $d = 3$ généré par le DGP₇.

		$\alpha = 0.05$							$\alpha = 0.10$						
$n = 100$	p_n	T_n			T_n^*			P_M	T_n			T_n^*			P_M
		BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	0.8	1.0	3.1	9.0	9.9	10.8	36.9	1.2	1.7	4.6	16.5	20.4	20.4	51.3
		(15.1)	(21.3)	(16.2)	(9.1)	(11.8)	(12.8)	(16.0)	(23.2)	(31.0)	(23.9)	(15.6)	(22.4)	(20.6)	(23.5)
	3	0.7	0.8	18.7	9.7	9.4	36.0	60.7	1.7	1.9	27.0	18.3	17.6	50.3	73.4
		(18.2)	(19.0)	(43.9)	(11.1)	(11.6)	(39.4)	(44.1)	(26.1)	(26.6)	(56.3)	(18.1)	(19.4)	(53.3)	(56.8)
	5	3.2	4.4	17.3	17.5	17.0	26.6	47.0	5.7	6.8	26.3	30.0	29.1	39.8	61.8
		(31.3)	(31.5)	(37.6)	(23.1)	(24.4)	(34.2)	(37.7)	(40.4)	(40.5)	(52.3)	(31.4)	(32.7)	(49.0)	(51.6)
	9	10.1	13.4	13.6	24.5	22.4	13.3	30.4	16.3	19.4	20.6	40.2	36.9	22.2	42.8
		(40.1)	(39.8)	(29.6)	(34.5)	(36.2)	(27.9)	(29.7)	(50.0)	(51.4)	(39.6)	(45.1)	(48.5)	(37.9)	(39.7)
	12	12.7	14.6	12.7	23.4	19.5	9.6	24.6	18.8	22.5	19.0	38.2	31.7	15.1	35.8
		(38.8)	(38.0)	(24.8)	(33.7)	(35.1)	(24.1)	(24.1)	(49.4)	(51.8)	(34.5)	(46.4)	(47.8)	(33.2)	(33.1)
$n = 200$	p_n	T_n			T_n^*			P_M	T_n			T_n^*			P_M
		BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	1.8	3.3	10.5	15.1	23.1	27.7	58.3	3.3	7.0	16.7	25.2	37.7	40.7	72.9
		(24.1)	(39.6)	(38.2)	(16.7)	(27.3)	(32.8)	(38.2)	(34.5)	(53.0)	(47.6)	(26.2)	(38.4)	(42.6)	(47.7)
	3	3.5	3.8	68.3	19.0	19.4	84.2	95.2	5.7	6.0	76.2	31.6	31.7	91.3	98.0
		(30.6)	(31.5)	(88.3)	(19.6)	(21.7)	(87.3)	(88.3)	(44.0)	(46.8)	(94.4)	(34.3)	(35.4)	(92.6)	(94.5)
	5	17.6	22.1	63.0	53.2	55.5	76.9	89.9	26.6	30.8	74.6	71.9	72.8	87.7	95.8
		(72.7)	(75.6)	(85.4)	(62.6)	(68.6)	(83.3)	(85.0)	(81.5)	(83.7)	(90.5)	(74.6)	(78.1)	(89.7)	(90.5)
	10	51.6	60.4	45.4	76.6	76.7	51.3	67.6	64.4	70.9	58.9	87.2	87.4	65.0	79.3
		(86.1)	(86.6)	(63.4)	(83.4)	(83.3)	(62.5)	(63.0)	(93.1)	(93.5)	(74.7)	(91.4)	(91.9)	(74.0)	(74.1)
	15	55.3	59.0	36.0	73.9	67.5	33.1	52.8	68.2	69.6	48.3	84.2	79.1	47.3	66.6
		(84.2)	(80.4)	(51.5)	(80.2)	(78.8)	(50.1)	(50.6)	(91.8)	(89.5)	(64.9)	(90.4)	(88.2)	(64.1)	(62.8)

TAB. A.6. Puissances empiriques basées sur les valeurs critiques asymptotiques et empiriques (entre parenthèses) de la statistique de test transformée T_n^* définie par (2.0.16), de la statistique de test originale T_n donnée par (2.0.6), pour différents noyaux et différents paramètres de lissage ou points de troncature p_n , et de la statistique de test portemanteau de Hosking's (1980) (2.0.5) utilisant $M = p_n$, pour le processus VARMA(3,1) de dimension $d = 2$ généré par le DGP₈.

		$\alpha = 0.05$						$\alpha = 0.10$							
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
$=100$	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	0.1	0.1	0.8	16.5	14.5	22.8		0.1	0.3	1.1	25.6	25.8	36.5	
		(10.0)	(10.5)	(24.9)	(16.7)	(17.6)	(26.9)	(26.5)	(16.8)	(18.7)	(37.5)	(25.9)	(29.4)	(38.1)	(38.2)
	3	0.2	0.2	0.8	19.6	22.2	22.3	65.2	0.3	0.3	1.6	32.6	35.3	34.9	81.5
		(15.3)	(18.3)	(23.7)	(20.9)	(25.0)	(25.2)	(23.8)	(26.8)	(33.3)	(36.8)	(33.4)	(39.5)	(38.2)	(36.7)
	5	0.4	0.5	2.1	24.9	23.6	35.3	45.0	0.7	0.9	4.3	40.5	38.1	51.6	61.3
		(26.3)	(28.0)	(35.7)	(29.6)	(31.4)	(40.2)	(36.1)	(36.8)	(39.3)	(53.2)	(42.3)	(43.2)	(56.3)	(54.1)
	9	0.9	1.2	4.1	34.6	31.7	25.2	30.7	1.3	2.3	7.4	50.4	47.1	38.4	44.8
		(39.3)	(41.8)	(29.9)	(42.7)	(43.0)	(31.5)	(28.4)	(52.1)	(54.1)	(45.0)	(56.1)	(57.6)	(45.9)	(42.7)
	12	1.2	1.4	4.8	33.7	29.2	18.5	24.3	2.2	3.6	9.6	50.2	44.7	29.3	38.4
		(38.5)	(40.6)	(27.9)	(41.6)	(42.1)	(28.7)	(26.7)	(55.1)	(55.0)	(41.0)	(58.8)	(58.5)	(42.4)	(39.9)
n		T_n			T_n^*			P_M	T_n			T_n^*			P_M
$=200$	p_n	BAR	DAN	TR	BAR	DAN	TR		BAR	DAN	TR	BAR	DAN	TR	
	2	0.0	0.0	4.2	39.1	44.6	54.7		0.0	0.0	6.7	50.5	57.1	66.4	
		(25.0)	(31.7)	(57.2)	(40.0)	(49.4)	(56.6)	(56.0)	(39.5)	(46.3)	(69.5)	(52.1)	(61.8)	(70.7)	(69.2)
	3	0.1	0.1	5.3	47.5	58.1	58.3	83.8	0.3	0.9	8.5	62.3	71.4	69.3	92.4
		(42.8)	(54.0)	(61.7)	(50.9)	(64.5)	(63.1)	(61.3)	(59.1)	(69.9)	(71.7)	(66.8)	(78.6)	(74.2)	(70.9)
	5	2.3	3.4	25.6	62.7	62.8	78.3	81.8	4.1	5.5	35.3	74.2	74.1	86.7	89.7
		(61.8)	(65.3)	(78.2)	(66.7)	(67.5)	(80.9)	(77.8)	(75.0)	(76.9)	(88.3)	(80.0)	(80.4)	(89.6)	(87.9)
	10	13.7	19.7	28.9	77.5	77.8	63.7	66.8	21.7	28.1	39.0	86.7	86.7	76.5	78.3
		(81.1)	(83.1)	(72.1)	(84.3)	(85.2)	(73.2)	(70.8)	(87.9)	(88.8)	(79.6)	(90.3)	(92.0)	(80.4)	(78.6)
	15	21.9	26.4	27.5	77.9	75.4	52.2	57.5	31.5	38.8	37.5	87.7	85.6	64.8	69.0
		(83.7)	(84.6)	(60.7)	(85.6)	(85.9)	(60.9)	(59.3)	(90.2)	(89.9)	(71.8)	(91.6)	(91.7)	(72.5)	(70.5)

Annexe B

DÉMONSTRATIONS DES PRINCIPAUX RÉSULTATS

Dans cette section, quelques démonstrations des principaux résultats obtenus sont présentées. Une argumentation alternative est présentée dans l'article mais des calculs plus directs peuvent être d'intérêt en eux-mêmes.

Obtention de la forme quadratique de \mathcal{Q}_n

Afin d'obtenir de meilleures expressions de la moyenne et variance de la variable aléatoire \mathcal{Q}_n définie par (2.0.9) intervenant dans la définition des statistiques de test de Duchesne et Roy (2004) et Duchesne (2005a), nous avons exprimé \mathcal{Q}_n comme une forme quadratique. Le résultat matriciel suivant (Harville, 1997, théorème 16.2.2) est nécessaire au développement de la forme quadratique de \mathcal{Q}_n :

$$\text{tr}(\mathbf{A}^\top \mathbf{B} \mathbf{C} \mathbf{D}^\top) = (\text{vec}(\mathbf{A}))^\top (\mathbf{D} \otimes \mathbf{B}) (\text{vec}(\mathbf{C})),$$

où les matrices \mathbf{A} , \mathbf{B} , \mathbf{C} et \mathbf{D} sont des matrices de dimensions appropriées. Ainsi, il découle directement en utilisant cette propriété les égalités suivantes :

$$\begin{aligned} \mathcal{Q}_n &= n \sum_{j=1}^{n-1} k^2(j/p_n) \text{tr}\{\mathbf{C}_{\hat{\mathbf{a}}}^\top(j) \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \mathbf{C}_{\hat{\mathbf{a}}}(j) \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0)\} \\ &= n \sum_{j=1}^{n-1} k^2(j/p_n) \mathbf{c}_{\hat{\mathbf{a}}}^\top(j) \underbrace{\{\mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0)\}}_{\mathbf{W}} \mathbf{c}_{\hat{\mathbf{a}}}(j) \\ &= n \mathbf{c}_{\hat{\mathbf{a}}}^\top \text{diag}(k^2(1/p_n) \mathbf{W}, \dots, k^2((n-1)/p_n) \mathbf{W}) \mathbf{c}_{\hat{\mathbf{a}}} \\ &= n \mathbf{c}_{\hat{\mathbf{a}}}^\top \mathbf{K}^* \mathbf{c}_{\hat{\mathbf{a}}}, \end{aligned}$$

où $\mathbf{c}_{\hat{\mathbf{a}}}(j) = \text{vec}\{\mathbf{C}_{\hat{\mathbf{a}}}(j)\}$, $\mathbf{c}_{\hat{\mathbf{a}}} = (\mathbf{c}_{\hat{\mathbf{a}}}^\top(1), \dots, \mathbf{c}_{\hat{\mathbf{a}}}^\top(n-1))^\top$, $\mathbf{K}^* = \mathbf{K}^2 \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0)$ et $\mathbf{K} = \text{diag}(k_1, \dots, k_{n-1})$.

□

Afin d'obtenir la moyenne et la variance de la forme quadratique de la variable aléatoire \mathcal{Q}_n , nous avons besoin de la distribution des autocovariances résiduelles pour un processus VAR(p). La distribution des autocovariances résiduelles dans le cas de séries chronologiques stationnaires et partiellement non-stationnaires ont été obtenues respectivement par Li et McLeod (1981) et par Brüggemann *et al.* (2006). Nous pouvons formuler de manière générale la distribution des autocovariances résiduelles par :

$$\sqrt{n}\mathbf{c}_{\hat{\mathbf{a}}} \xrightarrow{\mathcal{L}} N_{d^2M}(\mathbf{0}, \mathbf{\Omega}) = \mathbf{Z}_M = (\mathbf{z}_1^\top, \dots, \mathbf{z}_M^\top)^\top,$$

où $\mathbf{\Omega}$ est soit donné par (2.0.3) si le processus est stationnaire ou soit donné par (2.0.12) si le processus est partiellement non-stationnaire.

Obtention de la moyenne de la variable aléatoire \mathcal{Q}_n

Nous calculons ci-après le premier moment d'une forme idéalisée de la forme quadratique \mathcal{Q}_n , que nous notons \mathcal{Q}_n^0 et que nous définissons par $\mathcal{Q}_n^0 = \mathbf{Z}_{n-1}^\top \mathbf{K}_0^* \mathbf{Z}_{n-1}$, où \mathbf{K}_0^* est une matrice fixée qui peut être interprétée comme la limite probabiliste de la matrice aléatoire \mathbf{K}^* , et où nous présumons que $n^{1/2}\mathbf{c}_{\hat{\mathbf{a}}_i} = \mathbf{z}_i + o_p(1)$ uniformément. Maintenant, pour n fixé,

$$\begin{aligned} E(\mathcal{Q}_n^0) &= E\{\mathbf{Z}_{n-1}^\top \mathbf{K}_0^* \mathbf{Z}_{n-1}\} \\ &= E\{\text{tr}[\mathbf{Z}_{n-1}^\top \mathbf{K}_0^* \mathbf{Z}_{n-1}]\} \\ &= E\{\text{tr}[\mathbf{K}_0^* \mathbf{Z}_{n-1} \mathbf{Z}_{n-1}^\top]\} \\ &= \text{tr}\{E[\mathbf{K}_0^* \mathbf{Z}_{n-1} \mathbf{Z}_{n-1}^\top]\} \\ &= \text{tr}\{\mathbf{K}_0^* E[\mathbf{Z}_{n-1} \mathbf{Z}_{n-1}^\top]\} \\ &= \text{tr}[\mathbf{K}_0^* \mathbf{\Omega}]. \end{aligned}$$

□

La dérivation précédente suggère d'estimer le premier moment de \mathcal{Q}_n par $\text{tr}[\mathbf{K}^* \hat{\mathbf{\Omega}}]$, où $\hat{\mathbf{\Omega}}$ est un estimateur convergent de la matrice $\mathbf{\Omega}$.

Obtention de la variance de la variable aléatoire Q_n

Le même genre de calculs permettent d'obtenir la variance de la forme quadratique idéalisée Q_n^0 .

$$\begin{aligned}
 \text{var}(Q_n^0) &= \text{var}\{n\mathbf{c}_{\hat{\mathbf{a}}}^\top \mathbf{K}_0^* \mathbf{c}_{\hat{\mathbf{a}}}\} \\
 &= \text{var} \left[n \sum_{i=1}^{d^2(n-1)} \sum_{j=1}^{d^2(n-1)} k_{ij0}^* \mathbf{c}_{\hat{\mathbf{a}}_i} \mathbf{c}_{\hat{\mathbf{a}}_j} \right] \\
 &= \text{cov} \left[n \sum_{i=1}^{d^2(n-1)} \sum_{j=1}^{d^2(n-1)} k_{ij0}^* \mathbf{c}_{\hat{\mathbf{a}}_i} \mathbf{c}_{\hat{\mathbf{a}}_j}, n \sum_{r=1}^{d^2(n-1)} \sum_{s=1}^{d^2(n-1)} k_{rs0}^* \mathbf{c}_{\hat{\mathbf{a}}_r} \mathbf{c}_{\hat{\mathbf{a}}_s} \right] \\
 &= n^2 \sum_{i=1}^{d^2(n-1)} \sum_{j=1}^{d^2(n-1)} \sum_{r=1}^{d^2(n-1)} \sum_{s=1}^{d^2(n-1)} k_{ij0}^* k_{rs0}^* \text{cov}[\mathbf{c}_{\hat{\mathbf{a}}_i} \mathbf{c}_{\hat{\mathbf{a}}_j}, \mathbf{c}_{\hat{\mathbf{a}}_r} \mathbf{c}_{\hat{\mathbf{a}}_s}].
 \end{aligned}$$

Or, présumant encore que $n^{1/2}\mathbf{c}_{\hat{\mathbf{a}}_i} = \mathbf{z}_i + o_p(1)$ uniformément, nous avons que $\text{cov}[\mathbf{z}_i \mathbf{z}_j, \mathbf{z}_k \mathbf{z}_l]$ est obtenu à partir d'un résultat de Anderson (1984, pp. 49-50) :

$$\begin{aligned}
 \text{cov}[\mathbf{z}_i \mathbf{z}_j, \mathbf{z}_k \mathbf{z}_l] &= E\{\mathbf{z}_i \mathbf{z}_j \mathbf{z}_k \mathbf{z}_l\} - E\{\mathbf{z}_i \mathbf{z}_j\} E\{\mathbf{z}_k \mathbf{z}_l\} \\
 &= \omega_{ir} \omega_{js} + \omega_{jr} \omega_{is} + \omega_{ij} \omega_{rs} - \omega_{ij} \omega_{rs} \\
 &= \omega_{ir} \omega_{js} + \omega_{jr} \omega_{is}.
 \end{aligned}$$

Ainsi, à l'ordre $o(1)$, les relations suivantes sont valides :

$$\begin{aligned}
 \text{var}(Q_n^0) &= n^2 \sum_{i=1}^{d^2(n-1)} \sum_{j=1}^{d^2(n-1)} \sum_{r=1}^{d^2(n-1)} \sum_{s=1}^{d^2(n-1)} k_{ij0}^* k_{rs0}^* \left(\frac{\omega_{ir} \omega_{js}}{n} + \frac{\omega_{jr} \omega_{is}}{n} \right) \\
 &= \sum_{i=1}^{d^2(n-1)} \sum_{j=1}^{d^2(n-1)} \sum_{r=1}^{d^2(n-1)} \sum_{s=1}^{d^2(n-1)} k_{ij0}^* k_{rs0}^* \omega_{ir} \omega_{js} \\
 &\quad + \sum_{i=1}^{d^2(n-1)} \sum_{j=1}^{d^2(n-1)} \sum_{r=1}^{d^2(n-1)} \sum_{s=1}^{d^2(n-1)} k_{ij0}^* k_{rs0}^* \omega_{jr} \omega_{is}.
 \end{aligned}$$

Selon le résultat algébrique bien connu sur les produits de kronecker : $(\mathbf{A} \otimes \mathbf{B})^\top = \mathbf{A}^\top \otimes \mathbf{B}^\top$ (voir par exemple Lütkepohl 1993, p. 464), nous notons que la

matrice \mathbf{K}^* et la limite \mathbf{K}_0^* sont symétriques. Par exemple, dans le cas de \mathbf{K}^* :

$$\begin{aligned}\mathbf{K}^{*\top} &= [\mathbf{K}^2 \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0)]^\top \\ &= \mathbf{K}^2 \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \otimes \mathbf{C}_{\hat{\mathbf{a}}}^{-1}(0) \\ &= \mathbf{K}^*.\end{aligned}$$

Conséquemment, nous déduisons de ce dernier résultat que :

$$\begin{aligned}\text{var}(\mathcal{Q}_n^0) &= \sum_{i=1}^{d^2(n-1)} \sum_{j=1}^{d^2(n-1)} \sum_{r=1}^{d^2(n-1)} \sum_{s=1}^{d^2(n-1)} k_{ij0}^* \omega_{js} k_{sr0}^* \omega_{ri} \\ &\quad + \sum_{i=1}^{d^2(n-1)} \sum_{j=1}^{d^2(n-1)} \sum_{r=1}^{d^2(n-1)} \sum_{s=1}^{d^2(n-1)} k_{ij0}^* \omega_{jr} k_{rs0}^* \omega_{si}.\end{aligned}$$

Or, par un résultat élémentaire sur la trace d'un produit de matrices nous obtenons :

$$\text{tr}(\mathbf{ABCD}) = \sum_i \sum_j \sum_k \sum_l a_{ij} b_{jk} c_{kl} d_{li}.$$

Finalement, nous obtenons

$$\begin{aligned}\text{var}(\mathcal{Q}_n^0) &= \text{tr}[\mathbf{K}_0^* \mathbf{\Omega} \mathbf{K}_0^* \mathbf{\Omega}] + \text{tr}[\mathbf{K}_0^* \mathbf{\Omega} \mathbf{K}_0^* \mathbf{\Omega}] \\ &= 2\text{tr}[\mathbf{K}_0^* \mathbf{\Omega} \mathbf{K}_0^* \mathbf{\Omega}].\end{aligned}$$

□

Cette dernière dérivation suggère d'estimer le second moment de la forme quadratique par $2\text{tr}[\mathbf{K}^* \hat{\mathbf{\Omega}} \mathbf{K}^* \hat{\mathbf{\Omega}}]$, où encore une fois $\hat{\mathbf{\Omega}}$ est un estimateur convergent de la matrice $\mathbf{\Omega}$.

BIBLIOGRAPHIE

- [1] AHN, S. K. ET REINSEL, G. C. (1990), Estimation for Partially Nonstationary Multivariate Autoregressive Models, *Journal of the American Statistical Association*, **85**, 813–823.
- [2] ANDERSON, T. W. (1984), *An Introduction to Multivariate Statistical Analysis*, 2^e édition, Wiley : New York.
- [3] BOX, G. E. P. ET TIAO, G. C. (1977), A Canonical Analysis of Multiple Time Series, *Biometrika*, **64**, 355–365.
- [4] BRILLINGER, D. R. (1981), *Time Series : Data Analysis and Theory*, édition rallongée, Holden-Day : San Francisco.
- [5] BROCKWELL, P. J. ET DAVIS, R. A. (1991), *Time Series : Theory and Methods*, 2^e édition, Springer-Verlag : New York.
- [6] BROCKWELL, P. J. ET DAVIS, R. A. (2002), *Introduction to Time Series and Forecasting*, 2^e édition, Springer : New York.
- [7] BRÜGGEMANN, R., LÜTKEPOHL, H. ET SAIKKONEN, P. (2006), Residual Autocorrelation Testing for Vector Error Correction Models, *Journal of Econometrics*, **134**, 579–604.
- [8] CÉDRAS, L. (1997), *Test de non corrélation de deux séries chronologiques multivariées non stationnaires*, Mémoire de maîtrise, Département de mathématiques et de statistique, Université de Montréal, Montréal.
- [9] CHEN, W. W. ET DEO, R. S. (2004a), Power Transformations to Induce Normality and their Applications, *Journal of the Royal Statistical Society, Ser. B*, **66**, 117–130.
- [10] CHEN, W. W. ET DEO, R. S. (2004b), A Generalized Portmanteau Goodness-of-Fit Test for Time Series Models, *Econometric Theory*, **20**, 382–416.

- [11] CHITTURI, R. V. (1974), Distribution of Residual Autocorrelations in Multiple Autoregressive Schemes, *Journal of the American Statistical Association*, **69**, 928–934.
- [12] CHITTURI, R. V. (1976), Distribution of Multivariate White Noise Autocorrelations, *Journal of the American Statistical Association*, **71**, 223–226.
- [13] DUCHESNE, P. (2005a), Testing for Serial Correlation of Unknown Form in Cointegrated Time Series Models, *Annals of the Institute of Statistical Mathematics*, **57**, 575–595.
- [14] DUCHESNE, P. (2005b), On the Asymptotic Distribution of Residual Autocovariances in VARX Models with Applications, *Test*, **14**, 449–473.
- [15] DUCHESNE, P. ET ROY, R. (2004), On Consistent Testing for Serial Correlation of Unknown Form in Vector Time Series Models, *Journal of Multivariate Analysis*, **89**, 148–180.
- [16] FULLER, W. A. (1996), *Introduction to Statistical Time Series*, 2^e édition, Wiley : New York.
- [17] HANNAN, E. J. (1970), *Multiple Time Series*, Wiley : New York.
- [18] HARVILLE, D. A. (1997), *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag : Berlin.
- [19] HILLMER, S. C. ET TIAO, G. C. (1979), Likelihood Function of Stationary Multiple Autoregressive Moving Average Models, *Journal of the American Statistical Association*, **74**, 652–660.
- [20] HONG, Y. (1996), Consistent Testing for Serial Correlation of Unknown Form, *Econometrica*, **64**, 837–864.
- [21] HOSKING, J. (1980), The Multivariate Portmanteau Statistic, *Journal of the American Statistical Association*, **75**, 602–608.
- [22] HOSKING, J. (1981a), Equivalent Forms of the Multivariate Portmanteau Statistic, *Journal of the Royal Statistical Society, Ser. B*, **43**, 261–262, (Corr : V51/1989, p. 303).
- [23] HOSKING, J. (1981b), Lagrange-Multiplier Tests of Multivariate Time Series Models, *Journal of the Royal Statistical Society, Ser. B*, **43**, 219–230.

- [24] LI, W. K. ET MCLEOD, A. I. (1981), Distribution of the Residual Auto-correlations in Multivariate ARMA Time Series Models, *Journal of the Royal Statistical Society, Ser. B*, **43**, 231–239.
- [25] LÜTKEPOHL, H. (1993), *Introduction to Multiple Time Series Analysis*, 2^e édition, Springer-Verlag : Berlin.
- [26] PHAM, D. T., ROY, R. ET CÉDRAS L. (2000), Tests for Non-Correlation of Two Cointegrated ARMA Time Series, Rapport de recherche CRM-2649, Université de Montréal, Janvier 2000.
- [27] PHAM, D. T., ROY, R. ET CÉDRAS L. (2003), Tests for Non-Correlation of Two Cointegrated ARMA Time Series, *Journal of Time Series Analysis*, **24**, 553-577.
- [28] PRIESTLEY, M. B. (1981), *Spectral Analysis and Time Series*, Academic Press : New York.
- [29] REINSEL, G. C. (1993), *Elements of Multivariate Time Series Analysis*, Springer-Verlag : New York.
- [30] SHUMWAY, R. H. ET STOFFER, D. S. (2006), *Time Series Analysis and its Applications : With R Examples* , 2^e édition, Springer : New York.
- [31] TSAY, R. S. (2005), *Analysis of Financial Time Series*, 2^e édition, Wiley : Hoboken.
- [32] WEI, W. W.-S. (2006), *Time Series Analysis : Univariate and Multivariate Methods*, 2^e édition, Pearson Addison Wesley : Boston.