

Université de Montréal

Sur la validation des modèles de séries
chronologiques spatio-temporelles multivariées

par

Robinson Saint-Frard

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

juin 2011

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

**Sur la validation des modèles de séries
chronologiques spatio-temporelles multivariées**

présenté par

Robinson Saint-Frard

a été évalué par un jury composé des personnes suivantes :

Martin Bilodeau

(président-rapporteur)

Pierre Duchesne

(directeur de recherche)

David Haziza

(membre du jury)

Mémoire accepté le:

23 juin 2011

SOMMAIRE

Le présent mémoire porte sur les séries chronologiques qui en plus d'être observées dans le temps, présentent également une composante spatiale. Plus particulièrement, nous étudions une certaine classe de modèles, les modèles autorégressifs spatio-temporels généralisés, ou GSTAR. Dans un premier temps, des liens sont effectués avec les modèles vectoriels autorégressifs (VAR). Nous obtenons explicitement la distribution asymptotique des autocovariances résiduelles pour les modèles GSTAR en supposant que le terme d'erreur est un bruit blanc gaussien, ce qui représente une première contribution originale. De ce résultat, des tests de type portemanteau sont proposés, dont les distributions asymptotiques sont étudiées. Afin d'illustrer la performance des statistiques de test, une étude de simulations est entreprise où des modèles GSTAR sont simulés et correctement ajustés. La méthodologie est illustrée avec des données réelles. Il est question de la production mensuelle de thé en Java occidental pour 24 villes, pour la période janvier 1992 à décembre 1999.

Mots-clés : Séries chronologiques multivariées ; statistiques portemanteaux ; autocovariances résiduelles matricielles ; paramétrisation structurée.

SUMMARY

In this master thesis, time series models are studied, which have also a spatial component, in addition to the usual time index. More particularly, we study a certain class of models, the *Generalized Space-Time AutoRegressive* (GSTAR) time series models. First, links are considered between *Vector AutoRegressive* models (VAR) and GSTAR models. We obtain explicitly the asymptotic distribution of the residual autocovariances for the GSTAR models, assuming that the error term is a Gaussian white noise, which is a first original contribution. From that result, test statistics of the portmanteau type are proposed, and their asymptotic distributions are studied. In order to illustrate the behaviour of the test statistics, a simulation study is conducted where GSTAR models are simulated and correctly fitted. The methodology is illustrated with monthly real data concerning the production of tea in west Java for 24 cities from the period January 1992 to December 1999.

Key words : Multivariate time series ; Portmanteau test statistic ; Residual autocovariance matrices ; Structured parameterization.

TABLE DES MATIÈRES

Sommaire	iii
Summary	iv
Liste des figures	vii
Liste des tableaux	viii
Remerciements	1
Introduction	2
Chapitre 1. Préliminaires sur les processus multivariés	6
1.1. Processus multivariés	6
1.1.1. Rappels sur les processus multivariés.....	6
1.1.1.1. Analyse du processus VAR(1).....	9
1.1.1.2. La fonction d'autocovariance d'un processus VAR(p).....	13
1.1.2. Estimation des paramètres d'un processus stationnaire	18
Chapitre 2. Analyse de données spatio-temporelles multivariées.	21
2.1. Concepts généraux sur les modèles spatiaux.....	21
2.2. Modèles spatio-temporels multivariés.....	26
2.2.1. Analyse du processus STAR(1_1).....	28
2.2.2. Analyse du processus GSTAR(1_1).....	29
2.3. Méthodes d'identification d'un processus GSTAR(p_λ)	30
2.3.1. Fonctions d'autocovariance d'un processus GSTAR(p_λ).....	34

2.4.	Estimation des paramètres d'un modèle $\text{GSTAR}(p_\lambda)$	34
Chapitre 3. Statistiques de test pour diagnostiquer des modèles		
	$\text{GSTAR}(p_\lambda)$	38
3.1.	Autocovariances résiduelles matricielles.....	38
3.1.1.	Statistiques de test basées sur les matrices d'autocovariances résiduelles pour les modèles VARMA.....	41
3.2.	Tests de type portemanteau pour un modèle GSTAR	43
3.3.	Distribution des autocovariances résiduelles pour un modèle GSTAR estimé par moindres carrés ordinaires	45
Chapitre 4. Étude de simulations		
4.1.	Algorithmes pour l'étude de simulations.....	50
4.2.	Description de la simulation.....	51
4.3.	Résultats de simulation	53
Chapitre 5. Analyse de données réelles		
5.1.	Analyse et modélisation de la série.....	57
5.1.1.	Estimations des paramètres des modèles retenus.....	58
5.1.2.	Choix du modèle suivant le test portemanteau	61
Conclusion		
Annexe A. Simulations complémentaires		
Annexe B. Données sur la production de thé		
Annexe C. Programme développé en R pour les simulations		
Bibliographie		

LISTE DES FIGURES

2.1	Types de contiguïté conventionnels	22
2.2	Exemples d'ordres spatiaux d'ordre un, deux, trois et quatre.....	23
5.1	Différents sites de la production de thé en Java occidental, Indonésie .	62
B.1	Graphiques des 24 séries chronologiques portant sur la production de thé, pour chacun des sites en Java occidental	B-ii

LISTE DES TABLEAUX

4.1	Paramètres Φ_0 et Φ_1 pour générer la série chronologique multivariée $\mathbf{Y}_t, t = 1, \dots, n$.	52
4.2	Matrice de poids \mathbf{W} pour $N = 10$.	52
4.3	Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour trois sites ($N = 3$).	55
4.4	Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour dix sites ($N = 10$).	56
5.1	Estimateurs des paramètres des modèles $\text{GSTAR}(1_{1_1})$ et $\text{GSTAR}(2_{1_1})$.	59
5.2	Estimateurs des paramètres du modèle $\text{GSTAR}(3_{1_1})$.	60
5.3	Résultats du test d'adéquation pour la série de 24 sites.	61
A.1	Matrice de poids \mathbf{W} pour $N = 10$.	A-ii
A.2	Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour trois sites ($N = 3$) pour \mathbf{W}_1 .	A-iii
A.3	Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour trois sites ($N = 3$) pour \mathbf{W}_2 .	A-iv
A.4	Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour dix sites ($N = 10$) pour \mathbf{W}_1 .	A-v
A.5	Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour dix sites ($N = 10$) pour \mathbf{W}_2 .	A-vi

REMERCIEMENTS

J'aimerais tout d'abord remercier Dieu pour tout en particulier la santé et l'intelligence qu'il m'a données tout au long des études. Je tiens aussi à remercier mon directeur de recherche, Pierre Duchesne, pour ses conseils, sa patience et son support financier. Je le remercie également pour sa continuelle disponibilité, la confiance et l'aide indispensable qu'il m'a apportées au cours des deux années d'études.

J'exprime mon entière reconnaissance envers tous les professeurs du Département de mathématiques et de statistique, en particulier Christian Léger et David Haziza. Je remercie également Miguel Chagnon.

Finalement, une attention toute particulière est dirigée vers ma famille particulièrement ma femme, pour son courage.

INTRODUCTION

Depuis les années soixante-dix, les chercheurs analysent la prise en compte dans les modèles économétriques des données géographiques, c'est-à-dire des observations d'une ou plusieurs variables mesurées dans le temps pour des localisations particulières réparties dans l'espace. Ainsi, cette prise en considération du temps et de l'espace a nécessité dans ce domaine le développement de certains concepts dont l'autocorrélation spatiale, qui se réfère à la mesure potentielle de dépendance entre des observations géographiques observées dans l'espace, voir Le Gallo (2000). En fait, Le Gallo (2000) présente le contexte historique qui a mené aux développements en statistique et en économétrie spatiales. Elle présente également les outils statistiques pertinents dans ce contexte.

Cliff *et* Ord (1973) ont présenté les premières analyses en statistique et en économétrie spatiales. Depuis, plusieurs auteurs se sont intéressés à ces modèles. D'un point de vue strictement théorique, les modèles multivariés courants, tels les modèles vectoriels autorégressifs (VAR) ou encore les modèles autorégressifs moyennes mobiles multivariés (VARMA) peuvent être utilisés dans le contexte de données spatio-temporelles. Cependant, d'un point de vue pratique, notre contexte se situe dans des situations où il y a relativement peu de données, comparativement au nombre de paramètres impliqué par la formulation d'un modèle multivarié global tenant compte de l'ensemble des séries chronologiques en cause. Par exemple, un simple modèle VAR(1) fait intervenir un maximum de N^2 paramètres indépendants dans la matrice autorégressive, où N est le nombre de séries chronologiques univariées. Avec des données macro-économiques (par exemple, faisant intervenir des données mensuelles ou annuelles), la taille de l'échantillon,

notée n , peut être petite par rapport à N^2 . Par exemple, si nous disposons de données annuelles sur une période de cent ans, pour chacun des 50 états américains, nous trouvons $N^2 = 2500$ qui est beaucoup plus grand que $n = 100$. Afin de modéliser de telles données, il importe d'avoir des modèles avec une structure tenant compte de la composante temporelle et spatiale. Pfeifer *et* Deutsch (1980) ont présenté la théorie de l'estimation et des tests d'hypothèses pour les modèles multivariés spatiaux. Plus particulièrement, la classe des modèles autorégressifs moyennes mobiles spatio-temporels (STARMA) a été étudiée. Di Giacinto (2006) *et* Terzi (1997) ont généralisé le modèle présenté par ces derniers et ont étudié les conditions de stationnarité tout en montrant que le modèle généralisé peut s'écrire sous la forme d'un modèle multivarié VARMA. Borovkova *et al.* (2008) ont étudié le comportement des estimateurs des moindres carrés dans le cadre de certains modèles spatiaux généralisés. Leurs travaux sont particulièrement généraux, permettant un terme d'erreur pas forcément constitué de vecteurs aléatoires indépendants, mais satisfaisant une condition du type différence de martingales.

Dans cette étude, nous nous intéressons plus particulièrement aux autocovariances résiduelles. Ces mesures de dépendance représentent des outils de base dans la validation des modèles de séries chronologiques. Une difficulté dans notre contexte concerne le fait que lors de la phase d'estimation, les modèles ajustés ont une structure particulière. Ainsi, les résultats déjà existants pour les modèles VARMA ne sont pas directement applicables car il s'agit de modèles VARMA que nous qualifierons de structurés ou encore possédant une paramétrisation structurée. À notre connaissance, les travaux de Ahn (1988) sont parmi les premiers à s'être penchés sur le comportement des autocovariances résiduelles pour une classe particulière de modèles VAR avec une structure générale. Plus récemment, Boubacar Mainassara *et* Francq (2009) ont étudié l'estimation pour une classe générale de modèles VARMA structurés, où les hypothèses sur le terme d'erreur sont générales. Plus particulièrement, la méthode d'estimation considérée est la méthode du quasi maximum de vraisemblance. Plus précisément, ces auteurs étudient la convergence et la normalité asymptotique des estimateurs des paramètres

dans leur contexte. Boubacar Mainassara (2009) a étudié quant à lui des statistiques de test de type portemanteau pour des modèles VARMA structurés lorsque le terme d'erreur est un bruit blanc faible.

Dans ce mémoire, nous montrons dans un premier temps que le modèle spatial généralisé GSTAR peut être reformulé comme un modèle VAR structuré. Afin de diagnostiquer le modèle, nous tentons alors de récupérer le cadre de travail de Ahn (1988), afin de pouvoir appliquer directement ses résultats. Des liens entre les méthodes d'estimation considérées par Ahn (1988) et les techniques générales de Borovkova *et al.* (2008) sont également effectués. Sous certaines conditions, les méthodes sont équivalentes. À partir de notre approche, nous développons des statistiques de test de type portemanteau.

Dans un second temps, nous illustrons nos résultats théoriques en effectuant une étude de simulations où notre principal objectif est le niveau empirique de la statistique portemanteau en considérant un cas particulier du modèle GSTAR, soit le modèle GSTAR(1₁), sous l'hypothèse de normalité et l'indépendance du terme d'erreur.

Ce mémoire est constitué de cinq chapitres. Dans le premier chapitre, nous introduisons certaines notions et définitions importantes sur les séries chronologiques multivariées. Les méthodes d'identification et d'estimation des modèles VARMA sont aussi analysées dans ce chapitre. Les notions sur la statistique spatiale ainsi que les modèles spatio-temporels sont décrits dans le deuxième chapitre. Au Chapitre 3, nous proposons des tests de type portemanteau pour diagnostiquer un processus GSTAR. Au quatrième chapitre, nous décrivons les différentes étapes de nos simulations afin d'illustrer nos résultats théoriques. Les méthodes utilisées pour générer les séries chronologiques, pour estimer les paramètres du modèle simulé et pour calculer les statistiques de test sont décrites dans cette section. Ces méthodes ont toutes été programmées en **R** et le code utilisé est fourni en annexe. Les résultats de simulation du niveau de notre test portemanteau sont aussi analysés dans ce chapitre. Finalement au dernier chapitre, nous effectuons

une analyse avec des données réelles portant sur la production de thé en Java occidental, Indonésie, pour vingt-quatre sites. Les sites sont identifiés sur une carte géographique, qui a été gracieusement fournie par S. Borovkova. Les résultats de cette analyse de données ainsi que l'application de notre test portemanteau sont présentés dans ce chapitre.

Chapitre 1

PRÉLIMINAIRES SUR LES PROCESSUS MULTIVARIÉS

Dans ce chapitre, nous présentons en premier lieu une brève revue des notions sur les processus multivariés. Dans la première partie, des notions générales sur les processus linéaires multivariés (ou vectoriels) sont présentées. Plusieurs notions fondamentales introduites dans le cas d'une seule variable, c'est-à-dire dans le cas univarié, trouvent leurs extensions dans le cas de plusieurs variables, que nous appelons multivarié. Pour une exposition des notions sur les processus linéaires univariés, voir par exemple Brockwell *et* Davis (2002). Nous exposons également une brève description des propriétés de stationnarité et de stabilité des processus vectoriels. Nous introduisons également les autocovariances échantillonales et résiduelles, et discutons l'estimation des processus vectoriels.

1.1. PROCESSUS MULTIVARIÉS

Dans ce chapitre, on considère le processus stochastique $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$, où \mathbb{Z} est l'ensemble des entiers naturels, et $\mathbf{Y}_t = (Y_t(1), \dots, Y_t(N))^\top$ est un vecteur aléatoire de dimension N . On observe une série chronologique $\mathbf{Y}_t, t = 1, \dots, n$. Ceci est précisé dans la définition présentée dans la prochaine sous-section.

1.1.1. Rappels sur les processus multivariés

Définition 1.1.1. *Une série temporelle multivariée est une réalisation finie de taille n d'un processus stochastique multivarié $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$, où le vecteur aléatoire*

\mathbf{Y}_t est de dimension N et $Y_t(i)$ représente la i ème composante de \mathbf{Y}_t . Les propriétés du second ordre d'un processus stochastique sont spécifiées par le vecteur des moyennes :

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

et les matrices de covariance sont définies comme suit :

$$\boldsymbol{\Gamma}(t, t-h) = E\{(\mathbf{Y}_t - \boldsymbol{\mu}_t)(\mathbf{Y}_{t-h} - \boldsymbol{\mu}_{t-h})^\top\} = (\gamma_{ij}(t, t-h))_{i,j=1}^N. \quad (1.1.1)$$

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 $i \neq j$).

Une classe de processus importante possède une propriété qui garantit une certaine uniformité dans le temps. Cette propriété est celle dite de stationnarité.

Définition 1.1.2. *Un processus est stationnaire au second ordre si les deux premiers moments sont invariants par rapport à l'indice temporel t . Plus précisément :*

$$E(\mathbf{Y}_t) = \boldsymbol{\mu}, \quad \forall t,$$

$$\boldsymbol{\Gamma}(t, t-h) = E\{(\mathbf{Y}_t - \boldsymbol{\mu}_t)(\mathbf{Y}_{t-h} - \boldsymbol{\mu}_{t-h})^\top\} = \boldsymbol{\Gamma}(h), \quad \forall h.$$

De plus, on présume que les espérances $E\{Y_t^2(i)\}$ existent, $i = 1, \dots, N$.

On note le léger abus de notation, car dans la définition précédente $\boldsymbol{\Gamma}(\cdot, \cdot)$ est une fonction de deux variables, alors que dans le cas stationnaire $\boldsymbol{\Gamma}(\cdot)$ n'est fonction que du délai h . La matrice d'autocovariance mesure l'intensité de la dépendance linéaire entre les vecteurs définissant le processus. Un cas particulier de processus stationnaire est le processus bruit blanc. Ce processus, qui est très utilisé en analyse des séries chronologiques, mérite d'être discuté en détails. Ceci fait l'objet de la Définition 1.1.3.

Définition 1.1.3. *Le processus $\{\mathbf{a}_t\}$ est un bruit blanc si et seulement si :*

- (i) $E(\mathbf{a}_t) = \mathbf{0}$,
- (ii) $E(\mathbf{a}_t \mathbf{a}_t^\top) = \text{Var}(\mathbf{a}_t) = \boldsymbol{\Sigma}_a$,
- (iii) $E(\mathbf{a}_t \mathbf{a}_s^\top) = \mathbf{0}, \forall t \neq s$.

En général, il n'est pas nécessaire que les vecteurs constituant le processus bruit blanc soient indépendants mais seulement non-corrélés. Lorsque les vecteurs aléatoires \mathbf{a}_t sont non-corrélés mais dépendants, nous parlons alors de bruit blanc faible. Si les éléments de la séquence sont aussi indépendants alors on dit que le bruit blanc est fort.

Pour caractériser la dépendance linéaire d'un processus, on utilise la matrice d'autocorrélation qui est essentiellement une version standardisée de la matrice d'autocovariance.

Définition 1.1.4. *La matrice d'autocorrélation au délai h , $\mathbf{R}(h)$, est donnée par :*

$$\mathbf{R}(h) = \mathbf{D}^{-1}\mathbf{\Gamma}(h)\mathbf{D}^{-1},$$

où

$$\mathbf{D} = \text{diag}(\gamma_{11}^{-1/2}(0), \dots, \gamma_{NN}^{-1/2}(0)),$$

et la matrice $\mathbf{\Gamma}(h)$ est définie dans (1.1.1). De façon plus précise, les éléments de la matrice $\mathbf{R}(h) = (\rho_{ij}(h))_{i,j=1}^N$ sont :

$$\rho_{ij}(h) = \frac{\gamma_{ij}(h)}{\{\gamma_{ii}(0)\gamma_{jj}(0)\}^{1/2}}.$$

Ayant exposé la notion de stationnarité, on discute certains modèles et on énonce des résultats qui assurent la stationnarité lorsque certaines conditions sont satisfaites. On présente maintenant la définition d'un processus stationnaire autorégressif vectoriel. On introduit B l'opérateur de retard temporel, c'est-à-dire $B\mathbf{Y}_t = \mathbf{Y}_{t-1}$.

Définition 1.1.5. *Un processus stationnaire $\{\mathbf{Y}_t\}$ autorégressif de dimension N et d'ordre p , noté $\text{VAR}(p)$, est défini par :*

$$\mathbf{\Phi}(B)\mathbf{Y}_t = \boldsymbol{\mu} + \mathbf{a}_t, \quad t \in \mathbb{Z}, \quad (1.1.2)$$

où $\mathbf{\Phi}(B) = \mathbf{I}_N - \sum_{i=1}^p \mathbf{\Phi}_i B^i$ est l'opérateur matriciel de dimension $N \times N$, qui est l'opérateur autorégressif multivarié tel que $\mathbf{\Phi}_p \neq \mathbf{0}$, $\boldsymbol{\mu}$ est un vecteur constant de dimension $N \times 1$ et $\{\mathbf{a}_t\}$ est un bruit blanc tel que $\text{Var}(\mathbf{a}_t) = \boldsymbol{\Sigma}_a$.

Nous parlons de $\text{VAR}(p)$ fort si le terme d'erreur $\{\mathbf{a}_t\}$ est un bruit blanc fort. On note qu'un $\text{VAR}(p)$ avec un terme d'erreur gaussien sera forcément un $\text{VAR}(p)$

fort. Sinon, il s'agira d'un VAR(p) faible. Dans la littérature portant sur les processus multivariés, des notions comme la causalité et l'inversibilité jouent un rôle crucial. Dans la littérature économétrique, la notion de stationnarité causale est appelée la stabilité d'un processus. Cette notion est intuitivement raisonnable, car une observation au temps t est alors expliquée par le passé de cette variable. On note qu'il existe des processus stationnaires dont \mathbf{Y}_t est expliquée par le futur. Ceci est discuté par exemple dans Shumway *et* Stoffer (2006, pp. 88-89), dans le cas d'un processus autorégressif univarié explosif. La notion de stabilité joue un rôle important dans les applications économétriques, plus particulièrement dans l'étude de méthodes qui trouveraient des applications dans l'élaboration d'un système de prévision. Voir par exemple Lütkepohl (2005). Le théorème suivant donne les conditions pour qu'un processus soit stable.

Théorème 1.1.1. *Le processus VAR(p) représenté par l'équation (1.1.2) est stable si :*

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

Dans ce cas, le processus VAR(p) admet la représentation linéaire suivante :

$$\mathbf{Y}_t = \sum_{i=0}^{\infty} \Psi_i \mathbf{a}_{t-i} = \Psi(B) \mathbf{a}_t,$$

où les matrices Ψ_i sont déterminées par :

$$\Psi(B) = \sum_{i=0}^{\infty} \Psi_i B^i = \Phi^{-1}(B),$$

et les coefficients $\{\Psi_i\}$ sont absolument sommables, c'est-à-dire que :

$$\sum_{i=0}^{\infty} \|\Psi_i\| < \infty,$$

avec $\|\cdot\|$ correspondant à la norme euclidienne d'une matrice.

La preuve de ce théorème se trouve dans Brockwell *et* Davis (1991, Section 11.3).

1.1.1.1. Analyse du processus VAR(1)

Examinons le cas où le processus $\{\mathbf{Y}_t\}$ est un VAR(1) admettant une écriture de la forme :

$$\mathbf{Y}_t = \boldsymbol{\nu} + \boldsymbol{\Phi}_1 \mathbf{Y}_{t-1} + \mathbf{a}_t. \quad (1.1.3)$$

En procédant à des remplacements successifs, il est possible d'obtenir une expression pour $\{\mathbf{Y}_t\}$ en fonction des erreurs $\{\mathbf{a}_t\}$. Ainsi, pour $t = 1$, on a l'expression suivante :

$$\mathbf{Y}_1 = \boldsymbol{\nu} + \boldsymbol{\Phi}_1 \mathbf{Y}_0 + \mathbf{a}_1.$$

Un premier remplacement donne :

$$\begin{aligned} \mathbf{Y}_2 &= \boldsymbol{\nu} + \boldsymbol{\Phi}_1 \mathbf{Y}_1 + \mathbf{a}_2, \\ &= \boldsymbol{\nu} + \boldsymbol{\Phi}_1(\boldsymbol{\nu} + \boldsymbol{\Phi}_1 \mathbf{Y}_0 + \mathbf{a}_1) + \mathbf{a}_2, \\ &= (\mathbf{I}_N + \boldsymbol{\Phi}_1)\boldsymbol{\nu} + \boldsymbol{\Phi}_1 \mathbf{a}_1 + \mathbf{a}_2 + \boldsymbol{\Phi}_1^2 \mathbf{Y}_0. \end{aligned}$$

En continuant les récursions, $\{\mathbf{Y}_t\}$ s'écrit sous cette nouvelle forme :

$$\mathbf{Y}_t = (\mathbf{I}_N + \boldsymbol{\Phi}_1 + \dots + \boldsymbol{\Phi}_1^{t-1})\boldsymbol{\nu} + \boldsymbol{\Phi}_1^t \mathbf{Y}_0 + \sum_{i=0}^{t-1} \boldsymbol{\Phi}_1^i \mathbf{a}_{t-i}. \quad (1.1.4)$$

À la limite, il est raisonnable de penser que le processus $\{\mathbf{Y}_t\}$ écrit en (1.1.4) admet l'écriture suivante, sous certaines conditions :

$$\mathbf{Y}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\Phi}_1^i \mathbf{a}_{t-i}, \quad (1.1.5)$$

où $\boldsymbol{\mu} = (\mathbf{I}_N - \boldsymbol{\Phi}_1)^{-1}\boldsymbol{\nu}$ représente l'espérance du processus $\{\mathbf{Y}_t\}$. En fait, il peut être montré que $\sum_{i=0}^{\infty} \boldsymbol{\Phi}_1^i \mathbf{a}_{t-i}$ existe en moyenne quadratique, voir Lütkepohl (2005, p. 14). Cette représentation comme somme infinie d'un bruit blanc est la limite de l'équation (1.1.4) et cette écriture est valide sous la condition de stabilité donnée dans le Théorème 1.1.1.

La condition de stabilité dans le Théorème 1.1.1 est valide pour les processus VAR(p). Dans le cas $p = 1$, il est particulièrement simple de donner la condition de stabilité en fonction des valeurs propres de $\boldsymbol{\Phi}_1$. Ceci fait l'objet de la proposition suivante.

Proposition 1.1.1. *Un processus VAR(1) est stable si les valeurs propres de $\boldsymbol{\Phi}_1$ sont toutes inférieures à un en module.*

Afin de montrer la proposition précédente, notons que selon le Théorème 1.1.1, nous devons avoir

$$\det(\mathbf{I}_N - \Phi_1 z) \neq 0, \text{ pour } |z| \leq 1.$$

Afin de montrer la proposition, nous devons distinguer deux cas, à savoir $\det(\Phi_1) \neq 0$ et $\det(\Phi_1) = 0$. Si $\det(\Phi_1) \neq 0$, alors $\lambda \neq 0$ n'est pas solution de l'équation $\det(\lambda \mathbf{I}_N - \Phi_1) = 0$. De plus, les valeurs propres de Φ_1 satisfont l'équation suivante :

$$\det(\lambda \mathbf{I}_N - \Phi_1) = \lambda^N \det(\mathbf{I}_N - \lambda^{-1} \Phi_1) = 0.$$

De cette relation nous concluons que les solutions satisfont $|\lambda^{-1}| > 1$ ou encore que les valeurs propres doivent être strictement inférieures à un. Si $\det(\Phi_1) = 0$, alors $\lambda = 0$ est une solution de l'équation $\det(\lambda \mathbf{I}_N - \Phi_1) = 0$ et cette valeur propre satisfait la condition. Pour les autres valeurs non nulles, la même argumentation que précédemment est valide, d'où le résultat.

Un résultat qui s'avérera important pour la suite concerne l'écriture d'un processus VAR(p) sous la forme d'un certain processus VAR(1) de grande dimension. Ceci permet d'appliquer les résultats obtenus dans le cas particulier d'un VAR(1) au processus plus général VAR(p). Ainsi, le processus VAR(p) peut se représenter comme suit :

$$\mathbf{Y}_t^* = \boldsymbol{\mu}^* + \Phi^* \mathbf{Y}_{t-1}^* + \mathbf{a}_t^*, \quad (1.1.6)$$

où

$$\mathbf{Y}_t^* = \begin{pmatrix} \mathbf{Y}_t \\ \mathbf{Y}_{t-1} \\ \vdots \\ \mathbf{Y}_{t-p+1} \end{pmatrix}, \quad \mathbf{a}_t^* = \begin{pmatrix} \mathbf{a}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix},$$

$$\boldsymbol{\mu}_t^* = \begin{pmatrix} \boldsymbol{\mu}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}, \quad \Phi^* = \begin{pmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\ \mathbf{I}_N & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_N & \mathbf{0} \end{pmatrix}.$$

Ainsi, \mathbf{Y}_t^* , \mathbf{a}_t^* et $\boldsymbol{\mu}^*$ sont de dimension $(Np) \times 1$, et Φ^* est une matrice de dimension $(Np) \times (Np)$. L'équation précédemment citée est la représentation

VAR(1) d'un VAR(p). Ainsi, d'un point de vue théorique, la condition de stabilité d'un VAR(p) peut être reformulée en fonction des valeurs propres de la matrice Φ^* .

Proposition 1.1.2. *Un processus VAR(p) est stable si toutes les valeurs propres de Φ^* sont inférieures à un en module ou de façon équivalente si :*

$$\det(\mathbf{I}_{Np} - \Phi^* z) \neq 0, \text{ pour } |z| \leq 1.$$

Voir aussi Lütkepohl (2005, pp. 15-16). Pour avoir une expression du vecteur original \mathbf{Y}_t , on introduit l'opérateur \mathbf{E}_i tel que $\mathbf{E}_i = \mathbf{e}_i \otimes \mathbf{I}_N$, où \mathbf{e}_i est un vecteur dont les composantes sont zéro sauf en position i qui est un. L'opérateur " \otimes " est le produit de Kronecker, dont les propriétés fondamentales se trouvent dans Lütkepohl (2005, Annexe A.11). La matrice $\mathbf{E}_1^\top = (\mathbf{I}_N : \mathbf{0} : \dots : \mathbf{0})$ permet d'extraire les N premières composantes du vecteur auquel il est multiplié. Ainsi, on utilise les formules de la représentation générale VAR(1) d'un processus VAR(p) de l'équation (1.1.6) afin d'obtenir une représentation de \mathbf{Y}_t en fonction des erreurs, en invoquant l'équation (1.1.5) :

$$\mathbf{Y}_t^* = \boldsymbol{\mu}^* + \sum_{i=0}^{\infty} \Phi^{*i} \mathbf{a}_{t-i}^*.$$

En multipliant \mathbf{E}_1^\top de chaque côté de l'équation précédente, nous obtenons :

$$\mathbf{E}_1^\top \mathbf{Y}_t^* = \mathbf{E}_1^\top \boldsymbol{\mu}^* + \mathbf{E}_1^\top \sum_{i=0}^{\infty} \Phi^{*i} \mathbf{a}_{t-i}^*,$$

et donc :

$$\begin{aligned} \mathbf{Y}_t &= \boldsymbol{\mu} + \sum_{i=0}^{\infty} \mathbf{E}_1^\top \Phi^{*i} \mathbf{a}_{t-i}^*, \\ &= \boldsymbol{\mu} + \sum_{i=0}^{\infty} \mathbf{E}_1^\top \Phi^{*i} \mathbf{E}_1 \mathbf{E}_1^\top \mathbf{a}_{t-i}^*, \\ &= \boldsymbol{\mu} + \sum_{i=0}^{\infty} \phi_i^* \mathbf{a}_{t-i}, \end{aligned}$$

où $\phi_i^* = \mathbf{E}_1^\top \Phi^{*i} \mathbf{E}_1$ et $\mathbf{a}_{t-i} = \mathbf{E}_1^\top \mathbf{a}_{t-i}^*$.

Le résultat précédent est obtenu en notant que :

$$\begin{aligned} \mathbf{E}_1 \mathbf{E}_1^\top &= (\mathbf{e}_1 \otimes \mathbf{I}_N)(\mathbf{e}_1^\top \otimes \mathbf{I}_N), \\ &= (\mathbf{e}_1 \mathbf{e}_1^\top) \otimes \mathbf{I}_N, \end{aligned}$$

et repose sur la structure toute particulière de \mathbf{a}_{t-i}^* . Voir aussi Lütkepohl (2005, p. 18). Cette équation permet de voir qu'un processus VAR(p) peut s'écrire sous la forme d'une moyenne mobile :

$$\mathbf{Y}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \phi_i^* \mathbf{a}_{t-i}, \quad (1.1.7)$$

où $\phi_i^* = \mathbf{E}_1^\top \Phi^{*i} \mathbf{E}_1$ et $\mathbf{a}_{t-i} = \mathbf{E}_1^\top \mathbf{a}_{t-i}^*$.

1.1.1.2. La fonction d'autocovariance d'un processus VAR(p)

La fonction d'autocovariance d'un processus VAR(p) peut s'écrire de la façon suivante en considérant l'équation (1.1.6) précédente :

$$\begin{aligned} \Gamma(h) &= E\{(\mathbf{Y}_t - \boldsymbol{\mu})(\mathbf{Y}_{t-h} - \boldsymbol{\mu})^\top\}, \\ &= E\left\{\left(\sum_{i=0}^{h-1} \phi_i^* \mathbf{a}_{t-i} + \sum_{i=0}^{\infty} \phi_{h+i}^* \mathbf{a}_{t-h-i}\right)\left(\sum_{i=0}^{\infty} \phi_i^* \mathbf{a}_{t-h-i}\right)^\top\right\}, \\ &= \sum_{i=0}^{\infty} \phi_{h+i}^* \boldsymbol{\Sigma}_a \phi_i^{*\top}. \end{aligned}$$

La condition de stabilité assure que la fonction d'autocovariance existe et ainsi que la somme infinie devient sommable. La fonction d'autocovariance $\{\Gamma(h), h \in \mathbb{Z}\}$ d'un VAR(p) peut également s'écrire sous la forme récursive suivante :

$$\Gamma(h) = \sum_{i=1}^p \Phi_i \Gamma(h-i), \quad (1.1.8)$$

ou sous la forme du système de type Yule-Walker :

$$\begin{cases} \Gamma(0) = \sum_{i=1}^p \Phi_i \Gamma^\top(i) + \boldsymbol{\Sigma}_a, & h = 0, \\ \Gamma(h) = \sum_{i=1}^p \Phi_i \Gamma(h-i), & h > 0. \end{cases} \quad (1.1.9)$$

Les démonstrations des résultats (1.1.8) et (1.1.9) s'obtiennent en prenant l'espérance du produit $\mathbf{Y}_t - \boldsymbol{\mu}$ avec $\dot{\mathbf{Y}}_{t-h}$, où $\dot{\mathbf{Y}}_{t-h} = \boldsymbol{\Phi}_1 \dot{\mathbf{Y}}_{t-h-1} + \boldsymbol{\Phi}_2 \dot{\mathbf{Y}}_{t-h-2} + \dots + \boldsymbol{\Phi}_p \dot{\mathbf{Y}}_{t-h-p-1} + \mathbf{a}_{t-h}$, $\dot{\mathbf{Y}}_{t-h} = \mathbf{Y}_{t-h} - \boldsymbol{\mu}$. On note pour $h = 0$ que $\boldsymbol{\Gamma}(0) = \boldsymbol{\Gamma}^\top(0)$, et que $E[(\mathbf{Y}_t - \boldsymbol{\mu})\mathbf{a}_t^\top] = \boldsymbol{\Sigma}_a$, en utilisant (1.1.5). On note que pour (1.1.9) la stabilité du processus est invoquée, et le résultat est en fait obtenu en effectuant des manipulations analogues à ce qui est fait dans le cas univarié, en prenant cependant en considération que $\boldsymbol{\Gamma}(h) = \boldsymbol{\Gamma}^\top(-h)$ pour $h < 0$.

Nous introduisons dans ce qui suit les processus VARMA, qui sont une extension des processus VAR présentés dans la section précédente. Cette classe de modèles est fort importante dans la modélisation et la prévision des séries chronologiques réelles vectorielles. Tout comme dans le cas univarié, l'intérêt premier de l'introduction d'une moyenne mobile est la parcimonie, car souvent dans les applications une réduction non-négligeable en paramètres est possible, comparativement à l'utilisation de modèles VAR(p) avec p grand.

Définition 1.1.6. *Un processus stationnaire $\{\mathbf{Y}_t\}$ autorégressif moyenne mobile de dimension N et d'ordres p et q , noté VARMA(p, q), est défini par :*

$$\boldsymbol{\Phi}(B)\mathbf{Y}_t = \boldsymbol{\mu} + \boldsymbol{\Theta}(B)\mathbf{a}_t, \quad t \in \mathbb{Z}, \quad (1.1.10)$$

où B est l'opérateur de retard temporel, $\boldsymbol{\Phi}(B) = \mathbf{I}_N - \sum_{i=1}^p \boldsymbol{\Phi}_i B^i$ et $\boldsymbol{\Theta}(B) = \mathbf{I}_N + \sum_{i=1}^q \boldsymbol{\Theta}_i B^i$ sont des opérateurs matriciels de dimension $N \times N$, qui sont respectivement les opérateurs autorégressifs et moyenne mobile multivariés tels que $\boldsymbol{\Phi}_p \neq \mathbf{0}$ et $\boldsymbol{\Theta}_q \neq \mathbf{0}$, $\boldsymbol{\mu}$ est un vecteur constant de dimension $N \times 1$ et $\{\mathbf{a}_t\}$ est un bruit blanc tel que $\text{Var}(\mathbf{a}_t) = \boldsymbol{\Sigma}_a$.

Lorsque $p = 0$, le processus VARMA(p, q) devient 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 VARMA(p, q) est le processus autorégressif multivarié d'ordre p VAR(p); dans un tel cas c'est l'opérateur moyenne mobile qui est la matrice identité.

La notion de causalité d'un processus s'applique également aux processus VARMA(p, q). Un moyen pour vérifier si un processus est causal est d'étudier la nature des racines du déterminant de l'opérateur autorégressif, c'est-à-dire étudier les racines de $\det\{\boldsymbol{\Phi}(z)\}$, voir le Théorème 1.1.1. De plus, un processus

VARMA(p,q) pouvant s'exprimer comme une combinaison linéaire infinie des valeurs passées de \mathbf{Y}_t fait intervenir le concept d'inversibilité. Le théorème suivant traite de l'inversibilité.

Théorème 1.1.2. *Le processus VARMA(p,q) représenté par l'équation (1.1.10) est inversible si :*

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

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

$$\mathbf{a}_t = \sum_{i=0}^{\infty} \mathbf{\Pi}_i \mathbf{Y}_{t-i} = \mathbf{\Pi}(B) \mathbf{Y}_t,$$

et les matrices $\mathbf{\Pi}_i$ sont uniquement déterminées par :

$$\mathbf{\Pi}(B) = \sum_{i=0}^{\infty} \mathbf{\Pi}_i B^i = \Theta^{-1}(B) \Phi(B).$$

Une démonstration de ce théorème est présentée dans Brockwell et Davis (1991, p. 419).

Nous abordons brièvement une extension possible des modèles VAR ou VARMA, obtenue par l'inclusion de variables exogènes (ou encore variables explicatives). Un exemple de tel modèle est le modèle VAR avec variables exogènes :

$$\Phi(B) \mathbf{Y}_t = \boldsymbol{\mu} + \mathbf{V}(B) \mathbf{X}_t + \mathbf{a}_t, \quad (1.1.11)$$

où $\mathbf{V}(B) = \sum_{i=0}^s \mathbf{V}_i B^i$, $\mathbf{V}_s \neq \mathbf{0}$, $\{\mathbf{a}_t\}$ est le terme d'erreur et $\{\mathbf{X}_t\}$ est un processus stationnaire au second ordre. Ce modèle est parfois appelé VARX, voir Lütkepohl (2005). Il est aussi possible de considérer un VARMA avec variables exogènes, qui est alors appelé VARMAX. Les modèles VARX sont particulièrement intéressants, car ils permettent d'inclure des variables explicatives, parfois également appelées variables d'input, sur des variables de sortie, ou d'output. Lütkepohl (2005) discute de l'estimation et de la prévision dans ces modèles. Il est intéressant de noter que ces modèles font partie de la catégorie de modèles dit dynamiques, mais ils demeurent pratiquement aussi facile à estimer que les modèles VAR, puisqu'il est possible de ré-écrire un VARX comme un certain modèle de régression linéaire multiple. Duchesne (2005) discute de la validation des modèles VARX avec des statistiques de test de type portemanteau.

En général, la modélisation d'une série chronologique univariée selon l'approche de Box et Jenkins est un processus itératif en trois étapes principales. L'étape d'identification consiste à déterminer, à partir de l'observation des comportements des autocorrélations et autocorrélations partielles échantillonnales, les ordres autorégressif et moyenne mobile p et q du modèle. Vient ensuite l'étape d'estimation, que ce soit par la méthode du maximum de vraisemblance (MV) ou par la méthode des moindres carrés, des paramètres autorégressifs et moyennes mobiles, de la moyenne et de la variance du processus. La dernière étape est celle relative à la validation du modèle estimé par le biais d'une analyse de la dépendance restante dans les résidus. Dans la modélisation de séries chronologiques multivariées, les grandes étapes générales sont les mêmes. Ainsi, les autocovariances matricielles peuvent être utilisées afin d'identifier l'ordre moyenne mobile q . Cependant, une différence notable concerne l'identification de l'ordre autorégressif, où la généralisation n'est pas immédiate. En fait, l'utilisation de l'autocorrélation partielle dans le cas univarié sert principalement à identifier l'ordre autorégressif, car si une série est $AR(p)$, la fonction d'autocorrélation partielle s'annule après le délai p . Un concept possible utilisé dans le cas multivarié est la fonction matricielle autorégressive partielle introduite par Tiao *et* Box (1981). L'autocovariance matricielle et la fonction matricielle autorégressive partielle sont alors des outils utiles pour identifier les ordres p et q d'un processus VARMA. Les autres étapes sont essentiellement les mêmes que dans le cas univarié, à savoir l'estimation et la validation du modèle. On note que cet outil d'identification est raisonnable et s'inscrit dans la lignée de la modélisation de Box *et* Jenkins. Cependant, ce concept ne semble pas faire l'unanimité. En effet, Brockwell *et* Davis (2002) n'abordent pas ce concept afin de modéliser des séries multivariées.

Tiao *et* Box (1981) ont défini la matrice autorégressive des autocorrélations partielles d'ordre h comme le dernier coefficient de la régression de \mathbf{Y}_t sur $\mathbf{Y}_{t-1}, \dots, \mathbf{Y}_{t-h}$. Ils notent cette matrice $\mathcal{P}(h)$. La description la plus simple se trouve dans l'ouvrage de Wei (1994, pp. 351-354). Il s'agit dans ce cas d'un prolongement direct d'une définition possible de la fonction d'autocorrélation partielle pour les séries chronologiques univariées de Box *et* Jenkins. Ainsi, pour le modèle $VAR(p)$

de l'équation (1.1.2), $\mathcal{P}(h) = \Phi_p$, pour $h = p$. Ce résultat est obtenu en minimisant l'équation (1.1.12) :

$$E\{|\mathbf{Y}_{t+h} - \Phi_{h,1}\mathbf{Y}_{t+h-1} - \Phi_{h,2}\mathbf{Y}_{t+h-2} - \dots - \Phi_{h,h}\mathbf{Y}_t|^2\}. \quad (1.1.12)$$

La minimisation de (1.1.12) nous conduit à une généralisation multivariée des équations de Yule-Walker sous la forme suivante :

$$\begin{bmatrix} \Gamma(0) & \Gamma^\top(1) & \dots & \Gamma^\top(h-1) \\ \Gamma(1) & \Gamma(0) & \dots & \Gamma^\top(h-2) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma(h-1) & \Gamma(h-2) & \dots & \Gamma(0) \end{bmatrix} \begin{bmatrix} \Phi_{h,1}^\top \\ \Phi_{h,2}^\top \\ \vdots \\ \Phi_{h,h}^\top \end{bmatrix} = \begin{bmatrix} \Gamma(1) \\ \Gamma(2) \\ \vdots \\ \Gamma(h) \end{bmatrix},$$

ou sous cette forme alternative :

$$\begin{bmatrix} \mathbf{A}(h) & \mathbf{b}(h) \\ \mathbf{b}^\top(h) & \Gamma(0) \end{bmatrix} \begin{bmatrix} \Phi_*^\top(h-1) \\ \Phi_{h,h}^\top \end{bmatrix} = \begin{bmatrix} \mathbf{c}(h) \\ \Gamma(h) \end{bmatrix}.$$

Ainsi, nous déduisons après quelques manipulations algébriques le vecteur $\Phi_{h,h}^\top$ défini par :

$$\Phi_{h,h}^\top = \{\Gamma(0) - \mathbf{b}^\top(h)[\mathbf{A}(h)]^{-1}\mathbf{b}(h)\}^{-1}\{\Gamma(h) - \mathbf{h}^\top(h)[\mathbf{A}(h)]^{-1}\mathbf{c}(h)\}.$$

De ce résultat, nous trouvons la fonction matricielle autorégressive partielle :

$$\mathcal{P}(h) = \begin{cases} \Gamma^\top(1)[\Gamma(0)]^{-1}, & h = 1, \\ \{\Gamma^\top(h) - \mathbf{c}^\top(p)[\mathbf{A}(h)]^{-1}\mathbf{b}(h)\}^{-1}\{\Gamma(0) - \mathbf{b}^\top(h)[\mathbf{A}(h)]^{-1}\mathbf{b}(h)\}, & h \geq 1, \end{cases}$$

où

$$\mathbf{A}(h) = \begin{bmatrix} \Gamma(0) & \Gamma^\top(1) & \dots & \Gamma^\top(h-2) \\ \Gamma(1) & \Gamma(0) & \dots & \Gamma^\top(h-3) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma(h-2) & \Gamma(h-3) & \dots & \Gamma(0) \end{bmatrix}, \mathbf{b}(h) = \begin{bmatrix} \Gamma^\top(h-1) \\ \Gamma^\top(h-2) \\ \vdots \\ \Gamma^\top(1) \end{bmatrix},$$

$$\mathbf{c}(h) = \begin{bmatrix} \Gamma^\top(h-1) \\ \Gamma^\top(h-2) \\ \vdots \\ \Gamma^\top(1) \end{bmatrix} \text{ et } \Phi_*^\top(h-1) = \begin{bmatrix} \Phi_{h,1}^\top \\ \Phi_{h,2}^\top \\ \vdots \\ \Phi_{h,h-1}^\top \end{bmatrix}.$$

La matrice $\mathcal{P}(h)$ définie par l'équation précédente a la propriété suivante pour le processus VAR(p) :

$$\mathcal{P}(h) = \begin{cases} \Phi_p, & h = p, \\ \mathbf{0}, & h > p. \end{cases} \quad (1.1.13)$$

Plus de détails sont également fournis dans Tiao *et* Box (1981, p. 805). La propriété (1.1.13) est remarquable et suggère que $\mathcal{P}(h)$ est une définition appropriée afin d'identifier l'ordre autorégressif. Ainsi, nous retrouvons un outil d'identification d'un ordre autorégressif dans le cas multivarié.

1.1.2. Estimation des paramètres d'un processus stationnaire

Afin d'étudier un processus stochastique vectoriel stationnaire, il est utile d'estimer sa fonction moyenne et sa fonction d'autocovariance matricielle. C'est l'objet de cette section. Ayant observé une série chronologique \mathbf{Y}_t , $t = 1, \dots, n$, la fonction moyenne est simplement estimée par la moyenne échantillonnale, c'est-à-dire que : $\hat{\boldsymbol{\mu}} = \bar{\mathbf{Y}} = n^{-1} \sum_{t=1}^n \mathbf{Y}_t$.

La fonction d'autocovariance échantillonnale $\{\hat{\boldsymbol{\Gamma}}_{\mathbf{Y}}(h), |h| < n\}$ est définie quant à elle par l'équation suivante :

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

Les propriétés générales de ces statistiques sont décrites dans Hannan (1970). Il est à noter que la fonction d'autocovariance échantillonnale définie précédemment repose sur une réalisation d'un processus stochastique. Il est naturel d'introduire des mesures reposant sur des résidus de modèles. On parle alors de fonction d'autocovariance résiduelle. Ces mesures peuvent alors être utilisées afin de diagnostiquer des modèles. Les distributions asymptotiques des autocovariances résiduelles ont été étudiées abondamment et dans plusieurs contextes. On peut citer McLeod *et* Li (1981), Hosking (1980), Ahn (1988) ou Duchesne (2005), par exemple.

Dans la littérature portant sur la modélisation de séries chronologiques multivariées, plusieurs méthodes d'estimation sont envisageables pour estimer les paramètres d'un processus VARMA. Parmi ces méthodes, on peut citer par exemple, le maximum de vraisemblance, le quasi-maximum vraisemblance et les méthodes des moindres carrés. La procédure la plus largement étudiée est la méthode de maximum de vraisemblance (MV) où les erreurs des modèles sont gaussiennes, voir Brockwell *et* Davis (1991, p. 430). Cependant, la maximisation de cette méthode requiert des calculs très lourds puisque pour chaque ordre autorégressif et moyenne mobile (p, q) une maximisation non quadratique satisfaisant des contraintes d'inégalité doit être réalisée à l'aide d'algorithmes itératifs, voir Tiao *et* Box (1981, p. 809). L'estimation par la méthode du maximum de vraisemblance dans les VARMA est discutée dans Brockwell *et* Davis (1991). Une difficulté non-négligeable réside dans le fait qu'en général les VARMA ne sont pas identifiables. Autrement dit, deux ensembles de paramètres différents peuvent correspondre au même processus. Un exemple simple se trouve dans Brockwell *et* Davis (2002, p. 243). Le problème ne survient pas pour les VAR(p), ce qui explique probablement leur popularité dans les applications statistiques et économétriques. On note que Brockwell *et* Davis (1991, p. 432) discutent de l'estimation des VAR(p) reposant sur l'utilisation d'une version multivariée de l'algorithme de Durbin-Levinson. Ils suggèrent de trouver p en utilisant le critère AIC. Voir Brockwell *et* Davis (1991, pp. 432-434). Tenant compte des difficultés dans les VARMA, d'autres méthodes d'estimation ont été étudiées dans la littérature. Par exemple, Pukkila *et al.* (1990) ont proposé une procédure de moindres carrés généralisés (MCG). Cette technique est plus simple numériquement que celle du maximum de vraisemblance et il a été démontré théoriquement qu'elle fournit des estimateurs asymptotiquement efficaces, voir Ahn *et al.* (1990). Dans ce qui suit, nous présentons les différentes étapes de cette méthode pour un VAR(p).

La première étape de cette méthode consiste à réécrire le modèle VAR(p) défini en (1.1.2) sous la forme compacte suivante : $\mathbf{H} = \mathbf{BZ} + \mathbf{A}$, où :

$$\begin{aligned}
\mathbf{H} &= [\mathbf{Y}_1, \dots, \mathbf{Y}_n], & (N \times n); \\
\mathbf{h} &= \text{vec}(\mathbf{Y}), & (Nn \times 1); \\
\mathbf{B} &= [\mu, \Phi_1, \dots, \Phi_p], & (N \times (Np + 1)); \\
\boldsymbol{\beta} &= \text{vec}(\mathbf{B}), & ((N^2p + N) \times 1); \\
\mathbf{Z}_t &= [1, \mathbf{Y}_t^\top, \dots, \mathbf{Y}_{t-p+1}^\top]^\top, & ((Np + 1) \times n); \\
\mathbf{Z} &= [\mathbf{Z}_0, \dots, \mathbf{Z}_{n-1}], & ((Np + 1) \times n); \\
\mathbf{A} &= [\mathbf{a}_1, \dots, \mathbf{a}_n], & (N \times n); \\
\mathbf{a} &= \text{vec}(\mathbf{A}), & (Nn \times 1);
\end{aligned}$$

où l'opérateur $\text{vec}(\cdot)$ désigne l'empilement des colonnes \mathbf{c}_i de dimension $m \times 1$ d'une matrice $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$ de dimension $m \times n$, c'est-à-dire $\text{vec}(\mathbf{C}) = (\mathbf{c}_1^\top, \dots, \mathbf{c}_n^\top)^\top$. Dans un second temps, on empile les colonnes et on retrouve la forme standard suivante sur laquelle les méthodes des moindres carrés peuvent être utilisées pour estimer les paramètres, soit : $\mathbf{h} = (\mathbf{Z}^\top \otimes \mathbf{I}_n) \boldsymbol{\beta} + \mathbf{a}$.

Les estimateurs par la méthode MCG s'obtiennent en minimisant le critère : $\mathbf{S} = \mathbf{a}^\top \{\mathbf{I}_n \otimes \boldsymbol{\Sigma}_a\}^{-1} \mathbf{a} = \text{tr}\{\mathbf{A}^\top \boldsymbol{\Sigma}_a^{-1} \mathbf{A}\}$,

où $\mathbf{I}_n \otimes \boldsymbol{\Sigma}_a$ est la matrice de variance du vecteur \mathbf{a} , $E(\mathbf{a}_t \mathbf{a}_t^\top) = \boldsymbol{\Sigma}_a$, et $\text{tr}(\cdot)$ représente la trace d'une matrice, voir Lütkepohl (2005, p. 71) pour plus de détails. Ainsi en présumant que $\{\mathbf{a}_t\}$ est un bruit blanc, on trouve les estimateurs sous la forme : $\hat{\boldsymbol{\beta}} = \{(\mathbf{Z}\mathbf{Z}^\top)^{-1} \mathbf{Z} \otimes \mathbf{I}_n\} \mathbf{h}$. Les propriétés statistiques des moindres carrés généralisés se trouvent dans Kariya et Kurata (2004).

Les propriétés asymptotiques de $\hat{\boldsymbol{\beta}}$ se trouvent dans Lütkepohl (2005). En effet, $\hat{\boldsymbol{\beta}}$ est sous certaines conditions de loi asymptotiquement normale, et il est possible d'explicitier la variance asymptotique de $\hat{\boldsymbol{\beta}}$ pour des fins inférentielles. L'objectif de cette section était de présenter les résultats fondamentaux sur les processus stochastiques multivariés. Le prochain chapitre introduit les concepts fondamentaux entourant la modélisation des données spatiales.

Chapitre 2

ANALYSE DE DONNÉES SPATIO-TEMPORELLES MULTIVARIÉES

L'objectif principal de ce chapitre est de présenter, dans un premier temps, les concepts généraux sur la statistique spatiale. Ensuite, nous accordons une attention particulière à la classe des modèles GSTAR et à la méthode d'estimation de leurs paramètres.

2.1. CONCEPTS GÉNÉRAUX SUR LES MODÈLES SPATIAUX

Les processus stochastiques multivariés permettent de décrire plusieurs séries chronologiques. Les mesures de corrélation permettent de décrire plusieurs régions spatiales, leurs liens et les dépendances temporelles. Cependant, un modèle multivarié général entraînera typiquement un grand nombre de paramètres, surtout si le nombre de séries chronologiques est relativement important.

Dans le contexte de données spatio-temporelles, un raffinement des modèles de séries chronologiques consiste à incorporer explicitement les dépendances systématiques entre les observations pour une région donnée, ainsi que les dépendances d'une région avec les régions avoisinantes. Ceci se fait en introduisant des matrices de poids. Le Gallo (2000) distingue deux grandes catégories de matrices de poids : les matrices de contiguïté d'une part, et les matrices de poids généralisés d'autre part.

La contiguïté entre deux régions survient si elles ont une frontière commune. À l'ordre un, la matrice de contiguïté est une matrice carrée, symétrique et de

dimension $N \times N$, où N correspond maintenant au nombre de sites. On a alors $\mathbf{W} = (w_{ij})$ où les éléments de la matrice sont :

$$w_{ij} = \begin{cases} 1, & \text{si les sites } i \text{ et } j \text{ sont contiguës,} \\ 0, & \text{sinon.} \end{cases}$$

Par convention, $w_{ii} = 0$. Le nombre total de zones contiguës à une zone i est égale à la somme de ligne i des éléments de la matrice carrée symétrique \mathbf{W} , et correspond donc à $\sum_{j=1}^N w_{ij}$. Ces matrices peuvent être standardisées pour que la somme de chaque ligne soit égale à un, en utilisant la formule $\tilde{w}_{ij} = \frac{w_{ij}}{\sum_{j=1}^N w_{ij}}$. Dans leur ouvrage sur les méthodes spatiales, Cliff et Ord (1981) ont distingué trois types de contiguïté conventionnels. La Figure 2.1 présente visuellement ces concepts.

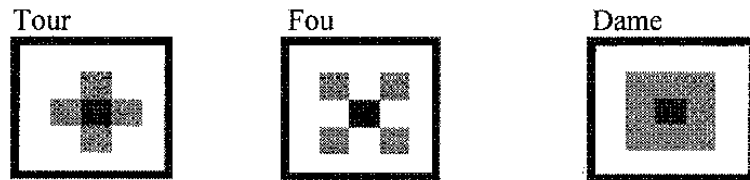


FIGURE 2.1. Types de contiguïté conventionnels

Ainsi, on trouve la Tour, le Fou et la Dame.

- (1) Tour : la zone centrale a quatre voisins dans les directions Nord-Sud et Est-Ouest.
- (2) Fou : La zone centrale a quatre voisins dans les directions Nord-Est, Sud-Ouest, Sud-Est et Nord-Ouest.
- (3) Dame : La zone centrale a huit voisins dans les directions Nord-Sud, Est-Ouest, Nord-Est, Sud-Ouest, Sud-Est et finalement Nord-Ouest.

Le concept de contiguïté peut être généralisé à un ordre supérieur à un. À l'ordre k , les régions i et j sont contiguës si k est le nombre minimal de frontières à traverser pour aller de la région i à la région j .

Les définitions de contiguïté telles que discutées précédemment sont plutôt compliquées, particulièrement si on désire inclure ce type d'information dans des modèles statistiques. Ainsi, d'autres notions ont été introduites afin d'incorporer

la composante spatiale dans les modèles de séries chronologiques. Les notions de contiguïté ont été exposées afin de comprendre les difficultés entourant le choix des poids dans les matrices de poids. De manière générale, ces poids doivent représenter un ordre hiérarchique parmi des voisins spatiaux. Dans un sens à préciser, il faut donc que les voisins du premier ordre soient plus près que les voisins du second ordre, par exemple. Et naturellement, les voisins au premier ordre devraient en fait être les plus près possibles. Lorsque les régions sont sur une grille cartésienne, il semble convenable qu'un tel ordre puisse être établi. La Figure 2.2 permet de mieux comprendre la notion d'ordre des matrices de poids. Cette définition d'ordre spatial utilise un ordonnancement en utilisant le concept naturel de distance euclidienne de tous les sites entourant une région donnée. Les matrices de poids ont systématiquement été étudiées par Bavaud (1998). Ainsi,

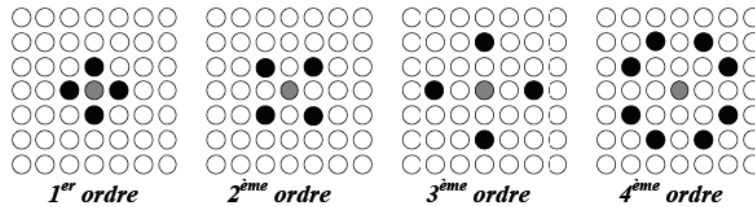


FIGURE 2.2. Exemples d'ordres spatiaux d'ordre un, deux, trois et quatre.

dans le cas général de sites pouvant se trouver sur une carte, l'utilisation de mesures de distance peut sembler naturel (on note alors que la nature irrégulière des régions sur les cartes se distingue nettement du cas cartésien très régulier). Les poids sont alors fonction de ces distances. Ainsi, dans ce cas, on mesure la proximité (Le Gallo (2000) parle plutôt de l'intensité de l'interaction car le poids n'est pas forcément directement une distance) entre deux régions i et j par la distance entre les centres de ces régions et des poids sont par la suite calculés. Dans le cas d'états ou de pays, on pourrait considérer la distance entre les capitales. Diverses formes fonctionnelles sont également possibles. Pour deux régions i et j , soit d_{ij} la distance euclidienne entre ces deux sites. Une première proposition est donnée par :

$$c(d_{ij} + 1)^{-\alpha}, \quad (2.1.1)$$

où c et α sont à spécifier. Dans le même ordre d'idée, on pourrait également mesurer la proximité avec la fonction exponentielle et considérer :

$$c \exp(d_{ij} + 1)^{-\alpha}. \quad (2.1.2)$$

Dans tous les cas, la matrice de poids, \mathbf{W} , est une matrice carrée dont chaque élément w_{ij} représente l'intensité de l'interaction entre les deux régions. Ainsi, les éléments w_{ij} de \mathbf{W} satisfont :

- a) $w_{ij} \geq 0, \forall i, j$ et $w_{ii} = 0$;
- b) $\sum_{j=1}^N w_{ij} = 1$, pour tous les sites $i = 1, \dots, N$.

On note que la matrice de poids défini ainsi permet de considérer aisément une distance de premier ordre. Il n'est alors pas immédiatement évident de généraliser ce type de concept à des matrices de poids d'ordre supérieur.

Dans leur application, Borovkova et al. (2008) ont considéré dans un premier temps des seuils afin de mesurer les distances entre les villes et afin de considérer des matrices de poids d'ordre supérieur. Ils discutent aussi des choix de c et α dans les formules (2.1.1) et (2.1.2). Ayant une carte géographique, il devient possible de calculer les distances et l'inverse de ces distances. Soient $0 = d_0 < d_1 < \dots < d_\lambda$, où les d_i sont les seuils. La matrice de poids à l'ordre spatial k repose sur le poids inverse $\frac{1}{(d_{ij}+1)}$, pour deux sites i et j , où d_{ij} est la distance euclidienne se trouvant dans l'intervalle (d_{k-1}, d_k) . Le poids $w_{ij}^{(k)}$ est donc $\frac{1}{(d_{ij}+1)}$ dans $\mathbf{W}^{(k)}$, et les entrées (i, j) sont donc 0 dans les matrices $\mathbf{W}^{(l)}$, $l \neq k$. Un second type de matrices de poids a été considéré par Borovkova et al. (2008). Pour un site i , tous les sites j à une distance appartenant à l'intervalle (d_{k-1}, d_k) sont encore considérés comme des sites voisins à l'ordre k du site i , et les poids correspondants sont cependant égaux. Toutes les matrices de poids sont normalisées de telle sorte que pour tous les lieux leurs poids somment à un.

Comme pour les matrices de poids, l'opérateur de retard spatial est un outil fondamental en statistique spatiale ou du moins dans un contexte de séries chronologiques spatio-temporelles. Ainsi $L^{(l)}Y_t(i)$ nous indique comment au temps t le i ème site peut être expliqué par les régions avoisinantes, pour $i = 1, \dots, N$. On note que certaines hypothèses sont formulées dans l'utilisation de $L^{(l)}$. On note

aussi que $L^{(l)}$ est un opérateur linéaire, et qu'il est indépendant de t . Ainsi, dans le système vectoriel, $L^{(l)}\mathbf{Y}_t$ indique comment les régions avoisinantes expliquent \mathbf{Y}_t , pour tout $t \in \mathbb{Z}$. L'importante réduction en paramètres repose sur le caractère déterministe des matrices de poids $\mathbf{W}^{(l)}$, qui doivent être cependant spécifiées par l'analyste. Techniquement, il est similaire à l'opérateur de retard temporel B que l'on a déjà introduit dans la Section 1.1. L'opérateur spatial défini plus bas peut être utilisé dans la construction de nombreux modèles spatio-temporels multivariés. Une caractéristique commune de ces modèles est qu'ils constituent un système à plusieurs variables présentant la dépendance systématique entre les observations à chaque unité spatiale (régions, villes, pays, etc) et les observations faites à des unités spatiales voisines (autocorrélation spatiale). L'opérateur spatial est une façon parcimonieuse d'exprimer l'ordre spatial dans ce système à plusieurs variables. À un temps t et pour un site i , on exprime $Y_t(i)$ comme une combinaison linéaire d'observations avoisinantes au temps t . Plus précisément :

$$L^{(l)}Y_t(i) = \begin{cases} Y_t(i), & \text{si } l = 0, \\ \sum_{j=1}^N w_{ij}^{(l)} Y_t(j), & \text{lorsque } l > 0, \end{cases}$$

où $w_{ij}^{(l)}$ sont des poids à l'ordre l . Ainsi, à l'ordre l , les poids déterministes permettent de formuler comment agit l'ordre spatial l au temps t sur le i ème site $Y_t(i)$. On rappelle que les poids spatiaux satisfont la condition :

$$\sum_{j=1}^N w_{ij}^{(l)} = 1.$$

Sous forme matricielle, on peut écrire :

$$L^{(l)}\mathbf{Y}_t = \begin{cases} \mathbf{W}^{(0)}\mathbf{Y}_t = \mathbf{I}_N\mathbf{Y}_t, & \text{si } l = 0, \\ \mathbf{W}^{(l)}\mathbf{Y}_t, & \text{pour } l > 0. \end{cases} \quad (2.1.3)$$

L'indépendance temporelle de $L^{(l)}$ est encore plus visible en (2.1.3), dans le sens que $L^{(l)}\mathbf{Y}_t$ et $L^{(l)}\mathbf{Y}_s$, $t \neq s$, font intervenir la même matrice déterministe $\mathbf{W}^{(l)}$.

2.2. MODÈLES SPATIO-TEMPORELS MULTIVARIÉS

Nous étudions dans ce qui suit une classe particulière de modèles spatiaux appelée modèles autorégressifs spatio-temporels (*Space-Time AutoRegressive models*), dont l'abréviation est STAR. Ce processus a été généralisé, ce qui a amené au modèle GSTAR (*Generalized Space-Time AutoRegressive model*). Le modèle STAR a été introduit dans la littérature au début de l'année 1970 par les auteurs Cliff *et* Ord (1975), Martin *et* Oppen (1975) et Pfeifer *et* Deutsch (1980). Le modèle STAR($p_{\lambda_1, \dots, \lambda_p}$) pour N sites fait l'objet de la Définition 2.2.1.

Définition 2.2.1. *Un processus STAR($p_{\lambda_1, \dots, \lambda_p}$) à N dimensions est défini comme suit :*

$$\mathbf{Y}_t = \sum_{s=1}^p \sum_{k=0}^{\lambda_s} \phi_{sk} \mathbf{W}^{(k)} \mathbf{Y}_{t-s} + \mathbf{a}_t, \quad t = 0, \pm 1, \pm 2, \dots, \quad (2.2.1)$$

où $\{\mathbf{Y}_t\}$ est une série temporelle multivariée de dimension $N \times 1$, λ_s est l'ordre spatial pour le $s^{\text{ième}}$ terme autorégressif, p est l'ordre autorégressif du modèle, ϕ_{sk} est un paramètre scalaire et est appelé le paramètre autorégressif temporel décalé d'ordre k et spatial d'ordre s , $\mathbf{W}^{(k)}$ est une matrice $N \times N$ de poids spatiaux d'ordre k , pour $k = 0, 1, \dots, \lambda_s, s = 1, \dots, p$, avec $\mathbf{W}^{(0)} = \mathbf{I}_N$ et \mathbf{a}_t est le terme d'erreur.

Les ordres $\lambda_s, s = 1, \dots, p$, permettent de spécifier, pour un décalage temporel s donné, comment les décalages spatiaux $\mathbf{W}^{(k)} \mathbf{Y}_{t-s}, k = 0, 1, \dots, \lambda_s$, expliquent \mathbf{Y}_t . Ainsi, on constate que les modèles STAR forment une dépendance linéaire dans le temps et l'espace par l'introduction de valeurs retardées de \mathbf{Y}_t , sur lesquelles les matrices de poids $\mathbf{W}^{(k)}$ agissent, permettant de faire intervenir une dépendance spatiale. Cette classe de modèles est un cas particulier des modèles VAR qui sont plus généraux. Les matrices de poids $\mathbf{W}^{(k)}$ étant fixées, les seuls paramètres à estimer sont les constantes $\phi_{sk}, k = 0, 1, \dots, \lambda_s, s = 1, \dots, p$. On constate que si les matrices de poids sont bien choisies, il est possible d'ordonner de manière hiérarchisée les voisins de chaque position spatiale, et d'expliquer adéquatement $Y_t(i)$ par $L^{(k)} Y_{t-s}(i), k = 0, 1, \dots, \lambda_s$. Dans cette formulation, on note que les retards temporels $L^{(k)} Y_{t-1}(i), \dots, L^{(k)} Y_{t-p}(i)$, expliquent $Y_t(i)$ de la

même manière quant à la structure spatiale. Ceci peut faire du sens, car intuitivement les régions ne sont pas supposées changer dans l'espace : ainsi les voisins de \mathbf{Y}_t retardés expliquent précisément \mathbf{Y}_t de la même façon quant à la structure spatiale. Ainsi, dans cette formulation, si deux sites sont très voisins du site i les valeurs retardées des observations voisines agissent de la même manière dans la structure spatiale pour expliquer $Y_t(i)$. Il semble clair qu'une stratégie efficace fera intervenir des ordres p et $\lambda_1, \dots, \lambda_p$ plutôt petits. Le cas le plus simple est lorsque $p = 1$ et $\lambda_1 = 1$. Le terme d'erreur dans la Définition 2.2.1 est général. Il pourrait être d'une distribution multivariée arbitraire. Dans un premier temps, le cas qui a été retenu l'attention dans la littérature a été le cas où $\mathbf{a}_t \sim \mathcal{N}_N(\mathbf{0}, \sigma^2 \mathbf{I}_N)$. Voir par exemple Pfeifer *et al.* (1980). Assez rapidement, une structure plus générale des variances-covariances a été étudiée, permettant $\mathbf{a}_t \sim \mathcal{N}_N(\mathbf{0}, \Sigma_{\mathbf{a}})$, où $\Sigma_{\mathbf{a}}$ est définie positive. Il est intéressant de noter que la littérature géographique qui utilise les modèles espace-temps accorde beaucoup d'attention aux problèmes potentiels d'hétéroscédasticité, voir par exemple Di Giacinto (2003, 2006), en spécifiant $\Sigma_{\mathbf{a}}$ comme une matrice diagonale avec variances inégales pour les N sites. D'un point de vue statistique, il semble que le fait de présumer $\Sigma_{\mathbf{a}}$ diagonale ou non change vraiment peu l'étude statistique du problème, du moins avec ce que l'on connaît sur l'estimation des VAR structurés. Les hypothèses les plus générales sur le terme d'erreur sont dans Borovkova *et al.* (2008), où $\{\mathbf{a}_t\}$ pourrait être de distribution autre que normale, avec une matrice de variance $\Sigma_{\mathbf{a}_t}$ qui est maintenant indexée dans le temps. Il faut noter cependant que cette généralité accrue ne semble pas vraiment compatible avec les applications issues de la géographie qui utilisent somme toute assez peu de données temporelles. Par exemple, Di Giacinto (2003) utilise comme sites les états américains, et la période couvre 1958 à 2000, observées de façon trimestrielles. Ceci donne environ 172 trimestres, ce qui est nettement inférieur aux applications financières impliquant des milliers d'observations, qui pourraient alors être utilisées pour estimer des variances conditionnelles. Ainsi, dans la section suivante, nous examinons en détail le cas du processus STAR(1₁).

2.2.1. Analyse du processus STAR(1₁)

On considère le modèle (2.2.1) avec $p = 1$ et $\lambda = 1$. On explicite ce modèle comme suit :

$$\begin{aligned}\mathbf{Y}_t &= \phi_{10} \mathbf{W}^{(0)} \mathbf{Y}_{t-1} + \phi_{11} \mathbf{W}^{(1)} \mathbf{Y}_{t-1} + \mathbf{a}_t, \\ &= \phi_{10} \mathbf{I}_N \mathbf{Y}_{t-1} + \phi_{11} \mathbf{W}^{(1)} \mathbf{Y}_{t-1} + \mathbf{a}_t, \\ &= \mathbf{\Phi}_1 \mathbf{Y}_{t-1} + \mathbf{a}_t,\end{aligned}$$

où $\mathbf{\Phi}_1 = \phi_{10} \mathbf{I}_N + \phi_{11} \mathbf{W}^{(1)}$. Ainsi, on constate que le modèle STAR(1₁) peut s'écrire comme un modèle VAR(1). Cependant, pour un modèle incluant N sites, le nombre total de paramètres est réduit de N^2 à seulement deux, à savoir ϕ_{10} et ϕ_{11} , la matrice de poids $\mathbf{W}^{(1)}$ étant considérée fixée. Pour revenir à l'exemple incluant 50 états américains décrits par un VAR(1), il y a donc une réduction de 2498 paramètres.

La réduction en terme de paramètres est importante dans les modèles STAR. Cependant, le modèle est possiblement trop rigide, supposant que les paramètres ϕ_{sk} sont les mêmes pour tous les sites. Des généralisations sont issues de la littérature géographique, voir Di Giacinto (2006). Puisque ces paramètres sont vraisemblablement différents pour des sites différents, les modèles STAR généralisés ont été proposés. Les modèles GSTAR font l'objet de la Définition 2.2.2.

Définition 2.2.2. *Le processus GSTAR($p_{\lambda_1, \dots, \lambda_p}$) représente une généralisation du processus STAR($p_{\lambda_1, \dots, \lambda_p}$), permettant que les paramètres autorégressifs varient selon le site. Ils s'écrivent comme suit :*

$$\mathbf{Y}_t = \sum_{s=1}^p \{ \mathbf{\Phi}_{s0} \mathbf{Y}_{t-s} + \sum_{k=1}^{\lambda_s} \mathbf{\Phi}_{sk} \mathbf{W}^{(k)} \mathbf{Y}_{t-s} \} + \mathbf{a}_t, \quad t = 0, \pm 1, \pm 2, \dots, \quad (2.2.2)$$

où $\phi_{sk}^{(i)}, i = 1, 2, \dots, N$ varient maintenant selon le site, étant indexés par i , $i = 1, \dots, N$. On regroupe ces paramètres dans les matrices $\mathbf{\Phi}_{sk}$, où $\mathbf{\Phi}_{sk} = \text{diag}(\phi_{sk}^{(1)}, \dots, \phi_{sk}^{(N)})$ et $\text{diag}(a_1, \dots, a_N)$ est une matrice diagonale $N \times N$. Le rôle de la matrice $\mathbf{W}^{(k)}$ est le même que celui défini dans le processus STAR de la Définition 2.2.1. On présume que $\{\mathbf{a}_t\}$ est un bruit blanc tel que $\text{Var}(\mathbf{a}_t) = \mathbf{\Omega}$.

Les hypothèses sur le terme d'erreur $\{\mathbf{a}_t\}$ présument que $\text{Var}(\mathbf{a}_t) = \mathbf{\Omega}$ est définie positive. La littérature géographique présume souvent que $\mathbf{\Omega}$ est diagonale. De plus, si $\mathbf{\Omega} = \sigma^2 \mathbf{I}_N$, la situation est alors dite homogène, sinon il est question d'hétérogénéité car les variances entre les sites sont différentes.

On note que le modèle (2.2.2) admet une écriture sous la forme d'un VAR. En effet,

$$\begin{aligned}
\mathbf{Y}_t &= \sum_{s=1}^p \left\{ \mathbf{\Phi}_{s0} \mathbf{Y}_{t-s} + \sum_{k=1}^{\lambda_s} \mathbf{\Phi}_{sk} \mathbf{W}^{(k)} \mathbf{Y}_{t-s} \right\} + \mathbf{a}_t, \\
&= \sum_{s=1}^p \left\{ \mathbf{\Phi}_{s0} + \sum_{k=1}^{\lambda_s} \mathbf{\Phi}_{sk} \mathbf{W}^{(k)} \right\} \mathbf{Y}_{t-s} + \mathbf{a}_t, \\
&= \sum_{s=1}^p \left\{ \sum_{k=0}^{\lambda_s} \mathbf{\Phi}_{sk} \mathbf{W}^{(k)} \right\} \mathbf{Y}_{t-s} + \mathbf{a}_t, \\
&= \sum_{s=1}^p \mathbf{\Phi}_s^* \mathbf{Y}_{t-s} + \mathbf{a}_t, \tag{2.2.3}
\end{aligned}$$

où $\mathbf{\Phi}_s^* = \sum_{k=0}^{\lambda_s} \mathbf{\Phi}_{sk} \mathbf{W}^{(k)}$. Dans la section suivante, nous analysons en détail le cas du processus GSTAR(1₁).

2.2.2. Analyse du processus GSTAR(1₁)

Le processus GSTAR(1₁) admet la représentation suivante :

$$Y_t(i) = \phi_{0i} Y_{t-1}(i) + \phi_{1i} \sum_{j=1}^N w_{ij} Y_{t-1}(j) + a_t(i),$$

pour les sites $i = 1, 2, \dots, N$ et $t \in \mathbb{Z}$. On note que comparativement à la Définition 2.2.2, nous avons simplifié l'écriture des coefficients ϕ_{0i} et ϕ_{1i} , $i = 1, \dots, N$. Sous forme matricielle, le modèle s'écrit :

$$\mathbf{Y}_t = (\mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{W}^{(1)}) \mathbf{Y}_{t-1} + \mathbf{a}_t, \tag{2.2.4}$$

où $E(\mathbf{Y}_t) = \mathbf{0}_N$, $\mathbf{Y}_t = (Y_t(1), \dots, Y_t(N))^T$, $\mathbf{a}_t = (a_t(1), \dots, a_t(N))^T$, $\mathbf{\Phi}_0 = \text{diag}(\phi_{01}, \phi_{02}, \dots, \phi_{0N})$, $\mathbf{\Phi}_1 = \text{diag}(\phi_{11}, \phi_{12}, \dots, \phi_{1N})$ et $\mathbf{W} = (w_{ij})_{i,j=1,\dots,N}$ avec $w_{ii} = 0$, $\forall i$, et $\sum_{j=1}^N w_{ij} = 1$.

Ainsi, on constate que le modèle $\text{GSTAR}(1_1)$ de l'équation (2.2.4) peut aussi s'écrire comme un modèle $\text{VAR}(1)$. Cependant, comme pour les processus STAR , pour un modèle incluant N sites et une matrice de poids $\mathbf{W}^{(1)}$ fixée, le nombre total de paramètres est réduit de N^2 à seulement $2N$, à savoir les éléments diagonaux $\{\phi_{01}, \phi_{11}, \phi_{02}, \phi_{12}, \dots, \phi_{0N}, \phi_{1N}\}$ des matrices $\mathbf{\Phi}_0$ et $\mathbf{\Phi}_1$. Considérant à nouveau l'exemple portant sur les 50 états américains modélisés par un $\text{VAR}(1)$, il y a donc une réduction de 2400 paramètres. La notion de stabilité d'un processus s'applique également aux processus GSTAR . Ainsi, il est possible de vérifier si le processus $\text{GSTAR}(1_1)$ est causal en étudiant la nature des racines du déterminant de l'opérateur autorégressif, c'est-à-dire en calculant les racines de $\det\{\mathbf{I}_N - (\mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{W})z\}$, voir le Théorème 1.1.1.

2.3. MÉTHODES D'IDENTIFICATION D'UN PROCESSUS $\text{GSTAR}(p_\lambda)$

La méthodologie usuelle entourant la construction de modèles de séries chronologiques repose sur l'étude des autocorrélations et autocorrélations partielles. La méthodologie est bien comprise dans le cas univarié, et a été généralisée dans le cas multivarié. On a noté cependant des difficultés à étendre au cadre multivarié le concept d'autocorrélation partielle, voir Tiao *et* Box (1981). Compte tenu qu'un modèle GSTAR admet une écriture VARMA , nous pourrions en théorie considérer les matrices d'autocovariance ou encore les matrices d'autorégression partielles afin de procéder à l'identification des ordres maximaux dans l'écriture VARMA . Cependant cette stratégie ne permet pas d'identifier les ordres λ_s , $s = 1, \dots, p$. Afin de définir une mesure espace-temps, on utilise les opérateurs retards temporels et spatiaux. La fonction d'autocovariance espace-temps est donnée dans la définition suivante. Elle est due à Pfeifer *et al.* (1980a).

Définition 2.3.1. *La fonction d'autocovariance espace-temps de délai temporel s est définie par :*

$$\gamma_{lk}(s) = E \left\{ N^{-1} (\mathbf{W}^{(l)} \mathbf{Y}_t)^\top (\mathbf{W}^{(k)} \mathbf{Y}_{t+s}) \right\}, \quad (2.3.1)$$

où $l, k \in \{0, 1, \dots, N\}$.

Dans la définition précédente nous présumons que $E(\mathbf{Y}_t) = \mathbf{0}$. Le cas non-centré se traite de façon naturelle. Cette définition de $\gamma_{lk}(s)$ donne lieu à une

mesure scalaire. On note qu'une façon alternative d'écrire $\gamma_{lk}(s)$ est comme suit :

$$\gamma_{lk}(s) = E \left[N^{-1} \sum_{i=1}^N \{L^{(l)}Y_t(i)\} \{L^{(k)}Y_{t+s}(i)\} \right].$$

Ainsi, cette quantité est une covariance pondérée entre les voisins d'ordre l de n'importe quel site avec les voisins d'ordre k du même site, avec un retard temporel de s . La mesure est moyenne calculée par rapport aux N sites. Autrement formulé, $\gamma_{lk}(s)$ représente une moyenne à travers les sites, de la mesure de covariance $Cov(L^{(l)}Y_t(i), L^{(k)}Y_{t+s}(i))$: pour un site i donné, cette mesure décrit la dépendance entre les retards spatiaux d'ordre l et k avec un délai temporel s . La mesure $\gamma_{lk}(s)$ traduit donc le degré de dépendance moyenne entre les retards spatiaux d'ordre l et k , entre les variables espacées de s espaces temporels. On note que :

$$\begin{aligned} \gamma_{lk}(s) &= E \left\{ N^{-1} (\mathbf{W}^{(l)}\mathbf{Y}_t)^\top (\mathbf{W}^{(k)}\mathbf{Y}_{t+s}) \right\}, \\ &= E \left\{ N^{-1} \mathbf{Y}_t^\top (\mathbf{W}^{(l)})^\top \mathbf{W}^{(k)}\mathbf{Y}_{t+s} \right\}, \\ &= N^{-1} E \left[tr \left\{ \mathbf{Y}_t^\top (\mathbf{W}^{(l)})^\top \mathbf{W}^{(k)}\mathbf{Y}_{t+s} \right\} \right], \\ &= N^{-1} E \left[tr \left\{ (\mathbf{W}^{(l)})^\top \mathbf{W}^{(k)}\mathbf{Y}_{t+s}\mathbf{Y}_t^\top \right\} \right], \\ &= N^{-1} tr \left\{ (\mathbf{W}^{(l)})^\top \mathbf{W}^{(k)}\mathbf{\Gamma}_Y(s) \right\}. \end{aligned} \tag{2.3.2}$$

La troisième égalité dans le développement précédent repose sur le fait que la trace d'un scalaire est précisément cette valeur scalaire. La dernière égalité s'obtient en invoquant la linéarité de la trace et de l'espérance. De cette expression algébrique, il ressort que $\gamma_{lk}(s)$ est une fonction scalaire de la fonction d'autocovariance du processus $\{\mathbf{Y}_t\}$. On note que $\gamma_{lk}(s)$ est définie pour $l, k \in \{0, 1, \dots, N\}$. Puisque $\mathbf{W}^{(0)} = \mathbf{I}_N$, on note que $\gamma_{00}(s) = N^{-1}tr\{\mathbf{\Gamma}(s)\}$, qui est donc une moyenne (à travers les sites) des fonctions d'autocovariance univariées de délai s , $i = 1, \dots, N$. Une autre expression utile est $\gamma_{10}(s) = N^{-1}tr\{\mathbf{W}^{(1)}\mathbf{\Gamma}(s)\}$, qui est la moyenne à travers les N sites de l'autocovariance (univariée) du site i avec son voisin de premier ordre (pondéré), de délai temporel s . Pfeifer *et al.* (1980a) donnent d'autres propriétés de $\gamma_{lk}(s)$.

Tout comme l'autocorrélation est plus commode que l'autocovariance, il est souhaitable d'avoir une fonction d'autocorrélation espace-temps. Ceci fait l'objet de la Définition 2.3.2.

Définition 2.3.2. *La fonction d'autocorrélation espace-temps entre les voisins pondérés d'ordre l et k , de délai temporel s , est définie par :*

$$\rho_{lk}(s) = \frac{\gamma_{lk}(s)}{\{\gamma_{ll}(0)\gamma_{kk}(0)\}^{\frac{1}{2}}}, \quad (2.3.3)$$

où $\gamma_{lk}(s)$ est fonction d'autocovariance espace-temps, $l, k \in \{0, 1, \dots, N\}$.

Ayant la fonction d'autocovariance espace-temps, il est possible de se définir un système de type Yule-Walker (voir par exemple Shumway *et* Stoffer (2006, p. 122)). Ce système peut être utilisé afin de définir une fonction d'autocorrélation partielle espace-temps. Les détails se trouvent dans Pfeifer *et al.* (1980a, p. 39). Des formules exactes pour les autocorrélations espaces-temps sont obtenues dans le cas $p = 1$ dans Pfeifer *et al.* (1980b). Il est clair que la littérature portant sur la modélisation des séries chronologiques utilise beaucoup les propriétés des autocorrélations et autocorrélations partielles comme outils d'identification. Dans le cas des modèles GSTAR, Pfeifer *et al.* (1980a) procèdent ainsi afin de se définir une fonction d'autocorrélation partielle espace-temps. Il est utile de consulter Wei (1994, pp. 12-16) pour comprendre comment la démarche est calquée sur ce qui est connu dans les modèles ARMA univariés classiques.

Dans le cas univarié classique, rappelons qu'une façon d'introduire la fonction d'autocorrélation partielle consiste à considérer l'autorégression suivante :

$$Y_{t+h} = \phi_{h1}Y_{t+h-1} + \dots + \phi_{hh}Y_t + a_{t+h}.$$

Si on multiplie par Y_{t+h-j} et que l'on prend l'espérance, nous obtenons le système :

$$\gamma(j) = \phi_{h1}\gamma(j-1) + \dots + \phi_{hh}\gamma(j-h), \quad j = 1, \dots, h.$$

En résolvant le système possédant h équations et h inconnues, le coefficient ϕ_{hh} est précisément l'autocorrélation partielle de délai h . On note que ceci n'est pas une approche permettant de donner une définition à l'autocorrélation partielle, mais

bien une méthode calculatoire qui permet de déterminer une mesure de corrélation conditionnelle entre Y_{t+h} et Y_t , conditionnellement aux variables aléatoires intermédiaires $Y_{t+h-1}, \dots, Y_{t+1}$. Voir Wei (1994, pp. 12-16).

Afin de se définir une fonction d'autocorrélation partielle espace-temps, considérons maintenant l'autorégression suivante :

$$\mathbf{Y}_t = \sum_{j=1}^k \sum_{l=0}^{\lambda} \phi_{jl}^{(k\lambda)} \{\mathbf{W}^{(l)} \mathbf{Y}_{t-j}\} + \mathbf{u}_t. \quad (2.3.4)$$

Si on multiplie (2.3.4) par $\{\mathbf{W}^{(h)} \mathbf{Y}_{t-s}\}^\top$ et que l'on prend l'espérance, nous obtenons le système suivant :

$$\gamma_{h0}(s) = \sum_{j=1}^k \sum_{l=0}^{\lambda} \phi_{jl}^{(k\lambda)} \gamma_{hl}(s-j), \quad s = 1, \dots, k, \quad (2.3.5)$$

où l'on a utilisé l'hypothèse de stationnarité causale du processus sous-jacent, à savoir $E(\mathbf{Y}_{t-s}^\top \mathbf{u}_t) = 0$, $s > 0$. Nous obtenons ainsi un système pour $s = 1, \dots, k$ et $h = 0, 1, \dots, \lambda$, où k et λ sont fixes. Nous avons ainsi un système possédant $k(\lambda + 1)$ équations et $k(\lambda + 1)$ inconnues. On note que Pfeifer *et al.* (1980a) dans leur Figure 1 représentent explicitement le système à résoudre, qui peut être utile afin de fixer les idées. Pfeifer *et al.* (1980a) définissent la fonction d'autocorrélation espace-temps partielle comme le dernier coefficient $\phi_{jl}^{(k\lambda)}$, pour $l = 0, 1, \dots, \lambda$. Ainsi, la fonction d'autocorrélation espace-temps partielle d'ordre $\lambda = 2$ contient les termes $\phi_{10}^{(12)}$, $\phi_{11}^{(12)}$, $\phi_{12}^{(12)}$ lorsque $k = 1$, $\phi_{20}^{(22)}$, $\phi_{21}^{(22)}$, $\phi_{22}^{(22)}$ lorsque $k = 2$, etc.. Pfeifer *et al.* (1980a) discutent du choix de λ , qui ne devrait pas être trop grand.

Une propriété intéressante pour un modèle STAR est que la fonction d'autocorrélation partielle espace-temps s'annule après p délais temporels et λ_p délais dans l'espace. Ainsi, cet outil est un outil d'identification qui peut s'avérer utile. Voir Pfeifer *et al.* (1980a, p. 39) pour plus de détails.

Suivant les procédures de Box *et* Jenkins, la méthode populaire d'identification pour une série chronologique est d'utiliser la fonction d'autocorrélation temporelle et la fonction d'autocorrélation partielle. Les outils statistiques pour identifier les modèles spatiaux s'inscrivent donc dans ce même ordre d'idée. Dans ce cas, l'ordre autorégressif temporel et spatial des modèles spatiaux représentatifs se

déterminent par l'étude de la forme appropriée des fonctions d'autocorrélation et d'autocorrélation partielle espace-temps.

2.3.1. Fonctions d'autocovariance d'un processus GSTAR(p_λ)

Compte tenu qu'un modèle GSTAR(p_λ) admet une représentation VARMA, nous pouvons utiliser l'écriture des fonctions d'autocovariance et d'autocorrélation partielle d'un processus VAR(p) afin de déterminer la forme de la fonction d'autocovariance et la fonction d'autocorrélation d'un processus GSTAR(p_λ). Ainsi, se basant sur (1.1.9) et (2.2.3), nous pouvons déduire que le système d'équations de Yule-Walker d'un processus GSTAR(p_λ) stable est :

$$\begin{cases} \mathbf{\Gamma}(0) = \sum_{s=1}^p \mathbf{\Phi}_s^* \mathbf{\Gamma}^\top(s) + \mathbf{\Sigma}_a, & h = 0, \\ \mathbf{\Gamma}(h) = \sum_{s=1}^p \mathbf{\Phi}_s^* \mathbf{\Gamma}(h-s), & h > 0, \end{cases} \quad (2.3.6)$$

où $\mathbf{\Phi}_s^* = \sum_{k=0}^{\lambda_s} \mathbf{\Phi}_{sk} \mathbf{W}^{(k)}$.

De ce résultat, nous pouvons dégager la fonction d'autocovariance espace-temps en utilisant (2.3.2). en effet :

$$\begin{cases} \gamma_{lk}(s) = N^{-1} tr \left\{ (\mathbf{W}^{(l)})^\top \mathbf{W}^{(k)} \left(\sum_{s=1}^p \mathbf{\Phi}_s^* \mathbf{\Gamma}^\top(s) + \mathbf{\Sigma}_a \right) \right\}, & h = 0, \\ \gamma_{lk}(s) = N^{-1} tr \left\{ (\mathbf{W}^{(l)})^\top \mathbf{W}^{(k)} \sum_{s=1}^p \mathbf{\Phi}_s^* \mathbf{\Gamma}(h-s) \right\}, & h \geq 0, \end{cases} \quad (2.3.7)$$

où $\mathbf{\Phi}_s^* = \sum_{k=0}^{\lambda_s} \mathbf{\Phi}_{sk} \mathbf{W}^{(k)}$. Cette dernière forme (2.3.7) peut être utilisée pour obtenir les fonctions d'autocorrélation espace-temps.

2.4. ESTIMATION DES PARAMÈTRES D'UN MODÈLE GSTAR(p_λ)

Comme pour les modèles VARMA, la méthode du maximum de vraisemblance et la méthode des moindres carrés généralisés peuvent être aussi utilisées pour estimer les paramètres d'un processus GSTAR. Voir par exemple, Borovkova *et al.* (2008) et Di Giacinto (2006). Dans cette section, nous discutons uniquement de la méthode des moindres carrés, une procédure étudiée par Borovkova *et al.* (2008). Comme dans son article, cette méthode sera présentée pour un GSTAR(1_1) stationnaire, puis on la généralisera pour un modèle GSTAR(p_λ) stationnaire. Nous présentons en détails cette technique car elle sera utilisée dans nos études de simulations.

Pour estimer les paramètres Φ_0 et Φ_1 du GSTAR(1₁) de l'équation (2.4.1) suivante,

$$\mathbf{Y}_t = (\Phi_0 + \Phi_1 \mathbf{W}) \mathbf{Y}_{t-1} + \mathbf{a}_t, \quad (2.4.1)$$

Borovkova *et al.* (2008) présument que les observations sont centrées ($\mathbf{Y}_0, \dots, \mathbf{Y}_n$), que $\{\mathbf{a}_t\}$ est un terme d'erreur pas forcément constitué de vecteurs aléatoires indépendants, mais satisfaisant une condition du type différence de martingales. Afin de procéder à l'estimation, il est utile de ré-écrire le modèle comme un modèle usuel de régression linéaire multiple. Afin de construire un seul modèle de régression linéaire, il est utile de construire dans un premier temps un modèle linéaire pour chaque site i , $i = 1, \dots, N$. On note que l'écriture est différente de ce qui a été présenté précédemment, par exemple \mathbf{Y}_t dans le modèle GSTAR de la Définition 2.2.2 contient tous les sites au même instant t . Dans ce qui suit, les vecteurs \mathbf{h}_i , $i = 1, \dots, N$ contiennent tous les instants t , $t = 1, \dots, n$, pour un site donné i . Plus précisément, pour chaque site i , on peut écrire :

$$\mathbf{h}_i = \begin{pmatrix} Y_1(i) \\ Y_2(i) \\ \vdots \\ Y_n(i) \end{pmatrix} = \mathbf{X}_i \boldsymbol{\beta}_i + \mathbf{u}_i, \quad (2.4.2)$$

où

$$\mathbf{X}_i = \begin{pmatrix} Y_0(i) & V_0(1) \\ Y_1(i) & V_1(i) \\ \vdots & \vdots \\ Y_{n-1}(i) & V_{n-1}(i) \end{pmatrix}, \quad (2.4.3)$$

$\boldsymbol{\beta}_i = (\phi_{0i}, \phi_{1i})^\top$ et $\mathbf{u}_i = (a_1(i), \dots, a_n(i))^\top$. On a défini $V_t(i) = \sum_{i \neq j} w_{ij} Y_t(j)$ dans la formule (2.4.3). Par la suite, on regroupe dans un seul modèle les N sites,

ce qui donne :

$$\begin{pmatrix} Y_1(1) \\ Y_2(1) \\ \vdots \\ Y_n(1) \\ \vdots \\ Y_1(N) \\ Y_2(N) \\ \vdots \\ Y_n(N) \end{pmatrix} = \begin{pmatrix} Y_0(1) & V_0(1) & \cdots & 0 & 0 \\ Y_1(1) & V_1(1) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_{n-1}(1) & V_{n-1}(1) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & Y_0(N) & V_0(N) \\ 0 & 0 & \cdots & Y_1(N) & V_1(N) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & Y_{n-1}(N) & V_{n-1}(N) \end{pmatrix} \begin{pmatrix} \phi_{01} \\ \phi_{11} \\ \vdots \\ \phi_{0N} \\ \phi_{1N} \end{pmatrix} + \begin{pmatrix} a_1(1) \\ a_2(1) \\ \vdots \\ a_n(1) \\ \vdots \\ a_1(N) \\ a_2(N) \\ \vdots \\ a_n(N) \end{pmatrix}.$$

De manière compacte, on écrit :

$$\mathbf{h} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad (2.4.4)$$

où $\mathbf{h} = (\mathbf{h}_1^\top, \mathbf{h}_2^\top, \dots, \mathbf{h}_N^\top)^\top$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top, \dots, \boldsymbol{\beta}_N^\top)^\top$ et $\mathbf{u} = (\mathbf{u}_1^\top, \mathbf{u}_2^\top, \dots, \mathbf{u}_N^\top)^\top$. La matrice de design \mathbf{X} est une matrice diagonale par blocs et peut s'écrire comme :

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_N \end{pmatrix}.$$

Ainsi, les vecteurs \mathbf{h} et \mathbf{u} jouent les rôles de la variable réponse et du terme d'erreur et sont des vecteurs de dimension $(Nn) \times 1$. Les dimensions de \mathbf{X} sont $(Nn) \times (2pN)$. Finalement, $\boldsymbol{\beta}$ est de dimension $(2pN) \times 1$, avec $p = 1$. Il est à noter que pour chaque i , $i = 1, \dots, N$, nous avons un modèle linéaire séparé, de la forme $\mathbf{h}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{u}_i$. Comme noté dans Borovkova *et al.* (2008, p. 488), ceci est particulièrement commode car les estimations peuvent être faites séparément, site par site. Il est à noter que les estimateurs $\hat{\boldsymbol{\beta}}_i$, $i = 1, \dots, N$, ne seront cependant pas indépendants, car les matrices de design sont stochastiques et font intervenir les autres sites à travers les variables aléatoires $V_t(i) = \sum_{j \neq i} w_{ij} Y_t(j)$. Si l'estimation est faite globalement, l'estimateur des moindres carrés de $\boldsymbol{\beta}$ est tout de même :

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{h}. \quad (2.4.5)$$

Dans le cas général, la démarche est en tout point la même, mais des complications d'écriture apparaissent. Ainsi, pour un $\text{GSTAR}(p_{\lambda_1, \lambda_2, \dots, \lambda_p})$, on procède ainsi. On commence par calculer les matrices spatiales $\mathbf{W}^{(\lambda_i)}$ pour les différents ordres $\lambda_i, i = 1, 2, \dots, p$, puis on détermine les variables dépendantes spatiales en appliquant la formule : $V_t^{(k)}(i) = \sum_{j \neq i}^N w_{ij}^{(k)} Y_t(j)$, où $k = \lambda_1, \lambda_2, \dots, \lambda_p$ et $V_t^{(0)}(i) = Y_t(i)$. Finalement, on estime les paramètres du modèle général (2.4.4) après avoir déterminé les matrices respectives $\mathbf{h}_i, \mathbf{X}_i$, où $\mathbf{h} = (\mathbf{h}_1^\top, \dots, \mathbf{h}_N^\top)^\top$, $\mathbf{X} = \text{diag}(\mathbf{X}_1, \dots, \mathbf{X}_N)$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \dots, \boldsymbol{\beta}_N^\top)^\top$ et $\mathbf{u}_i = (a_1(i), \dots, a_n(i))^\top$, avec $\mathbf{h}_i = (Y_1(i), \dots, Y_n(i))^\top$, $\boldsymbol{\beta}_i = (\phi_{10}^{(i)}, \dots, \phi_{1\lambda_1}^{(i)}, \phi_{20}^{(i)}, \dots, \phi_{2\lambda_2}^{(i)}, \dots, \phi_{p0}^{(i)}, \dots, \phi_{p\lambda_p}^{(i)})^\top$ et

$$\mathbf{X}_i = \begin{pmatrix} V_{p-1}^{(0)}(i) & \dots & V_{p-1}^{(\lambda_1)}(i) & \dots & V_0^{(0)}(i) & \dots & V_{p-1}^{(\lambda_1)}(i) \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ V_{n-1}^{(0)}(i) & \dots & V_{n-1}^{(\lambda_1)}(i) & \dots & V_{n-p}^{(0)}(i) & \dots & V_{n-p}^{(\lambda_1)}(i) \end{pmatrix}.$$

Pour plus de détails entourant cette technique d'estimation et les résultats asymptotiques concernant l'estimation de $\boldsymbol{\beta}$, voir Borovkova *et al.* (2008, pp. 492-493).

L'objectif principal de cette section était de présenter les concepts fondamentaux en statistique spatiale et quelques modèles spatio-temporels multivariés. Le prochain chapitre introduit des tests de validation en séries chronologiques multivariées basés sur les autocovariances résiduelles matricielles, et des méthodes sont présentées pour les modèles GSTAR .

Chapitre 3

STATISTIQUES DE TEST POUR DIAGNOSTIQUER DES MODÈLES $GSTAR(P_\lambda)$

Le chapitre précédent a présenté la classe des modèles $GSTAR$. De plus, nous avons discuté de l'estimation des paramètres pour ces modèles. Un modèle correctement ajusté devrait avoir des résidus compatibles avec l'hypothèse nulle de bruit blanc du terme d'erreur. Dans un contexte de modèles de séries chronologiques multivariées, les matrices d'autocovariances résiduelles peuvent être utilisées à cette fin. Ce qui distingue notre contexte du cadre des modèles $VARMA$ est le fait que nos modèles $GSTAR$ présentent une structure particulière, reposant sur des matrices de poids possédant des interprétations bien précises. Ainsi, dans un premier temps, nous allons présenter les résultats généraux de Ahn (1988), qui portent sur l'étude de la distribution des autocovariances résiduelles dans des modèles VAR structurés. Ahn (1988) a également discuté des statistiques de test de type portemanteau dans ce contexte. De ce point de vue, nos résultats sont de nouvelles applications du cadre général de Ahn (1988). Nous étudions également les statistiques portemanteaux, et nous prenons soin de fournir les degrés de liberté appropriés pour les statistiques de test possédant approximativement des lois de distribution khi-carré.

3.1. AUTOCOVARIANCES RÉSIDUELLES MATRICIELLES

Considérons un certain modèle de séries chronologiques de dimension N . À ce stade, il pourrait être question d'un modèle VAR ou $VARMA$. La validation d'un

modèle qui a été ajusté pour les données \mathbf{Y}_t , $t = 1, \dots, n$, repose généralement sur l'étude des mesures de corrélation résiduelle, comme les autocovariances ou autocorrélations matricielles résiduelles. Formellement, soit $\{\mathbf{a}_t\}$ le terme d'erreur d'un certain modèle. L'adéquation revient donc à tester l'hypothèse nulle : $\mathbf{\Gamma}_a(h) = \mathbf{0}$, $h = 1, 2, \dots$, où $\{\mathbf{\Gamma}_a(h), h \in \mathbb{Z}\}$ est la fonction d'autocovariance matricielle de $\{\mathbf{a}_t\}$. Puisque la fonction d'autocovariance est un outil théorique, introduisons les résidus $\hat{\mathbf{a}}_t$, $t = 1, \dots, n$, de ce modèle. Les autocovariances résiduelles sont définies par :

$$\mathbf{C}_{\hat{\mathbf{a}}}(h) = \begin{cases} n^{-1} \sum_{t=h+1}^n \hat{\mathbf{a}}_t \hat{\mathbf{a}}_{t-h}^\top, & h = 0, 1, \dots, n-1, \\ \mathbf{C}_{\hat{\mathbf{a}}}^\top(-h), & h = -1, \dots, -n+1. \end{cases}$$

On pose $\mathbf{c}_{\hat{\mathbf{a}}}(h) = \text{vec}\{\mathbf{C}_{\hat{\mathbf{a}}}(h)\}$, $\mathbf{c}_{\hat{\mathbf{a}}}^{(M)} = (\mathbf{c}_{\hat{\mathbf{a}}}^\top(1), \dots, \mathbf{c}_{\hat{\mathbf{a}}}^\top(M))^\top$, où M est un entier fixé à spécifier, tel que $M < n$. Les propriétés des autocovariances et autocorrélations résiduelles ont été étudiées par plusieurs auteurs, particulièrement Li *et* McLeod (1981) qui ont établi la distribution asymptotique normale de $\mathbf{c}_{\hat{\mathbf{a}}}$ dans les modèles VAR et VARMA. Duchesne (2005) donne la distribution asymptotique des autocovariances et autocorrélations résiduelles dans les modèles avec variables exogènes de type VARX

On note que les autocovariances résiduelles satisfont $\mathbf{C}_{\hat{\mathbf{a}}}(h) = (C_{\hat{\mathbf{a}},ij}(h))$ pour $i, j = 1, \dots, N$, $|h| < n$. On peut alors définir les matrices d'autocorrélations résiduelles comme : $\mathbf{R}_{\hat{\mathbf{a}}}(h) = (R_{\hat{\mathbf{a}},ij}(h))_{i,j=1,\dots,N}$, où

$$R_{\hat{\mathbf{a}},ij}(h) = \frac{C_{\hat{\mathbf{a}},ij}(h)}{\{C_{\hat{\mathbf{a}},ii}(0)C_{\hat{\mathbf{a}},jj}(0)\}^{\frac{1}{2}}}, \quad i, j \in \{1, \dots, N\}.$$

Si on pose $\mathbf{r}_{\hat{\mathbf{a}}}(h) = \text{vec}\{\mathbf{R}_{\hat{\mathbf{a}}}(h)\}$ et $\mathbf{r}_{\hat{\mathbf{a}}}^{(M)} = (\mathbf{r}_{\hat{\mathbf{a}}}^\top(1), \dots, \mathbf{r}_{\hat{\mathbf{a}}}^\top(M))^\top$, $M < n$, Li *et* McLeod (1981) ont montré pour les modèles VAR et VARMA que l'expansion de Taylor suivante est valide :

$$\mathbf{r}_{\hat{\mathbf{a}}}^{(M)} = \mathbf{r}_{\hat{\mathbf{a}}}^{(M)} - \mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + \mathbf{o}_p(n^{-1}), \quad (3.1.1)$$

où $\mathbf{r}_{\hat{\mathbf{a}}}^{(M)}$ représente le vecteur contenant la fonction d'autocorrélation échantillonnale de $\{\mathbf{a}_t\}$, \mathbf{X} est une matrice de dérivées définies dans Li *et* McLeod (1981), $\hat{\boldsymbol{\beta}}$

est l'estimateur à vraisemblance maximale de $\boldsymbol{\beta}$, contenant les paramètres autorégressifs moyennes mobiles du modèle VARMA. Ainsi, il peut être montré que :

$$\sqrt{n}\hat{\mathbf{r}}_{\mathbf{a}}^{(M)} \rightarrow_{\mathcal{L}} \mathcal{N}_{N^2M}(\mathbf{0}, \mathbf{I}_M \otimes \boldsymbol{\Sigma}_{\mathbf{a}} \otimes \boldsymbol{\Sigma}_{\mathbf{a}} - \mathbf{X}\mathcal{I}^{-1}\mathbf{X}^{\top}), \quad (3.1.2)$$

où \mathcal{I} est la matrice d'information de Fisher associée à $\hat{\boldsymbol{\beta}}$. L'utilisation du résultat (3.1.2) est particulièrement commode dans les applications. Pour un délai h particulier, il est pertinent de vérifier si les autocovariances ou autocorrélations résiduelles semblent significativement différentes du vecteur nul. Afin de vérifier cette hypothèse, il est possible d'utiliser les variances asymptotiques contenues dans (3.1.2). Si $\sqrt{n}\hat{\mathbf{c}}_{\mathbf{a}}(h) \rightarrow_{\mathcal{L}} \mathcal{N}_{N^2}(\mathbf{0}, \boldsymbol{\Omega}(h))$ où $\boldsymbol{\Omega}(h)$ est non-singulière, on peut définir des statistiques de test de la forme :

$$H(h) = n\mathbf{c}_{\hat{\mathbf{a}}}^{\top}(h)\hat{\boldsymbol{\Omega}}^{-1}(h)\mathbf{c}_{\hat{\mathbf{a}}}(h),$$

$$H^*(h) = \frac{n}{n - |h|}H(h),$$

où $\hat{\boldsymbol{\Omega}}(h)$ représente un estimateur convergeant de $\boldsymbol{\Omega}(h)$. Ce type de statistique a été considéré pour des VAR avec variables exogènes dans Duchesne (2005). Les grandes valeurs de $H(h)$ ou $H^*(h)$ sont celles qui entraînent le rejet de l'hypothèse nulle. La version modifiée $H^*(h)$ a souvent de meilleures propriétés en échantillons finis, dans le sens que la distribution asymptotique approxime mieux la loi exacte de $H^*(h)$ que celle de $H(h)$.

Pour des modèles VAR ou VARMA, la matrice des variances et covariances asymptotique est approximativement singulière, et il a été montré que les statistiques de test sont de la forme suivante :

$$Q(M) = n \sum_{h=1}^M \mathbf{c}_{\hat{\mathbf{a}}}^{\top}(h) \{ \hat{\boldsymbol{\Omega}}^{-1}(h) \otimes \hat{\boldsymbol{\Omega}}^{-1}(h) \} \mathbf{c}_{\hat{\mathbf{a}}}(h), \quad (3.1.3)$$

$$Q^*(M) = n^2 \sum_{h=1}^M (n - h)^{-1} \mathbf{c}_{\hat{\mathbf{a}}}^{\top}(h) \{ \hat{\boldsymbol{\Omega}}^{-1}(h) \otimes \hat{\boldsymbol{\Omega}}^{-1}(h) \} \mathbf{c}_{\hat{\mathbf{a}}}(h), \quad (3.1.4)$$

suivent approximativement des lois χ^2 avec $N^2(M - p - q)$ degrés de liberté, où p et q sont les ordres autorégressif et moyenne mobile, respectivement. Ces statistiques ont été étudiées par Hosking (1980), voir aussi Li et McLeod (1981).

Duchesne (2005) discute comment des statistiques portemanteaux peuvent être développées dans des modèles VAR avec des variables exogènes.

3.1.1. Statistiques de test basées sur les matrices d'autocovariances résiduelles pour les modèles VARMA

La littérature portant sur les séries chronologiques a accordé beaucoup d'attention sur l'étude des autocorrélations et autocovariances résiduelles. Dans un cadre multivarié, des extensions ont également été considérées par de nombreux auteurs. Cependant, les modèles vectoriels sont souvent accompagnés d'un grand nombre de paramètres, et nous avons vu que le phénomène est amplifié dans les modèles spatio-temporels. Cependant, une structure dans les matrices autorégressives et moyennes mobiles risquent d'affecter les propriétés asymptotiques des autocorrélations résiduelles et donc des statistiques de type portemanteau.

À notre connaissance, Ahn (1988) a été le premier à s'intéresser aux autocovariances résiduelles dans les modèles autorégressifs vectoriels où les matrices autorégressives ont une structure particulière. En fait, son cadre consiste à étudier le modèle suivant :

$$\mathbf{Y}_t = \sum_{j=1}^p \Phi_j(\boldsymbol{\beta}) \mathbf{Y}_{t-j} + \mathbf{a}_t, \quad t \in \mathbb{Z}, \quad (3.1.5)$$

où les matrices autorégressives $\Phi_j(\boldsymbol{\beta})$, $j = 1, \dots, p$, de dimension $N \times N$, sont des fonctions continues et dérivables du vecteur $\boldsymbol{\beta}$ qui est supposé être de dimension $\mathcal{B} \times 1$, où $\mathcal{B} \leq N^2 p$. Dans ce qui suit, le terme d'erreur $\{\mathbf{a}_t\}$ est un bruit blanc fort, avec $\text{Var}(\mathbf{a}_t) = \boldsymbol{\Sigma}_a$, où $\boldsymbol{\Sigma}_a$ est une matrice de variance définie positive. On note que le modèle VAR structuré inclut le modèle VAR usuel, si $\boldsymbol{\beta} = \text{vec}\{(\Phi_1, \dots, \Phi_p)^\top\}$. Cependant, dans les applications, il est attendu que \mathcal{B} sera beaucoup plus petit que $N^2 p$.

Ahn (1988) considère l'étude de modèles VAR structurés gaussiens. Ainsi, l'estimateur à vraisemblance maximale $\hat{\boldsymbol{\beta}}$ reposant sur la série chronologique $\mathbf{Y}_1, \dots, \mathbf{Y}_n$, maximise la vraisemblance :

$$l(\boldsymbol{\beta}, \boldsymbol{\Omega}) = -\frac{n}{2} \log \det(\boldsymbol{\Sigma}_a) - \frac{1}{2} \sum \mathbf{a}_t^\top \boldsymbol{\Omega}_a^{-1} \mathbf{a}_t. \quad (3.1.6)$$

Il peut alors être montré sous certaines conditions que :

$$\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \mathcal{I}^{-1}(\boldsymbol{\beta})\{n^{-1}\frac{\partial l}{\partial \boldsymbol{\beta}}\} + \mathbf{o}_p(n^{-\frac{1}{2}}), \quad (3.1.7)$$

où $\mathcal{I}(\boldsymbol{\beta})$ est la matrice d'information de Fisher associée au vecteur de paramètres $\boldsymbol{\beta}$. Il peut être montré que :

$$\begin{aligned} \mathcal{I}(\boldsymbol{\beta}) &= \sum_{i=1}^p \sum_{j=1}^p \frac{\partial \phi_i^\top}{\partial \boldsymbol{\beta}} \{\boldsymbol{\Omega}_a^{-1} \otimes \boldsymbol{\Gamma}(i-j)\} \frac{\partial \phi_j}{\partial \boldsymbol{\beta}^\top}, \\ &= \mathbf{H}^\top \mathbf{G} \mathbf{H}, \end{aligned} \quad (3.1.8)$$

où $\phi_j = \text{vec}\{\boldsymbol{\Phi}_j^\top(\boldsymbol{\beta})\}$, $j = 1, \dots, p$, et $\boldsymbol{\Gamma}(\cdot)$ correspond à la fonction d'autocovariance du processus $\{\mathbf{Y}_t\}$. Il est possible de décrire explicitement les matrices \mathbf{G} et \mathbf{H} . La matrice \mathbf{G} est explicitée dans Ahn *et* Reinsel (1988). La matrice \mathbf{H} est une matrice de dérivées qui est détaillée dans Ahn (1988). Utilisant des résultats de McLeod (1978) et un théorème limite central pour différences de martingales, la normalité asymptotique jointe de $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}, \mathbf{c}_a^{(M)})$ peut être établie. Utilisant l'expansion de Taylor suivante :

$$\mathbf{c}_a^{(M)} = \mathbf{c}_a^{(M)} - \mathbf{U}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + \mathbf{o}_p(n^{-\frac{1}{2}}), \quad (3.1.9)$$

il ressort que :

$$\sqrt{n}\mathbf{c}_a^{(M)} \rightarrow_{\mathcal{L}} \mathcal{N}_{N^2M}(\mathbf{0}, \mathbf{I}_M \otimes \boldsymbol{\Sigma}_a \otimes \boldsymbol{\Sigma}_a - \mathbf{X}\{\mathcal{I}^{-1}(\boldsymbol{\beta})\}\mathbf{X}^\top). \quad (3.1.10)$$

La matrice \mathbf{X} dans le résultat (3.1.10) est telle que :

$$\frac{\partial \mathbf{c}_a^{(M)}}{\partial \boldsymbol{\beta}^\top} = -\mathbf{X} + \mathbf{o}_p(1). \quad (3.1.11)$$

Considérons maintenant les statistiques de test (3.1.3) et (3.1.4). En effectuant des analyses similaires à ce qui est fait pour des modèles VAR et VARMA, il peut être montré que les distributions asymptotiques de ces statistiques de test sont toujours approximativement de lois du khi-carré, avec cependant $N^2M - \mathcal{B}$ degrés de liberté. Ainsi dans le cas classique des modèles VAR, $\mathcal{B} = N^2p$ et on retrouve le résultat stipulant que (3.1.3) et (3.1.4) sont approximativement de lois $\chi_{N^2(M-p)}^2$. Les modèles avec des contraintes sur les paramètres représentent une autre façon de formuler le problème. Les résultats portant sur les VAR avec des contraintes sur les paramètres sont discutés dans Lütkepohl (2005, chapitre 5).

3.2. TESTS DE TYPE PORTEMANTEAU POUR UN MODÈLE GSTAR

Dans la Section 2.2, nous avons montré que le modèle GSTAR peut être reformulé sous la forme d'un modèle VARMA structuré, voir l'équation (2.2.3). En effet, en posant $\Phi_j(\beta) = \sum_{k=0}^{\lambda_j} \Phi_{jk} \mathbf{W}^{(k)}$ où $\Phi_{jk} = \text{diag}(\phi_{jk}^{(1)}, \dots, \phi_{jk}^{(N)})$, le modèle GSTAR(p_λ) de (2.2.3) s'écrit sous la forme d'un VAR(p) structuré et admet une écriture sous la forme du modèle (3.1.5). Dans notre situation β s'écrit comme :

$$\beta = \left(\phi_{10}^{(1)}, \dots, \phi_{10}^{(N)}, \dots, \phi_{1\lambda_1}^{(1)}, \dots, \phi_{1\lambda_1}^{(N)}, \dots, \phi_{p0}^{(1)}, \dots, \phi_{p0}^{(N)}, \dots, \phi_{p\lambda_p}^{(1)}, \dots, \phi_{p\lambda_p}^{(N)} \right)^\top$$

et est donc de dimension $\sum_{j=1}^p N(\lambda_j + 1) \times 1$. Si $\lambda_j \equiv 1$, on trouve que β est de dimension $(2pN) \times 1$. Ce cas sera traité plus loin lorsque $p = 1$. On note que notre modèle s'inscrit dans les hypothèses de Ahn (1988). En effet, les matrices $\Phi_j(\beta)$, $j = 1, \dots, p$, sont bien des matrices dont les éléments sont des fonctions deux fois continûment dérivables du vecteur β , qui est composé d'un nombre totalisant $N \sum_{j=1}^p (\lambda_j + 1)$ éléments indépendants. Ainsi, lorsque $p = 1$, $\lambda_1 = 1$, on trouve $\beta = (\phi_{10}^{(1)}, \dots, \phi_{10}^{(N)}, \phi_{11}^{(1)}, \dots, \phi_{11}^{(N)})$ et l'expansion de Taylor suivante est valide :

$$\mathbf{c}_a(h) = \mathbf{c}_a(h) - \mathbf{E}_h^\top (\hat{\beta} - \beta) + \mathbf{o}_p(n^{-\frac{1}{2}}),$$

et la matrice \mathbf{E}_h est de dimension $\mathcal{B} \times N^2$. En fait, la matrice \mathbf{E}_h satisfait la relation :

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{cov} \left(\frac{\partial l}{\partial \beta}, \mathbf{c}_a(h) \right) &= \frac{\partial \phi^\top}{\partial \beta} (\mathbf{I}_N \otimes \Psi_{h-1} \Omega_a), \\ &= \mathbf{E}_h, \end{aligned} \quad (3.2.1)$$

où Ψ_h correspond au poids Ψ multivarié dans la représentation VMA(∞) du processus GSTAR :

$$\mathbf{Y}_t = \mathbf{a}_t + \sum_{i=1}^{\infty} \Psi_i \mathbf{a}_{t-i}. \quad (3.2.2)$$

On rappelle que l'existence d'une telle représentation découle de notre hypothèse de stationnarité du processus $\{\mathbf{Y}_t\}$. La matrice $\partial \phi^\top / \partial \beta$ peut être évaluée explicitement, puisque dans le modèle GSTAR(1_1) on a déjà noté que :

$$\Phi = \Phi_0 + \Phi_1 \mathbf{W}^{(1)}.$$

Ainsi,

$$\begin{aligned}
\boldsymbol{\phi} &= \text{vec}(\boldsymbol{\Phi}_0^\top + \mathbf{W}^{(1)\top} \boldsymbol{\Phi}_1^\top), \\
&= \text{vec}(\boldsymbol{\Phi}_0) + \text{vec}(\mathbf{W}^{(1)\top} \boldsymbol{\Phi}_1^\top), \\
&= \text{vec}(\boldsymbol{\Phi}_0) + (\mathbf{I}_N \otimes \mathbf{W}^{(1)\top}) \text{vec}(\boldsymbol{\Phi}_1^\top), \\
&= \boldsymbol{\phi}_0 + (\mathbf{I}_N \otimes \mathbf{W}^{(1)\top}) \boldsymbol{\phi}_1,
\end{aligned}$$

où l'on note que $\boldsymbol{\phi}_0 = \text{vec}(\boldsymbol{\Phi}_0)$, $\boldsymbol{\phi}_1 = \text{vec}(\boldsymbol{\Phi}_1)$ avec $\boldsymbol{\Phi}_0$ et $\boldsymbol{\Phi}_1$ des matrices diagonales, donc symétriques. De plus, l'on a utilisé la relation suivante : $\text{vec}(\mathbf{A}\mathbf{B}) = (\mathbf{I} \otimes \mathbf{A})\text{vec}(\mathbf{B})$, voir Lütkepohl (2005, p. 662). On note que :

$$\begin{aligned}
\text{vec}(\boldsymbol{\Phi}_0) &= \left(\phi_{10}^{(1)} \mathbf{e}_1^\top, \phi_{10}^{(2)} \mathbf{e}_2^\top, \dots, \phi_{10}^{(N)} \mathbf{e}_N^\top \right)^\top, \\
\text{vec}(\boldsymbol{\Phi}_1) &= \left(\phi_{11}^{(1)} \mathbf{e}_1^\top, \phi_{11}^{(2)} \mathbf{e}_2^\top, \dots, \phi_{11}^{(N)} \mathbf{e}_N^\top \right)^\top,
\end{aligned}$$

où les $\mathbf{e}_i, i = 1, \dots, N$ sont des vecteurs de dimension $N \times 1$ avec une valeur unitaire en position i , et zéro sinon. Ainsi, le vecteur de dérivées :

$$\begin{cases} \partial \boldsymbol{\phi} / \partial \phi_{10}^{(i)} = \partial \boldsymbol{\phi}_0 / \partial \phi_{10}^{(i)} + (\mathbf{I}_N \otimes \mathbf{W}^{(1)\top}) \partial \boldsymbol{\phi}_1 / \partial \phi_{10}^{(i)}, \\ \partial \boldsymbol{\phi} / \partial \phi_{11}^{(i)} = \partial \boldsymbol{\phi}_0 / \partial \phi_{11}^{(i)} + (\mathbf{I}_N \otimes \mathbf{W}^{(1)\top}) \partial \boldsymbol{\phi}_1 / \partial \phi_{11}^{(i)}, \end{cases}$$

pour $i = 1, \dots, N$. Nous trouvons de plus que :

$$\begin{cases} \partial \boldsymbol{\phi}_0 / \partial \phi_{10}^{(i)} = (\mathbf{0}^\top, \dots, \mathbf{0}^\top, \mathbf{e}_i^\top, \mathbf{0}^\top, \dots, \mathbf{0}^\top)^\top, \\ \partial \boldsymbol{\phi}_1 / \partial \phi_{10}^{(i)} = \mathbf{0}, \end{cases}$$

alors que :

$$\begin{cases} \partial \boldsymbol{\phi}_0 / \partial \phi_{11}^{(i)} = \mathbf{0}, \\ \partial \boldsymbol{\phi}_1 / \partial \phi_{11}^{(i)} = (\mathbf{0}^\top, \dots, \mathbf{0}^\top, \mathbf{e}_i^\top, \mathbf{0}^\top, \dots, \mathbf{0}^\top)^\top. \end{cases}$$

Ainsi, la matrice \mathbf{E}_h définie par la relation (3.2.1) est particulièrement facile à calculer. De même, le calcul des dérivées que nous venons de réaliser permet de détailler assez facilement la matrice d'information de Fisher associée à $\boldsymbol{\beta}$ dans le modèle GSTAR(1₁), car la relation suivante est valide :

$$\mathcal{I}(\boldsymbol{\beta}) = \frac{\partial \boldsymbol{\phi}^\top}{\partial \boldsymbol{\beta}} \{ \boldsymbol{\Omega}_a^{-1} \otimes \boldsymbol{\Gamma}(0) \} \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\beta}^\top}.$$

Ainsi, on trouve que :

$$\sqrt{n}\mathbf{c}_a(h) \rightarrow_{\mathcal{L}} \mathcal{N}_{N^2}(\mathbf{0}, \mathbf{\Omega}(h)), \quad (3.2.3)$$

où $\mathbf{\Omega}(h) = \mathbf{\Sigma}_a \otimes \mathbf{\Sigma}_a - \mathbf{E}_h\{\mathcal{I}(\boldsymbol{\beta})\}^{-1}\mathbf{E}_h^\top$. Le résultat précédent est une nouvelle application des résultats de Ahn (1988). Il permet de faire des tests à chaque délai h , $h = 1, \dots, M$.

Compte tenu du fait que nous supposons $\mathbf{\Sigma}_a$ définie positive, nous pouvons conclure que les tests portemanteaux définis par (3.1.3) et (3.1.4) pour un GSTAR(1₁) suivent approximativement des lois de distribution khi-carrée avec $N^2M - 2N$ degrés de liberté. Plus généralement, les statistiques portemanteaux pour un modèle GSTAR($p_{\lambda_1}, \dots, p_{\lambda_p}$) font intervenir un nombre total de $N \sum_{j=1}^p (\lambda_j + 1)$ paramètres. Il en ressort que les statistiques portemanteaux (3.1.3) et (3.1.4) suivent approximativement des distributions khi-carré avec $N^2M - N \sum_{j=1}^p (\lambda_j + 1)$ degrés de liberté. Notons que pour que l'approximation soit bonne, il importe que M soit assez grand. De plus, il faut évidemment que $NM > \sum_{j=1}^p (\lambda_j + 1)$, ce qui ne devrait pas être difficile à satisfaire en pratique. Par exemple, si $\lambda_j \equiv 1$, la condition est $NM > 2p$, et elle est particulièrement facile à satisfaire si le nombre de sites est modéré et si p est relativement petit.

3.3. DISTRIBUTION DES AUTOCOVARIANCES RÉSIDUELLES POUR UN MODÈLE GSTAR ESTIMÉ PAR MOINDRES CARRÉS ORDINAIRES

Les résultats obtenus dans la section précédente sont en fait des applications des résultats généraux de Ahn (1998). On rappelle que ces résultats sont valides pour un VAR structuré dont le terme d'erreur est un bruit blanc fort. L'objectif de cette section vise à établir des liens avec les résultats de Borovkova *et al.* (2008). Soit le modèle GSTAR issu de la définition 2.2.2 :

$$\mathbf{Y}_t = \sum_{s=1}^p \{\Phi_{s0} \mathbf{Y}_{t-s} + \sum_{k=1}^{\lambda_s} \Phi_{sk} \mathbf{W}^{(k)} \mathbf{Y}_{t-s}\} + \mathbf{a}_t, \quad (3.3.1)$$

que nous rappelons ici afin de faciliter la compréhension. On rappelle que l'on peut formuler le modèle précédent comme un modèle de régression linéaire multiple de

la forme $\mathbf{h} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$, voir la Section 2.4. Borovkova *et al.* (2008) ont considéré les hypothèses suivantes :

BLR_1 : Le modèle GSTAR est stationnaire.

BLR_2 : Le processus $\{\mathbf{a}_t\}$ est telle que la variable aléatoire est mesurable par rapport à \mathcal{F}_t où $\mathcal{F}_t = \sigma(\mathbf{a}_t, \mathbf{a}_{t-1}, \dots)$ est la σ -algèbre associée à $\{\mathbf{a}_s, s \leq t\}$.

De plus $E(\mathbf{a}_t/\mathcal{F}_{t-1}) = \mathbf{0}$.

On note que pour un GSTAR(1), la condition BLR_1 peut être reformulée en étudiant les valeurs propres de $\boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \mathbf{W}$, qui doivent être strictement inférieures à un en valeurs absolues. Puisqu'un GSTAR admet une écriture VAR et que tout VAR peut s'écrire comme un VAR(1), il est toujours possible de formuler la condition de stationnarité en fonction de valeurs propres d'une certaine matrice. Voir la condition (C1) dans Borovkova *et al.* (2008, p. 489) et la version généralisée (C1*) de Borovkova *et al.* (2008, p. 493). On note que la condition BLR_2 est plus générale que de présumer un bruit blanc fort. En fait, présumer que le terme d'erreur satisfait la condition $E(\mathbf{a}_t/\mathcal{F}_{t-1}) = \mathbf{0}$ revient à présumer que $\{\mathbf{a}_t\}$ est une différence de martingales.

BLR_3 : $E\{\mathbf{a}_t \mathbf{a}_t^\top / \mathcal{F}_{t-1}\} = \boldsymbol{\Sigma}_t$, presque sûrement, et $n^{-1} \sum_{t=1}^n \boldsymbol{\Sigma}_t \rightarrow_p \boldsymbol{\Sigma}$.

BLR_4 : $\sup_{t \geq 1} E\{\mathbf{a}_t^\top \mathbf{a}_t \mathbf{1}[\mathbf{a}_t^\top \mathbf{a}_t > \delta] / \mathcal{F}_{t-1}\}$ converge vers zéro en probabilité, lorsque δ tend vers l'infini.

Soit maintenant

$$\boldsymbol{\Gamma} = \sum_{j=0}^{\infty} \mathbf{A}^j \boldsymbol{\Sigma} (\mathbf{A}^\top)^j,$$

où $\mathbf{A} = \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \mathbf{W}$ et $\boldsymbol{\Sigma}$ est défini dans BLR_3 . Borovkova *et al.* (2008) montrent si $\boldsymbol{\Gamma}$ est inversible que l'estimateur des moindres carrés classique satisfaisant (2.4.5) converge en probabilité vers $\boldsymbol{\beta}$. Sous des conditions plus fines, il est également possible de montrer la convergence presque sûre mais nous référons à l'article de Borovkova *et al.* (2008) pour ce genre de considérations.

Les conditions BLR_1 - BLR_4 sont satisfaites si les erreurs sont indépendantes avec une matrice de variance constante et des moments d'ordre quatre finis. On considère maintenant l'hypothèse :

$$BLR_5 : n^{-1} \sum_{t=\max(r,s)+2}^n \{\boldsymbol{\Sigma}_t \otimes (\mathbf{a}_{t-1-r} \mathbf{a}_{t-1-s}^\top)\} \rightarrow_p (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) \delta_{rs},$$

où $r, s \in \{1, \dots, N\}$, $\delta_{rs} = 1$ si $r = s$, et $\delta_{rs} = 0$ si $r \neq s$.

Sous les hypothèses BLR_1 - BLR_5 , Borovkova *et al.* (2008) ont établi la normalité asymptotique de l'estimateur des moindres carrés, et ont donné explicitement la matrice des variances et covariances asymptotique. Voir leur Théorème 3 et leur Théorème 6.

Dans l'étude des modèles VAR :

$$\Phi(B)\mathbf{Y}_t = \mathbf{a}_t,$$

où $E(\mathbf{a}_t) = \mathbf{0}$ et $\text{Var}(\mathbf{a}_t) = \Sigma_{\mathbf{a}}$, le critère des moindres carrés $\sum_{t=1}^n \mathbf{a}_t^\top \mathbf{a}_t$ ou moindres carrés généralisés $\sum_{t=1}^n \mathbf{a}_t^\top \Sigma_{\mathbf{a}}^{-1} \mathbf{a}_t$ mènent à des estimateurs asymptotiquement équivalents (en effet, chaque "observation" \mathbf{a}_t a la même "variance" $\Sigma_{\mathbf{a}}$; c'est une situation "homoscédastique" dans un cadre multivarié). Voir Lütkepohl (2005, p. 71). Le résultat est connu depuis Zellner (1962) qui a montré que dans un modèle d'équations multiples le critère des moindres carrés et le critère des moindres carrés généralisés mènent exactement au même estimateur. Ainsi, on note que le critère des moindres carrés $S(\boldsymbol{\beta}) = (\mathbf{h} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{h} - \mathbf{X}\boldsymbol{\beta})$ adopté par Borovkova *et al.* (2008) diffère quelque peu du critère adopté dans Ahn (1988), qui est davantage un critère de type moindres carrés généralisés. Voir l'expression (3.1.6). Dans le modèle de régression linéaire multiple de la forme :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

où $E(\boldsymbol{\epsilon}) = \mathbf{0}$, $\text{Var}(\boldsymbol{\epsilon}) = \Sigma$, minimiser le critère des moindres carrés $S_{MCO}(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ ou celui des moindres carrés généralisés $S_{MCG}(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top \Sigma^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ ne mènent pas à la même solution, à savoir $\hat{\boldsymbol{\beta}}_{MCO} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$ dans le premier cas, et $\hat{\boldsymbol{\beta}}_{MCG} = (\mathbf{X}^\top \Sigma^{-1} \mathbf{X})^{-1} (\mathbf{X}^\top \Sigma^{-1} \mathbf{y})$ dans le second. Cependant, ce n'est plus vrai dans les VAR avec contraintes et ce ne sera plus vrai dans les modèles VAR structurés. Ce point est discuté dans Lütkepohl (2005, p. 196) concernant les VAR avec contraintes sur les paramètres. Cependant, moindres carrés généralisés et estimation à vraisemblance maximale sont en général plus efficaces que l'estimation par moindres carrés ordinaires. On note que

l'estimateur des moindres carrés satisfait :

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{h}, \\ &= \boldsymbol{\beta} + (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \boldsymbol{\epsilon}, \\ &= \boldsymbol{\beta} + (n^{-1} \mathbf{X}^\top \mathbf{X})^{-1} (n^{-1} \mathbf{X}^\top \boldsymbol{\epsilon}).\end{aligned}$$

Borovkova *et al.* (2008) ont montré dans un GSTAR(1₁) que :

$$n^{-1} \mathbf{X}^\top \mathbf{X} \rightarrow_p \mathbf{M}(\mathbf{I} \otimes \boldsymbol{\Gamma}) \mathbf{M}^\top,$$

où \mathbf{M} est une matrice diagonale par blocs. Il peut être montré que :

$$\mathbf{X}^\top = \mathbf{M} \{ \mathbf{I}_N \otimes [\mathbf{Y}_0 \ \mathbf{Y}_1 \ \cdots \ \mathbf{Y}_{n-1}] \},$$

où $\mathbf{M} = \text{diag}(\mathbf{M}_1, \dots, \mathbf{M}_N)$, et \mathbf{M}_i correspond à une matrice de dérivées par rapport aux coefficients. Les matrices \mathbf{M}_i sont définies dans Borovkova *et al.* (2008, p. 488). Il découle que :

$$\mathbf{X}^\top \mathbf{X} = \mathbf{M} \left\{ \mathbf{I} \otimes \sum_{t=1}^n \mathbf{Y}_{t-1} \mathbf{Y}_{t-1}^\top \right\} \mathbf{M}^\top,$$

où il est à noter que pour des fins techniques \mathbf{Y}_0 est supposée disponible. Ainsi, $n^{-1} \mathbf{X}^\top \mathbf{X} \rightarrow_p \mathbf{M}(\mathbf{I} \otimes \boldsymbol{\Gamma}) \mathbf{M}^\top$ et $n^{-1/2} \sum_{t=1}^n \text{vec}\{\mathbf{Y}_{t-1} \mathbf{a}_t^\top\}$ converge vers une loi multivariée normale, ce qui entraîne que :

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \rightarrow_{\mathcal{L}} \mathcal{N}_{2N}(\mathbf{0}, \boldsymbol{\Sigma}_\beta),$$

où $\boldsymbol{\Sigma}_\beta = \{\mathbf{M}(\mathbf{I} \otimes \boldsymbol{\Gamma}) \mathbf{M}^\top\}^{-1} \{\mathbf{M}(\boldsymbol{\Sigma} \otimes \boldsymbol{\Gamma}) \mathbf{M}^\top\} \{\mathbf{M}(\mathbf{I} \otimes \boldsymbol{\Gamma}) \mathbf{M}^\top\}^{-1}$.

Supposons que le terme d'erreur du modèle GSTAR est constitué de variables aléatoires indépendantes. On trouve alors que $\text{Var}(\mathbf{a}_t) = \boldsymbol{\Sigma}_t \equiv \boldsymbol{\Sigma}_a$. On a déjà vu que dans le cadre de Ahn (1988), la variance asymptotique de $\hat{\boldsymbol{\beta}}_{MV}$ devient $\mathcal{I}^{-1}(\boldsymbol{\beta})$, où $\mathcal{I}(\boldsymbol{\beta}) = \mathbf{H}^\top \mathbf{G} \mathbf{H}$. Ici, la forme compliquée de la forme sandwich est due au fait qu'un critère de type moindres carrés a été utilisé, plutôt qu'un critère de type moindres carrés généralisés. Si on suppose que $\{\mathbf{a}_t\}$ est indépendant, et que de plus $\text{Var}(\mathbf{a}_t) = \sigma^2 \mathbf{I}$, alors les résultats suivant l'approche de Ahn (1988) ou encore de Borovkova *et al.* (2008) donnent exactement la même variance asymptotique.

On note que les résultats généraux trouvés par Borovkova *et al.* (2008) pour Σ_β incluent le cas où $\mathbf{a}_t \sim \mathcal{N}_N(\mathbf{0}, \sigma^2 \mathbf{I}_N)$, et dans ce cas précis, on déduit que $\Sigma_\beta = \{\mathcal{I}(\hat{\beta}_{MCO})\}^{-1}$ et $\Omega(h) = \Sigma_{\mathbf{a}} \otimes \Sigma_{\mathbf{a}} - \mathbf{E}_h \{\mathcal{I}(\hat{\beta}_{MCO})\}^{-1} \mathbf{E}_h^\top$. Ainsi, on retrouve que le résultat précédent est une nouvelle application des résultats de Ahn (1988). Ainsi, sous les hypothèses de normalité et d'homoscédasticité du terme d'erreur $\{\mathbf{a}_t\}$ du modèle (3.1.5), les statistiques (3.1.3) et (3.1.4) peuvent être utilisées pour diagnostiquer le modèle $\text{GSTAR}(p_{\lambda_1}, \dots, p_{\lambda_p})$ et suivent approximativement les mêmes distributions que dans la section précédente. En fait, sous ces conditions les résultats de Ahn (1988) peuvent être utilisés pour diagnostiquer un modèle GSTAR estimé par les techniques de Borovkova *et al.* (2008). On note aussi que le cas où $\mathbf{a}_t \sim \mathcal{N}_N(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ a été déjà considéré par Pfeifer *et al.* (1980a) lors de l'estimation des paramètres du modèle STARMA , voir Pfeifer *et al.* (1980a, p. 37) et il est souvent commode de présumer ces hypothèses.

Le chapitre précédent a introduit les tests de validation en séries chronologiques multivariées sur les autocovariances résiduelles matricielles. Nous présentons dans le chapitre suivant une brève étude de simulations Monte Carlo.

Chapitre 4

ÉTUDE DE SIMULATIONS

Dans le chapitre 2, nous avons exposé le modèle GSTAR. Dans le chapitre précédent, nous avons présenté des statistiques de test de type portemanteau afin de diagnostiquer ces modèles. L'objectif du présent chapitre est de vérifier si les distributions approximatives reposant sur une étude asymptotique approchent bien les distributions exactes des statistiques de test (3.1.3) et (3.1.4). Pour ce faire, nous allons étudier les niveaux exacts des tests portemanteaux pour des niveaux nominaux couramment utilisés en pratique. Toute la programmation a été réalisée avec le logiciel **R** et les codes de programmation sont en annexe.

4.1. ALGORITHMES POUR L'ÉTUDE DE SIMULATIONS

Pour générer une série chronologique multivariée de dimension N et de longueur n d'un processus autorégressif multivarié avec des erreurs gaussiennes, nous avons créé une fonction sur le logiciel **R** à partir de la librairie MASS. Cette fonction, qui décrit le processus multivarié de l'équation (2.2.4), fait appel aux matrices de poids \mathbf{W} et de paramètres Φ_0 et Φ_1 . On note que les matrices de paramètres ont été choisies de telle sorte que la série multivariée générée soit stationnaire.

L'estimation des paramètres du modèle GSTAR(1₁) par la méthode d'estimation présentée dans la Section 2.4, le calcul des résidus ainsi que les autocorrélations résiduelles sont obtenus à l'aide d'une routine créée en **R**. Ainsi, en considérant $\beta = (\phi_{01}, \phi_{11}, \phi_{02}, \phi_{12}, \dots, \phi_{0N}, \phi_{1N})^\top$ comme le vecteur des paramètres

du modèle (2.2.4), on retrouve alors dans un premier temps $\hat{\beta}$, qui permet d'en déduire les matrices des estimateurs $\hat{\Phi}_0$ et $\hat{\Phi}_1$. Ensuite, les résidus sont obtenus par la formule suivante :

$$\hat{\mathbf{a}}_t = \mathbf{Y}_t - \left(\hat{\Phi}_0 + \hat{\Phi}_1 \mathbf{W} \right) \mathbf{Y}_{t-1}, \quad t = 1, \dots, n,$$

avec $\mathbf{Y}_0 = \mathbf{0}$. Suite à l'estimation des erreurs et au calcul des autocorrélations résiduelles, les statistiques du test de type portemanteau définies par (3.1.3) et (3.1.4) ont été calculées pour les différents délais M . Les valeurs de M choisies sont $M = 1, \dots, 10$. On a considéré aussi $M = 12, 18$. La routine complète pour calculer ces statistiques de test est présentée en annexe du document.

4.2. DESCRIPTION DE LA SIMULATION

Nous voulons comparer la distribution de la statistique portemanteau $Q(M)$ avec la distribution approximative khi-deux, sous l'hypothèse nulle de l'adéquation du modèle. Pour ce faire, nous avons généré deux séries multivariées de longueurs $n = 100$ et $n = 600$ provenant d'un processus GSTAR(1₁). Pour chaque série multivariée de trois et de dix sites c'est-à-dire $N = 3$ et $N = 10$, 1000 réalisations indépendantes ont été générées. Les statistiques $Q(M)$ et $Q^*(M)$ ont été calculées pour différentes valeurs de M . Nous avons calculé les niveaux empiriques, c'est-à-dire les fréquences de rejet de l'hypothèse nulle, pour les trois niveaux nominaux de signification suivants : 1%, 5% et 10%. Pour chaque réalisation, les étapes suivantes ont été exécutées lors de la programmation :

- 1) Nous avons généré la série multivariée \mathbf{Y}_t , $t = 1, \dots, n$ du GSTAR(1₁) suivant, à partir de la fonction VAR.sim du code présenté en annexe. Les arguments de cette fonction à savoir les matrices des paramètres sont présentées dans le Tableau 4.1 ci-dessous. Les trois premières valeurs des deux lignes du tableau représentent les éléments diagonaux des paramètres Φ_0 et Φ_1 , respectivement, pour une série chronologique de trois sites. L'ensemble des données du tableau ont servi lorsque le nombre de sites était égal à 10.

Le modèle généré est donc :

$$\mathbf{Y}_t = (\Phi_0 + \Phi_1 \mathbf{W}) \mathbf{Y}_{t-1} + \mathbf{a}_t, \quad (4.2.1)$$

où

$$\mathbf{W} = \begin{pmatrix} 0.00 & 0.48 & 0.52 \\ 0.90 & 0.00 & 0.10 \\ 0.85 & 0.15 & 0.00 \end{pmatrix},$$

lorsque $N = 3$. Pour $N = 10$, les éléments de la matrice de poids \mathbf{W} se trouvent dans le Tableau 4.2. Pour les nombres de sites respectifs 3 et 10, les valeurs propres maximales de la matrice $(\Phi_0 + \Phi_1 \mathbf{W})$, sont 0.436 et 0.939, et elles sont toutes inférieures à un en module, garantissent la stationnarité du modèle estimé par l'équation (4.2.1). Des simulations complémentaires sont effectuées en Annexe A pour d'autres matrices de poids pour un nombre respectif de 3 et de 10 sites.

TABLE 4.1. Paramètres Φ_0 et Φ_1 pour générer la série chronologique multivariée \mathbf{Y}_t , $t = 1, \dots, n$.

Position	1	2	3	4	5	6	7	8	9	10
$diag(\Phi_0)$	0.30	0.15	0.10	0.25	0.20	0.40	0.40	0.60	0.60	0.70
$diag(\Phi_1)$	0.40	0.30	0.30	0.50	0.50	0.25	0.25	0.45	0.45	0.65

TABLE 4.2. Matrice de poids \mathbf{W} pour $N = 10$.

w_{ij}	1	2	3	4	5	6	7	8	9	10
1	0.00	0.11	0.12	0.13	0.14	0.13	0.11	0.15	0.06	0.05
2	0.24	0.00	0.03	0.05	0.13	0.13	0.03	0.09	0.15	0.15
3	0.25	0.03	0.00	0.03	0.12	0.13	0.04	0.09	0.16	0.16
4	0.24	0.04	0.03	0.00	0.10	0.11	0.05	0.10	0.16	0.16
5	0.21	0.10	0.09	0.09	0.00	0.02	0.08	0.14	0.13	0.15
6	0.21	0.10	0.10	0.10	0.02	0.00	0.08	0.13	0.12	0.14
7	0.24	0.03	0.04	0.06	0.11	0.11	0.00	0.10	0.16	0.15
8	0.24	0.07	0.07	0.09	0.14	0.14	0.07	0.00	0.10	0.08
9	0.10	0.12	0.13	0.15	0.14	0.12	0.12	0.10	0.00	0.03
10	0.08	0.12	0.13	0.15	0.16	0.14	0.11	0.09	0.03	0.00

- 2) Pour une série chronologique donnée \mathbf{Y}_t , $t = 1, \dots, n$, nous avons estimé les paramètres du modèle $\text{GSTAR}(1_1)$ par la méthode décrite dans la Section 2.4 afin d'obtenir les matrices $\hat{\Phi}_0$ et $\hat{\Phi}_1$.
- 3) Les résidus $\hat{\mathbf{a}}_t$ sont estimés tenant compte de l'estimateur obtenu à l'étape 2.
- 4) Les valeurs des statistiques de test de type portemanteau $Q(M)$ et $Q^*(M)$ ont été obtenues pour $M = 1, \dots, 10, 12$ et 18.
- 5) Pour chaque M , $Q(M)$ et $Q^*(M)$ ont été comparées à la valeur critique de la distribution khi-deux correspondante aux niveaux nominaux respectifs 1%, 5% et 10%.
- 6) Finalement, pour les 1000 réalisations indépendantes, nous avons calculé la proportion de rejet pour chaque statistique du test et pour chaque niveau nominal.

Les résultats sont analysés dans la prochaine section.

4.3. RÉSULTATS DE SIMULATION

Les résultats de l'étude de niveau pour le modèle $\text{GSTAR}(1_1)$ sont présentés dans les Tableaux 4.3 et 4.4 ci-dessous. Ces résultats permettent de constater que la distribution exacte de la statistique du test portemanteau est en général bien approximée par sa distribution asymptotique pour des délais supérieurs à un. Mais lorsque le délais augmente, nous constatons une détérioration de l'approximation pour la statistique $Q(M)$, particulièrement au niveau nominal 1%. Ceci est d'autant plus vrai quand la taille de la série est grande. Aux premiers délais quelque soit la taille de la série, le taux de rejet de l'hypothèse nulle est très élevé, spécifiquement quand N est grand.

En effet, quand $N = 3$, le taux de sur-rejet au premier délai est deux fois en moyenne plus grand que les niveaux nominaux pour les deux tailles échantillonnelles choisies (100 et 600). Par ailleurs, plus la taille de la série augmente, moins on retrouve des valeurs éloignées des limites de signification. Par contre lorsque le nombre de sites augmente, par exemple $N = 10$, les taux de rejet de la statistique du test portemanteau sont plus éloignés des niveaux nominaux, spécifiquement au niveau de 1%. Dans ce cas, le taux de sur-rejet au premier délai est plus élevé, soit 3.5 fois en moyenne plus grand que les niveaux nominaux.

Concernant la statistique de test de type portemanteau corrigée $Q^*(M)$, les résultats montrent que l'approximation affiche une certaine stabilité par rapport à $Q(M)$. En effet, tous les niveaux empiriques sont en général assez proches des limites de signification des niveaux nominaux de 1%, 5% et 10% et ceci pour les deux tailles échantillonnelles étudiées. Pour une série de taille 600 et de 10 sites, il y a moins de valeurs à l'extérieur des limites de signification pour $Q^*(M)$.

Pour les trois niveaux nominaux, les statistiques du test portemanteau tendent à sur-rejeter pour les petites valeurs de M l'hypothèse nulle d'adéquation du modèle $\text{GSTAR}(1_1)$ à mesure que le nombre de site augmente. Notons qu'il était attendu qu'un grand nombre de sites soit plus difficile à diagnostiquer comparativement à un N plus petit. En effet, un grand nombre de sites implique un modèle de grande dimension. Ainsi, nous pouvons affirmer que les résultats sont en général assez satisfaisants, surtout avec l'utilisation de $Q^*(M)$ avec un M modéré à grand.

TABLE 4.3. Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour trois sites ($N = 3$).

M	$\alpha = 0.01$				$\alpha = 0.05$				$\alpha = 0.10$			
	$n = 100$		$n = 600$		$n = 100$		$n = 600$		$n = 100$		$n = 600$	
	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$
1	2.1	2.1	2.2	2.2	8.4	8.4	8.2	8.2	14.6	14.6	16.0	16.0
2	0.5	0.5	1.2	1.2	4.5	5.0	6.4	6.4	10.7	11.3	12.0	12.2
3	0.6	0.6	0.9	0.9	4.4	5.1	4.7	4.8	9.0	10.0	9.9	10.0
4	0.5	0.6	1.0	1.0	3.4	4.2	5.0	5.2	8.1	9.8	9.9	10.2
5	0.6	0.6	1.0	1.0	3.8	4.8	4.3	4.6	7.6	9.2	9.6	9.7
6	0.7	0.8	0.7	0.7	3.6	4.7	4.6	4.7	7.0	9.1	9.0	9.4
7	0.1	1.0	1.1	1.1	2.8	4.6	5.3	5.9	6.3	8.2	10.0	11.1
8	0.4	0.7	1.0	1.2	3.1	4.7	4.8	5.2	6.1	8.9	9.9	10.6
9	0.3	1.1	1.3	1.4	2.2	4.4	5.3	5.6	5.3	9.6	10.8	11.6
10	0.4	1.0	1.1	1.2	1.8	4.4	5.2	6.0	4.2	8.6	10.3	11.2
12	0.3	1.3	0.5	0.9	0.6	4.4	4.8	5.4	3.7	8.6	10.7	12.3
18	0.2	0.9	0.9	1.1	0.6	4.2	3.0	4.4	1.6	8.8	8.1	9.8

TABLE 4.4. Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour dix sites ($N = 10$).

M	$\alpha = 0.01$				$\alpha = 0.05$				$\alpha = 0.10$			
	$n = 100$		$n = 600$		$n = 100$		$n = 600$		$n = 100$		$n = 600$	
	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$
1	5.9	5.9	4.1	4.1	15.6	15.6	16.0	16.0	27.3	27.3	27.5	27.5
2	1.3	1.4	2.6	2.6	8.4	9.0	9.5	9.6	15.2	16.4	17.3	17.5
3	0.9	1.2	1.7	1.8	4.4	6.1	7.6	7.7	9.9	12.4	14.3	14.6
4	0.6	1.2	1.1	1.1	2.8	5.6	6.0	6.3	7.9	10.9	10.8	11.8
5	0.5	1.2	1.2	1.4	2.1	4.7	5.1	5.7	5.0	10.6	10.4	11.6
6	0.2	0.6	1.0	1.2	1.7	4.7	4.0	4.5	3.4	10.1	9.0	10.2
7	0.1	0.7	0.8	1.0	0.8	4.1	3.3	3.7	2.7	9.5	7.3	9.5
8	0.1	0.8	0.8	1.1	0.7	4.7	2.8	4.1	1.9	8.9	8.5	10.3
9	0.0	0.9	1.0	1.1	0.4	4.7	3.3	4.4	1.1	8.2	7.3	9.4
10	0.0	1.0	0.7	1.2	0.3	4.3	3.2	4.5	0.9	9.7	6.1	8.2
12	0.0	0.7	0.7	1.2	0.0	4.6	2.8	4.2	0.1	8.5	5.5	9.3
18	0.0	0.6	0.4	1.0	0.0	4.4	1.7	5.3	0.0	9.0	4.6	9.9

Chapitre 5

ANALYSE DE DONNÉES RÉELLES

Dans ce chapitre, nous appliquons la théorie vue aux deux premiers chapitres sur des données réelles que nous présenterons dans la section suivante. Nous considérons une série multivariée de vingt-quatre sites ($N = 24$) analysée par Borovkova *et al.* (2008). Ces données sont mensuelles et proviennent de l'Ouest de Jakarta en Indonésie. La période d'observation s'étend de janvier 1992 à décembre 1999.

5.1. ANALYSE ET MODÉLISATION DE LA SÉRIE

Pour modéliser la série de vingt-quatre sites, il est important de déterminer les matrices de poids \mathbf{W} associées. La méthode utilisée pour le calcul des matrices de poids est identique à celle utilisée par Borovkova *et al.* (2008), voir (2.1.1) et (2.1.2). Ainsi, les différentes entrées dans la matrice de poids reposent sur l'utilisation de la distance euclidienne et nous avons utilisé (2.1.1) avec $c = 1$ et $\alpha = 1$. Le code de l'analyse de données se trouve également en annexe. Selon la figure 5.1 suivante, nous avons calculé la distance entre les différents sites utilisés. L'ordre autorégressif p du GSTAR est obtenu en analysant les autocorrélations et les autocovariances résiduelles. Nous avons trouvé l'ordre maximal $p = 3$. Ainsi, les modèles retenus sont $\text{GSTAR}(1_{\lambda_1})$, $\text{GSTAR}(2_{\lambda_i, i=1,2})$ et $\text{GSTAR}(3_{\lambda_i, i=1,2,3})$.

L'estimation des modèles pour les différents ordres p et pour les ordres spatiaux λ_p a été effectuée en utilisant la méthode de Borovkova *et al.* (2008). Une fois les modèles estimés, nous avons pu déterminer les séries résiduelles et donc

les autocovariances résiduelles $\mathbf{\Gamma}_{\hat{\mathbf{a}}}(h)$, $|h| < n$. Par la suite, les statistiques portemanteaux $Q(M)$ et $Q^*(M)$ ont été calculées.

5.1.1. Estimations des paramètres des modèles retenus

Pour estimer les paramètres des différents modèles retenus, nous avons appliqué la méthode d'estimation discutée dans la Section 2.4. On note que lors de nos estimations, nous avons utilisé uniquement la matrice de poids d'ordre un. Ainsi, nous avons trouvé les résultats respectifs suivants pour :

GSTAR(1_{1_1}) :

$$\hat{\mathbf{Y}}_t = \hat{\mathbf{\Phi}}_{10} \mathbf{Y}_{t-1} + \hat{\mathbf{\Phi}}_{11} \mathbf{W}^{(1)} \mathbf{Y}_{t-1},$$

GSTAR(2_{1_1}) :

$$\hat{\mathbf{Y}}_t = (\hat{\mathbf{\Phi}}_{10} + \hat{\mathbf{\Phi}}_{20}) \mathbf{Y}_{t-1} + (\hat{\mathbf{\Phi}}_{11} \mathbf{W}^{(1)} + \hat{\mathbf{\Phi}}_{21} \mathbf{W}^{(1)}) \mathbf{Y}_{t-1},$$

et GSTAR(3_{1_1}) :

$$\hat{\mathbf{Y}}_t = (\hat{\mathbf{\Phi}}_{10} + \hat{\mathbf{\Phi}}_{20} + \hat{\mathbf{\Phi}}_{30}) \mathbf{Y}_{t-1} + (\hat{\mathbf{\Phi}}_{11} \mathbf{W}^{(1)} + \hat{\mathbf{\Phi}}_{21} \mathbf{W}^{(1)} + \hat{\mathbf{\Phi}}_{31} \mathbf{W}^{(1)}) \mathbf{Y}_{t-1},$$

où $\mathbf{W}^{(1)}$ est définie suivant l'équation (2.1.1), et les valeurs des matrices des estimateurs $\hat{\mathbf{\Phi}}_{s0}$, $\hat{\mathbf{\Phi}}_{s1}$ pour $s = 1, \dots, 3$ des modèles GSTAR(1_{1_1}), GSTAR(2_{1_1}) et GSTAR(3_{1_1}) respectives sont définies dans les Tableaux 5.1 et 5.2.

TABLE 5.1. Estimateurs des paramètres des modèles GSTAR(1_{1_1})
et GSTAR(2_{1_1}).

i	GSTAR(1_{1_1})		GSTAR(2_{1_1})			
	$diag(\hat{\Phi}_{10})$	$diag(\hat{\Phi}_{11})$	$diag(\hat{\Phi}_{10})$	$diag(\hat{\Phi}_{11})$	$diag(\hat{\Phi}_{20})$	$diag(\hat{\Phi}_{21})$
1	0,3596	0,6544	0,2769	0,6154	0,2587	-0,1394
2	0,0165	0,9974	0,0811	0,2279	-0,1759	0,8833
3	0,3048	0,7084	0,2531	0,4708	0,1489	0,1385
4	0,2445	0,7238	0,2259	0,3881	0,0366	0,3187
5	0,1943	0,7656	0,1980	0,8317	-0,0530	-0,0201
6	0,0645	0,9653	0,0330	0,8122	0,1606	0,0217
7	0,0313	0,9917	0,0302	0,8957	-0,0400	0,1397
8	0,0663	0,9166	0,0928	1,0946	0,0059	-0,2097
9	0,1456	0,8731	0,1494	0,8340	0,0003	0,0384
10	0,2294	0,7370	0,2643	0,5693	-0,0981	0,2282
11	0,0256	0,9171	0,0262	0,9260	-0,1442	0,1274
12	0,2744	0,6716	0,2501	0,4534	0,1213	0,1287
13	0,4251	0,5568	0,2786	-0,0166	0,1135	0,6055
14	-0,0811	1,1100	-0,1115	0,8460	-0,1155	0,4146
15	-0,1030	1,0899	-0,0900	1,2818	0,0853	-0,2887
16	0,0595	0,9743	0,0963	0,6254	-0,1125	0,4289
17	0,0460	0,9836	0,0546	0,8118	-0,0527	0,2173
18	0,4350	0,5876	0,4218	0,3970	-0,0010	0,2065
19	0,1537	0,8934	0,2114	1,0831	-0,1033	-0,1421
20	0,0168	0,9563	0,0183	0,5263	0,0671	0,3655
21	-0,0104	0,9958	0,0092	0,5926	-0,0366	0,4206
22	0,0754	0,8825	0,1080	0,6129	-0,1538	0,3850
23	0,1592	0,8333	0,1474	0,7969	0,0902	-0,0420
24	0,4081	0,5882	0,3395	0,6679	0,1934	-0,2050

TABLE 5.2. Estimateurs des paramètres du modèle GSTAR(3₁).

	GSTAR(3 ₁)					
i	$diag(\hat{\Phi}_{10})$	$diag(\hat{\Phi}_{11})$	$diag(\hat{\Phi}_{20})$	$diag(\hat{\Phi}_{21})$	$diag(\hat{\Phi}_{30})$	$diag(\hat{\Phi}_{31})$
1	0,3281	0,5930	0,2808	-0,2882	-0,1909	0,2907
2	0,1014	0,1997	-0,1907	1,0573	-0,0701	-0,0797
3	0,2335	0,3762	0,1101	-0,0908	0,1299	0,2513
4	0,2135	0,3486	0,0104	0,1808	0,1115	0,1074
5	0,1970	0,8038	-0,0523	-0,1849	-0,0358	0,2274
6	0,0550	0,7640	0,1695	-0,1495	-0,0558	0,2441
7	0,0461	0,8155	-0,0611	-0,1532	0,1588	0,2158
8	0,1270	1,0342	-0,0219	-0,3765	-0,1008	0,3195
9	0,1765	0,7658	-0,0085	-0,1848	-0,0121	0,2849
10	0,2646	0,5686	-0,1004	0,0995	0,0536	0,0802
11	0,0337	0,9189	-0,1460	0,1321	0,0540	-0,0540
12	0,2365	0,4194	0,1108	-0,1198	0,0830	0,2281
13	0,2532	-0,0745	0,0533	0,5749	0,1630	0,0138
14	-0,1301	0,7511	-0,1592	0,1587	-0,1185	0,5362
15	-0,0868	1,3115	0,0665	-0,3111	-0,1443	0,1495
16	0,1278	0,5625	-0,1125	0,3447	0,2248	-0,1185
17	0,0754	0,7708	-0,0571	0,0703	0,3691	-0,2094
18	0,4477	0,4230	-0,0272	0,3984	0,0271	-0,2480
19	0,2379	1,0144	-0,1322	-0,2405	-0,1125	0,2888
20	0,0152	0,5116	0,1080	0,1130	0,0215	0,2087
21	0,0096	0,5921	-0,0511	0,5179	-0,1907	0,1044
22	0,0857	0,6210	-0,1463	0,4172	-0,1694	0,1356
23	0,1472	0,7727	0,0606	-0,1386	0,0629	0,0885
24	0,3969	0,6395	0,2477	-0,3127	-0,2102	0,2358

5.1.2. Choix du modèle suivant le test portemanteau

Les valeurs des statistiques de test de type portemanteau pour $M = 1, \dots, \sqrt{n}$ calculées à partir de la série résiduelle pour les différents modèles estimés, en faisant une transformation logarithmique sur les données, permettent de voir que les modèles $\text{GSTAR}(1_{1_1})$ et $\text{GSTAR}(2_{1_1})$ décrivent mieux la série utilisée. D'après le Tableau 5.3, nous pouvons voir qu'on ne rejette pour aucun délai l'hypothèse nulle. On note que dans ce tableau, les valeurs entre parenthèses sont les probabilités correspondantes des valeurs des statistiques $Q(M)$ et $Q^*(M)$. Ce résultat n'est pas surprenant puisque Borovkova *et al.* (2008) ont trouvé comme modèle adéquat pour ces données le $\text{GSTAR}(2_{1_1})$. Les intervalles de confiance à 95% des paramètres des modèles $\text{GSTAR}(1_{1_1})$ et $\text{GSTAR}(2_{1_1})$ trouvés lors de notre estimation sont légèrement différents de ceux trouvés par Borovkova *et al.* (2008) car nous les avons déterminés suivant la variance de l'estimateur $\{\mathcal{I}(\boldsymbol{\beta})\}^{-1}$ présenté dans Ahn (1988) en supposant que $\mathbf{a}_t \sim \mathcal{N}_N(\mathbf{0}, \sigma^2 \mathbf{I}_N)$. De plus, il n'est pas exclu que les matrices de poids diffèrent un peu.

TABLE 5.3. Résultats du test d'adéquation pour la série de 24 sites.

	GSTAR(1 _{1₁})		GSTAR(2 _{1₁})		GSTAR(3 _{1₁})	
M	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$
1	559.452	565.404	519.884	525.474	508.721	414.250
	(0.166)	(0.126)	(0.101)	(0.074)	(0.006)	(0.004)
2	1103.648	1121.303	1058.463	1075.761	1032.322	1049.359
	(0.497)	(0.352)	(0.473)	(0.329)	(0.290)	(0.178)
3	1620.753	1655.269	1585.782	1620.464	1564.638	1599.419
	(0.847)	(0.662)	(0.790)	(0.576)	(0.631)	(0.388)
6	3294.350	3421.678	3251.498	3379.899	3203.470	3331.350
	(0.917)	(0.431)	(0.908)	(0.401)	(0.909)	(0.403)
12	6448.493	6927.193	6390.856	6873.379	6313.405	6796.077
	(1.000)	(0.293)	(1.000)	(0.310)	(0.999)	(0.402)
18	9399.402	10452.471	9358.190	10425.296	9219.276	10282.416
	(1.000)	(0.178)	(1.000)	(0.143)	(1.000)	(0.340)

FIGURE 5.1. Différents sites de la production de thé en Java occidentale, Indonésie



CONCLUSION

Dans ce mémoire, nous nous sommes intéressés aux modèles spatio-temporels pour des séries chronologiques multivariées. Dans un premier temps, nous avons exposé des définitions sur la statistique spatiale et sur le modèle GSTAR présenté par Borovkova *et al.* (2008). Nous avons analysé aussi le lien qui existe entre les modèles GSTAR et VAR, suivi de la méthode d'estimation des moindres carrés présentée par Borovkova *et al.* (2008) pour estimer les paramètres du modèle GSTAR. Ensuite, nous avons discuté des statistiques de test de type portemanteau pour les modèles VARMA. Inspiré des résultats dans Ahn (1988), nous avons déterminé la distribution asymptotique des autocovariances résiduelles pour les modèles GSTAR en présumant que le terme d'erreur est un bruit blanc gaussien. Une fois trouvé les résultats théoriques, nous avons étudié les statistiques de test de type portemanteau $Q(M)$ et $Q^*(M)$ pour vérifier l'adéquation du modèle GSTAR.

Des études de simulations ont été réalisées pour illustrer nos résultats théoriques. Ainsi, nous avons effectué des études de simulations pour les niveaux 1%, 5% et 10% en considérant un modèle GSTAR(1₁₁) pour deux tailles échantillonales différentes et de nombre de sites différents. Ces études visaient à montrer que, sous l'hypothèse de normalité des erreurs et de la stationnarité du modèle, les niveaux empiriques des statistiques de test de type portemanteau étaient tous raisonnablement proches des niveaux théoriques, comme attendu. Les résultats montrent aussi qu'à mesure que les délais M augmentent les niveaux empiriques se rapprochent de plus en plus des niveaux théoriques. Nous avons constaté aussi que les niveaux empiriques de $Q^*(M)$ sont plus stables que $Q(M)$ particulièrement pour de grandes valeurs de M et d'un nombre élevé de sites. Nous avons

aussi fait une analyse des données de la production du thé en Java occidental. Grâce à la carte géographique de cette ville, nous avons calculé les différentes matrices de poids pour les 24 sites, puis nous avons estimé les modèles GTSAR(1₁) et GTSAR(2₁) suivant la méthode d'estimation de Borovkova *et al.* (2008). Les résultats trouvés montrent que les erreurs du modèle GTSAR(1₁) et GTSAR(2₁) sont compatibles avec l'hypothèse nulle de bruit blanc pour tous les délais M .

Annexe A

SIMULATIONS COMPLÉMENTAIRES

Dans ce chapitre, nous avons effectué des simulations supplémentaires pour d'autres matrices de poids \mathbf{W} afin d'étudier le comportement des distributions exactes des statistiques de test (3.1.3) et (3.1.4). Pour ce faire, nous avons considéré le même modèle dans la Section 4.2 avec différentes matrices de poids \mathbf{W} . Plus précisément, nous avons utilisé le modèle suivant :

$$\mathbf{Y}_t = (\Phi_0 + \Phi_1 \mathbf{W}) \mathbf{Y}_{t-1} + \mathbf{a}_t, \quad t = 1, \dots, n, \quad (\text{A.0.1})$$

où $\mathbf{Y}_0 = \mathbf{0}$, et les matrices de paramètres Φ_0 et Φ_1 sont définies dans le Tableau 4.1. Pour les matrices de poids, nous avons considéré deux matrices distinctes pour $N = 3$ et $N = 10$. Pour $N = 3$, nous avons les deux matrices de poids suivantes :

$$\mathbf{W}_1 = \begin{pmatrix} 0.00 & 0.60 & 0.40 \\ 0.70 & 0.00 & 0.30 \\ 0.67 & 0.33 & 0.00 \end{pmatrix} \quad \text{et} \quad \mathbf{W}_2 = \begin{pmatrix} 0.00 & 0.48 & 0.52 \\ 0.80 & 0.00 & 0.20 \\ 0.81 & 0.19 & 0.00 \end{pmatrix}.$$

Tandis que pour $N = 10$, les matrices de poids \mathbf{W}_1 et \mathbf{W}_2 se trouvent dans le Tableau A.1 ci-dessous.

Comme dans la Section 4.2, nous avons étudié les niveaux empiriques des statistiques $Q(M)$ et $Q^*(M)$ aux niveaux nominaux 1%, 5% et 10% pour des tailles échantillonales différentes $n = 100$, $n = 600$. Pour chaque série multivariée de trois ($N = 3$) et de dix sites ($N = 10$), et pour chacune des matrices de poids ($\mathbf{W}_1, \mathbf{W}_2$), 1000 réalisations indépendantes ont été générées. Les statistiques

portemanteaux définies par (3.1.3) et (3.1.4) ont été calculées pour les valeurs $M = 1, \dots, 10, 12$ et 18. Pour $N = 3$, les valeurs propres maximales des matrices $(\Phi_0 + \Phi_1 \mathbf{W}_1)$ et $(\Phi_0 + \Phi_1 \mathbf{W}_2)$, sont 0.658 et 0.848, toutes inférieures à un en module, garantissent la stationnarité du modèle (A.0.1). De même, pour $N = 10$ les valeurs propres maximales respectives pour $(\Phi_0 + \Phi_1 \mathbf{W}_1)$ et $(\Phi_0 + \Phi_1 \mathbf{W}_2)$ sont 0.907 et 0.936.

Une analyse des résultats présentés dans les Tableaux A.2, A.3, A.4 et A.5 permettent de voir que la distribution exacte des statistiques de test de type portemanteau est en général bien approximée par sa distribution asymptotique pour des délais supérieurs à un. De plus, la statistique corrigée $Q^*(M)$ du test portemanteau affiche de meilleurs niveaux empiriques que celle de $Q(M)$.

TABLE A.1. Matrice de poids \mathbf{W} pour $N = 10$.

	\mathbf{W}_1									
w_{ij}	1	2	3	4	5	6	7	8	9	10
1	0.00	0.08	0.09	0.13	0.16	0.20	0.07	0.07	0.10	0.11
2	0.08	0.00	0.15	0.16	0.15	0.13	0.13	0.11	0.04	0.04
3	0.03	0.13	0.00	0.15	0.15	0.14	0.13	0.11	0.08	0.07
4	0.14	0.07	0.05	0.00	0.10	0.12	0.07	0.12	0.16	0.17
5	0.22	0.07	0.06	0.05	0.00	0.09	0.07	0.12	0.15	0.16
6	0.22	0.07	0.06	0.04	0.09	0.00	0.08	0.11	0.16	0.16
7	0.20	0.07	0.06	0.05	0.09	0.10	0.00	0.11	0.15	0.16
8	0.20	0.08	0.07	0.06	0.10	0.11	0.08	0.00	0.15	0.16
9	0.20	0.09	0.08	0.06	0.10	0.11	0.09	0.12	0.00	0.16
10	0.19	0.09	0.08	0.07	0.10	0.11	0.09	0.12	0.15	0.00
	\mathbf{W}_2									
1	0.00	0.04	0.06	0.07	0.09	0.12	0.14	0.12	0.17	0.19
2	0.04	0.00	0.04	0.06	0.09	0.11	0.13	0.13	0.20	0.22
3	0.06	0.04	0.00	0.04	0.06	0.09	0.11	0.11	0.23	0.25
4	0.08	0.06	0.04	0.00	0.04	0.07	0.09	0.10	0.26	0.27
5	0.10	0.09	0.06	0.04	0.00	0.04	0.06	0.08	0.26	0.28
6	0.11	0.10	0.08	0.06	0.03	0.00	0.03	0.06	0.25	0.26
7	0.12	0.11	0.09	0.07	0.05	0.03	0.00	0.06	0.24	0.24
8	0.11	0.11	0.09	0.08	0.07	0.06	0.06	0.00	0.21	0.21
9	0.09	0.10	0.11	0.12	0.13	0.14	0.15	0.12	0.00	0.03
10	0.10	0.11	0.12	0.12	0.13	0.14	0.14	0.12	0.03	0.00

TABLE A.2. Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour trois sites ($N = 3$) pour \mathbf{W}_1 .

M	$\alpha = 0.01$				$\alpha = 0.05$				$\alpha = 0.10$			
	$n = 100$		$n = 600$		$n = 100$		$n = 600$		$n = 100$		$n = 600$	
	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$
1	1.5	1.5	1.6	1.6	6.4	6.4	8.7	8.7	11.9	11.9	16.6	16.6
2	0.8	0.8	0.8	0.8	5.1	5.1	4.7	5.0	10.7	10.8	10.7	11.2
3	0.8	0.8	0.6	1.0	5.0	5.0	3.9	4.6	8.9	9.0	9.5	10.0
4	0.6	0.6	0.4	0.5	4.6	4.9	3.3	4.0	10.0	10.4	7.6	9.5
5	0.7	0.8	0.3	0.3	4.9	5.0	3.0	4.0	9.3	9.7	7.0	8.9
6	0.9	1.0	0.4	0.6	3.7	4.2	3.0	3.8	7.5	7.9	5.6	7.2
7	0.7	0.9	0.4	0.9	3.4	3.6	2.7	3.6	8.3	8.5	5.0	7.6
8	0.8	0.9	0.4	1.0	3.4	4.0	2.5	3.5	7.4	8.4	4.4	7.3
9	0.7	0.8	0.3	0.8	2.7	3.2	1.9	3.0	7.2	8.0	3.6	6.4
10	0.7	0.7	0.4	1.0	2.7	3.3	2.2	3.6	7.6	8.6	3.8	6.8
12	0.7	0.9	0.2	1.2	3.0	3.3	1.8	4.1	7.3	8.6	3.4	7.5
18	0.6	0.8	0.1	1.2	3.0	4.0	0.7	4.5	6.5	8.7	1.2	8.5

TABLE A.3. Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour trois sites ($N = 3$) pour \mathbf{W}_2 .

M	$\alpha = 0.01$				$\alpha = 0.05$				$\alpha = 0.10$			
	$n = 100$		$n = 600$		$n = 100$		$n = 600$		$n = 100$		$n = 600$	
	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$
1	1.6	1.6	1.3	1.3	7.8	7.8	7.1	7.1	14.7	14.7	14.0	14.0
2	0.7	0.8	1.0	1.0	4.3	4.6	4.1	4.1	9.7	9.9	8.9	8.9
3	0.3	0.4	1.1	1.2	3.3	3.6	5.2	5.4	8.9	9.5	8.3	8.3
4	0.7	0.7	0.9	0.9	3.0	3.5	5.2	5.4	7.0	8.2	9.6	9.6
5	0.3	0.5	0.8	0.8	3.0	3.7	4.3	4.4	6.6	8.6	8.3	8.7
6	0.2	0.7	1.1	1.2	2.9	3.9	4.8	4.8	6.4	8.6	8.6	9.4
7	0.2	0.6	1.2	1.2	2.8	4.7	4.9	5.4	6.3	8.0	9.4	9.6
8	0.3	0.7	0.8	0.8	2.5	4.8	5.0	5.2	5.8	7.5	8.6	9.3
9	0.2	0.7	0.9	0.9	1.9	3.8	4.7	4.8	4.9	8.6	8.3	9.0
10	0.1	0.8	0.7	0.8	1.3	3.9	3.6	4.5	2.8	8.7	8.1	8.9
12	0.2	0.8	0.9	1.0	0.5	3.9	3.8	4.6	1.6	8.7	8.4	9.4
18	0.2	1.4	0.1	0.2	0.5	3.7	3.4	5.2	9.9	7.4	8.6	10.9

TABLE A.4. Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour dix sites ($N = 10$) pour \mathbf{W}_1 .

M	$\alpha = 0.01$				$\alpha = 0.05$				$\alpha = 0.10$			
	$n = 100$		$n = 600$		$n = 100$		$n = 600$		$n = 100$		$n = 600$	
	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$
1	3.6	3.6	3.3	3.3	13.9	13.9	15.2	15.2	26.2	26.2	25.0	25.0
2	0.8	0.8	1.2	1.2	6.5	7.4	7.5	7.5	13.4	14.5	14.4	14.7
3	0.4	0.7	0.8	0.8	3.9	5.2	5.1	5.2	8.5	11.5	10.7	11.4
4	0.4	0.8	0.5	0.5	2.3	4.1	4.5	5.4	6.2	9.7	10.3	11.2
5	0.4	0.9	0.8	0.8	1.8	4.4	3.9	4.2	4.5	8.9	8.8	10.6
6	0.3	0.9	0.7	0.7	1.6	3.7	4.4	5.4	3.3	8.7	8.5	10.0
7	0.2	1.2	0.4	0.7	1.2	4.3	3.6	4.4	2.2	7.5	7.2	9.1
8	0.1	0.9	0.3	0.3	0.9	4.0	3.1	3.9	1.6	8.6	6.7	9.1
9	0.1	0.9	0.4	0.8	0.8	3.8	3.0	4.2	1.3	8.1	6.6	8.7
10	0.1	1.2	0.6	0.8	0.1	4.9	2.9	4.0	0.2	7.6	6.3	9.2
12	0.0	1.5	0.5	0.7	0.0	4.9	2.9	4.3	0.0	7.6	5.6	8.3
18	0.0	1.2	0.0	0.5	0.0	5.3	1.5	4.5	14.5	8.9	3.9	7.7

TABLE A.5. Niveaux empiriques (en pourcentage) des statistiques de test portemanteau pour dix sites ($N = 10$) pour \mathbf{W}_2 .

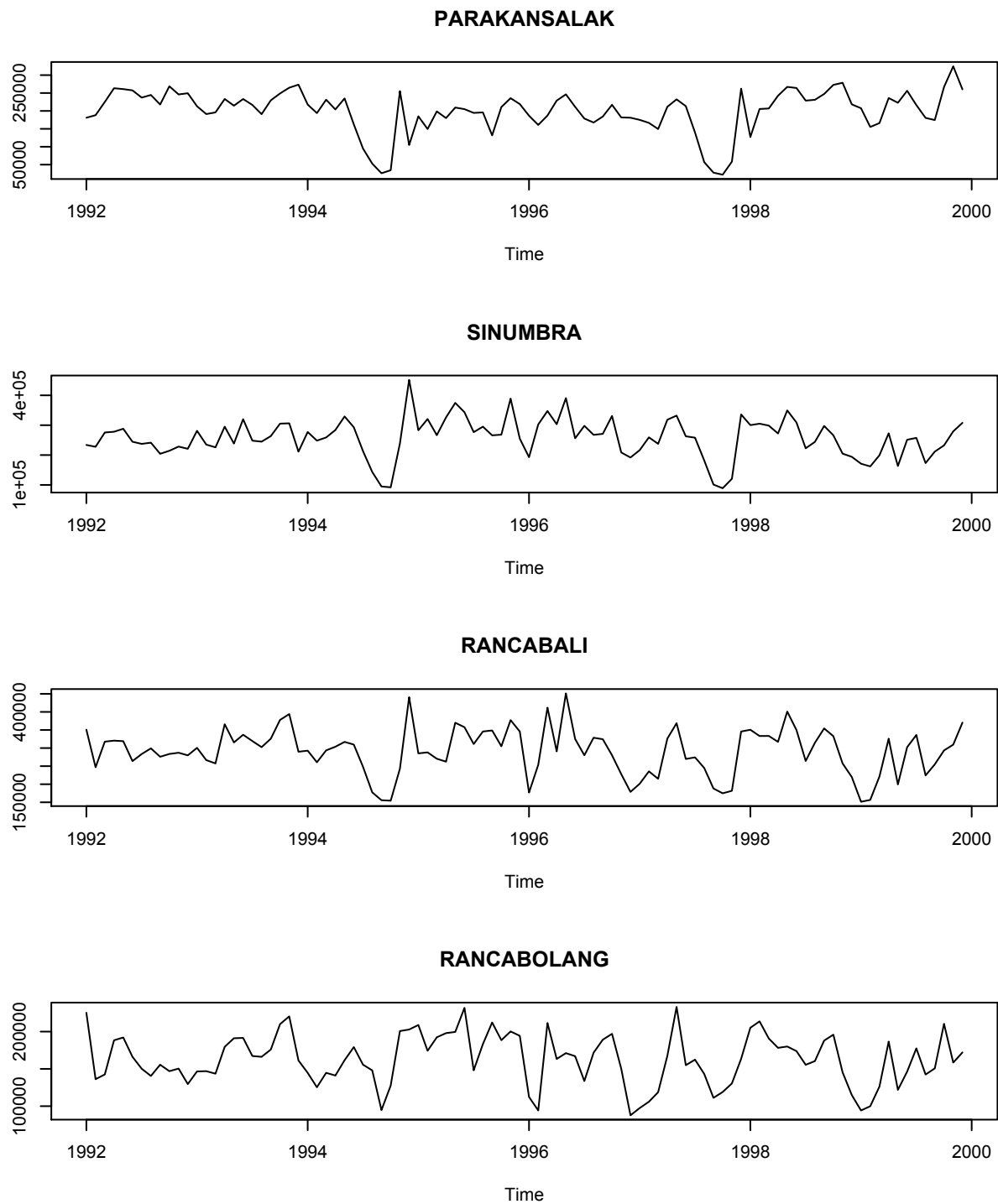
M	$\alpha = 0.01$				$\alpha = 0.05$				$\alpha = 0.10$			
	$n = 100$		$n = 600$		$n = 100$		$n = 600$		$n = 100$		$n = 600$	
	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$	$Q(M)$	$Q^*(M)$
1	3.6	3.6	4.0	4.0	16.3	16.3	15.4	15.4	28.8	28.8	26.8	26.8
2	1.1	1.2	1.8	1.8	7.8	8.9	9.5	9.5	16.4	17.8	17.8	18.2
3	0.7	1.3	1.2	1.2	6.2	7.9	7.3	7.5	10.7	13.4	14.0	14.7
4	0.3	0.6	0.6	1.0	4.0	6.5	6.4	6.9	8.6	13.2	11.7	12.4
5	0.4	1.4	1.3	1.6	2.7	5.4	5.7	6.7	6.3	11.5	11.5	12.7
6	0.2	1.2	1.5	1.9	1.9	6.3	6.6	7.5	5.0	11.9	10.4	11.7
7	0.1	0.9	1.0	1.6	1.4	5.5	5.9	6.7	3.0	11.5	10.7	12.5
8	0.1	0.7	1.1	1.4	0.5	5.0	5.0	6.9	1.9	11.2	10.6	13.2
9	0.0	0.8	0.6	0.8	0.4	5.5	4.4	6.1	1.0	11.1	10.2	12.7
10	0.0	0.7	0.9	1.0	0.0	5.0	3.4	6.3	0.0	9.8	9.1	12.3
12	0.0	0.6	0.7	1.4	0.0	5.0	3.5	5.8	0.0	9.8	7.5	12.4
18	0.0	0.9	0.4	1.7	0.0	4.7	2.8	5.6	0.0	9.7	4.7	10.9

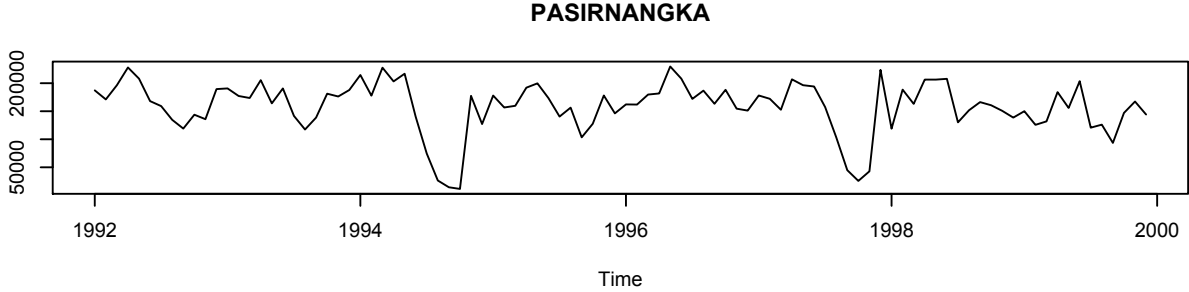
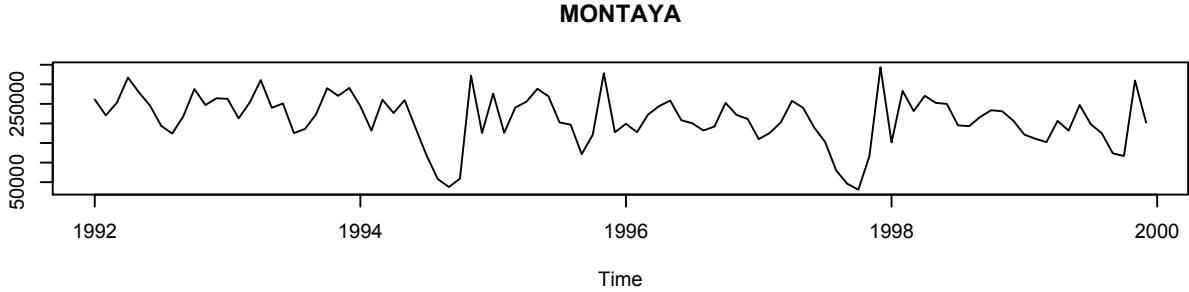
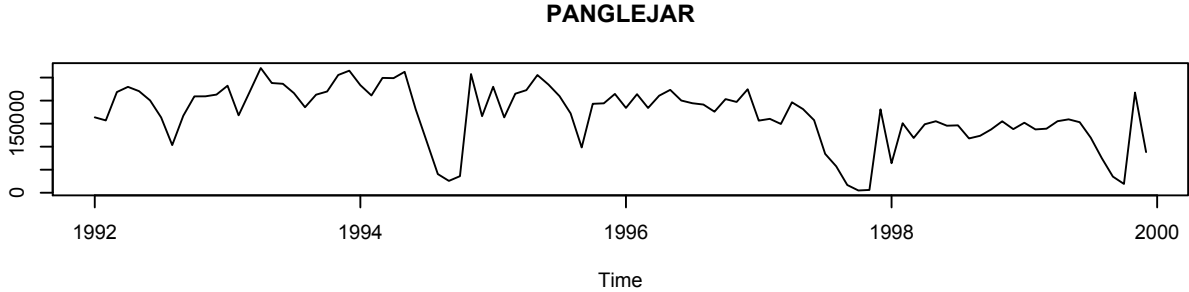
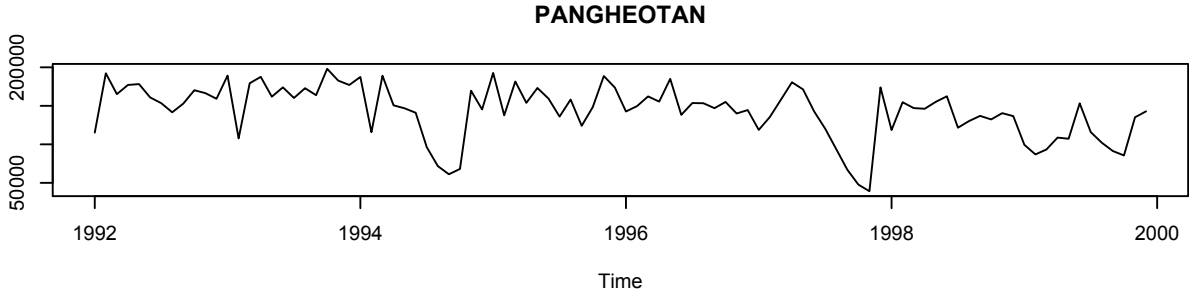
Annexe B

DONNÉES SUR LA PRODUCTION DE THÉ

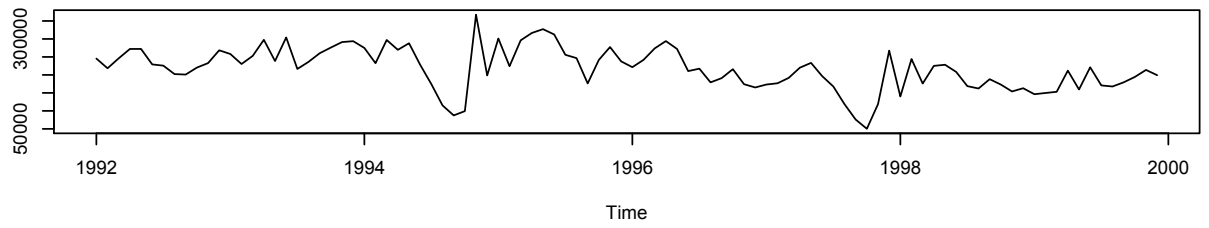
Nous présentons dans cette partie le graphique de la série des différentes sites de production de thé sur la période de janvier 1992 à décembre 1999.

FIGURE B.1. Graphiques des 24 séries chronologiques portant sur la production de thé, pour chacun des sites en Java occidental

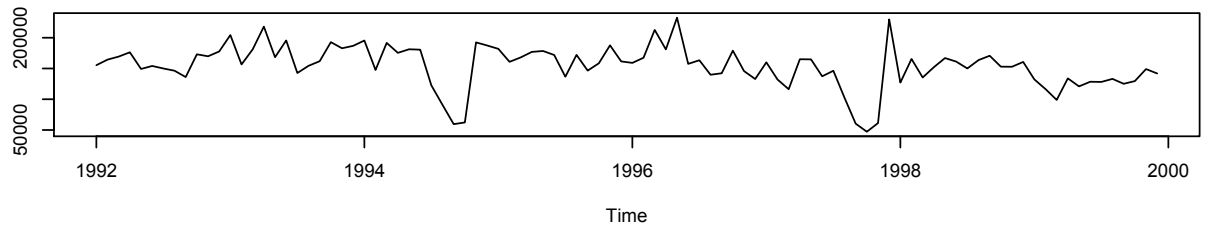




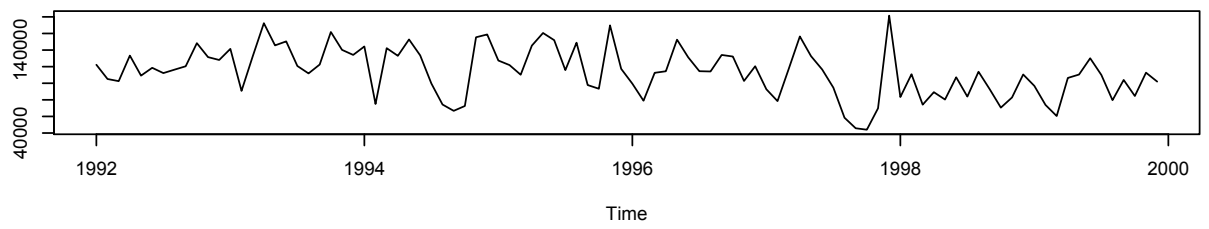
GEDEH



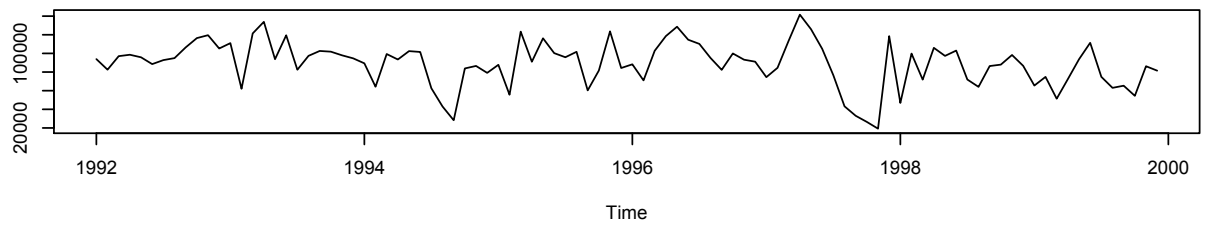
GOALPARA



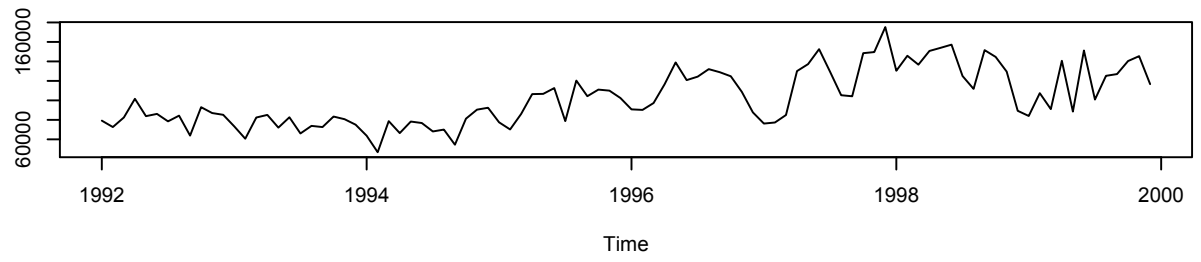
PANYAIRAN



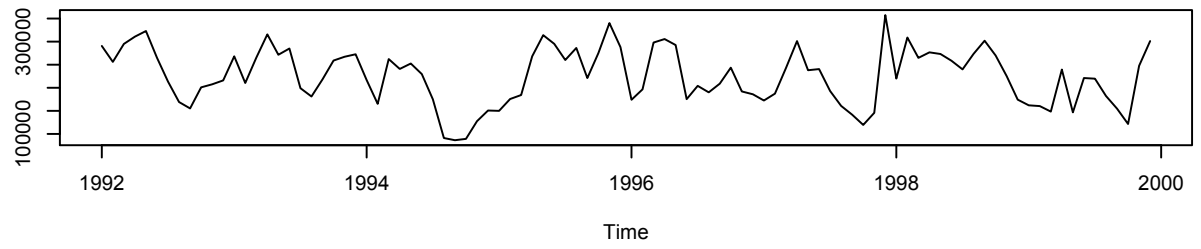
GUNUNMAS



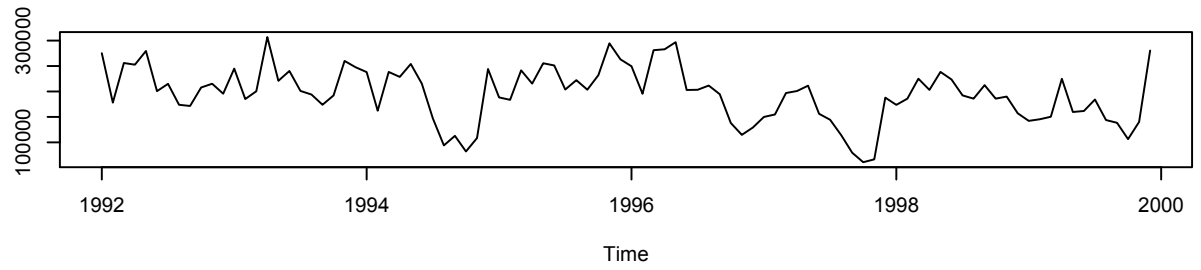
CIATEN



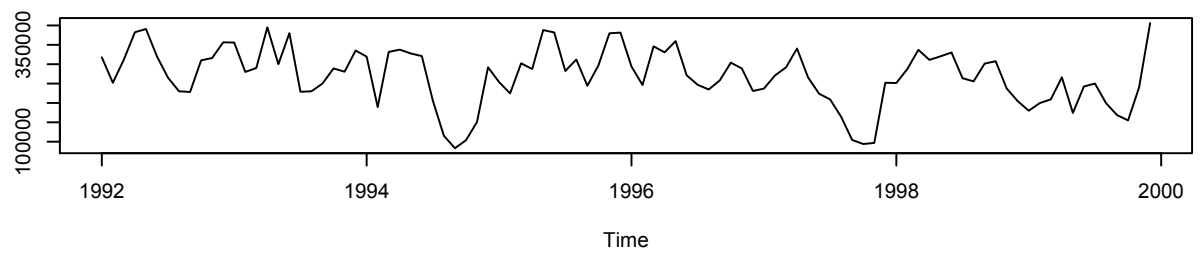
PASIRMALANG



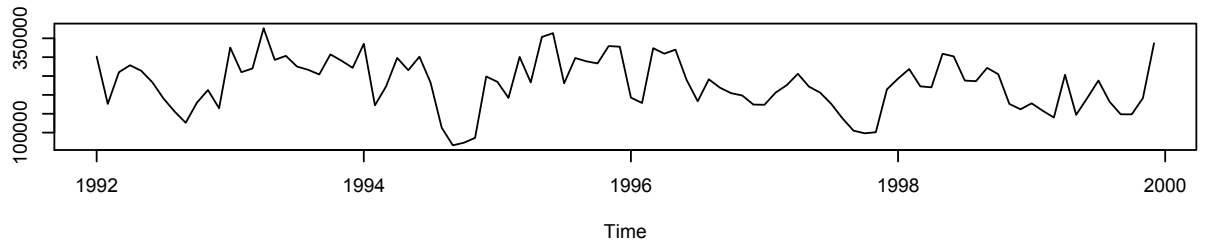
KERTAMANAH



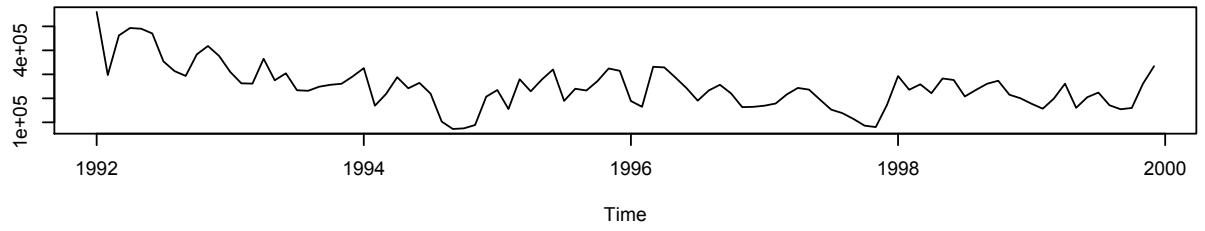
MALABAR



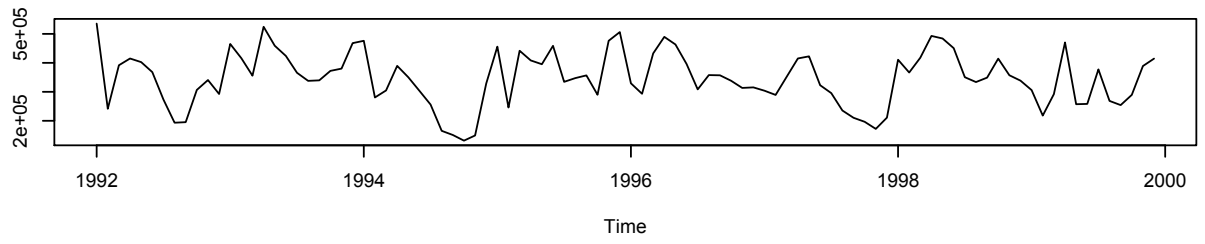
PURBASARI



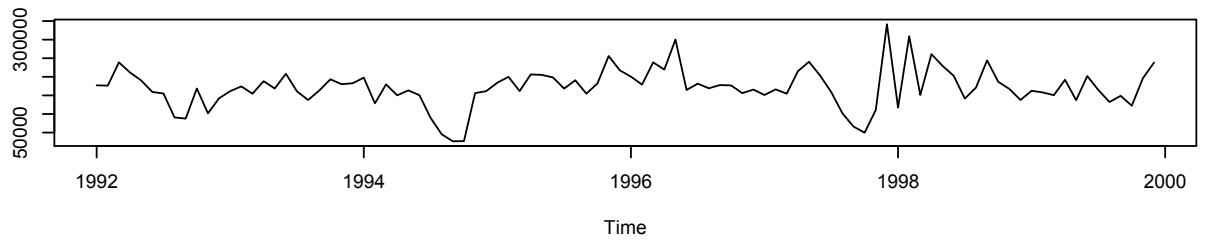
SANTOSA



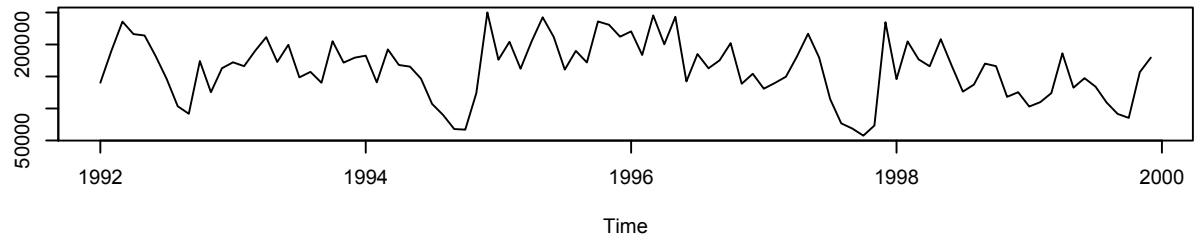
SEDEP



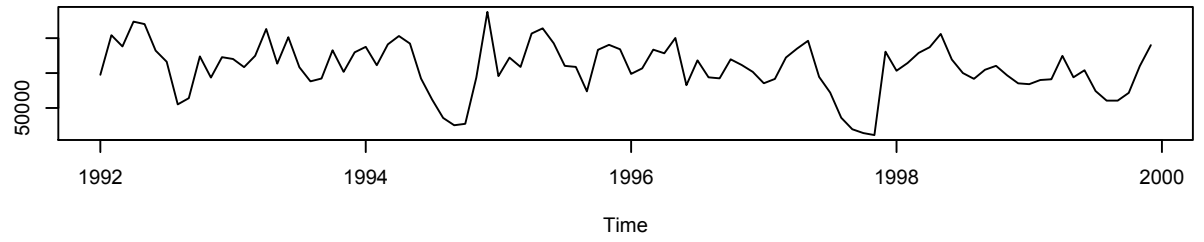
PAPANDAYAN



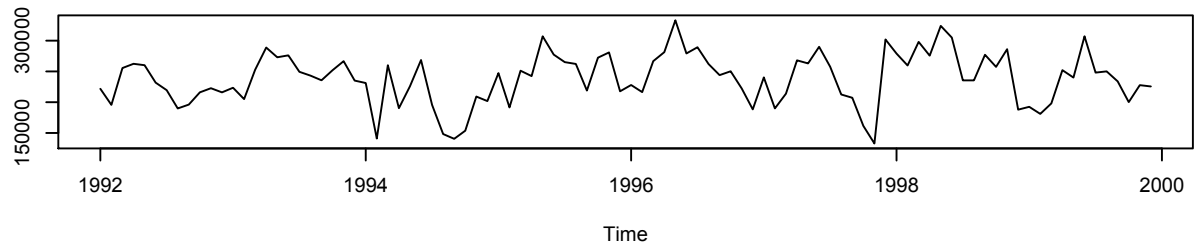
CISARUNI



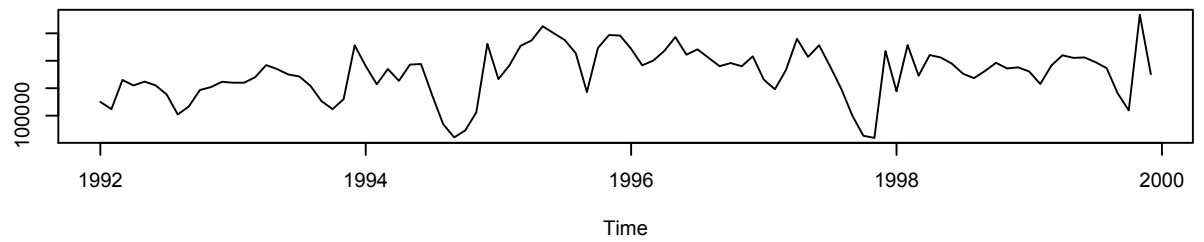
DAYEUEHMANGGUNG



CIATER



TAMBAKSARI



Annexe C

PROGRAMME DÉVELOPPÉ EN R POUR LES SIMULATIONS

```
# Quelques libraries importantes
#####
library(gdata)
library(fBasics)
library(MASS)
library(magic)
library(psych)
#library(mAr)
#library(tsDyn)

# Voici le code pour le GSTAR
#####
library(gplots)
dis=read.xls("/Users/dagrobesson/Desktop/Modele_SIM/distance.xls") # donnee
pour determiner la matrice de poids
#attach(tea)
#head(tea)
attach(dis)
head(dis)
# Distance pour (d1,d2)=(3,inf)
dis3=matrix(NA,nrow=nrow(dis),ncol=ncol(dis))
for(i in 1:ncol(dis)){
```

```

for(j in 1:ncol(dis)){
dis3[i,j]=0+ifelse(dis[i,j]>=3,dis[i,j],0)
}
}
# Distance pour (d1,d2,d3)=(3,4.5,inf)
#dis4=matrix(NA,nrow=ncol(dis),ncol=ncol(dis))
#for(i in 1:ncol(dis)){
#for(j in 1:ncol(dis)){
#dis4[i,j]=0+ifelse(dis3[i,j]>=4.5,dis3[i,j],0)
#}
# }
#####
# Fonction pour calculer la Matrice de poids distance inverse

inv.dis=function(dis){
output=matrix(NA,nrow=nrow(dis),ncol=ncol(dis))
a=matrix(NA,nrow=nrow(dis),ncol=ncol(dis))
b=matrix(NA,nrow=nrow(dis),ncol=ncol(dis))
for (i in 1:nrow(dis)){
for (j in 1:ncol(dis))
a[i,j]=1/(dis[i,j]+1)
diag(a)=0
b[i,]=(a[i,]/rowSums(a)[i])
}
output=b
}
ind=inv.dis
#####
# Les matrices de poids (d'ordre 1,2 et 3)
W1=inv.dis(dis[1:10,1:10]) # matrice de poids pour 10 sites
#####
# Fonction pour calculer les Vi de Borovkova(2008)
Vcal=function(W,Z){
output1=matrix(NA,nrow=nrow(Z),ncol=ncol(W))

```

```

output=matrix(NA,nrow=nrow(Z),ncol=ncol(W))
WZ=matrix(NA,nrow=nrow(Z),ncol=ncol(W))
for(i in 1:ncol(W)){
for(j in 1:ncol(W)){
WZ[1:nrow(Z),j]=W[i,j]*Z[1:nrow(Z),j]
}
output1[,i]=matrix(rowSums(WZ[1:nrow(Z),j]))
}
output=output1
}
V=Vcal
# On initialise le nombre d'iteration (B), le nombre de delais
# individuel (k) pour le test de Portemanteau,
# le nombre de sites (N) et La periode (T)
#####
B=1000
N=10
T=100
k=2*as.integer(sqrt(T))
P=vector(length=k)
probcPort1=vector(length=(length(P)))
probcPort5=vector(length=(length(P)))
probcPort10=vector(length=(length(P)))
for(i in 1:k){
P[i]=i
probcPort1[i]=qchisq(0.99,P[i]*N*N-2*N)
probcPort5[i]=qchisq(0.95,P[i]*N*N-2*N)
probcPort10[i]=qchisq(0.90,P[i]*N*N-2*N)
}
#####
# Initialisation des vecteurs et matrices utilises
#####
Qk_test=matrix(0,nrow=B,ncol=k)
Qk1=matrix(0,ncol=1,nrow=(length(P)))

```



```

Qk5=matrix(0,ncol=1,nrow=(length(P)))
Qk10=matrix(0,ncol=1,nrow=(length(P)))
#####
Qkmod_test=matrix(0,nrow=B,ncol=k)
Qkmod1=matrix(0,ncol=1,nrow=(length(P)))
Qkmod5=matrix(0,ncol=1,nrow=(length(P)))
Qkmod10=matrix(0,ncol=1,nrow=(length(P)))
#####
Q1k_test=matrix(0,nrow=B,ncol=k)
Q1k1=matrix(0,ncol=1,nrow=(length(P)))
Q1k5=matrix(0,ncol=1,nrow=(length(P)))
Q1k10=matrix(0,ncol=1,nrow=(length(P)))
#####
Q1kmod_test=matrix(0,nrow=B,ncol=k)
Q1kmod1=matrix(0,ncol=1,nrow=(length(P)))
Q1kmod5=matrix(0,ncol=1,nrow=(length(P)))
Q1kmod10=matrix(0,ncol=1,nrow=(length(P)))
Beta_est=matrix(NA,nrow=B,ncol=2*N)
#####
QM1=matrix(0,ncol=1,nrow=(length(P)))
QM5=matrix(0,ncol=1,nrow=(length(P)))
QM10=matrix(0,ncol=1,nrow=(length(P)))
QM_test=matrix(0,nrow=B,ncol=k)
#Qk=vector(length=(length(P)))
Z=matrix(0,nrow=T,ncol=N)
#c=vector(length=k*N*N)
#####
# Valeurs des parametres du modele GSTAR(1_1)
#####
#####
#Donnees generees
PHI_0=adiag(0.3,0.1,0.1,0.2,0.2,0.4,0.4,0.6,0.6,0.7) # considerant 10 sites
PHI_1=adiag(0.4,0.3,0.3,0.5,0.5,0.25,0.25,0.45,0.45,0.65)
#Matrice de poids uniforme avec signe differente

```

```

A=PHI_0+PHI_1%*%W1

for(a in 1:B){
#####
#Donnees multivariees par VAR$(1)$
VAR1.sim=function(N,A,T){
mu=rep(0,N)
T1=2*T+1
Z0=rep(0,N)
Sigma=diag(N)
output=matrix(0,nrow=T1,ncol=N)
for(i in 1:T1){
if(i==1){
e1=mvrnorm(1,mu=mu,Sigma=Sigma)
Z1=A%*%Z0+e1
output[1,]=Z1
}
else{
Ztm1=output[i-1,]
et=mvrnorm(1,mu=mu,Sigma=Sigma)
Zt=A%*%Ztm1+et
output[i,]=Zt
}
}
output[(T+2):T1,]
}
Zt=VAR1.sim(N,A,T)
Z=matrix(NA, nrow(Zt), ncol(Zt))
for (i in 1:ncol(Zt)){
Z[,i]=Zt[,i]-mean(Zt[,i])
}

V=Vcal(W1,Z)
#V1=Vcal(W1,Z)

```

```

#V2=Vcal(W2,Z) # V2 pas encore utilisé
# Calcul des Xi en supposant que le modele est GSTAR(2_1_2)
Xical=function(i){
p=1
  output1=matrix(NA,nrow=nrow(Z),ncol=2*p)
output=matrix(NA,nrow=nrow(Z),ncol=2*p)
output1=cbind(Z[(p-1+1):(nrow(Z)-1),i],V[(p-1+1):(nrow(Z)-1),i])
  }
  output=output1
}
Xi=Xical
# calcul de la matrix Bloc diagonale de X
X.cal=function (Z,p){
output1=matrix(0,nrow=nrow(Xical(1))*ncol(Z),ncol=ncol(Z)*2*p)
for(i in 1:ncol(Z)){
output1[(nrow(Xical(1))*(i-1)+1):(nrow(Xical(1))*(i-1)+nrow(Xical(1))),
(2*p*(i-1)+1):(2*p*(i-1)+2*p)]=Xical(i)
}
return(output1)
}
X=X.cal
X=X.cal(Z,1)
# Fonction pour Calculer la matrice Y
Y.cal=function(Z,p){
output1=matrix(NA,nrow=ncol(Z)*(nrow(Z)-p),ncol=1)
output=matrix(NA,nrow=ncol(Z)*(nrow(Z)-p),ncol=1)
for(i in 1:ncol(Z)){
output1[((nrow(Z)-p)*(i-1)+1):((nrow(Z)-p)*(i-1)+(nrow(Z)-p)),]
=matrix(Z[((p+1):nrow(Z)),i])
}
output=output1
}
y=Y.cal
Y=Y.cal(Z,1)

```

```

# Calcul de M POUR l'interval de confiance selon Ahn
M.cal=function (W,p){
output1=matrix(0,nrow=(p+1)*ncol(W),ncol=ncol(W)*ncol(W))
a=diag(ncol(W))
for(i in 1:ncol(W)){
output1[(2*(i-1)+1):(2*(i-1)+2),(ncol(W)*(i-1)+1):(i*ncol(W))]=
  t(cbind(a[i,],W[i,]))
}
return(output1)
}
M=M.cal
M=M.cal(W1,1)
# les dimensions de Y et de X
#dim(Y)
#dim(X)
# le modèle
EE=lm(Y~X-1)
E=summary(EE)
# Intervalle de confiance pour les param.
# les matrices (PHi)
PHI_01est=matrix(0,nrow=ncol(Z),ncol=ncol(Z))
PHI_11est=matrix(0,nrow=ncol(Z),ncol=ncol(Z))
for(i in 1:ncol(Z)){
PHI_01est[i,i]=coef(E)[(2*(i-1)+1),1]
PHI_11est[i,i]=coef(E)[(2*(i-1)+2),1]
}
A1=PHI_01est+PHI_11est%*%W1
Beta_est[a,]=matrix(coef(E)[,1])
#fonction pour calculer lag d'une variable
lag.sim=function (x, lag = 1)
{
  totcols <- ncol(x)
  nas <- matrix(NA, nrow = lag, ncol = totcols)

```

```

    x <- rbind(nas, x)
    totrows <- nrow(x)
    x <- x[-c((totrows - lag + 1):totrows), ]
    return(x)
}

#calcul de Z(t-1)
Zt_1=lag.sim(Z,1)
#Zt_12=lag.sim(Z,2)

#Calcul des Residus (R_t)
p=1
B1=matrix(NA,nrow=(nrow(Z)-p)*ncol(Z),ncol=ncol(Z))
R_t=matrix(NA,nrow=(nrow(Z)-p),ncol=ncol(Z))
for(j in 1:ncol(Z)){
  for(i in 1:ncol(Z)){
    b1=(nrow(Z)-p)*(j-1)+1
    b2=(nrow(Z)-p)*(j-1)+(nrow(Z)-(p))
    B1[b1:b2,i]=matrix(A1[j,i]*Zt_1[(p+1):nrow(Z),i])
  }
  R_t[,j]=matrix(Z[(p+1):nrow(Z),j]-rowSums(B1[b1:b2,]))
}

#Test de Portemanteau pour validation du modele
#####
Gamma_est=(1/(T+1))*((t(Z))%*%Z)
#####
# Calcul de C et chat
#####
#Calcul de C et chat
#####
res=R_t-mean(R_t)
co=t(res)%*%res
c_ii=diag(co)
c_iir=sqrt(c_ii%*%t(c_ii))
co=co/c_iir

```

```

co_inv=solve(co)
rr=kronecker(co_inv,co_inv)
Q=c(0)
Q1=c(0)
Qk=rep(NA,k)
  Qkmod=rep(NA,k)
QM=rep(NA,k)
cl=matrix(0,nrow=N,ncol=N)
x=matrix(0,nrow=N*N,ncol=1)
#####
#Facon 2
#####
Sigma <- crossprod(res)/(T-1)
  Sigma_inv <- solve(Sigma)
  tracesum <- rep(NA, k)
  tracesum1 <- rep(NA, k)
  Q1k=rep(NA,k)
  Q1kmod=rep(NA,k)
#####
for(i in 1:k){
cl=(t(res[1:((T-1)-i),])%*%res[(i+1):(T-1),])/c_iir
x=vec(cl)
Q[i]=(T-1)*(t(x)%*%rr%*%x)
Q1[i]=((T-1)/(T-i))*(t(x)%*%rr%*%x)
Qk[i]=sum(Q[1:i])
Qk_test[a,i]=Qk[i]
Qkmod[i]=(T-1)*sum(Q1[1:i])
Qkmod_test[a,i]=Qkmod[i]
#####
  Qk1[i]=Qk1[i]+ifelse(Qk[i]>probcPort1[i],1,0)
Qk5[i]=Qk5[i]+ifelse(Qk[i]>probcPort5[i],1,0)
Qk10[i]=Qk10[i]+ifelse(Qk[i]>probcPort10[i],1,0)
#####
  Qkmod1[i]=Qkmod1[i]+ifelse(Qkmod[i]>probcPort1[i],1,0)

```

```

Qkmod5[i]=Qkmod5[i]+ifelse(Qkmod[i]>probcaport5[i],1,0)
Qkmod10[i]=Qkmod10[i]+ifelse(Qkmod[i]>probcaport10[i],1,0)
#####
      Ut_1 <- lag.sim(res, lag = i)[-c(1:i), ]
      Ut <- res[-c(1:i), ]
      Cl <- crossprod(Ut, Ut_1)/(T-1)
tracesum[i] <- (T-1)*sum(diag(t(Cl) %*% Sigma_inv %*% Cl %*% Sigma_inv))
tracesum1[i] <- ((T-1)/(T-i))*sum(diag(t(Cl) %*% Sigma_inv %*% Cl %*% Sigma_inv))
      Q1k[i]=sum(tracesum[1:i])
      Q1k_test[a,i]=Q1k[i]
      Q1kmod[i]=(T-1)*sum(tracesum1[1:i])
      Q1kmod_test[a,i]=Qkmod[i]
#####
      Q1k1[i]=Q1k1[i]+ifelse(Q1k[i]>probcaport1[i],1,0)
Q1k5[i]=Q1k5[i]+ifelse(Q1k[i]>probcaport5[i],1,0)
Q1k10[i]=Q1k10[i]+ifelse(Q1k[i]>probcaport10[i],1,0)
#####
      Q1kmod1[i]=Q1kmod1[i]+ifelse(Q1kmod[i]>probcaport1[i],1,0)
Q1kmod5[i]=Q1kmod5[i]+ifelse(Q1kmod[i]>probcaport5[i],1,0)
Q1kmod10[i]=Q1kmod10[i]+ifelse(Q1kmod[i]>probcaport10[i],1,0)
#####
QM[i]=(Q1k[i] - (N*N*i-2*N))/sqrt(2*(N*N*i-2*N))
QM1[i]=QM1[i]+ifelse(QM[i]>qnorm(0.99),1,0)
QM5[i]=QM5[i]+ifelse(QM[i]>qnorm(0.95),1,0)
QM10[i]=QM10[i]+ifelse(QM[i]>qnorm(0.90),1,0)
QM_test[a,i]=QM[i]
}
}
      Probcap=cbind(probcaport1,probcaport5,probcaport10)
Resul1=cbind(P,Qk1,Qk5,Qk10)
Resul2=cbind(P,Qkmod1,Qkmod5,Qkmod10)
Resul1
Resul2

```

```

colSums(Beta_est)/B #Comparaison des estimateurs
diag(PHI_0)
diag(PHI_1)
CC=cbind(Qk,Q1k,Qkmod,Q1kmod)
par(mfrow=c(1,3))
x <- Q1k_test
hist(x, prob=TRUE,main = paste("Hist. of x pour n=10"))
#curve( dchisq(x, df=180), col='green', add=TRUE)
curve(x, col='red', add=TRUE )

x <- Q1kmod_test
hist(x, prob=TRUE,main = paste("Hist. of x pour n=100"))
#curve( dchisq(x, df=180), col='green', add=TRUE)
curve( dchisq(x, df=180), col='red', add=TRUE )

x <- QM_test
hist(x, prob=TRUE,main = paste("Hist. of x pour n=1000"))
#curve( dchisq(x, df=180), col='green', add=TRUE)
curve(x, col='red', add=TRUE )
#####
Code pour l'analyse de données réelles.
GSTAR(2)
#####
library(gdata)
library(fBasics)
library(gplots)
library(MASS)
tea=read.xls("/Users/dagrobesson/Desktop/Modele_SIM/meantea.xls")
dis=read.xls("/Users/dagrobesson/Desktop/Modele_SIM/distance.xls")
attach(tea)
head(tea)
attach(dis)
head(dis)
# Distance pour (d1,d2)=(3,inf)

```



```

dis3=matrix(NA,nrow=nrow(dis),ncol=ncol(dis))
for(i in 1:ncol(dis)){
for(j in 1:ncol(dis)){
dis3[i,j]=0+ifelse(dis[i,j]>=3,dis[i,j],0)
}
}
# Distance pour (d1,d2,d3)=(3,4.5,inf)
dis4=matrix(NA,nrow=ncol(dis),ncol=ncol(dis))
for(i in 1:ncol(dis)){
for(j in 1:ncol(dis)){
dis4[i,j]=0+ifelse(dis3[i,j]>=4.5,dis3[i,j],0)
}
}
#####
# Function pour calculer la Matrice de poids distance inverse

inv.dis=function(dis){
output=matrix(NA,nrow=nrow(dis),ncol=ncol(dis))
a=matrix(NA,nrow=nrow(dis),ncol=ncol(dis))
b=matrix(NA,nrow=nrow(dis),ncol=ncol(dis))
for (i in 1:nrow(dis)){
for (j in 1:ncol(dis))
a[i,j]=1/(dis[i,j]+1)
diag(a)=0
b[i,]=(a[i,]/rowSums(a)[i])
}
output=b
}
ind=inv.dis
#####
# Les matrices de poids (d'ordre 1,2 et 3)
W2=inv.dis(dis3)
W1=inv.dis(dis)
#####

```

```

# Fonction pour calculer les Vi qui se trouve dans V suivant
#W d'ordre 1 (matrice de poids)
Vcal=function(W,Z){
output1=matrix(NA,nrow=nrow(Z),ncol=ncol(W))
output=matrix(NA,nrow=nrow(Z),ncol=ncol(W))
WZ=matrix(NA,nrow=nrow(Z),ncol=ncol(W))
for(i in 1:ncol(W)){
for(j in 1:ncol(W)){
WZ[1:nrow(Z),j]=W[i,j]*Z[1:nrow(Z),j]
}
output1[,i]=matrix(rowSums(WZ[1:nrow(Z),j]))
}
output=output1
}
V=Vcal
#####
#calcul respective des Vi suivant les differentes matrices
#de poids d'ordre 1,2 et 3
Z=ts(log(tea),start=1996,frequency=12)
V=Vcal(W1,Z)
V1=Vcal(W1,Z)
#V2=Vcal(W2,Z) # V2 pas encore utilisé
# Calcul des Xi en supposant que le modele est GSTAR(2_1_2)
Xical=function(i){{
p=2
output1=matrix(NA,nrow=nrow(Z),ncol=2*p)
output=matrix(NA,nrow=nrow(Z),ncol=2*p)
output1=cbind(Z[(p-1+1):(nrow(Z)-1),i],V[(p-1+1):(nrow(Z)-1),i],
Z[1:(nrow(Z)-p),i],V1[1:(nrow(Z)-p),i])
}
output=output1
}
Xi=Xical

```

```

# calcul de la matrix Bloc diagonale de X
X.cal=function (Z,p){
output1=matrix(0,nrow=nrow(Xical(1))*ncol(Z),ncol=ncol(Z)*2*p)
for(i in 1:ncol(Z)){
output1[(nrow(Xical(1))*(i-1)+1):(nrow(Xical(1))*(i-1)+
nrow(Xical(1))), (2*p*(i-1)+1):(2*p*(i-1)+2*p)]=Xical(i)
}
return(output1)
}
X=X.cal
X=X.cal(Z,2)
# Fonction pour Calculer la matrice Y
Y.cal=function(Z,p){
output1=matrix(NA,nrow=ncol(Z)*(nrow(Z)-p),ncol=1)
output=matrix(NA,nrow=ncol(Z)*(nrow(Z)-p),ncol=1)
for(i in 1:ncol(Z)){
output1[((nrow(Z)-p)*(i-1)+1):((nrow(Z)-p)*(i-1)+
(nrow(Z)-p)),]=matrix(Z[((p+1):nrow(Z)),i])
}
output=output1
}
y=Y.cal
Y=Y.cal(Z,2)
# Calcul de M POUR l'intervale de confiance selon Ahn
M.cal=function (W,p){
output1=matrix(0,nrow=(p+1)*ncol(W),ncol=ncol(W)*ncol(W))
a=diag(ncol(W))
for(i in 1:ncol(W)){
output1[(2*(i-1)+1):(2*(i-1)+2), (ncol(W)*(i-1)+1):(i*ncol(W))]=
t(cbind(a[i,],W[i,]))
}
return(output1)
}

```

```

M=M.cal
M=M.cal(W1,1)

# les dimensions de Y et de X
#dim(Y)
#dim(X)

# le modèle
EE=lm(Y~X-1)
E=summary(EE)

# les matrices (PHi)
PHI_01=matrix(0,nrow=ncol(Z),ncol=ncol(Z))
PHI_11=matrix(0,nrow=ncol(Z),ncol=ncol(Z))
PHI_02=matrix(0,nrow=ncol(Z),ncol=ncol(Z))
PHI_22=matrix(0,nrow=ncol(Z),ncol=ncol(Z))
for(i in 1:ncol(Z)){
#####
PHI_01[i,i]=coef(E)[(4*(i-1)+1),1]
PHI_11[i,i]=coef(E)[(4*(i-1)+2),1]
PHI_02[i,i]=coef(E)[(4*(i-1)+3),1]
PHI_22[i,i]=coef(E)[(4*(i-1)+4),1]
}
A1=PHI_01+PHI_11%*%W1
A2=PHI_02+PHI_22%*%W1

#fonction pour calculer lag d'une variable
lag.sim=function (x, lag = 1)
{
  totcols <- ncol(x)
  nas <- matrix(NA, nrow = lag, ncol = totcols)
  x <- rbind(nas, x)
  totrows <- nrow(x)
  x <- x[-c((totrows - lag + 1):totrows), ]
  return(x)
}

#calcul de Z(t-1)

```

```

Zt_1=lag.sim(Z,1)
Zt_12=lag.sim(Z,2)

#Calcul des Residus (R_t)
p=2
B=matrix(NA,nrow=(nrow(Z)-p)*ncol(Z),ncol=ncol(Z))
R_t=matrix(NA,nrow=(nrow(Z)-p),ncol=ncol(Z))
for(j in 1:ncol(Z)){
for(i in 1:ncol(Z)){
b1=(nrow(Z)-p)*(j-1)+1
b2=(nrow(Z)-p)*(j-1)+(nrow(Z)-(p))
B[b1:b2,i]=matrix(A1[j,i]*Zt_1[(p+1):nrow(Z),i]+A2[j,i]*Zt_12[(p+1):nrow(Z),i])
}
R_t[,j]=matrix(Z[(p+1):nrow(Z),j]-rowSums(B[b1:b2,]))
}
#Test de Portemanteau pour validation du modele
N=ncol(Z)
T=nrow(Z)
k=2*as.integer(sqrt(T))
P=vector(length=k)
probcPort1=vector(length=(length(P)))
probcPort5=vector(length=(length(P)))
probcPort10=vector(length=(length(P)))
for(i in 1:k){
P[i]=i
probcPort1[i]=qchisq(0.99,P[i]*N*N-2*p*N)
probcPort5[i]=qchisq(0.95,P[i]*N*N-2*p*N)
probcPort10[i]=qchisq(0.90,P[i]*N*N-2*p*N)
}
*****
# Initialisation des vecteurs et matrices utilises
*****
#Qk_test=matrix(0,nrow=B,ncol=k)
Qk1=matrix(0,ncol=1,nrow=(length(P)))

```

```

Qk5=matrix(0,ncol=1,nrow=(length(P)))
Qk10=matrix(0,ncol=1,nrow=(length(P)))
#####
#Qkmod_test=matrix(0,nrow=B,ncol=k)
Qkmod1=matrix(0,ncol=1,nrow=(length(P)))
Qkmod5=matrix(0,ncol=1,nrow=(length(P)))
Qkmod10=matrix(0,ncol=1,nrow=(length(P)))
#####
#Q1k_test=matrix(0,nrow=B,ncol=k)
Q1k1=matrix(0,ncol=1,nrow=(length(P)))
Q1k5=matrix(0,ncol=1,nrow=(length(P)))
Q1k10=matrix(0,ncol=1,nrow=(length(P)))
#####
#Q1kmod_test=matrix(0,nrow=B,ncol=k)
Q1kmod1=matrix(0,ncol=1,nrow=(length(P)))
Q1kmod5=matrix(0,ncol=1,nrow=(length(P)))
Q1kmod10=matrix(0,ncol=1,nrow=(length(P)))
QM1=matrix(0,ncol=1,nrow=(length(P)))
QM5=matrix(0,ncol=1,nrow=(length(P)))
QM10=matrix(0,ncol=1,nrow=(length(P)))
QM_test=matrix(0,nrow=B,ncol=k)
#####
Gamma_est=(1/(T+1))*((t(Z))%*%Z)
#####
# Calcul de C et chat
#####
res=R_t-mean(R_t)
co=t(res)%*%res
c_ii=diag(co)
c_iir=sqrt(c_ii%*%t(c_ii))
co=co/c_iir
co_inv=solve(co)
rr=kronecker(co_inv,co_inv)
Q=c(0)

```

```

Q1=c(0)
Qk=rep(NA,k)
  Qkmod=rep(NA,k)
cl=matrix(0,nrow=N,ncol=N)
x=matrix(0,nrow=N*N,ncol=1)
#####
#Facon 2
#####
Sigma <- crossprod(res)/(T-p)
  Sigma_inv <- solve(Sigma)
  tracesum <- rep(NA, k)
  tracesum1 <- rep(NA, k)
  Q1k=rep(NA,k)
QM=rep(NA,k)
  Q1kmod=rep(NA,k)
  Pval=rep(NA,k)
  Pval1=rep(NA,k)
PvalM=rep(NA,k)
#####
#calcul de la variance des paramètres
I_B=M%*(kronecker(Sigma_inv,Gamma_est))%*(t(M))
I_B_inv=solve(I_B)/(T-p)
sigm=matrix(sqrt(diag(I_B_inv)))
# les intervalles de confiance (CIi)
CI_01=matrix(0,nrow=ncol(Z),ncol=2)
CI_11=matrix(0,nrow=ncol(Z),ncol=2)
CI_02=matrix(0,nrow=ncol(Z),ncol=2)
CI_22=matrix(0,nrow=ncol(Z),ncol=2)
for(i in 1:ncol(Z)){
CI_01[i,]=cbind((coef(E)[(4*(i-1)+1),1]-1.96*sigm[(2*(i-1)+1),1]),
(coef(E)[(4*(i-1)+1),1]+1.96*sigm[(2*(i-1)+1),1]))
CI_11[i,]=cbind((coef(E)[(4*(i-1)+2),1]-1.96*sigm[(2*(i-1)+2),1]),
(coef(E)[(4*(i-1)+2),1]+1.96*sigm[(2*(i-1)+2),1]))
CI_02[i,]=cbind((coef(E)[(4*(i-1)+3),1]-1.96*sigm[(2*(i-1)+1),1]),

```

```

(coef(E)[(4*(i-1)+3),1]+1.96*sigm[(2*(i-1)+1),1]))
CI_22[i,]=cbind((coef(E)[(4*(i-1)+4),1]-1.96*sigm[(2*(i-1)+2),1]),
(coef(E)[(4*(i-1)+4),1]+1.96*sigm[(2*(i-1)+2),1]))
}

#####
for(i in 1:k){
cl=(t(res[1:(T-2)-i,])%*%res[(i+1):(T-2),])/c_iir
x=vec(cl)
Q[i]=(T-p)*(t(x)%*%rr%*%x)
Q1[i]=Q1[i]+((T-p)/(T-p-i))*(t(x)%*%rr%*%x)
Qk[i]=sum(Q[1:i])
#Qk_test[a,i]=Qk[i]
Qkmod[i]=(T-p)*sum(Q1[1:i])
#Qkmod_test[a,i]=Qkmod[i]

#####
      Qk1[i]=Qk1[i]+ifelse(Qk[i]>probcPort1[i],1,0)
Qk5[i]=Qk5[i]+ifelse(Qk[i]>probcPort5[i],1,0)
Qk10[i]=Qk10[i]+ifelse(Qk[i]>probcPort10[i],1,0)
#####
      Qkmod1[i]=Qkmod1[i]+ifelse(Qkmod[i]>probcPort1[i],1,0)
Qkmod5[i]=Qkmod5[i]+ifelse(Qkmod[i]>probcPort5[i],1,0)
Qkmod10[i]=Qkmod10[i]+ifelse(Qkmod[i]>probcPort10[i],1,0)
#####
      Ut_1 <- lag.sim(res, lag = i)[-c(1:i), ]
      Ut <- res[-c(1:i), ]
      Cl <- crossprod(Ut, Ut_1)/(T-p)
      tracesum[i] <- (T-p)*sum(diag(t(Cl) %*% Sigma_inv %*% Cl %*% Sigma_inv))
tracesum1[i] <- ((T-p)/(T-p-i))*sum(diag(t(Cl) %*% Sigma_inv %*% Cl %*% Sigma_inv))
      Q1k[i]=sum(tracesum[1:i])
      #Q1k_test[a,i]=Q1k[i]
      Q1kmod[i]=(T-p)*sum(tracesum1[1:i])
      #Q1kmod_test[a,i]=Q1kmod[i]

#####
      Q1k1[i]=Q1k1[i]+ifelse(Q1k[i]>probcPort1[i],1,0)

```



```

Q1k5[i]=Q1k5[i]+ifelse(Q1k[i]>probcapPort5[i],1,0)
Q1k10[i]=Q1k10[i]+ifelse(Q1k[i]>probcapPort10[i],1,0)
#####
      Q1kmod1[i]=Q1kmod1[i]+ifelse(Q1kmod[i]>probcapPort1[i],1,0)
Q1kmod5[i]=Q1kmod5[i]+ifelse(Q1kmod[i]>probcapPort5[i],1,0)
Q1kmod10[i]=Q1kmod10[i]+ifelse(Q1kmod[i]>probcapPort10[i],1,0)
QM[i]=(Q1k[i]-(N*N*i-2*p*N))/sqrt(2*(N*N*i-2*p*N))
QM1[i]=QM1[i]+ifelse(QM[i]>qnorm(0.99),1,0)
QM5[i]=QM5[i]+ifelse(QM[i]>qnorm(0.95),1,0)
QM10[i]=QM10[i]+ifelse(QM[i]>qnorm(0.90),1,0)
Pval[i]=1-pchisq(Q1k[i],N*N*i-2*p*N)
Pval1[i]=1-pchisq(Q1kmod[i],N*N*i-2*p*N)
PvalM=1-pnorm(QM[i])
}

#graphique des intervalles de confiance des Phi
par(mfrow=c(2,2))
plotCI(diag(PHI_01), li=CI_01[,1],ui=CI_01[,2],ylim=c(-0.5,1.5),
col="black", barcol="black",gap=0, lwd=1,
main = expression(paste(plain() * phi[10]^i),
" for GSTAR",plain((2[11])))),xlab="Location",ylab="")
abline(a=.38,b=0,col = "red", lty=10)
abline(a=.0,b=0,col = "black", lty=10)

plotCI(diag(PHI_02), li=CI_02[,1],ui=CI_02[,2],ylim=c(-0.5,1.5),
col="black", barcol="black",gap=0, lwd=1,
main = expression(paste(plain() * phi[20]^i),
" for GSTAR",plain((2[11])))),xlab="Location",ylab="")
abline(a=.20,b=0,col = "red", lty=10)
abline(a=.0,b=0,col = "black", lty=10)
plotCI(diag(PHI_11), li=CI_11[,1],ui=CI_11[,2],ylim=c(-1.0,2.5),
col="black", barcol="black",gap=0, lwd=1,
main = expression(paste(plain() * phi[11]^i),
" for GSTAR",plain((2[11])))),xlab="Location",ylab="")
abline(a=.33,b=0,col = "red", lty=10)

```

```

abline(a=.0,b=0,col = "black", lty=10)
plotCI(diag(PHI_22), li=CI_22[,1],ui=CI_22[,2], ylim=c(-1.5,1.0),
col="black", barcol="black",gap=0, lwd=1,
main = expression(paste(plain() * phi[21]^i,
" for GSTAR",plain((2[11])))),xlab="Location",ylab="")
abline(a=-.32,b=0,col = "red", lty=10)
abline(a=.0,b=0,col = "black", lty=10)
Aic=log(det((1/(T-2*(p+2)-1))*((t(R_t))%*%R_t)))+(2*(p+2+1+1)+N*(N+1))/(T-p)
Aic

Resul1=cbind(P,"1% facon I"=Qk1,"5% facon I"=Qk5,"10% facon I"=Qk10)
Resul2=cbind(P,"1% facon I (mod)"=Qkmod1,"5% facon I (mod)"=Qkmod5,"10%
facon I (mod)"=Qkmod10)
Resul3=cbind(P,"1% facon II"=Q1k1,"5% facon II"=Q1k5,"10%
facon II"=Q1k10)
Resul4=cbind(P,"1% facon II (mod)"=Q1kmod1,"5%
facon II (mod)"=Q1kmod5,"10% facon II (mod)"=Q1kmod10)
Resul5=cbind(P,"1% facon II (mod)"=QM1,"5%
facon II (mod)"=QM5,"10% facon II (mod)"=QM10)
Resul1
Resul2
Resul3
Resul4
Resul5
CCC=cbind(P,as.numeric(sprintf("%.3f",Q1k)),as.numeric(sprintf("%.3f",Q1kmod)),
as.numeric(sprintf("%.3f",Pval)),as.numeric(sprintf("%.3f",Pval1)))
CCC
# estimation des paramètres du modèle GSTAR(2)
tab1=cbind(diag(PHI_01),diag(PHI_11),diag(PHI_02),diag(PHI_22))
tab1

```

BIBLIOGRAPHIE

- [1] Ahn, S. K. (1988), 'Distribution for residual autocovariances in multivariate autoregressive models with structured parameterization'. *Biometrika* **75**, 590-593.
- [2] Ahn, S. K. et Reinsel, G. C. (1990), 'Estimation of partially nonstationary multivariate autoregressive models'. *Journal of the American Statistical Association* **85**, 813-823.
- [3] Ahn, S. K. et Reinsel, G. C. (1988), 'Nested reduced rank autoregressive models for multiple time series'. *Journal of the American Statistical Association* **83**, 849-856.
- [4] Bavaud, F. (1998), 'Models for spatial weights : a systematic look'. *Geographical Analysis* **30**, 153-171.
- [5] Borovkova, S., Lopuhaä, H. P. et Ruchjana, B. N. (2008), 'Consistency and asymptotic normality of least squares estimators in generalized STAR models'. *Statistica Neerlandica* **62**, 482-508.
- [6] Boubacar Mainassara, Y. et Francq, C. (2009), 'Estimating structural VARMA models with uncorrelated but non-independent error terms'. *Document de travail, Université Lille III, France*.
- [7] Boubacar Mainassara, Y. (2009), 'Multivariate portmanteau test for structural VARMA models with uncorrelated but non-independent errors terms'. *MPRA No 23371, Juin 2010*.
- [8] Brockwell, P. J. et Davis, R. A. (1991), *Time Series : Theory and Methods*, 2^e édition. Springer : New York.
- [9] Brockwell, P. J. et Davis, R. A. (2002), *Introduction to Time Series and Forecasting*, 2^e édition. Springer : New York.
- [10] Cliff, A. D. et Ord, J. (1973), *Spatial Autocorrelation*. Pioneer : London.

- [11] Cliff, A. D. *et* Ord, J. (1981), *Spatial Processes, Models and Applications*. Methuen : New York.
- [12] Di Giacinto, V. (2003), 'Differential regional effects of monetary policy : A geographical SVAR approach'. *International Regional Science Review* **26**, 313-341.
- [13] Di Giacinto, V. (2006), 'A generalized space-time ARMA model with an application to regional unemployment analysis in Italy'. *International Regional Science Review* **29**, 159-198.
- [14] Duchesne, P. (2005), 'On the asymptotic distribution of residual autocovariances in VARX models with applications'. *Test* **14**, 449-473.
- [15] Gaetan, C. *et* Guyon, X. (2008), *Modélisation et Statistique Spatiales*. Springer : New York.
- [16] Gatrell, A. C. (1979), 'Autocorrelation in spaces'. *Environment and Planning A* **11**, 507-516.
- [17] Hannan, E. J. (1970), *Multiple Time Series*. Wiley : New York.
- [18] Hordijk, L. (1974), 'Spatial correlation in the disturbances of a linear interregional model'. *Regional and Urban Economics* **4**, 117-140.
- [19] Hosking, J. R. M. (1980), 'The multivariate portmanteau statistic'. *Journal of the American Statistical Association* **75**, 602-608.
- [20] Karyia, T. *et* Kurata, H. (2004), *Generalized Least Squares*. Wiley : New York.
- [21] Krige, D. (1951), 'A statistical approach to some basic mine valuation problems on the witwatersrand'. *Journal of the Chemical, Metallurgical and Mining Society of South Africa* **52**, 119-139.
- [22] Le Gallo, J. (2000), 'Econométrie spatiale, parties 1 et 2'. *Document de travail, Université de Bourgogne, France*.
- [23] Ljung, G. M. *et* Box, G. E. P. (1978), 'On a measure of lack of fit in time series models'. *Biometrika* **65**, 297-304.
- [24] Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*. Springer-Verlag : Berlin.
- [25] Martin, R. L. *et* Oeppen, J. E. (1975), 'The identification of regional forecasting models using space-time correlation functions'. *Transactions of the Institute of British Geographers* **66**, 95-118.

- [26] McLeod, A. I. (1978), 'On the distribution of residual autocorrelations in Box-Jenkins models'. *Journal of the Royal Statistical Society, Series B* **40**, 296-302.
- [27] McLeod, A. I. et Li, W. K. (1981), 'Distribution of the residual autocorrelations in multivariate ARMA time series models'. *Journal of the Royal Statistical Society, Series B* **43**, 231-239.
- [28] Pfeifer, P. E. et Deutsch, S. J. (1980a), 'A three-stage iterative procedure for space-time modeling'. *Technometrics* **22**, 35-47.
- [29] Pfeifer, P. E. et Deutsch, S. J. (1980b), 'Identification and interpretation of first order space-time ARMA models'. *Technometrics* **22**, 397-408.
- [30] Pfeifer, P. E. et Deutsch, S. J. (1981), 'Space-time ARMA modeling with contemporaneously correlated innovations'. *Technometrics* **23**, 401-409.
- [31] Pukkila, T. S. K. et Kallinen, A. (1990), 'The identification of ARMA models'. *Biometrika* **77**, 537-548.
- [32] Reinsel, G. C. (2005), *Elements of Multivariate Time Series Analysis*. Springer-Verlag : Berlin.
- [33] Shumway, R. H. et Stoffer, D. S. (2006), *Time Series Analysis and Its Applications With R Examples*. Springer : New York.
- [34] Terzi, S. (1995), 'Maximum likelihood estimation of a generalized STAR($p; l_p$) model'. *Journal of the Italian Statistical Society* **3**, 377-393.
- [35] Terzi, S. et Giuseppina, G. (1997), 'Identifiability conditions for generalised STARMA models'. *Journal of the Italian Statistical Society* **3**, 245-255.
- [36] Tiao, G. C. et Box, G. E. P. (1981), 'Modeling multiple time series with applications'. *Journal of the American Statistical Association* **76**, 802-816.
- [37] Wei, W. W. (1994), *Time Series Analysis, Univariate and Multivariate Methods*. Addison-Wesley Company.
- [38] Zellner, A. (1962), 'An efficient method of estimating seemingly unrelated regressions, and tests for aggregation bias'. *Journal of the American Statistical Association* **57**, 348-368.