

N° d'ordre : 28 / 2011 – M / MT

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie «Houari Boumediene»
Faculté de Mathématiques



MEMOIRE

Présenté pour l'obtention du diplôme de **MAGISTER**

En : MATHEMATIQUES

Spécialité : **Recherche Opérationnelle** : Génie Mathématiques

Par : **GATT Fella**

Thème

Modèles à Seuil Classiques et Périodiques.

Soutenu publiquement le 07 /07/ 2011, devant le jury composé de :

M. MOULAI	Mustapha	Professeur,	à l'U. S. T. H. B.,	Président
M. BENTARZI	Mohamed	Professeur,	à l'U. S. T. H. B.,	Directeur de mémoire
M.GUERBYENNE	Hafida	MCA,	à l'U. S. T. H. B.,	Examinatrice
M. HAMDI	Fayçal	MCB,	à l'U. S. T. H. B.,	Invité

Modèles à Seuil Classiques et Périodiques

GATT Fella

Département de Recherche Opérationnelle

Faculté de Mathématiques,

Université des Sciences et de la Technologie Houari Boumediene,

U. S. T. H. B.

Table des matières

Introduction	ii
0.1 Apport et présentation de la thèse	v
1 Modèles à seuil classiques	1
1.1 Introduction	1
1.2 Processus TARMA	3
1.3 Processus TMA	5
1.4 Processus TAR	6
1.5 Processus SETAR	7
1.6 Processus MTAR	10
1.7 Processus STAR	11
1.8 Processus TARSO	14
2 Identification, Spécification et estimation du modèle SETAR	15
2.1 Identification d'un modèle SETAR	15
2.2 Spécification d'un modèle SETAR	16
2.2.1 Choix des régresseurs	16
2.2.2 Choix de la variable de transition	17
2.3 Estimation d'un modèle SETAR	18
2.3.1 Méthodologie de Tsay (1989) et Hansen (1996)	19
2.4 Tests de linéarité	23

2.4.1	Test Reset	23
2.4.2	Test du multiplicateur de Lagrange	24
2.4.3	Test du CUSUM	25
3	Estimation d'un modèle SETAR par la méthode MCO	26
3.1	Modèle Autorégressif à seuil SETAR et Modèle de régression linéaire	27
3.2	Estimations des moindres carrés d'un SETAR	29
3.3	Estimation et Test sous Contraintes Linéaires	33
3.4	Estimation des paramètres d'un modèle SETAR sous contraintes linéaires	35
3.5	Estimation sous l'hypothèse de normalité	38
3.5.1	Estimateurs du maximum de vraisemblance	38
3.5.2	Loi de probabilité de $Q \left(\underline{b} / Y, \underline{X} \right)_H$	39
3.6	Test du Rapport de la Vraisemblance	40
4	Simulation et Application	42
4.1	Estimation des paramètres d'un modèle SETAR	43
4.2	Test du rapport de vraisemblance	51
4.3	Application sur des données réelles	55
5	Modèles à seuil périodiques	62
5.1	Introduction	62
5.2	Notations et Définitions principales	63
5.2.1	Processus Autorégressif à seuil Self-Exciting Classique (<i>SETAR</i>)	63
5.2.2	Processus Autorégressif à seuil Self-Exciting Périodique (<i>PSETAR</i>)	64
5.2.3	Hypothèses de régularités techniques	65
5.3	Normalité Asymptotique Locale	66
5.3.1	Suite de Rapports de Vraisemblance	66

<i>TABLE DES MATIÈRES</i>	iii
5.3.2 Normalité Asymptotique Locale	68
5.4 Test Asymptotique Locale (le plus Stringent)	71
5.5 Test Adaptatif	73
Conclusion et Perspectives	77
Bibliographie	79
ANNEXE	82

Alà barakati Allah

Introduction

L'analyse des séries chronologiques (temporelles) occupe une phase prépondérante dans l'étude des phénomènes aléatoires, à savoir l'économie, l'automatique, la météorologie,...son importance est due au vaste champ d'application qu'elle procure. Elle permet, entre autres de prévoir le comportement futur d'un phénomène écoulé dans le temps.

Les modèles les plus répandus et les plus utilisés dans la classe des modèles de séries chronologiques classiques sont les modèles autorégressifs linéaires de type $ARMA(p, q)$. Bien qu'ils ne représentent souvent qu'une approximation brutale de la réalité, les modèles $ARIMA$, $SARIMA$, $VARMA$..., ont prouvés leurs utilités dans la pratique, justifiée théoriquement par le théorème de Wold (1954) qui montre que tout processus faiblement stationnaire peut être réécrit sous la forme d'une moyenne mobile infinie de processus de type bruits blancs, c'est à dire sous la forme d'une combinaison linéaire d'une séquence de variable aléatoires non corrélées dans le temps. Or, l'hypothèse de processus stationnaire qu'ils soient univariés ($ARMA$) ou multivariés (VAR) ne permet pas de prendre en compte d'une part les mécanismes d'asymétrie et d'autre part les ruptures de forte amplitude. D'où la nécessité d'aller vers des modélisations non linéaires et plus particulièrement les changements de régimes.

La classe des modèles à changements de régimes a été initiée à l'origine par Tong (1978) et Tong & Lim (1980). Leur principale propriété est de permettre à la série économique de posséder une dynamique différente suivant les états ou les régimes dans laquelle elle se trouve.

Dans les modèles à changements de régimes stochastiques se pose le problème de définir le mécanisme de transition entre les différents régimes. Ce dernier dépend soit d'une variable de transition inobservable ou observable. Ce qui permet de faire la distinction entre deux grandes

classes de modèles existantes en séries temporelles : il s'agit des modèles à changements de régimes markoviens et les modèles à seuil.

Les modèles markoviens ont été introduit par Hamilton (1988, 1989) et étendu par plusieurs auteurs, entre autres, Chang-Jin Kim (1994), le mécanisme de transition repose sur une variable d'état inobservable qui est supposée suivre une chaîne de Markov. Contrairement aux modèles à seuil, où le mécanisme de transition s'effectue à l'aide d'une variable de transition observable, d'un seuil et d'une fonction de transition. L'inconvénient de ce type de modèles réside donc en la définition de cette variable observable. Pour le bon choix de cette variable, on fait recours aux tests de linéarité.

De là on distingue deux mécanismes de transition, selon la forme de la fonction de transition utilisée. En effet, si la fonction utilisée est une fonction indicatrice, la transition est alors dite brutale et le passage d'un régime à l'autre pouvant s'effectuer en une période. Autrement dit, il s'effectue en comparant la variable de transition à un seuil, si cette dernière devient supérieure ou inférieure à ce seuil, la transition se réalise alors instantanément. Cependant, cette démarche a des inconvénients qui ne sont pas toujours faciles à surmonter. Puisque une variation de la variable de transition entraîne subitement le passage d'un régime à l'autre, même si cette évolution est minime. De plus, la distance par rapport au seuil n'influe pas sur le passage d'un régime à l'autre. Sans oublier que ce type de modèle ne peut contenir qu'un nombre limité de régimes comme dans le cas du modèle à changements de régimes markoviens, ce qui rend difficile la prise en compte des mécanismes non-linéaires complexes.

Pour pallier à ces problèmes, Chan & Tong (1987) et Luukkonen, Saikkonen & Teräsvirta (1988) ont proposé de remplacer la fonction indicatrice des modèles à transition brutale par une fonction continue et bornée entre 0 et 1 afin d'autoriser une transition lisse entre les régimes extrêmes. permettant ainsi d'introduire une autre classe de modèles à transition lisse qui a l'avantage de pouvoir être vu comme un modèle contenant un continuum de régimes. Par ailleurs, l'appartenance à un régime dépend de la fonction de transition et de la distance entre le seuil et la variable de transition.

L'ensemble des modèles à changements de régimes évoqué jusqu'à présent a permis dans la littérature empirique de modéliser des asymétries telles que des dynamiques distinctes dans les phases ascendantes et descendantes à l'aide de leurs différents régimes. Ils permettent également de s'interroger sur la stabilité temporelle des coefficients dans le temps. Cependant,

contrairement au modèle de rupture, le passage d'un régime à l'autre n'est ni daté ni définitif étant déterminé endogènement en fonction d'un seuil.

Il y a lieu de noter que l'existence de plusieurs régimes dans un même modèle permet au processus d'être globalement stationnaire. En effet, tous les régimes ne sont pas obligatoirement caractérisés par la présence de racine unitaire dans leur polynôme autorégressif et réciproquement tous ne sont pas contraints à être stationnaires. En séries temporelles, cette question de la non-stationnarité versus la non-linéarité est très importante, sachant que ces deux notions peuvent être confondues à l'issue d'un test de stationnarité classique.

Après cette brève présentation des modèles non-linéaires à changements de régimes dans le contexte des séries temporelles, intéressons nous à présent à la classe des modèles à seuil. Les modèles les plus couramment utilisés dans cette catégorie sont les modèles *SETAR* et *STAR* où le changement de régime est régi par un seuil. Parmi ces modèles à seuil, nous retenons ceux dans lesquels la transition entre les régimes est brutale : les modèles *Self-Exciting Threshold Autoregressive (SETAR)*. Ces modèles fournissent un cadre privilégié de l'étude des fluctuations cycliques asymétriques et permettent de révéler les cycles limites et les diverses ruptures. L'idée sous-jacente est qu'une relation peut être non linéaire sur la période globale considérée, mais linéaire par sous période. Un modèle *SETAR* est ainsi formé de relations linéaires par morceaux. Parallèlement à la modélisation des asymétries cycliques et des changements de comportements, ces modèles ont été utilisés pour traiter des enjeux de non-stationnarité. Ainsi, la non-stationnarité d'un processus (établie dans un cadre linéaire) peut être remise en cause dans un cadre non-linéaire. En ce qui concerne les transitions douces, on fait référence aux modèles *Smooth Transition AutoRegressive (STAR)*. En effet, Luukkonen, Saikkonen et Teräsvirta (1988), Luukkonen et Teräsvirta (1991), et Teräsvirta et Anderson (1992) ont proposé cette classe de modèles qui est une extension des modèles *TAR* à deux régimes. Ces modèles reposent sur l'idée selon laquelle l'économie peut se situer dans deux régimes différents. Toutefois, contrairement aux modèles *TAR* où le passage d'un régime à un autre se fait de façon brutale, la transition sera douce dans le cas des processus *STAR*. On introduit ainsi un lissage sur le temps permettant d'atténuer les changements de régimes. Il peut dès lors exister un continuum d'états entre les deux régimes extrêmes. Les modèles *STAR* sont principalement utilisés pour caractériser le cycle économique ou l'évolution de certaines variables macroéconomiques et financière. La spécification de ces

modèles nécessite l'utilisation de tests de linéarité contre l'alternative donnée par le modèle à seuil. D'où l'objectif de ce mémoire qui est de proposer un test basé sur le test du rapport de vraisemblance pour tester l'hypothèse de linéarité contre l'alternative fournie par le modèle à seuil *Self-Exciting Threshold Autoregressive (SETAR)*. Il faut noter que, sous l'hypothèse nulle de linéarité, le paramètre délai - ou la variable de transition - et le seuil ne sont pas identifiables. En effet, la valeur de la vraisemblance du modèle linéaire ne dépend pas de ces paramètres si nous supposons que la variance des deux bruits blancs est la même dans les deux régimes. Ainsi, ils peuvent prendre n'importe quelles valeurs sans changer la fonction de vraisemblance. Pour accomplir cet objectif nous avons suivi une méthodologie présentée dans la section suivante :

0.1 Apport et présentation de la thèse

Chapitre 1 : Modèles à seuil classiques

Dans ce chapitre, on présente les modèles non linéaires à changement de régimes dans le cadre des séries temporelles, globalement, et la classe des modèles à seuil particulièrement. On passe en revue les modèles les plus connus dans cette catégorie, à savoir : les modèles *TARMA*, *TMA*, *TAR*, *SETAR*, *TARSO*, *STAR*,....

Chapitre 2 : Identification, Spécification et Estimation du modèle *SETAR*

Dans la première section de ce chapitre, nous exposons la procédure d'identification de Tong et Lim (1980). La technique de Tong et Lim (1980) repose essentiellement sur l'utilisation du critère d'information d'Akaike. Trois étapes sont nécessaires à l'application de cette méthode. La spécification du modèle à seuil est présentée dans la section suivante, cette étape nécessite le choix de certains paramètres qui ne peuvent être estimés par des méthodes usuelles. Il s'agit du nombre de régresseurs ainsi que du paramètre définissant la variable de transition. Ils sont alors sélectionnés préalablement à l'estimation des autres paramètres du modèle. Les tests de linéarité contre l'alternative donnée par le modèle à seuil guident ces choix. L'estimation des modèles *SETAR* est présentée dans la dernière section où l'on distingue deux grandes techniques d'estimation des modèles *SETAR* : la méthode

de Tsay (1989) et la méthode de Hansen (1996).

Chapitre 3 : Estimation d'un modèle *SETAR* par la méthode *MCO*

Ce chapitre est consacré à l'estimation du modèle *SETAR* par la méthode des moindres carrés ordinaire (*MCO*). Nous introduisons ce chapitre par des notations et définitions nécessaire pour la suite de notre travail. Dans la deuxième section, nous exposons la méthode la plus courante, parmi les méthodes d'estimation statistique des paramètres d'un modèle de régression dite *Méthode des Moindres Carrées Ordinaire (MCO)*. Sous un système d'hypothèses bien définit (*Conditions de Gauss Markov*), l'estimateur $\hat{\underline{b}}$ des moindres carrés ordinaires du vecteur des paramètres de régression \underline{b} et ainsi que l'estimateur $\hat{\sigma}^2$ de la variance commune σ^2 de l'erreur sont obtenus. Dans la section suivante, on suppose, en plus des conditions de Gauss-Markov, que la famille de lois de la variable aléatoire à expliquée est normale. En effet, nous montrons que sous l'hypothèse de normalité l'estimateur linéaire sans biais de variance minimale dans la classe de tous les estimateurs linéaires sans biais devient un estimateur sans biais de variance minimale dans la classe de tous les estimateurs sans biais. De plus, les estimateurs des moindres carrés sont, dans ce cas confondus avec les estimateurs du maximum de vraisemblance, normalement distribués.

Chapitre 4 : Simulation et application

Ce chapitre est composé de trois sections. Nous présentons dans la première section les résultats d'estimation du vecteur des paramètres associé au modèle à seuil *SETAR*. La deuxième section est consacrée aux résultats du test du rapport de vraisemblance qui revient à tester l'hypothèse nulle de linéarité contre l'hypothèse alternative qui est le modèle à seuil *SETAR*. Dans la dernière section, nous présentons des applications sur des données réelles (lynx canadien, Sunspot,...), auxquelles nous estimons les paramètres et nous appliquons le test, une comparaison avec les résultats de Tong est présentée.

Chapitre 5 : Modèles à seuil périodiques

Dans ce chapitre, on présente la classe des modèles à seuil Self-Exciting Threshold Autoregressive à coefficients périodiques (*PSETAR*) pour lequel Bentarzi et Merzougui (2009) ont construit des procédures optimales pour tester l'hypothèse de périodicité. La propriété de normalité locale asymptotique est obtenue. Cette propriété, donné par Koul et Schick (1997)

dans le cas non périodique, est établie à l'aide du Lemme de Swensen(1985). Les conditions de ce lemme ont été aisément vérifiées, pour autant que la racine carrée de la fonction de densité caractérisant le modèle soit différentiable en moyenne quadratique. Cette dernière condition été le point crucial à vérifier. A partir de ce résultat, les tests paramétriques localement et asymptotiquement optimaux peuvent être construit.

Chapitre 1

Modèles à seuil classiques

1.1 Introduction

Comme évoqué lors de l'introduction générale. Les modèles de séries chronologiques linéaires sont incapable de capturer les différentes caractéristiques de cycles économiques. En effet, ces modèles linéaires sont tous de type régression construits à partir de la fonction somme. Contrairement aux modèles non linéaires qui sont construits à partir de classes de fonctions plus complexes. C'est le cas des processus bilinéaires, processus hétéroscédastiques, et les processus longue mémoire qui sont construit à partir des **fonctions polynômiales**. Un autre cas est celui des processus à seuil, l'objet de ce mémoire, ainsi que les processus à changements de régimes markoviens qui sont construit à partir des **fonctions discontinues** et particulièrement les processus *EXPAR* et certains processus hétéroscédastiques qui sont issus des **fonctions de type logarithme ou exponentiel**.

La modélisation de ces processus permet de mettre en évidence les caractéristiques spécifiques vérifiées par ces modèles non linéaires et que les processus linéaires ne peuvent pas détecter.

Comme première étape pour l'analyse de ces processus on examine la représentation graphique à partir de laquelle on essaye de détecter des explosions, des ruptures, des sauts, des agrégats, et des phénomènes de persistance.

En effet, détecter la non linéarité seule ne suffit pas car il existe dans la littérature

plusieurs modèles non linéaires comme les modèles *ARCH*, *GARCH*,

FARMA, *GARMA*, ... qui prennent en compte l'existence de la volatilité ou de la longue mémoire sur les séries, mais qui ne sont pas très adaptés pour modéliser, soit l'existence de sauts, soit la présence de plusieurs états dans le système. Pour cela, nous présentons dans ce chapitre une classe de modèles qui permettent de tenir compte de ces comportements. Il s'agit des modèles à seuil.

Il convient tout d'abord de rappeler la définition d'un processus linéaire dans la mesure où les principales modélisations non linéaires vont partir de cette représentation de base.

Définition 1.1.1

Un processus purement stochastique est dit linéaire s'il peut s'écrire comme somme pondérée d'un processus en bruits blancs, soit :

$$Y_t = \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} \quad (1.1.1)$$

où μ est une constante, ψ_i des coefficients réels avec $\psi_0 = 1$ et $\{\varepsilon_t\}$ des variables aléatoires ayant même distribution, centrées indépendantes et de variance $\sigma_\varepsilon^2 < \infty$ et telles que :

$$E[\varepsilon_t \mid Y_{t-1}, Y_{t-2}, \dots] = 0$$

Si de plus $\sum_{i=0}^{\infty} |\psi_i| < \infty$, alors les deux premiers moments de Y_t sont indépendants du temps et Y_t est dit alors un processus faiblement stationnaire.

Les modèles les plus répandus et les plus utilisés dans l'analyse des séries chronologiques sont les processus *ARMA*(p, q) popularisée par Box et Jenkins qui ont pour écriture :

$$\Phi(L) Y_t = \Theta(L) \varepsilon_t \quad (1.1.2)$$

où L est l'opérateur de décalage défini par $L^j y_t = y_{t-j}$, $\Phi(L)$ et $\Theta(L)$ sont des polynômes en L de degrés respectifs p et q tels que leurs racines sont à l'extérieures du cercle complexe unitaire ce qui assure pour $\Phi(L)$ la stationnarité du processus et pour $\Theta(L)$ son inversibilité.

L'équation (1.1.1) peut être considérée comme étant un cas particulier issu d'une structure

dans laquelle Y_t est simplement fonction des variables aléatoires ε présente et passées :

$$Y_t = f(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) \quad (1.1.3)$$

et dans laquelle on impose la linéarité de la fonction f .

Si la fonction f est considérée non linéaire et selon l'événement qu'il soit observables ou non, alors on peut définir les processus non linéaires à seuil de type *TAR* (*Threshold AutoRegression*) ou à changements markoviens envisagés par Tong (1978, 1983) et Hamilton (1989).

Tous les modèles paramétriques cités précédemment permettent de construire des prévisions à partir d'une structure donnée par :

$$Y_t = g(Y_t, Y_{t-1}, Y_t, \dots, \varepsilon_t, \varepsilon_{t-1}, \theta) + \varepsilon_{t+1}$$

où θ représente un vecteur de paramètres de dimension finie. On remarque que le choix d'une spécification pour la fonction f entraîne en théorie celle de g .

Notons que dans le cas des processus non linéaire le calcul des prévision à partir de l'espérance conditionnelle en t de Y_{t+h} devient complexe à un horizon de prévision h quelconque. Ceci est dû au fait que :

$$E[f(\cdot)] \neq f(E[\cdot])$$

Si on veut ne pas spécifier la fonction f , dans ce cas on est amené à utiliser des méthodes non paramétriques telles que : les réseaux de neurones, les régressions à noyaux, les régressions locales *LOESS* ou locales pondérées *LOWESS*.... En pratique, ces techniques nécessitent un nombre d'observations qui dépasse beaucoup la taille des échantillons disponibles sur données macroéconomiques [10].

1.2 Processus TARMA

Dans ce type de processus une linéarité par morceaux de la fonction f est envisagée définissant ainsi les différents régimes, à travers lesquels le processus passe d'un état à un autre. En effet, si les moments de changement ou de transition d'un régime à un autre sont

connus, alors il suffit tout simplement d'intégrer une fonction indicatrice dans l'écriture du modèle pour tenir compte de ce changement. Dans le cas où les moments de transition ne sont pas observés sur l'échelle des temps, le passage d'un régime à l'autre sera guidé par une variable dite "variable de transition". D'où la définition de ce processus.

Définition 1.2.1

Un processus non linéaire $\{Y_t, t \geq 0\}$ suit un modèle Autoregressive Moyenne mobile à seuil $TARMA(2; p_1, q_1, p_2, q_2, d)$ à deux régimes, s'il est solution de l'équation caractéristique aux différences stochastiques suivantes :

$$Y_t = \begin{cases} \phi_0^{(1)} + \sum_{i=1}^{p_1} \phi_i^{(1)} Y_{t-i} + \sum_{j=1}^{q_1} \theta_j^{(1)} \varepsilon_{t-j}^{(1)} + \varepsilon_t^{(1)} & \text{si } Y_{t-d} \leq r \\ \phi_0^{(2)} + \sum_{i=1}^{p_2} \phi_i^{(2)} Y_{t-i} + \sum_{j=1}^{q_2} \theta_j^{(2)} \varepsilon_{t-j}^{(2)} + \varepsilon_t^{(2)} & \text{si } Y_{t-d} > r \end{cases}$$

où $\varepsilon_t^{(1)}, \varepsilon_t^{(2)}$: deux bruits blancs de variance respectivement σ_1^2 et σ_2^2 .

d : entier appelé "paramètre de délai".

r : réel inconnu appelé paramètre de seuil.

– Lorsque $Y_{t-d} \leq r$, c'est à dire d périodes plus tard, la variable y sera générée par un certain processus $ARMA(p_1, q_1)$.

– Si $Y_{t-d} > r$, c'est à dire après d périodes, y obéira à un autre processus $ARMA(p_2, q_2)$.

Ce processus à seuil à deux régimes se généralise à un nombre quelconque K de régimes et se note : $TARMA(K; p_1, p_2, q_1, q_2)$, soit l'extension la plus générale du modèle $ARMA$.

En considérant une variable de transition réelle on a :

$$Y_t = \sum_{k=1}^K (\phi_0^{(k)} + \sum_{i=1}^{p_k} \phi_i^{(k)} Y_{t-i} + \sum_{j=1}^{q_k} \theta_j^{(k)} \varepsilon_{t-j}^{(k)} + \varepsilon_t^{(k)}) \cdot I_{(Y_{t-d} \in R_i)} \quad (1.1.4)$$

où $\varepsilon_t^{(k)}$: bruit blanc de variance σ_k^2 , $k = 1, \dots, K$.

Les segments $R_i =]r_{i-1}, r_i]$ forment une partition de l'espace des réels conformément aux valeurs des paramètres de seuil r_i telque $-\infty = r_0 < r_1 < \dots < r_K = +\infty$.

$I(\cdot)$: fonction indicatrice prenant 1 si $Y_{t-i} \in R_i$, 0 sinon.

Selon [10], Tong (1983) considère le cas où la variable de transition est l'expliquée elle-même obtenant alors le modèle *Self – exciting TARMA* noté *SETARMA* dans lequel il a imposé également l'absence de changement dans les processus en bruits blanc entre les régimes ; soit $\varepsilon_t^{(k)} = \varepsilon_t$.

On présente ainsi la structure d'un modèle $SETARMA(K; p_1, \dots, p_k; q_1, \dots, q_k)$ à K régimes, comme suit :

$$Y_t = \sum_{k=1}^K (\phi_0^{(k)} + \sum_{i=1}^{p_k} \phi_i^{(k)} Y_{t-i} + \sum_{j=1}^{q_k} \theta_j^{(k)} \varepsilon_{t-j}^{(k)} + \varepsilon_t) \cdot I_{(Y_{t-d} \in R_i)} \quad (1.1.5)$$

Selon [10], la construction de prévision à partir de cette structure est discutée notamment par Amendola et Niglio (2003) qui autorisent une modification des variances des bruits selon les régimes. Quand aux conditions de stationnarité et d'ergodicité, ils ont été obtenues (Brockwell, Liu et Tweedie 1992) dans le cas particulier où la composante *MA* n'est pas soumise aux effets de seuil.

Pour pallier aux difficultés rencontrées lors de l'utilisation de l'écriture générale de ces processus, il a été considéré des simplifications de ces structures générales pour ne retenir que des processus à seuil en moyenne mobiles (*TMA*) et surtout autorégressif (*TAR*).

1.3 Processus TMA

L'une des structures simplifiées d'un processus *TARMA*, est le processus *TMA*, où chaque régime est gouverné par un processus moyenne mobile *MA*. Ainsi, un modèle suit un processus $SETMA(q_1, q_2, d)$ à un seul changement de régime ($K = 1$) s'il est solution de l'équation caractéristique aux différences stochastiques suivante :

$$Y_t = \sum_{i=1}^{q_1} \theta_i^{(1)} \varepsilon_{t-i} + \sum_{j=1}^{q_2} \theta_j^{(2)} I_{(Y_{t-d} \leq r)} \varepsilon_{t-i} + \varepsilon_t \quad (1.1.6)$$

Le processus *TMA* est toujours stationnaire comme le processus $MA(q)$ sans imposer des contraintes sur les coefficients [10]. Comme pour le processus $MA(q)$ il est nécessaire d'imposer des conditions d'inversibilité. Selon [10], seules des conditions suffisantes ont été énoncées

par Ling et Tong (2007) pour un *TMA* quelconque et des conditions nécessaires et suffisantes pour un *TMA(1)* par Ling, Tong et Li (2005).

Exemple

Soit Y_t un processus strictement stationnaire et ergodique, définit par :

$$Y_t = \varepsilon_t + (\theta_1 + \theta_2 I_{(Y_{t-d} \leq r)}) \varepsilon_{t-i}$$

avec $E[\ln |Y_t|] < \infty$. Le processus est *inversible* si $|\theta_1|^{1-F_y(r)} |\theta_1 + \theta_2|^{F_y(r)} < 1$, où $F_y(\cdot)$ est la fonction de répartition de Y_t , et il est *non inversible* si $|\theta_1|^{1-F_y(r)} |\theta_1 + \theta_2|^{F_y(r)} > 1$.

Selon [10], le processus *SETMA* est également analysé par Degooijer (1998) qui propose en particulier un test de lagrange de l'hypothèse *MA* versus *SETMA*. Ces processus engendrent naturellement des réponses asymétriques aux chocs.

Notons que le poids de ε_{t-i} dans la détermination de Y_t est donné par $\theta_i^{(1)}$ lorsque $Y_{t-d} > r$, sinon le poids est égal à $\theta_i^{(1)} + \theta_i^{(2)}$.

1.4 Processus TAR

Les modèles Autorégressifs à seuil, en anglais "*Threshold Autoregressive model*" *TAR* ont été introduits par Tong en (1978). Cette classe de modèles est populaire dans la littérature des séries temporelles non-linéaires du fait que ces modèles sont relativement simples à spécifier et à estimer en comparaison avec d'autres modèles de séries chronologiques non-linéaires existant.

Dans le cas des modèles à seuil, la non linéarité entre les variables économiques peut être modélisée de façon simple. En effet, la linéarisation par morceaux est une représentation approximative du comportement non linéaire d'un processus.

1.5 Processus SETAR

Définition 1.5.1

Un processus non linéaire $\{Y_t, t \geq 0\}$ suit un modèle *Self-exciting threshold autoregressive models* $SETAR(k; p_1, \dots, p_k, d)$ à k régimes, s'il est solution de l'équation caractéristique aux différences stochastiques suivantes :

$$Y_t = \begin{cases} \phi_0^{(1)} + \sum_{i=1}^{p_1} \phi_i^{(1)} Y_{t-i} + \varepsilon_t^{(1)} & \text{si } Y_{t-d} < r_1 \\ \phi_0^{(2)} + \sum_{i=1}^{p_2} \phi_i^{(2)} Y_{t-i} + \varepsilon_t^{(2)}, & \text{si } r_1 \leq Y_{t-d} < r_2 \\ \vdots & \\ \phi_0^{(k)} + \sum_{i=1}^{p_k} \phi_i^{(k)} Y_{t-i} + \varepsilon_t^{(k)}, & \text{si } Y_{t-d} \geq r_{k-1} \end{cases} \quad (1.1.7)$$

avec : d entier positif appelé paramètre de délai (ou de retard), $r_j, j = 1, \dots, k-1$, sont les paramètres de seuil pour lesquels le système passe d'un régime à l'autre, Y_{t-d} est la variable de seuil et les $\{\varepsilon_t^{(1)}, \dots, \varepsilon_t^{(k)}\}$ sont des bruits blancs non corrélés.

La valeur des seuil donne une première interprétation économique des régimes définissant la dynamique du processus. En effet, dans le cas d'un modèle à deux régimes ($k = 2$), on constate que le processus $SETAR$ suit un certain régime $AR(p)$ quand $Y_{t-d} \leq r$, et tombe dans un régime $AR(p)$ différent à la période suivante si $Y_{t-d+1} > r$, et lorsque la valeur seuil r est nulle on définit deux états : un régime de croissance positive et un régime de croissance négative.

On notera que ces modèles permettent de prendre en considération des structures de retards différentes pour chaque sous période étudiée, et donc de mener une modélisation relativement souple.

Un cas particulier de ce type de modèle est lorsque les ordres de retards sont les mêmes pour toutes les équations, c'est à dire $p_i = p \forall i$, on dira que les modèles $SETAR$ se ramènent

aux modèles *Threshold AutoRegressive (TAR)* d'ordre p qui s'écrivent comme suit :

$$Y_t = \begin{cases} \phi_0^{(1)} + \sum_{i=1}^p \phi_i^{(1)} Y_{t-i} + \varepsilon_t^{(1)} & \text{si } Y_{t-d} < r_1 \\ \phi_0^{(2)} + \sum_{i=1}^p \phi_i^{(2)} Y_{t-i} + \varepsilon_t^{(2)}, & \text{si } r_1 \leq Y_{t-d} < r_2 \\ \vdots & \\ \phi_0^{(k)} + \sum_{i=1}^p \phi_i^{(k)} Y_{t-i} + \varepsilon_t^{(k)}, & \text{si } Y_{t-d} \geq r_{k-1} \end{cases} \quad (1.1.8)$$

Un autre cas particulier est lorsque l'ordre autorégressif p vaut 1 dans chaque régime, alors on peut énoncer le théorème suivant :

Théorème 1 (*Chan, Petrucelli, Tong et Woolford (1985)*)

Un processus $TAR(1)$ à K régimes est ergodique si et seulement si l'une des conditions suivantes est vérifiée :

- 1) $\phi_1^{(1)} < 1, \phi_1^{(k)} < 1, \phi_1^{(1)} \phi_1^{(k)} < 1$
- 2) $\phi_1^{(1)} = 1, \phi_1^{(k)} < 1, \phi_0^{(1)} > 0$
- 3) $\phi_1^{(1)} < 1, \phi_1^{(k)} = 1, \phi_0^{(k)} < 0$
- 4) $\phi_1^{(1)} = 1, \phi_1^{(k)} = 1, \phi_0^{(k)} < 0 < \phi_0^{(1)}$
- 5) $\phi_1^{(1)} \phi_1^{(k)} = 1, \phi_1^{(1)} < 0, \phi_0^{(k)} + \phi_1^{(k)} \phi_0^{(1)} > 0$

Remarque

On introduisant des termes MA dans chaque régime du modèle $SETAR$, on obtient les processus $SETARMA$, cités précédemment.

Exemples :

— Soit un modèle $SETAR(K; p; d)$ où tous les AR sont du même ordre p ;

La condition nécessaire et suffisante de stationnarité [10] est $\max_k \sum_i^p |\phi_{ki}| < 1$ et si $p = 1$

cette condition devient $\max_k |\phi_{k1}| < 1$; ce qui revient à imposer la stationnarité de chacun

des K processus $AR(1)$.

Dans un modèle à deux régimes, les conditions à la fois nécessaires et suffisantes sont connues et correspondent aux contraintes : $\phi_{11} < 1$; $\phi_{21} < 1$ et $\phi_{11}\phi_{21} < 1$.

Selon [10], la condition suffisante de Chang est alors trop restrictive et il est possible de construire un modèle *SETAR* stationnaire dans lequel un des régimes est gouverné par un *AR* non stationnaire.

— Soit un modèle *SETAR*(2; 1; 1)

$$Y_t = \begin{cases} -2Y_{t-1} + \varepsilon_t & \text{si } Y_{t-1} \leq 0 \\ 0.6 Y_{t-1} + \varepsilon_t & \text{si } Y_{t-1} > 0 \end{cases}$$

C'est un processus stationnaire malgré que l'un des régimes est gouverné par un *AR* explosif. Du fait que $\phi_{11} = -2$, Y_t sera plus souvent positif que négatif de sorte que $E(Y_t) > 0$. Ainsi, $E(Y_t)$ sans terme constant est $\neq 0$. Par contraste, on rappelle qu'un *AR* est d'espérance non nulle précisément si et seulement si son équation fait apparaître un terme constant non nul [10].

3) Un autre cas souvent employé est celui d'un *SETAR* à trois régimes dans lequel le processus central possède une racine unitaire. Lorsqu'on se limite à des processus *AR*(1) cela revient à imposer que dans le régime intermédiaire la variable obéit à une marche aléatoire.

On a donc :

$$Y_t = \begin{cases} \phi_1 Y_{t-1} + \varepsilon_t & \text{si } Y_{t-1} \leq r_1 \\ Y_{t-1} + \varepsilon_t & \text{si } r_1 < Y_{t-1} \leq r_2 \\ \phi_2 Y_{t-1} + \varepsilon_t & \text{si } Y_{t-1} > r_2 \end{cases}$$

Si $|\phi_1| < 1$ et $|\phi_2| < 1$ alors le processus est globalement ergodique ce qui implique l'existence d'une distribution stationnaire pour Y_t (Tong, 1990) même si la série est localement non stationnaire [10].

Note

— La considération de processus de type *SETAR* avec racine unitaire donne lieu à des extensions particulièrement intéressantes généralisant la théorie de la cointégration initiée par Engle et Granger (1987) et Johansen (1988, 1991, 1996) qui suppose la linéarité et la symétrie des ajustements.

— La recherche de la spécification optimale pour un nombre de régimes fixé a priori, s'effectue

généralement au moyen de critères de sélection de type *AIC* ou *BIC* adaptés par Wong et Li (1998) dont les valeurs sont données par les équations suivantes :

$$AIC(p_1, p_2) = \sum_{k=1}^2 [N_k \log \hat{\sigma}_k^2 + 2(p_k + 2)]$$

$$AIC_c(p_1, p_2) = \sum_{k=1}^2 \left[N_k \log \hat{\sigma}_k^2 + \frac{N_k(N_k + p_k + 1)}{N_k - p_k - 3} \right]$$

$$BIC(p_1, p_2) = \sum_{k=1}^2 [N_k \log \hat{\sigma}_k^2 + (p_k + 1) \log N_k]$$

Des simulations de Monte Carlo réalisées par ces auteurs paraissent indiquer que le critère *BIC* est préférable sur grands échantillons alors que sur petits échantillons l'emploi du critère d'Akaïke corrigé *AIC_c* serait préférable. Selon [10], des critères de validation de choix des paramètres d'un *SETAR* sont donnés dans De Gooijer (2001) et Galeano et Peña (2004).

1.6 Processus MTAR

Définition 1.6.1

Un processus non linéaire $\{Y_t, t \geq 0\}$ suit un modèle Momentum Threshold AutoRegression (*MTAR*), s'il est solution de l'équation caractéristique aux différences suivante :

$$Y_t = \sum_{k=1}^K (\phi_0^{(k)} + \sum_{i=1}^{p_k} \phi_i^{(k)} Y_{t-i} + \varepsilon_t^{(k)}) \cdot I_{(\Delta Y_{t-d} \in R_i)} \quad (1.1.9)$$

où $\varepsilon_t^{(k)}$, $k = 1, \dots, K$, suite de bruit blanc non corrélés,

$\phi_i^{(k)}$, $i = \overline{0, p_k}$, représentent les coefficients du modèle *MTAR*, et $\Delta Y_{t-d} = Y_t - Y_{t-d}$ est la variation décalée de Y .

Ce type de processus est préconisés pour la modélisation des taux d'intérêt, les taux de change et, plus généralement, les rendements des actifs financiers.

Dans un modèle *MTAR*, la variable de transition est la **variation retardée** de l'expliquée, contrairement au modèle *SETAR* où la variable de transition est constituée du **niveau ratardé** de la variable explicative.

1.7 Processus STAR

Parmi les inconvénients des modèles autorégressif à seuil *TAR*, nous citons la brutalité de la transition entre les régimes qui peut mettre en doute la crédibilité des prévisions qui en sont issus.

Définition 1.7.1

Un processus non linéaire $\{Y_t, t \geq 0\}$, suit un modèle *Smooth transition autoregressive model* (*STAR*(2; p, d)) s'il est solution de l'équation caractéristique aux différences stochastiques suivante :

$$Y_t + \phi_{10} + \sum_{i=1}^p \phi_{1i} Y_{t-i} + \left(\phi_{20} + \sum_{i=1}^p \phi_{2i} Y_{t-i} \right) F(z_t) = \varepsilon_t \quad (1.1.10)$$

où $F(z_t)$ est une fonction de transition continue et bornée entre 0 et 1, et $\varepsilon_t \approx iid(0, \sigma^2)$.

Ce type de modèle consiste une généralisation des modèles *SETAR*, par le fait que $F(\cdot)$ n'est plus une fonction spécifiée mais restreinte à certaines propriétés comme la croissance stricte, l'imparité...

La version logistique d'un modèle *STAR*, recommandée par Luukkonen et al. [1988a] est donnée par :

$$F(z_t) = (1 + \exp(-z_t))^{-1}$$

$$z_t = \gamma(Y_{t-d} - r);$$

où $F(\cdot)$ est la fonction logistique et γ facteur multiplicatif non négatif.

En considérant $F :] - \infty, +\infty[\rightarrow [0, 1]$, on déduit facilement que ce modèle généralise le premier par une transition continue entre les deux régimes où γ est le paramètre qui contrôle la pente de transition. Notons qu'un modèle *STAR* dont le γ tend vers l'infini tend vers un modèle *SETAR*.

Dans le cas du modèle *STAR*, la brutalité de la transition d'un régime à l'autre est contrôlée par le paramètre γ . Comme les deux régimes ont des structures dynamiques différentes, la probabilité conditionnelle au passé de tomber dans un régime, n'est pas nécessairement la

même que celle de tomber dans l'autre. D'où une modélisation de l'aspect asymétrique du cycle. Par ailleurs la discontinuité au niveau du seuil complique l'étape d'estimation. Pour cela, Chang et Tong (1986) ont proposé de retenir une *transition douce* entre les différents régimes définissant ainsi la classe des processus *STAR* (*Smooth Transition Autoregressions*) [10].

Dans le cas simple caractérisé par une seule fonction de transition où l'on suppose pour simplifier l'égalité des variances des résidus, on obtient l'écriture suivante :

$$Y_t = (\phi_0^{(1)} + \sum_{i=1}^p \phi_i^{(1)} Y_{t-i}) [1 - G(x_t; \theta)] + (\phi_0^{(2)} + \sum_{i=1}^p \phi_i^{(2)} Y_{t-i}) G(x_t; \theta) + \varepsilon_t \quad (1.1.11)$$

où $G(\cdot)$: fonction continue de paramètre θ telque $0 \leq G(\cdot) \leq 1$ et x_t une variable de transition stationnaire.

— Un modèle *STAR* peut recevoir deux interprétations possibles : soit celle d'un modèle à seuil à deux régimes chacun étant associé aux valeurs extrêmes de la fonction de transition $G(x_t; \theta) = 0$ et $G(x_t; \theta) = 1$ avec une transition continue de l'un des régimes vers l'autre. soit celle d'un continuum de régimes, étant déterminé par une valeur particulière de la fonction de transition et donc de la variable de transition x_t .

L'équation peut être réécrite de manière équivalente :

$$Y_t = \beta_{0t} + \beta_{1t} Y_{t-1} + \beta_{2t} Y_{t-2} + \dots + \beta_{pt} Y_{t-p} + \varepsilon_t$$

avec : $\beta_{it} = \phi_i^{(1)} \times [1 - G(x_t; \theta)] + \phi_i^{(2)} \times G(x_t; \theta)$.

Le modèle *STAR* peut être vu comme un *AR* à coefficients variables. En conséquence le modèle *STAR* est localement linéaire lorsque la variable de transition x_t est constante.

Selon la fonction de transition deux formulations sont envisageable :

— Si la fonction de transition est de type logistique, on introduit le modèle *LSTAR* (*Logistic STAR*); donné par :

$$G(x_t; \theta) = L(x_t; \gamma, r) = [1 + \exp(-\gamma(x_t - r))]^{-1}; \quad \gamma > 0$$

— Si la fonction de transition est de type exponentielle on obtient le modèle *ESTAR*

(Exponential *STAR*) définit par :

$$G(x_t; \theta) = E(x_t; \gamma, r) = [1 - \exp(-\gamma(x_t - r)^2)]; \gamma > 0$$

Notons que dans un modèle *LSTAR*, lorsque γ tend vers 0 alors la fonction $L(x_t; \gamma, r)$ tend vers 0.5 et le *SETAR* dégénère en un *AR* linéaire, et lorsque γ tend vers $+\infty$ alors la fonction $L(x_t; \gamma, r)$ tend vers 1 si $x_t > r$, et tend vers 0 si $x_t < r$.

En d'autres termes le processus *TAR* à deux régimes est un cas particulier du *LSTAR* obtenu avec des valeurs élevées de la pente de la fonction de transition. Compte-tenu de ces caractéristiques, le processus *LSTAR* avec une seule transition est adapté pour la modélisation des variables pouvant obéir à deux dynamiques différentes dans des régimes extrêmes avec un comportement asymétrique en fonction de l'écart à l'équilibre. C'est le cas des cycles d'activité avec alternance de phases de récession et d'expansion.

Concernant le processus *ESTAR*, le comportement de la fonction de transition dans le cas simple à deux régimes extrêmes et une seule valeur pour le paramètre de seuil r , est sensiblement différent du cas *LSTAR*.

Du fait de la forme quadratique retenue, seule la distance entre la variable de transition, x_t , et le paramètre de seuil, r , va intervenir sans considération de leurs positions relatives, puisque :

$$E(x_t; \gamma, r) = E(x_t; 0, r) = 0 \text{ et } \lim_{x \rightarrow \pm\infty} E(x_t; \gamma, r \mid \gamma \neq 0) = \lim_{\gamma \rightarrow \infty} E(x_t; \gamma, r \mid x_t \neq r) = 1$$

Ainsi, en fonction de la distance $x_t - r$, la transition à une forme en U: un régime extrême correspond au cas où x_t tend vers $+\infty$ ou $-\infty$. De ce fait, le processus *ESTAR* est plutôt adapté à l'ajustement de systèmes possédant un corridor central tel que vu plus haut avec les *SETAR* à trois régimes. Notons cependant que le processus *ESTAR* ne dégénère pas en un *TAR* à trois régimes selon les valeurs de γ puisqu'aux limites de son espace de définition le modèle devient linéaire.

Le processus *LSTAR* d'ordre n aura pour fonction de transition :

$$LSTAR(x_t; \gamma, r_1, r_2, \dots, r_n) = [1 + \exp(-\gamma \prod_{i=1}^n (x_t - r_i))]^{-1}, \gamma > 0$$

Un autre type de processus appelé *Multiple Regime STAR (MRSTAR)* a été introduit par Van Dijk et Franses (1999) constitué de plusieurs fonctions de transition faisant chacune référence à des variables de transition différentes. Ce modèle possède la particularité, au même titre que les réseaux de neurones, de pouvoir approximer toute fonction continue non linéaire.

Enfin, des travaux récents ont proposé des processus à transition douce avec mémoire longue. Dans un modèle *FISTAR*, la variable modélisée est la différence fractionnaire de Y_t , soit :

$$(1-L)^d Y_t = (\phi_0^{(1)} + \sum_{i=1}^p \phi_i^{(1)} (1-L)^d Y_{t-i}) [1-G(x_t; \theta)] + (\phi_0^{(2)} + \sum_{i=1}^p \phi_i^{(2)} (1-L)^d Y_{t-i}) G(x_t; \theta) + \varepsilon_t \quad (1.1.12)$$

Un tel processus est étudié par Van Dijk, Franses et Paap (2002) avec une fonction de transition logistique, et par Smallwood (2005) au moyen d'une fonction exponentielle [].

1.8 Processus TARSO

Si un processus de commande ou de contrôle intervient dans la définition du modèle à seuil, on définit le modèle à seuil avec boucle, en anglais, "*Open-loop threshold autoregressive model*" qu'on note *TARSO*. Sous des conditions identiques à celles définies précédemment, on dit que les processus (X_t, Y_t) suivent un modèle $TARSO(l; (p_1, p'_1), \dots, (p_k, p'_k))$ si :

$$X_t = \phi_0^{(j)} + \sum_{i=1}^{p_j} \phi_i^{(j)} X_{t-i} + \sum_{i=1}^{p'_j} \theta_i^{(j)} Y_{t-j} + \varepsilon_t^{(j)} \quad (1.1.13)$$

conditionnellement au fait que $X_{t-d} \in R_j, j = 1, \dots, k$. Si (X_t, Y_t) est un processus *TARSO* et si (Y_t, X_t) est aussi un processus *TARSO*, alors on dit que le processus (X_t, Y_t) est un modèle *TARSC* ("closed-loop threshold model").

Chapitre 2

Identification, Spécification et estimation du modèle SETAR

2.1 Identification d'un modèle SETAR

Dans cette section, nous exposons la procédure d'identification de Tong et Lim (1980). La technique de Tong et Lim (1980) repose essentiellement sur l'utilisation du critère d'information d'Akaike.

Considérons le modèle $SETAR(2; p, q)$; ce choix permet de simplifier la présentation dans la mesure où une seule valeur seuil r est à estimer :

$$Y_t = \begin{cases} \phi_0^{(1)} + \sum_{i=1}^p \phi_i^{(1)} y_{t-i} + \varepsilon_t^{(1)} & \text{si } y_{t-d} \leq r \\ \phi_0^{(2)} + \sum_{i=1}^q \phi_i^{(2)} y_{t-i} + \varepsilon_t^{(2)}, & \text{si } y_{t-d} > r \end{cases} \quad (2.2.1)$$

Trois étapes sont nécessaires à l'application de la méthode de Tong et Lim.

Étape1. On commence par se fixer arbitrairement un ensemble de valeurs pour le paramètre de délai d et pour le seuil r . on estime chacun des deux modèles par les méthodes des moindres carrés ordinaires (MCO) en choisissant les valeurs de p et q qui minimisent le critère d'information AIC. On calcule un critère AIC global égal à la somme des deux critères AIC déterminés pour chaque modèle.

Étape2. On cherche ici à identifier la valeur du seuil. On conserve fixée la valeur de d et

l'on fait varier la valeur de r . On répète la première étape pour les différentes valeurs de r et l'on retient comme valeur optimale du seuil celle qui minimise le critère AIC global.

Étape3. Cette étape a pour objet de déterminer la valeur optimale de d . À cette fin, on répète les deux étapes précédentes pour chaque valeur choisie de d et l'on retient comme valeur optimale du paramètre de délai celle qui minimise le critère AIC global.

Cette technique est très coûteuse en terme de temps de calcul. Il existe une autre procédure, plus directe, afin de choisir la variable de transition, et donc le paramètre de délai d . On utilise pour cela les résultats des tests de linéarité. L'idée générale sous-jacente est en effet que la linéarité sera d'autant plus fortement rejetée que le modèle *SETAR* (donc non linéaire) sera bien spécifier. En d'autres termes, le test de linéarité est d'autant plus puissant que la variable de transition est correctement choisie. Une stratégie consiste alors à appliquer les tests de linéarité pour plusieurs valeurs du paramètre de délai d ($1 \leq d \leq p$) et à retenir comme valeur de d celle pour laquelle la linéarité est la plus fortement rejetée (c'est à dire celle qui minimise la probabilité du test) [14].

2.2 Spécification d'un modèle SETAR

La spécification du modèle à seuil nécessite le choix de certains paramètres qui ne peuvent être estimés par des méthodes usuelles. Il s'agit du nombre de régresseurs ainsi que du paramètre définissant la variable de transition. Ils sont alors sélectionnés préalablement à l'estimation des autres paramètres du modèle. Pour faire ce choix on fait recours aux tests de linéarité contre l'alternative donnée par le modèle à seuil.

2.2.1 Choix des régresseurs

Le choix des régresseurs, c'est à dire des k variables explicatives et des p retards de l'endogène, du modèle à seuil s'effectue sur la base de critères d'information (*AIC*, *BIC*, *Hannan*), du test du portemanteau et/ ou test de significativité du retard le plus élevé présent dans la régression. Cette première étape est très importante puisque le modèle linéaire constitue l'hypothèse nulle du test de linéarité. Comme le fait remarquer Teräsvirta (1994) dans le cadre d'un modèle *STAR*, si le retard autorégressif est surestimé, cela peut avoir un effet

pervers sur la puissance du test (comparé au cas où p serait connu). Alternativement, si le retard sélectionné est trop faible et qu'il reste de l'autocorrélation résiduelle, les tests de linéarité sont biaisés (la linéarité est rejetée à tort).

2.2.2 Choix de la variable de transition

Le choix de la variable de transition, ou l'estimation du paramètre de délai dans un modèle *TAR*, ne relève pas des méthodes usuelles. En effet, la fonction de vraisemblance n'est pas dérivable par rapport à ce paramètre. Par conséquent, la variable de transition est choisie par une procédure particulière, basée sur les résultats des tests de linéarité. Sachant que cette variable est inconnue, on construit le test de linéarité pour toutes les variables de transition possibles i.e. pour l'ensemble des variables X , ou pour toutes les valeurs du paramètre de délai d telles que $1 \leq d \leq p$ dans le cas du modèle *TAR*. On retient ensuite la variable de transition pour laquelle la linéarité est la plus fortement rejetée, c'est -à-dire celle qui minimise la probabilité du test. En effet, le test de linéarité est d'autant plus puissant que la variable de transition est bien choisie.

Les méthodes de spécification du modèle à seuil proposées par Tsay (1989) et par Hansen (1996) incorporent un test de linéarité et une procédure de détection du seuil.

2.3 Estimation d'un modèle SETAR

L'estimation des modèles *SETAR* nécessite l'estimation de plusieurs paramètres : le nombre de régimes (c'est à dire le nombre de seuil), le paramètre de retard d , la valeur des seuil r_j , $j = 1, \dots, k - 1$, ainsi que les coefficients autorégressifs relatifs à chaque régime.

Selon [3], pour un ensemble donné d'observations $X_t = (Y_{t-1}, \dots, Y_{t-p})'$, et pour une valeur du paramètre de délai d et du seuil r , les estimateurs par moindres carrés ordinaires des coefficients autorégressifs $\phi_i^{(j)}$, $i = 0, \dots, p$, du régime j sont asymptotiquement normaux (Chan et Tong.1986). Ils minimisent la somme des carrés des résidus du régime j :

$$SSE^{(j)} = \sum_{i \in K(j)} (Y_t - \phi_0^{(j)} - \phi_i^{(j)} X_t)^2 \quad (2.2.2)$$

où $K(j)$ est l'ensemble $\{p < t \leq n : Z_t \text{ variable de transition appartenant au régime } j\}$.

L'estimateur de la matrice de variance-covariance des coefficients estimés est donné par :

$$\hat{V}^{(j)} = \hat{\sigma}_{(j)}^2 (X'X)^{-1} \quad (2.2.3)$$

où $X = (1, X_t')_{i \in K(j)}$, et

$$\hat{\sigma}_{(j)}^2 = \frac{SSE^{(j)}}{n_j - (p + 1)} \quad (2.2.4)$$

n_j est le nombre d'observations dans le régime j .

Dans cette section, on distingue deux grandes techniques de détection des modèles *SETAR* : la méthode de Tsay (1989) et la méthode de Hansen (1996). Quant à l'estimation des modèles *SETAR* par la méthode des moindres carrés ordinaires (*MCO*), elle sera traitée en détaille dans le chapitre suivant.

2.3.1 Méthodologie de Tsay (1989) et Hansen (1996)

Dans cette partie nous présentons la méthodologie proposée par Tsay, pour tester l'hypothèse de linéarité contre l'alternative fournie par le modèle à seuil, ainsi que la procédure graphique de détection du seuil.

Tsay (1989) ne traite pas le paramètre de seuil de la même manière que le paramètre définissant la variable de transition. Le paramètre de délai est fixé lors de la mise en oeuvre du test de linéarité, et, comme il a été mentionné précédemment, on retient la variable de transition pour laquelle la linéarité est la plus fortement rejetée.

L'originalité de la méthode de Tsay (1989) est de se débarrasser du paramètre de nuisance (le seuil), dans la logique des tests *CUSUM*, adaptés aux modèles à seuil par Petrucelli et Davies (1986). Tsay a réécrit l'alternative du test de linéarité en réarrangeant toutes les observations selon l'ordre croissant de la variable de transition. La régression ordonnée ainsi obtenue est constituée d'une première régression linéaire pour les r observations correspondant aux valeurs de la variable de transition inférieures au seuil. La deuxième régression correspond à l'autre régime (ou aux autres si le modèle est constitué de plus de deux régimes).

Procédure de spécification de Tsay (1989)

Afin de simplifier la présentation du test, considérons un processus $TAR(p)$ à deux régimes :

$$Y_t = \begin{cases} \phi_0^{(1)} + \sum_{i=1}^p \phi_i^{(1)} y_{t-i} + \varepsilon_t^{(1)} & \text{si } y_{t-d} \leq r \\ \phi_0^{(2)} + \sum_{i=1}^p \phi_i^{(2)} y_{t-i} + \varepsilon_t^{(2)}, & \text{si } y_{t-d} > r \end{cases} \quad (2.2.5)$$

soit encore :

$$Y_t = \begin{cases} \phi_0^{(1)} + \phi^{(1)} X_t + \varepsilon_t^{(1)} & \text{si } y_{t-d} \leq r \\ \phi_0^{(2)} + \phi^{(2)} X_t + \varepsilon_t^{(2)}, & \text{si } y_{t-d} > r \end{cases} \quad (2.2.6)$$

avec $X_t = (y_{t-1}, \dots, y_{t-p})'$, $\phi^{(j)} = (\phi_i^{(j)}, \dots, \phi_i^{(j)})$, pour $j = 1, 2$, et $\varepsilon_t, \varepsilon_t$ *iid* $(0, \sigma^2)$.

Étape 1. On estime un modèle autorégressif usuel d'ordre p sur la série et l'on se donne un ensemble de valeurs possibles pour le paramètre de retard d .

Étape 2. Cette étape consiste à effectuer le test de linéarité de Tsay (1989)

Ce test reprend la logique des tests de type *CUSUM* adaptés aux modèles à seuil par Petrucelli et Davies (1986). Pour mener à bien le test de linéarité, on ordonne les observations selon les valeurs croissantes de la variable de seuil. On obtient alors deux régressions : la première est relative aux r premières observations correspondant aux faibles valeurs de la variable de seuil, la seconde correspond à l'autre régime. Le modèle ordonné ainsi obtenu correspond au modèle (2.2.5) dans lequel le seuil se situe entre la $r^{\text{ème}}$ et la $(r + 1)^{\text{ème}}$ observation ordonnée et s'écrit :

$$Y_t = \begin{cases} \phi_0^{(1)} + \phi^{(1)} X_{(l)} + \varepsilon_{(l)}^{(1)}, & \text{pour les } r \text{ premières observations de } Z_L \\ \phi_0^{(2)} + \phi^{(2)} X_{(l)} + \varepsilon_{(l)}^{(2)}, & \text{pour les observations restantes} \end{cases} \quad (2.2.7)$$

où Z est la variable de seuil ($Z_t = y_{t-d}$) et L est l'ordre correspondant au classement des observations selon les valeurs croissantes de la variable de seuil ($Z_{(l)} \leq Z_{(l+1)}, (l) \in L$).

Cette régression ordonnée en (l) est estimée de façon récursive pour chaque valeur de d . Sur la base du modèle ordonné (2.2.7), l'hypothèse nulle testée est : $H_0 : \phi_i^{(1)} = \phi_i^{(2)}$ pour $i = 0, 1, \dots, p$.

La statistique de test de linéarité de Tsay (1989), notée $Q(p)$, est donnée par :

$$Q(p) = \frac{(\sum \hat{\varepsilon}_t^2 - \sum \hat{u}_t^2)}{\sum \hat{u}_t^2} \cdot \frac{N - b - 2p - 1}{p + 1}$$

où b est le nombre d'observations utilisées pour initialiser les estimations récursives, $\hat{\varepsilon}_t$ sont les résidus ε_t normalisés, et \hat{u}_t sont les résidus de la régression de $\hat{\varepsilon}_t$ sur $(1, X'_{(l)})$. Sous l'hypothèse nulle de linéarité, la statistique $Q(p)$ suit un Fisher à $(p + 1)$ et $(N - b - 2p - 1)$ degrés de liberté. Comme le conseille Tsay, on peut prendre $b = \frac{N}{10} + p$ dans le cas d'un

modèle autorégressif de type TAR . Dans le cas d'une régression, nous proposons $b = \text{Max}(\frac{N}{10}, p + 3)$.

La variable de transition choisie est celle qui maximise la statistique $Q(p)$ sachant que la linéarité est rejetée.

Étape 3. On cherche à déterminer la valeur du seuil de façon graphique. Une méthode couramment employée consiste à représenter graphiquement les t de Student des coefficients autorégressifs estimés de manière récursive et à repérer le moment où une rupture apparaît. L'idée sous-jacente est la suivante : dans le modèle ordonné (2.2.7), le processus associé aux r premières observations est linéaire. Dès que l'on introduit l'observation dont la valeur correspond au seuil, le processus devient une combinaison de deux régressions. La trajectoire des résidus récursifs et des autres statistiques (coefficients estimés, t de Student des coefficients estimés,...) va alors se déformer en passant par la valeur du seuil. On repère ainsi le seuil de façon graphique en portant sur l'axe des abscisses la variable de seuil ordonnée et sur l'axe des ordonnées la statistique considérée (notamment, les t de Student des coefficients autorégressifs estimés). Le seuil ainsi donné par la première rupture repérée.

Étape 4. On estime le modèle SETAR par les méthodes de régression usuelles, connaissant le seuil et le paramètre de délai.

Procédure de spécification de Hansen (1996)

Dans cette partie nous présentons la méthodologie proposée par Hansen, qui permet de tester l'hypothèse de linéarité contre l'alternative fournie par le modèle à seuil, ainsi que la procédure de sélection du seuil.

Face à la présence de paramètres de nuisance non identifiés sous l'hypothèse nulle de linéarité, Hansen (1996) construit le test de linéarité en considérant le paramètre qui définit la variable de transition (paramètre de délai dans le cas du modèle TAR) fixé, comme le fait Tsay. Mais contrairement à Tsay, il traite de la même façon le paramètre de seuil et le paramètre définissant la variable de transition. Ainsi, il ne réécrit pas l'alternative à la linéarité comme Tsay, mais construit un test qui dépend de ces deux paramètres.

Le test de linéarité qu'il propose est fonction de ces deux paramètres.

Étape 1. On estime un modèle autorégressif usuel d'ordre p sur la série et on note $\hat{\varepsilon}_t$ les résidus estimés. On se donne un ensemble de valeurs possibles pour le paramètre de délai d .

Étape 2. Pour chaque valeur du paramètre de délai on procède à un test de linéarité du type multiplicateur de Lagrange (LM). On calcul ainsi une séquence de statistiques $LM(r)$ pour toutes les valeurs possibles du seuil r , c'est à dire pour toutes les observations de la variable de seuil :

$$LM(r) = S(r)' \cdot I(r) \cdot S(r)$$

où $S(r) = \frac{1}{N} \sum_{i=1}^N X_t \hat{\varepsilon}_t(r)$ est le score du modèle à seuil estimé sous H_0 , et $I(r)$ est la matrice d'information de Fisher. On calcule les statistiques $SupLM(r)$, $\exp LM(r)$ et $moyLM(r)$; où $\exp LM(r)$ correspond à $\exp(\frac{1}{2}LM(r))$ et moy représente la moyenne des statistiques $LM(r)$.

Si l'hypothèse nulle de linéarité est rejetée, on poursuit la procédure et l'on retient comme valeur du paramètre de délai celle qui maximise les statistiques $SupLM(r)$, $\exp LM(r)$ et $moyLM(r)$.

Étape 3. La valeur du seuil est estimée en minimisant la variance résiduelle du modèle SETAR estimé parmi toutes les valeurs possibles du paramètre de délai.

Étape 4. On estime le modèle SETAR par les méthodes de régression usuelles, avec le seuil et le paramètre de délai connus.

2.4 Tests de linéarité

De nombreuses études récentes s'intéressent à détecter la non linéarité dans les séries économiques due à l'existence d'asymétries dans la dynamique des variables. Plusieurs procédures de tests ont été proposées pour prendre en compte ces caractéristiques.

Comme l'intérêt est autant de capturer des non-linéarités que de les caractériser et de les estimer, nous étudierons les tests de linéarité définis contre l'alternative donnée par les modèles à seuil. Il existe en effet deux possibilités pour classer les tests de linéarité, la première englobe les tests de linéarité qui ne reposent pas sur une alternative non-linéaire spécifiée. Le plus connu est le test *RESET*, pour Regression Error Specification Test. Dans ce cas, il est difficile de choisir un modèle pour les données étudiées quand l'hypothèse de linéarité est rejetée. Une seconde possibilité consiste à tester un modèle linéaire contre un modèle non linéaire donné. Dans ce cas on définit un ensemble de contraintes pour lesquelles le modèle devient linéaire. L'hypothèse de linéarité revient alors à tester cet ensemble de contraintes. Dans ce cas, si la linéarité est rejetée, on connaît l'alternative à considérer. Pour ces tests, la présence éventuelle de paramètres de nuisance non identifiables sous l'hypothèse nulle de linéarité conduit souvent à utiliser une approximation du modèle sous l'hypothèse alternative

Avant de procéder au test de linéarité, il faut vérifier qu'un modèle linéaire ne suffit pas à modéliser correctement la série en question, car il peut arriver (surtout si les séries temporelles sont courtes) que l'on estime -avec succès- un modèle non linéaire alors que la vraie relation sous-jacente est linéaire.

Il est important de noter également que le non-rejet de la linéarité ne signifie pas forcément que le processus est linéaire, l'alternative peut aussi être mal spécifiée.

2.4.1 Test Reset

Le test *RESET* ("*Regression Equation Specification Error Test*") a été proposé par Ramsey (1969). Il s'agit de tester un modèle linéaire contre un modèle augmenté de termes élevés à une puissance supérieure à deux. Ce qui revient à tester la nullité de ces derniers termes. L'inconvénient des tests sans alternative spécifiée se révèle dans le cas où la linéarité

est rejetée. On estime les paramètres du modèle linéaire :

$$(H_0) : X_t = \phi + \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t$$

et on calcule les résidus obtenus $\widehat{\varepsilon}_t$, les valeurs ajustées $\widehat{X}_t = \widehat{\phi} + \sum_{i=1}^p \widehat{\phi}_i X_{t-i} + \widehat{\varepsilon}_t = X_t - \widehat{\varepsilon}_t$ et

la somme des carrés des résidus $SSE_0 = \sum \widehat{\varepsilon}_t^2$. On estime alors, par la méthode des moindres carrés, les paramètres de :

$$\widehat{\varepsilon}_t = b_0 + \sum_{i=1}^p b_i X_{t-i} + \sum_{i=1}^p \theta_j \widehat{X}_{t-j} + \eta_t$$

et on calcule $SSE_1 = \sum \widehat{\eta}_t^2$. La statistique du test est :

$$RESET = \frac{(SSE_0 - SSE_1)/(h - 1)}{SSE_1/(n - (p + h))}$$

où $p + h$ est le nombre de paramètres du modèle de Ramsey augmenté, c'est à dire $p + 1$ paramètres du modèle d'origine plus les $h - 1$ variables rajoutées.

La statistique de ce test est un rapport de deux χ^2 indépendant. Donc, elle suit une loi de Fisher de degrés de liberté $(h - 1)$ et $(n - p - h)$.

2.4.2 Test du multiplicateur de Lagrange

Tester l'hypothèse de linéarité contre l'alternative fournie par les modèles à seuil revient à tester l'hypothèse nulle suivante : $H_0 : \phi_0^{(1)} = \phi_0^{(2)}$ et $\phi_i^{(1)} = \phi_i^{(2)}$ $i = 1, \dots, m$ contre l'alternative fournie par le modèle à seuil (2.2.1). Ainsi, le test requiert la spécification préalable du modèle linéaire, i.e. le choix des m régresseurs.

Il faut noter que, sous l'hypothèse nulle de linéarité, le paramètre de délai ou la variable de transition et le seuil ne sont pas identifiables. En effet, la valeur de la vraisemblance du modèle linéaire ne dépend pas de ces paramètres si on suppose que la variance des deux bruits blancs est la même dans les deux régimes. Ainsi, ils peuvent prendre n'importe quelles valeurs sans changer la fonction de vraisemblance. Les tests usuels d'un modèle contre un autre (statistiques de Wald, du ratio de vraisemblance ou du multiplicateur de lagrange) ne sont alors pas applicables. Face à ce problème, Tsay (1989) et Hansen (1996a) ont proposé

deux méthodes distinctes. Elles diffèrent notamment quant au traitement du seuil, la variable de transition étant traitée de la même manière. En effet, les deux procédures de tests sont développées pour une variable de transition choisie préalablement.

2.4.3 Test du CUSUM

Le test du *CUSUM* permet de détecter l'existence de régimes ou de seuil. Il consiste à reclasser les observations suivant les variables de transition. On estime alors les paramètres de manière récursive et on considère les sommes cummulées des résidus obtenus.

Dans le cas des modèles *SETAR*, on suppose que l'on connaît le délai d , et que l'on dispose des observations $X_{1-d}, \dots, X_0, \dots, X_T$. On ordonne les observations X_1, \dots, X_n selon les valeurs de X_{1-d}, \dots, X_{t-d} .

Soient $X_{(1)}^T, \dots, X_{(T)}^T$ les observations ainsi réordonnées. Pour tout j , on considère le modèle

linéaire suivant : $X_{(i)}^T = \phi_j + \sum_{k=1}^p \phi_k^j X_{(i)-k} + \varepsilon_{(j)}^T$ pour tout $i \leq j$

L'hypothèse (H_0) est alors $\phi_j = \phi$, $\phi_k^j = \phi_k$ pour tout k , et l'hypothèse alternative (H_1) est que (H_0) n'est pas vérifiée [9].

Chapitre 3

Estimation d'un modèle SETAR par la méthode MCO

La méthode des moindres carrés ordinaire (*MCO*) est une méthode d'estimation des paramètres visant à minimiser la somme des carrés des erreurs commises par le modèle sur l'ensemble des données.

Cette méthode d'estimation consiste à chercher les valeurs des paramètres qui minimisent le critère des moindres carré donné par $(Y - Xb)'(Y - Xb)$. Une technique de résolution consiste à dériver le critère par rapport aux paramètres et rechercher les solutions qui annulent ces dérivées. Dans le cas linéaire, les dérivées forment un système de n équations linéaires à n inconnues en b . La solution analytique existe, elle est donnée par $\hat{b} = (X' X)^{-1} X' Y$. Or, dans le cas des modèles non linéaires, le principe de minimisation des carrés des résidus peut encore être utilisé pour calculer les valeurs des paramètres du modèle. Toutefois, les dérivées partielles de la somme des carrés des résidus par rapport aux paramètres produisent un système d'équations non linéaires donc ils ne peuvent la plupart du temps se résoudre analytiquement. Quand les solutions analytiques des équations n'existent pas, minimiser la somme des carrés des résidus devient alors un problème d'optimisation qui ne peut être résolu que par des procédures numériques itératives comme la méthode de Gauss-Newton.

L'objet de ce chapitre est d'estimer les paramètres du modèle non linéaire à seuil *SETAR* par la méthode des moindres carrés ordinaire *MCO*. Sachant que la méthode des moindres carrés ordinaire n'est facilement applicable que dans le cas linéaire, nous avons exprimé le

modèle *SETAR* sous la forme d'un modèle de régression linéaire multiple. Nous commençons, tout d'abord, par l'estimation des paramètres du modèle *SETAR* par *MCO*, et par la suite, tester l'hypothèse de linéarité (modèle autorégressif) contre son alternative qui est le modèle *SETAR*.

3.1 Modèle Autorégressif à seuil SETAR et Modèle de régression linéaire

Nous avons vu qu'un modèle de séries chronologiques non linéaire de type Autorégressif à seuil *SETAR* (*Self-Exciting Threshold Autoregressive Model*), satisfaisant une équation aux différences stochastiques de la forme suivante :

$$Y_t = \begin{cases} \phi_0^{(1)} + \sum_{i=1}^p \phi_i^{(1)} Y_{t-i} + \varepsilon_t & \text{si } Y_{t-d} \leq r, \\ \phi_0^{(2)} + \sum_{i=1}^p \phi_i^{(2)} Y_{t-i} + \varepsilon_t & \text{si } Y_{t-d} > r, \end{cases} \quad (3.1.1)$$

avec p et d supposés connus, tels que $0 \leq d \leq p$.

peut s'écrire, par une opération de paramétrisation, sous le modèle linéaire suivant :

$$Y_t = \theta_0 + \theta_1 Y_{t-1} + \dots + \theta_p Y_{t-p} + \phi_0 I(Y_{t-d} \leq r) + \dots + \phi_q I(Y_{t-d} \leq r) Y_{t-q} + \varepsilon_t,$$

où les paramètres θ_i , $i = 0, 1, 2, \dots, p$ et ϕ_j , $j = 0, 1, 2, \dots, q$ sont donnés, en fonction des $\phi_i^{(1)}$,

$i = 1, 2, \dots, p$ et $\phi_j^{(2)}$, $j = 0, 1, 2, \dots, p$, comme suit :

$$Y_t = (\phi_0^{(1)} + \phi_1^{(1)} Y_{t-1} + \dots + \phi_p^{(1)} Y_{t-p}) I(Y_{t-d} \leq r) + (\phi_0^{(2)} + \phi_1^{(2)} Y_{t-1} + \dots + \phi_p^{(2)} Y_{t-p}).$$

$$(1 - I(Y_{t-d} \leq r)) + \varepsilon_t, \quad (3.1.2)$$

$$= \phi_0^{(2)} + \phi_1^{(2)} Y_{t-1} + \dots + \phi_p^{(2)} Y_{t-p} + (\phi_0^{(1)} - \phi_0^{(2)}) I(Y_{t-d} \leq r) + (\phi_1^{(1)} - \phi_1^{(2)}) I(Y_{t-d} \leq r) Y_{t-1} + \dots + (\phi_p^{(1)} - \phi_p^{(2)}) I(Y_{t-d} \leq r) Y_{t-p} + \varepsilon_t,$$

donc,

$$\begin{aligned} \theta_0 &= \phi_0^{(2)}, & \phi_0 &= \phi_0^{(1)} - \phi_0^{(2)} \\ \theta_1 &= \phi_1^{(2)}, & \phi_1 &= \phi_1^{(1)} - \phi_1^{(2)} \\ & \vdots & & \vdots \\ \theta_p &= \phi_p^{(2)}, & \phi_q &= \phi_p^{(1)} - \phi_p^{(2)} \end{aligned}$$

Ce dernier modèle peut s'écrire, encore, sous la forme d'un modèle de régression linéaire multiple, en les régresseurs,

$$1, Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}, I(Y_{t-d} \leq r), I(Y_{t-d} \leq r) Y_{t-1}, I(Y_{t-d} \leq r) Y_{t-2}, \dots, I(Y_{t-d} \leq r) Y_{t-q}$$

sous la forme suivante :

$$Y_t = b_0 + b_1 Y_{t-1} + \dots + b_p Y_{t-p} + b_{p+1} I(Y_{t-d} \leq r) + b_{p+2} I(Y_{t-d} \leq r) Y_{t-1} + \dots + b_{p+q+1} I(Y_{t-d} \leq r) Y_{t-q} + \varepsilon_t,$$

ou encore sous la forme équivalente suivante :

$$Y_t = b_0 + b_1 X_{1,t} + \dots + b_p X_{p,t} + b_{p+1} X_{p+1,t} + b_{p+2} X_{p+2,t} + \dots + b_{p+q+1} X_{p+q+1,t} + \varepsilon_t, \quad (3.1.3)$$

où les régresseurs $X_{i,t}$, $i = 1, 2, \dots, p + q + 1$ sont donnés par :

$$X_{i,t} = \begin{cases} Y_{t-i}, & i = 1, 2, \dots, p, \\ I(Y_{t-d} \leq r), & i = p + 1, \\ I(Y_{t-d} \leq r) Y_{t+p+1-i}, & i = p + 2, \dots, p + q + 1. \end{cases}$$

Le problème d'inférence statistique qui revient à tester un modèle autorégressif (AR) contre l'hypothèse alternative qui est le modèle *SETAR* (*Self-Exciting Threshold Autoregressive Model*) revient donc à tester la nullité des paramètres $b_{p+1}, b_{p+2}, \dots, b_{p+q+1}$.

C'est-à-dire on doit tester l'hypothèse nulle H_0 suivante :

$$H_0 : b_{p+1} = b_{p+2} = \dots = b_{p+q+1} = 0,$$

contre l'hypothèse alternative H_1 suivante :

$$H_1 : b_{p+i} \neq 0, i = 1, 2, \dots, q + 1.$$

Pour tester cette hypothèse, nous adoptons le test du rapport de vraisemblance sous l'hypothèse de normalité du processus de l'innovation ε_t .

Ce test exige, bien entendu, l'estimation des paramètres sous-jacents sous l'hypothèse nulle ainsi que sous l'hypothèse alternative. D'où la nécessité de l'obtention des estimations de la vraisemblance de ces paramètres sous chacune de ces deux hypothèses. Du fait que les estimateurs de la fonction de la vraisemblance, sous la normalité, sont eux mêmes les estimateurs de la *Méthode des Moindres Carrées*, alors nous procédons à l'obtention, sous chacune des deux hypothèses, des moindres carrées du modèle de régression linéaire multiple précédent.

Pour une série, de taille n , nous pouvons écrire :

$$Y_t = b_0 + b_1 X_{1,t} + \dots + b_p X_{p,t} + b_{p+1} X_{p+1,t} + \dots + b_{p+q+1} X_{p+q+1,t} + \varepsilon_t, \quad t = 1 \text{ à } n,$$

alors nous considérons les notations suivantes :

- $\underline{Y}' = (Y_1, Y_2, \dots, Y_t, \dots, Y_n)$, est le vecteur des n observations de la variable expliquée Y_t ,
- $\underline{\varepsilon}$ est un vecteur colonne des innovations défini par $\underline{\varepsilon}' = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t, \dots, \varepsilon_n)$
- \underline{b} est le vecteur colonne des paramètres de régression donné par $\underline{b}' = (b_0, b_1, \dots, b_p, b_{p+1}, \dots, b_{p+q+1})$

et

- \underline{X} est une matrice, de dimension, $n \times (p + q + 2)$, dite matrice des réalisations donnée par :

$$\underline{X} = \begin{pmatrix} 1 & X_{1,1} & X_{2,1} & \cdots & X_{p+q+1,1} \\ 1 & X_{1,2} & X_{2,2} & \cdots & X_{p+q+1,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1,n-1} & X_{2,n-1} & \cdots & X_{p+q+1,n-1} \\ 1 & X_{1,n} & X_{2,n} & \cdots & X_{p+q+1,n} \end{pmatrix}$$

En utilisant ces notations, on constate que le vecteur $y' = (y_1, \dots, y_n)$, est une réalisation du vecteur aléatoire Y satisfaisant le modèle linéaire :

$$Y = \underline{X} \underline{b} + \underline{\varepsilon}, \tag{3.1.4}$$

3.2 Estimations des moindres carrés d'un SETAR

L'estimateur $\hat{\underline{b}}$ des moindres carrés ordinaires du vecteur des paramètres de régression \underline{b} et ainsi que l'estimateur $\hat{\sigma}^2$ de la variance commune σ^2 de l'erreur seront obtenus sous le système d'hypothèses suivant :

$$(SH_1) : \begin{cases} H_0 : E(\underline{\varepsilon}) = 0; & \text{l'erreur est de moyenne nulle} \\ H_1 : E(\underline{\varepsilon} \underline{\varepsilon}') = I\sigma^2; & \text{les erreurs sont non corrélées de} \\ & \text{même variance} \end{cases} \tag{3.2.1}$$

qui est équivalent au système suivant :

$$(SH_2) : \begin{cases} H_0 : E(Y) = \underline{X} \underline{b}; & \text{la moyenne est égale à la fonction de} \\ & \text{régression} \\ H_1 : E(Y Y') = I \sigma^2; & \text{les variables endogènes sont} \\ & \text{non corrélées de même variance} \end{cases} \quad (3.2.2)$$

Rappelons que le principe de l'estimation par la méthode des moindres carrés consiste à trouver les estimations du vecteur des paramètres pour lesquelles la somme des erreurs au carrés $Q(\underline{b}/Y, \underline{X})$ suivante soit minimale :

$$\begin{aligned} Q(\underline{b}/Y, \underline{X}) &= \underline{\varepsilon} \underline{\varepsilon}' = (Y - \hat{Y})'(Y - \hat{Y}) = (Y - \underline{X} \underline{b})'(Y - \underline{X} \underline{b}); \\ &= Y'Y - 2\underline{b}' \underline{X}' Y + \underline{b}' (\underline{X}' \underline{X}) \underline{b} \end{aligned} \quad (3.2.3)$$

Alors, en dérivant partiellement, par rapport à \underline{b} , et en égalant à zéro, on obtient l'équation matricielle suivante dite équation normale :

$$(\underline{X}' \underline{X}) \underline{b} = \underline{X}' Y, \quad (3.2.4)$$

L'existence d'une solution unique à cette équation, autrement dit, l'existence de l'estimateur des moindres carrés ordinaires, dépend du rang de la matrice \underline{X} . En effet, si cette matrice n'est pas de rang plein alors la matrice carrée $\underline{X}' \underline{X}$ sera singulière et donc l'équation normale possède plus qu'une solution, en \underline{b} .

Donc, dans toute la suite, on suppose, sauf mention contraire, que la matrice \underline{X} est de rang plein et donc la matrice $\underline{X}' \underline{X}$ est inversible. Cette dernière matrice carrée est de la forme :

$$\underline{X}' \underline{X} = \begin{pmatrix} n & n\bar{X}_1 & n\bar{X}_2 & \cdots & n\bar{X}_{p+q+1} \\ n\bar{X}_1 & \sum_{i=1}^n X_{1,i}^2 & \sum_{i=1}^n X_{1,i} X_{2,i} & \cdots & \sum_{i=1}^n X_{1,i} X_{p+q+1,i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n\bar{X}_{p+q} & \sum_{i=1}^n X_{p+q,i} X_{1,i} & \sum_{i=1}^n X_{p+q,i} X_{2,i} & \cdots & \sum_{i=1}^n X_{p+q+1,i} X_{p+q+1,i} \\ n\bar{X}_{p+q+1} & \sum_{i=1}^n X_{p+q+1,i} X_{1,i} & \sum_{i=1}^n X_{p+q+1,i} X_{2,i} & \cdots & \sum_{i=1}^n X_{p+q+1,i}^2 \end{pmatrix}$$

Sous cette condition l'équation normale (3.2.4) possède une solution unique donnée par :

$$\hat{\underline{b}} = (\underline{X}' \underline{X})^{-1} \underline{X}' Y, \quad (3.2.5)$$

On montre facilement que la fonction objective à savoir la forme quadratique $Q(\underline{b}/Y, \underline{X})$

atteint vraiment sa valeur minimale pour cette solution. En effet, la somme des erreurs au carrés peut s'écrire sous la forme suivante :

$$\begin{aligned}
 Q(\underline{b}/Y, \underline{X}) &= (Y - \underline{X} \underline{b})'(Y - \underline{X} \underline{b}); \\
 &= \left((Y - \underline{X} \widehat{\underline{b}}) - \underline{X} (\underline{b} - \widehat{\underline{b}}) \right)' \left((Y - \underline{X} \widehat{\underline{b}}) - \underline{X} (\underline{b} - \widehat{\underline{b}}) \right); \\
 &= (Y - \underline{X} \widehat{\underline{b}})'(Y - \underline{X} \widehat{\underline{b}}) + (\underline{b} - \widehat{\underline{b}})' \underline{X}' \underline{X} (\underline{b} - \widehat{\underline{b}}) - 2(Y - \underline{X} \widehat{\underline{b}})' \underline{X} (\underline{b} - \widehat{\underline{b}}),
 \end{aligned}$$

On remarque que le troisième terme de l'expression précédente est nul, car on a

$$(Y - \underline{X} \widehat{\underline{b}})' \underline{X} = (\underline{X}' Y - \underline{X}' \underline{X} \widehat{\underline{b}})' = (\underline{X}' Y - \underline{X}' Y)' = 0,$$

alors l'expression de la somme des erreurs au carrés s'écrit comme suit

$$Q(\underline{b}/Y, \underline{X}) = (Y - \underline{X} \widehat{\underline{b}})'(Y - \underline{X} \widehat{\underline{b}}) + (\underline{b} - \widehat{\underline{b}})' \underline{X}' \underline{X} (\underline{b} - \widehat{\underline{b}}),$$

c'est une somme de deux quantités positives ou nulles, la première ne dépend pas du vecteur \underline{b} et la deuxième prend sa valeur minimale pour $\underline{b} = \widehat{\underline{b}}$, ce qui achève la preuve.

Après avoir présenter la méthode d'estimation des paramètres par moindres carrés dans le cas d'un modèle linéaire, en général, revenons à notre cas où le modèle est non linéaire à seuil *SETAR*. En effet, dans le cas des modèles *SETAR* le principe de la méthode des moindres carrés reste le même par le fait que ce modèle (*SETAR*) a été réécrit sous la forme d'un modèle linéaire (3.1.4). Donc, trouver les estimateurs des moindres carrés d'un modèle *SETAR* revient à minimiser la somme des erreurs au carré, notée $Q(\underline{b}/Y, \underline{X})$, correspondante. La seule différence qui s'impose est la forme de la matrice des réalisations \underline{X} . En effet, cette matrice devient la matrice des réalisations du modèle *SETAR*, définie par :

$$\underline{X}_{n \times (p+q+2)} = \begin{pmatrix} 1 & X_{1,1} & X_{2,1} & \cdots & X_{p,1} & X_{p+1,1} & X_{p+2,1} & \cdots & X_{p+q+1,1} \\ 1 & X_{1,2} & X_{2,2} & \cdots & X_{p,2} & X_{p+1,2} & X_{p+2,2} & \cdots & X_{p+q+1,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1,n-1} & X_{2,n-1} & \cdots & X_{p,n-1} & X_{p+1,n-1} & X_{p+2,n-1} & \cdots & X_{p+q+1,n-1} \\ 1 & X_{1,n} & X_{2,n} & \cdots & X_{p,n} & X_{p+1,n} & X_{p+2,n} & \cdots & X_{p+q+1,n} \end{pmatrix} \quad (3.2.6)$$

A partir de cette représentation, on s'aperçoit que la matrice \underline{X} , de dimension $n \times (p + q + 2)$, est composée de deux blocs, à savoir $(\underline{X}_1, \underline{X}_2)$, où \underline{X}_1 est représentée par les $(p + 1)^{\text{ème}}$ colonnes de la matrice (3.2.6), et \underline{X}_2 est constituée des $(q + 1)^{\text{ème}}$ dernières

colonnes, données par :

$$\underline{X}_1 = \begin{pmatrix} 1 & X_{1,1} & X_{2,1} & \cdots & X_{p,1} \\ 1 & X_{1,2} & X_{2,2} & \cdots & X_{p,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1,n-1} & X_{2,n-1} & \cdots & X_{p,n-1} \\ 1 & X_{1,n} & X_{2,n} & \cdots & X_{p,n} \end{pmatrix} \quad (3.2.7)$$

et

$$\underline{X}_2 = \begin{pmatrix} X_{p+1,1} & X_{p+2,1} & \cdots & X_{p+q+1,1} \\ X_{p+1,2} & X_{p+2,2} & \cdots & X_{p+q+1,2} \\ \vdots & \vdots & \vdots & \vdots \\ X_{p+1,n-1} & X_{p+2,n-1} & \cdots & X_{p+q+1,n-1} \\ X_{p+1,n} & X_{p+2,n} & \cdots & X_{p+q+1,n} \end{pmatrix} \quad (3.2.8)$$

Par conséquent, le vecteur des paramètres de régression \underline{b} est aussi composé de deux blocs, de la forme $\underline{b}' = (b'_1, b'_2)$ avec :

$$b'_1 = (b_0, b_1, \dots, b_p)'$$

$$b'_2 = (b_{p+1}, b_{p+2}, \dots, b_{p+q+1})'$$

Par la suite, on déduit facilement la matrice carrée $\underline{X}'\underline{X}$. Cette matrice est de la forme :

$$\underline{X}'\underline{X} = \begin{pmatrix} \underline{X}'_1\underline{X}_1 & \underline{X}'_1\underline{X}_2 \\ \underline{X}'_2\underline{X}_1 & \underline{X}'_2\underline{X}_2 \end{pmatrix}; \text{ donnée explicitement par :}$$

$$\underline{X}'\underline{X} =$$

$$\begin{pmatrix} n & \cdots & \sum_{i=p}^n Y_{i-p} & \sum_{i=p}^n I_r(Y_{i-d}) & \cdots & \sum_{i=p}^n I_r(Y_{i-d})Y_{i-q} \\ \sum_{i=p}^n Y_{i-1} & \cdots & \sum_{i=p}^n Y_{i-1}Y_{i-p} & \sum_{i=p}^n I_r(Y_{i-d})Y_{i-1} & \cdots & \sum_{i=p}^n I_r(Y_{i-d})Y_{i-1}Y_{i-q} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{i=p}^n Y_{i-p} & \cdots & \sum_{i=p}^n Y_{i-p}^2 & \sum_{i=p}^n I_r(Y_{i-d})Y_{i-p} & \cdots & \sum_{i=p}^n I_r(Y_{i-d})Y_{i-p}Y_{i-q} \\ \sum_{i=p}^n I_r(Y_{i-d}) & \cdots & \sum_{i=p}^n I_r(Y_{i-d})Y_{i-p} & \sum_{i=p}^n I_r^2(Y_{i-d}) & \cdots & \sum_{i=p}^n I_r^2(Y_{i-d})Y_{i-q} \\ \sum_{i=p}^n I_r(Y_{i-d})Y_{i-1} & \cdots & \sum_{i=p}^n I_r(Y_{i-d})Y_{i-1}Y_{i-p} & \sum_{i=p}^n I_r^2(Y_{i-d})Y_{i-1} & \cdots & \sum_{i=p}^n I_r^2(Y_{i-d})Y_{i-1}Y_{i-q} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{i=p}^n I_r(Y_{i-d})Y_{i-q} & \cdots & \sum_{i=p}^n I_r(Y_{i-d})Y_{i-q}Y_{i-p} & \sum_{i=p}^n I_r^2(Y_{i-d})Y_{i-q} & \cdots & \sum_{i=p}^n I_r^2(Y_{i-d})Y_{i-q}^2 \end{pmatrix}.$$

Or, le calcul de l'estimateur des moindres carrés du vecteur des paramètres \underline{b} , pour un modèle *SETAR*, noté $\widehat{\underline{b}}$, exige le calcul de la matrice $(\underline{X}'\underline{X})^{-1}$. Pour cela, nous énonçons le théorème (1), en annexe, qui établit l'inverse d'une matrice. Selon ce théorème, la matrice $\underline{X}'\underline{X}$ étant une matrice symétrique non singulière (régulière) et $\underline{X}'_1\underline{X}_1$ est bien une sous matrice inversible, par conséquent, $\underline{X}'\underline{X}$ est donnée sous la forme suivante :

$$(\underline{X}'\underline{X})^{-1} = \begin{pmatrix} (\underline{X}'_1\underline{X}_1)^{-1} + (\underline{X}'_1\underline{X}_1)^{-1}(\underline{X}'_1\underline{X}_2) \cdot V^{22} \cdot \underline{X}'_2\underline{X}_1(\underline{X}'_1\underline{X}_1)^{-1} & -(\underline{X}'_1\underline{X}_1)^{-1}(\underline{X}'_1\underline{X}_2)V^{22} \\ -V^{22}(\underline{X}'_2\underline{X}_1) \cdot (\underline{X}'_1\underline{X}_1)^{-1} & V^{22} \end{pmatrix} \quad (3.2.9)$$

où $V^{22} = [(\underline{X}'_2\underline{X}_2) - (\underline{X}'_2\underline{X}_1)(\underline{X}'_1\underline{X}_1)^{-1}(\underline{X}'_1\underline{X}_2)]^{-1}$

Donc, l'estimateur des moindres carrés ordinaires du vecteur des paramètres \underline{b} , soit $\widehat{\underline{b}}$ est donné par :

$$\begin{aligned} \widehat{\underline{b}} &= (\underline{X}'\underline{X})^{-1} \underline{X}'Y \\ &= \begin{pmatrix} (\underline{X}'_1\underline{X}_1)^{-1} + (\underline{X}'_1\underline{X}_1)^{-1}(\underline{X}'_1\underline{X}_2) \cdot V^{22} \cdot \underline{X}'_2\underline{X}_1(\underline{X}'_1\underline{X}_1)^{-1} & -(\underline{X}'_1\underline{X}_1)^{-1}(\underline{X}'_1\underline{X}_2)V^{22} \\ -V^{22}(\underline{X}'_2\underline{X}_1) \cdot (\underline{X}'_1\underline{X}_1)^{-1} & V^{22} \end{pmatrix} \begin{pmatrix} \underline{X}'_1 \\ \underline{X}'_2 \end{pmatrix} \cdot Y \\ &= \begin{pmatrix} (\underline{X}'_1\underline{X}_1)^{-1}\underline{X}_1 + (\underline{X}'_1\underline{X}_1)^{-1}(\underline{X}'_1\underline{X}_2) \cdot V^{22} \cdot \underline{X}'_2\underline{X}_1(\underline{X}'_1\underline{X}_1)^{-1}\underline{X}_1 - (\underline{X}'_1\underline{X}_1)^{-1}(\underline{X}'_1\underline{X}_2)V^{22}\underline{X}_1 \\ -V^{22}(\underline{X}'_2\underline{X}_1) \cdot (\underline{X}'_1\underline{X}_1)^{-1}\underline{X}_1 + V^{22}\underline{X}_2 \end{pmatrix} \cdot Y \end{aligned}$$

Notons $\widehat{\underline{b}} = \begin{pmatrix} \widehat{b}_1 \\ \widehat{b}_2 \end{pmatrix}$, avec :

$$\widehat{b}_1 = [(\underline{X}'_1\underline{X}_1)^{-1}\underline{X}_1 + (\underline{X}'_1\underline{X}_1)^{-1}(\underline{X}'_1\underline{X}_2) \cdot V^{22} \cdot \underline{X}'_2\underline{X}_1(\underline{X}'_1\underline{X}_1)^{-1}\underline{X}_1 - (\underline{X}'_1\underline{X}_1)^{-1}(\underline{X}'_1\underline{X}_2)V^{22}\underline{X}_1] \cdot Y$$

et

$$\widehat{b}_2 = [-V^{22}(\underline{X}'_2\underline{X}_1) \cdot (\underline{X}'_1\underline{X}_1)^{-1}\underline{X}_1 + V^{22}\underline{X}_2] \cdot Y$$

3.3 Estimation et Test sous Contraintes Linéaires

Souvent lorsque nous devons faire une étude, c'est que nous voulions confronté une théorie à la réalité, et dans la plupart du temps, nous sommes intéressés à savoir si une variable influence une autre ou pas. Pour répondre à ce genre de question, nous devons faire des tests d'hypothèses. Jusqu'à présent, on supposait que dans le modèle linéaire, b pouvait

prendre n'importe quelle valeur dans \mathbb{R}^k . Or on peut se demander s'il n'existe pas un ou plusieurs contraintes linéaires liant les coefficients b_j entre eux. La considération de cette relation fonctionnelle (généralement de type linéaire) entre quelques paramètres d'un modèle de régression linéaire multiple, se traduit par l'incorporation de cette contrainte ou cette hypothèse dans la définition du modèle de régression.

Une hypothèse linéaire (en les composantes du vecteur des paramètres) générale peut s'écrire sous la forme :

$$H : C' \underline{b} = \beta, \tag{3.3.1}$$

où C' est une matrice de rang plein connue, de dimension $(q+1) \times (p+q+2)$, ($q+1 \leq p+q+2$) et β un vecteur colonne, connu également, de dimension $(q+1)$.

Dans toute la suite, on désigne par SHP_2 (respectivement SHP_1) le système d'hypothèses qui est l'union du système d'hypothèses habituelles SH et de la nouvelle hypothèse linéaire ou contrainte linéaire, à savoir $H : C' \underline{b} = \beta$, imposée sur les paramètres du modèle de régression linéaire multiple. Plus précisément, le système d'hypothèses SHP_2 peut s'écrire sous la forme :

$$(SHP_2) : \begin{cases} H_0 : E(Y) = \underline{X} \underline{b}, & \text{la moyenne est égale à la fonction de régression} \\ H_1 : Var(Y) = I \sigma^2, & \text{non corrélées de même variance} \\ H : C' \underline{b} = \beta, & \text{contrainte linéaire imposée sur les paramètres} \end{cases}$$

Si de plus on ajoute l'hypothèse de normalité de la variable aléatoire expliquée, le système d'hypothèses SHP_2 (respectivement SHP_1) sera noté SHP_2^* (respectivement SHP_1^*). Le système d'hypothèse SHP_2^* dont on fera appelle, ultérieurement, en particulier à la section des tests des hypothèses linéaires s'écrit sous la forme suivante :

$$(SHP_2^*) : \begin{cases} H_0 : E(Y) = \underline{X} \underline{b}, & \text{la moyenne est égale à la fonction de régression} \\ H_1 : Var(Y) = I \sigma^2, & \text{non corrélées de même variance de loi normale} \\ H : C' \underline{b} = \beta, & \text{contrainte linéaire imposée sur les paramètres} \end{cases}$$

3.4 Estimation des paramètres d'un modèle SETAR sous contraintes linéaires

Considérons le modèle statistique non paramétrique de régression linéaire sous sa forme matricielle (3.1.4), dont on veut d'une part, obtenir les estimateurs des moindres carrés, de ses paramètres, à savoir les composantes du vecteur \underline{b} , des paramètres de régression, et la variance des erreurs σ^2 et d'autre part étudier leurs propriétés.

Les estimateurs des moindres carrés du modèle réduit (3.1.4) sont obtenus en minimisant l'erreur quadratique, par rapport aux composantes du vecteurs des paramètres \underline{b} , sous la contrainte (3.3.1) comme suit :

$$\min_{\underline{b}} Q(\underline{b}/Y, \underline{X}),$$

sous la contrainte $C' \underline{b} = \beta$

où $Q(\underline{b}/Y, \underline{X})$ est la forme quadratique habituelle, i.e :

$$\begin{aligned} Q(\underline{b}/Y, \underline{X}) &= \underline{\varepsilon} \underline{\varepsilon}' = (Y - \underline{X} \underline{b})' (Y - \underline{X} \underline{b}); \\ &= Y' Y - 2 \underline{b}' \underline{X}' Y - \underline{b}' (\underline{X}' \underline{X}) \underline{b}; \end{aligned}$$

Rappelons que notre objectif est de tester l'hypothèse nulle de linéarité $H_0 : b_{p+1} = b_{p+2} = \dots = b_{p+q+1} = 0$, contre l'hypothèse alternative $H_1 : b_{p+i} \neq 0, i = 1, \dots, q + 1$, ce qui revient à tester un modèle *AR* contre un modèle *SETAR* sous la contrainte linéaire : $C' \underline{b} = \underline{0}$, tel que :

$$C' \underline{b} = \underline{0} \iff \begin{pmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & 0 & 0 \\ \vdots & \dots & \vdots & \vdots & 0 & \dots & \vdots \\ \vdots & \dots & \vdots & \vdots & \vdots & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} b_0 \\ \vdots \\ b_p \\ b_{p+1} \\ \vdots \\ b_{p+q+1} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

où $C'_{(q+1) \times (p+q+2)} = (\underline{0}', I')$ et $\underline{b}'_{(p+q+2) \times 1} = (b_0, \dots, b_p, b_{p+1}, \dots, b_{p+q+1})$ avec :

$\underline{0}'$: matrice nulle de dimension $(q + 1) \times (p + 1)$

I' : matrice identité d'ordre $(q + 1)$

Pour résoudre le problème de minimisation de la forme quadratique $Q(\underline{b}/Y, \underline{X})$ sous la contrainte linéaire, on utilise la méthode de Lagrange. À cet effet, considérons le vecteur colonne $2\underline{\theta}$, de dimension $(q+1)$, (le réel 2 est pour une raison de commodité de calcul), alors le problème revient à minimiser, par rapport à \underline{b} et $\underline{\theta}$, la fonction objective suivante :

$$Q(\underline{b}/Y, \underline{X})_1 = Y' Y - 2 \underline{b}' \underline{X}' Y - \underline{b}' (\underline{X}' \underline{X}) \underline{b} + 2 \underline{\theta}' (C' \underline{b}), \quad (3.3.3)$$

par conséquent les estimations des moindres carrés, sous la contrainte (3.3.1), de \underline{b} notée $\widehat{\underline{b}}_H$, s'obtient en résolvant les équations, en $\widehat{\underline{b}}_H$ et $\widetilde{\underline{\theta}}$, obtenues en égalant à zéro les dérivées partielles de $Q(\underline{b}/Y, \underline{X})_1$, par rapport à \underline{b} et $\underline{\theta}$ (c'est la technique des multiplicateurs de Lagrange) :

$$\begin{aligned} 2 \underline{X}' Y - 2 (\underline{X}' \underline{X}) \widehat{\underline{b}}_H + 2 \widetilde{\underline{\theta}} &= 0 \\ C' \widehat{\underline{b}}_H &= \underline{0} \end{aligned}$$

Comme la matrice $\underline{X}' \underline{X}$ est supposée de rang plein, on obtient l'équation suivante :

$$\widehat{\underline{b}}_H = \widehat{\underline{b}} + (\underline{X}' \underline{X})^{-1} C' \widetilde{\underline{\theta}}, \quad (3.3.4)$$

où $\widehat{\underline{b}}$ désigne le vecteur des estimations habituels des moindres carrés (c'est-à-dire $\widehat{\underline{b}} = (\underline{X}' \underline{X})^{-1} \underline{X}' Y$)

En multipliant l'équation (3.3.4) par C' , on obtient l'équation suivante, en $\widetilde{\underline{\theta}}$:

$$\left(C' (\underline{X}' \underline{X})^{-1} C \right) \widetilde{\underline{\theta}} = C' (\widehat{\underline{b}}_H - \widehat{\underline{b}}),$$

Sachant que la matrice C' est de rang plein (ce qui implique que la matrice $C' (\underline{X}' \underline{X})^{-1} C$ est inversible), on obtient :

$$\begin{aligned} \widetilde{\underline{\theta}} &= \left(C' (\underline{X}' \underline{X})^{-1} C \right)^{-1} C' (\widehat{\underline{b}}_H - \widehat{\underline{b}}); \\ &= - \left(C' (\underline{X}' \underline{X})^{-1} C \right)^{-1} (C' \widehat{\underline{b}}); \quad (\text{car } C' \widehat{\underline{b}}_H = \underline{0}) \end{aligned}$$

et enfin en substituant dans (3.3.4) on obtient l'estimateur linéaire suivant :

$$\widehat{\underline{b}}_H = \widehat{\underline{b}} - (\underline{X}' \underline{X})^{-1} \left(C' (\underline{X}' \underline{X})^{-1} C \right)^{-1} (C' \widehat{\underline{b}}), \quad (3.3.6)$$

On peut facilement voir que cet estimateur est sans biais pour \underline{b} , et est donné par :

$$\widehat{\underline{b}}_H = \widehat{\underline{b}} - (\underline{X}' \underline{X})^{-1} \left(C' (\underline{X}' \underline{X})^{-1} C \right)^{-1} (C' \widehat{\underline{b}})$$

on procède, tout d'abord, au calcul de la quantité $\left(C' (\underline{X}' \underline{X})^{-1} C \right)^{-1}$:

$$\left(C' (\underline{X}' \underline{X})^{-1} C \right)^{-1} = \left[\begin{pmatrix} 0' & I' \end{pmatrix} (\underline{X}' \underline{X})^{-1} \cdot \begin{pmatrix} 0 \\ I \end{pmatrix} \right]^{-1} = (V^{22})^{-1}$$

d'où,

$$\widehat{\underline{b}}_H = \begin{pmatrix} \widehat{\underline{b}}_1 - (\underline{X}'_1 \underline{X}_1)^{-1} (\underline{X}'_1 \underline{X}_2) \widehat{\underline{b}}_2 \\ \underline{0} \end{pmatrix}$$

par conséquent, le vecteur prédicteur \widetilde{Y} , des valeur moyennes prédites \widetilde{y} , du vecteur aléatoire à expliquer Y est donné par :

$$\begin{aligned} \widetilde{Y} &= \underline{X} \widehat{\underline{b}}_H, \\ &= (\underline{X}_1, \underline{X}_2) \cdot \begin{pmatrix} \widehat{\underline{b}}_1 - (\underline{X}'_1 \underline{X}_1)^{-1} (\underline{X}'_1 \underline{X}_2) \widehat{\underline{b}}_2 \\ \underline{0} \end{pmatrix} \end{aligned}$$

d'où,

$$\widetilde{Y} = \underline{X}_1 \widehat{\underline{b}}_1 - \underline{X}_1 (\underline{X}'_1 \underline{X}_1)^{-1} (\underline{X}'_1 \underline{X}_2) \widehat{\underline{b}}_2 \quad (3.3.7)$$

il s'en suit alors que la somme des carrés des résidus notée, comme précédemment, SSE_H est donnée par :

$$\begin{aligned} SSE_H &= (Y - \widetilde{Y})' (Y - \widetilde{Y}); \\ &= \left((Y - \underline{X} \widehat{\underline{b}}) - \underline{X} M \widehat{\underline{b}} \right)' \left((Y - \underline{X} \widehat{\underline{b}}) - \underline{X} M \widehat{\underline{b}} \right), \end{aligned}$$

où la matrice M , de dimension $(p+q+2) \times (q+1)$, où $(q+1 \leq p+q+2)$, est donnée par :

$$M = (\underline{X}' \underline{X})^{-1} C \left(C' (\underline{X}' \underline{X})^{-1} C \right)^{-1}.$$

Après quelques manipulations simples, cette statistique peut s'écrire sous la forme :

$$SSE_H = (Y - \underline{X} \widehat{\underline{b}})' (Y - \underline{X} \widehat{\underline{b}}) - 2 (Y - \underline{X} \widehat{\underline{b}})' \underline{X} M + M' (\underline{X}' \underline{X})^{-1} M,$$

et du fait que

$$(Y - \underline{X} \widehat{\underline{b}})' \underline{X} = Y' \left(I - \underline{X} (\underline{X}' \underline{X})^{-1} \underline{X}' \right) \underline{X} = 0,$$

la forme quadratique précédente peut s'écrire sous la forme suivante :

$$\begin{aligned}
 SSE_H &= (Y - \underline{X}\hat{\underline{b}})' (Y - \underline{X}\hat{\underline{b}}) + (C'\hat{\underline{b}})' (C'(\underline{X}'\underline{X})^{-1}C)^{-1} (C'\hat{\underline{b}}); \\
 &= SSE + Q(\underline{b}/Y, \underline{X})_H,
 \end{aligned} \tag{3.3.8}$$

où SSE est la somme habituelle des carrés des résidus et la forme quadratique $Q(\underline{b}/Y, \underline{X})_H$ est donnée par :

$$\begin{aligned}
 Q(\underline{b}/Y, \underline{X})_H &= (C'\hat{\underline{b}})' (C'(\underline{X}'\underline{X})^{-1}C)^{-1} (C'\hat{\underline{b}}) \\
 &= \tilde{\underline{b}}_2' (V^{22})^{-1} \cdot \hat{\underline{b}}_2
 \end{aligned}$$

De l'expression (3.3.8) on voit clairement que la contrainte linéaire $C'\underline{b} = 0$ qui permet, d'une part, d'avoir un modèle réduit entraîne, d'autre part, une pénalité concernant l'augmentation de la somme des erreurs au carrés de la quantité positive $Q(\underline{b}/Y, \underline{X})_H$. De ce fait, on constate qu'une valeur assez grande de $Q(\underline{b}/Y, \underline{X})_H$ n'est pas à la faveur de l'hypothèse linéaire. Cependant, on a aucune idée sur la grandeur de $Q(\underline{b}/Y, \underline{X})_H$ à partir de laquelle on peut dire que $Q(\underline{b}/Y, \underline{X})_H$ est ou non à la faveur de l'hypothèse en question.

Néanmoins, on verra, ultérieurement que sous l'hypothèse de normalité, basé essentiellement sur cette quantité positive, on peut déterminer, à un seuil de signification fixé, la valeur maximale de $Q(\underline{b}/Y, \underline{X})_H$ pour laquelle elle soit à la faveur de l'hypothèse.

3.5 Estimation sous l'hypothèse de normalité

3.5.1 Estimateurs du maximum de vraisemblance

Sous le système d'hypothèses SHP_2^* , et la contrainte linéaire (3.3.1), l'estimateur du maximum de vraisemblance $\tilde{\underline{b}}_H$ est lui même l'estimateur linéaire sans biais des moindres carrés, sous la même contrainte linéaire. Cependant, l'estimateur du maximum de vraisemblance de la variance est biaisé et il est donné par :

$$\tilde{\sigma}^2 = \frac{(Y - \underline{X}\hat{\underline{b}})' (Y - \underline{X}\hat{\underline{b}})}{n} = \frac{SSE + Q(\underline{b}/Y, \underline{X})_H}{n}, \quad (3.4.1)$$

En effet, les estimations du maximum de vraisemblance de \underline{b} et σ^2 s'obtiennent en maximisant la fonction de vraisemblance

$$f(\underline{y}; \underline{b}, \sigma^2) = (2\pi\sigma^2)^{-n/2} \underline{\varepsilon}^{-\frac{1}{2\sigma^2}(\underline{y}-\underline{X} \underline{b})'(\underline{y}-\underline{X} \underline{b})},$$

sous la contrainte linéaire (3.3.1). Ce qui revient à minimiser la fonction

$$\begin{aligned} Q(\underline{b}/Y, \underline{X})_1 &= -\frac{1}{2\sigma^2}(\underline{y} - \underline{X} \underline{b})'(\underline{y} - \underline{X} \underline{b}) \\ &= -\frac{n}{2}\sigma^2 + \frac{1}{\sigma^2} Q(\underline{b}/Y, \underline{X}) \end{aligned}$$

sous la contrainte linéaire $C' \underline{b} = 0$. Ce problème peut être résolu par la technique des multiplicateurs de Lagrange. Ceci revient alors à résoudre le système d'équations (trois équations) obtenues en égalant à zéro les dérivés partielles, par rapport à \underline{b} , $\underline{\theta}$ et σ^2 , de la fonction suivante :

$$-\frac{n}{\sigma^2} + \frac{1}{\sigma^2} (Y' Y - 2 \underline{b}' \underline{X} Y - \underline{b}' (\underline{X}' \underline{X}) \underline{b} + 2 \underline{\theta}' C' \underline{b}).$$

Il est clair que les deux premières équations établies, sont celles (à un facteur près) obtenues lors de l'estimation par la méthodes des moindres carrés, d'où on obtient la même estimation $\tilde{\underline{b}}_H$ de \underline{b} et la troisième donne l'estimation de σ^2 .

3.5.2 Loi de probabilité de $Q(\underline{b}/Y, \underline{X})_H$

Pour le rôle important que joue la variable aléatoire $\frac{Q(\underline{b}/Y, \underline{X})_H}{\sigma^2}$ dans le problème de tests d'hypothèses linéaires dans les modèles de régression linéaires simples ou multiples sous le système d'hypothèses SHP_2^* , il s'avère utile d'établir, dès maintenant, sa loi de probabilité.

Propriété 1

Sous le système d'hypothèses SHP_2^* la statistique quadratique $Q(\underline{b}/Y, \underline{X})_H$ est indépendante de la statistique quadratique SSE et, de plus, la variable aléatoire $\frac{Q(\underline{b}/Y, \underline{X})_H}{\sigma^2}$ suit, une loi de khi-deux (centrée) à un $(q + 1)$ degrés de liberté, i. e.,

$$\frac{Q(\underline{b}/Y, \underline{X})_H}{\sigma^2} \sim \chi_{q+1}^2.$$

La preuve de cette propriété est donnée en annexe.

Propriété 2

La statistique $\frac{n-p-q-2}{q+1} \frac{Q(\underline{b}/Y, \underline{X})_H}{SSE}$ **suit, sous la condition de normalité et sous l'hypothèse nulle** H , **une loi de Fischer centrée à** $(q+1)$ **et** $(n-p-q-2)$ **degrés de liberté et suit, toujours sous l'hypothèse de normalité mais sous l'alternative de** H , **une loi de Fischer non centrée à** $(q+1)$ **et** $(n-p-q-2)$ **degrés de liberté avec un paramètre de non centralité égale à** z , **sous l'hypothèse alternative.**

3.6 Test du Rapport de la Vraisemblance

Supposons qu'on s'intéresse à tester l'hypothèse linéaire générale suivante :

$$H_0 : C' \underline{b} = \beta$$

contre l'hypothèse alternative consistant à ce que :

$$H_1 : \underline{b} \in \mathbb{R}^{k+1}$$

C'est à dire on regarde si les données statistiques prévoient (ou encore donnent) des raisons pour la rejeter ou non, selon un seuil, α , prédéterminé.

Un des tests les plus utilisés est le test du rapport de la fonction de vraisemblance, associé au modèle de régression adjacent et qui est, sous l'hypothèse de normalité, donnée par :

$$L(\underline{y} ; \underline{b}, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} (\underline{y} - \underline{X} \underline{b})' (\underline{y} - \underline{X} \underline{b}) \right\}$$

Ce test consiste en les étapes suivantes :

1. On calcul la valeur maximale de la fonction de la vraisemblance, $L(\cdot)$, sous l'hypothèse H_0 , dite hypothèse nulle. Cette valeur est notée $\max L_{H_0}$.
2. On calcul la valeur maximale de la fonction de la vraisemblance, $L(\cdot)$, sous l'hypothèse H_1 , dite hypothèse alternative. Cette valeur est notée $\max L_{H_1}$.

Le rapport de la vraisemblance est, par définition, donnée par :

$$L = \frac{\max L_{H_0}}{\max L_{H_1}}$$

et le test du rapport de la vraisemblance consiste à ne pas rejeter, au seuil α , l'hypothèse nulle si :

$$L = \frac{\max L_{H_0}}{\max L_{H_1}} > \lambda$$

pour une certaine valeur de λ à déterminer, à partir des données statistiques disponible.

Concernant la valeur $\max L_{H_0}$, est, d'après les résultats obtenus en estimations sous la contraintes, donnée par :

$$\begin{aligned} \max L_{H_0} &= (2\pi\tilde{\sigma}^2)^{-n/2} \exp \left\{ -\frac{1}{2\tilde{\sigma}^2} (\underline{y} - \underline{X} \tilde{b})' (\underline{y} - \underline{X} \tilde{b}) \right\} \\ &= \left(2\pi \frac{SSE + Q_H}{n} \right)^{-n/2} \exp\left(-\frac{n}{2}\right) \end{aligned}$$

et la valeur $\max L_{H_1}$ est donnée par :

$$\begin{aligned} \max L_{H_1} &= (2\pi\tilde{\sigma}^2)^{-n/2} \exp \left\{ -\frac{1}{2\tilde{\sigma}^2} (\underline{y} - \underline{X} \hat{b})' (\underline{y} - \underline{X} \hat{b}) \right\} \\ &= \left(2\pi \frac{SSE}{n} \right)^{-n/2} \exp\left(-\frac{n}{2}\right) \end{aligned}$$

Alors, le rapport de la vraisemblance est donné :

$$\begin{aligned} L &= \frac{\max L_{H_0}}{\max L_{H_1}} = \frac{\left(2\pi \frac{SSE + Q_H}{n} \right)^{-n/2} \exp\left(-\frac{n}{2}\right)}{\left(2\pi \frac{SSE}{n} \right)^{-n/2} \exp\left(-\frac{n}{2}\right)} \\ &= \frac{SSE}{SSE + Q_H} \end{aligned}$$

Chapitre 4

Simulation et Application

L'objectif de ce chapitre est de renforcer les résultats théoriques étudiés précédemment, par le biais des simulations. Trois processus générateurs de données de type $SETAR(2, p, d)$, sont utilisés pour simuler des séries temporelles de taille petite, modérée et relativement grande ($N = 30, 50, 80, 100; 150; 200; 300, 500, \text{ et } 1000$). L'ordre du modèle $SETAR$ est $p = 1, 2, 3$, et le paramètre de délai $d = 1$ (à titre d'essai). Le vecteur des erreurs $\{\varepsilon_t\}$ suit la loi normale de moyenne 0 et de variance 1. Pour chaque processus générateur on considère 1000 répétitions. Les résultats de l'estimation des paramètres des modèles $SETAR$, et les fréquences du rejet de l'hypothèse nulle, pour le test du rapport de vraisemblance, sont présentés dans des tables auxquelles on associe un tracé de trajectoires correspondant à chaque modèle.

Nous présentons dans la première section les résultats d'estimation du vecteur des paramètres associé au modèle à seuil $SETAR$. La deuxième section est consacrée aux résultats du test du rapport de vraisemblance qui revient à tester l'hypothèse nulle de linéarité contre l'hypothèse alternative qui est le modèle à seuil $SETAR$. Dans la dernière section, nous présentons des applications sur des données réelles (lynx canadien, Sunspot,...), auxquelles nous estimons les paramètres et nous appliquons le test, une comparaison avec les résultats de Tong est présentée. Nous avons utilisé le logiciel MATLAB, version 7.0 pour les simulations et les applications réelles.

4.1 Estimation des paramètres d'un modèle SETAR

Pour montrer la performance de la méthode d'estimation des paramètres d'un modèle *SETAR* développée par Tong (1990), nous avons simulé 1000 séries générées par le modèle (2.2.1), pour des tailles différentes $N = 50, 200, 500,$ et 1000 . nous avons retenu l'estimateur moyen pour chaque paramètre, ainsi que l'écart-type empirique associé. Les modèles choisis sont de type *SETAR*(2, p , 1) d'ordre $p = 1, 2$ et 3 .

Modèles SETAR d'ordre 1

- Modèle (1) :

Soit le modèle *SETAR*(2, 1, 1) suivant :

$$y_t = \begin{cases} 1 - 0.5y_{t-1} + \varepsilon_t & \text{si } y_{t-1} \leq 1 \\ 1 - 0.2y_{t-1} + \varepsilon_t & \text{si } y_{t-1} > 1 \end{cases}$$

où encore :

$$y_t = (1 - 0.5y_{t-1})I(y_{t-1} \leq 1) + (1 - 0.2y_{t-1})(1 - I(y_{t-1} \leq 1)) + \varepsilon_t \quad (4.1.1)$$

avec

$$I(y_{t-1} \leq 1) = \begin{cases} 1 & \text{si } y_{t-1} \leq 1 \\ 0 & \text{si } y_{t-1} > 1 \end{cases}$$

Selon Tong, ce modèle peut être réécrit de façon équivalente comme suit :

$$y_t = (1 - 0.2y_{t-1}) + I(y_{t-1} \leq 1) (-0.3y_{t-1}) + \varepsilon_t \quad (4.1.2)$$

Les résultats de simulation du modèle (4.1.2) sont présentés dans les tableaux (de 1 à 4) suivants :

Tableau 1.1

$N = 50$	θ_0	θ_1	ϕ_0	ϕ_1
Vraie valeur	1	-0.2	0	-0.3
Estimateur moyen (\hat{b}_{moy})	0.9487	-0.1781	0.0572	-0.3246
Ecart-type empirique	0.0696	0.0367	0.0725	0.0473

Tableau 1.2

$N = 200$	θ_0	θ_1	ϕ_0	ϕ_1
Vraie valeur	1	-0.2	0	-0.3
Estimateur moyen (\widehat{b}_{moy})	0.9996	-0.2007	-0.0025	-0.3067
Ecart-type empirique	0.0346	0.0184	0.0358	0.0234

Tableau 1.3

$N = 500$	θ_0	θ_1	ϕ_0	ϕ_1
Vraie valeur	1	-0.2	0	-0.3
Estimateur moyen (\widehat{b}_{moy})	0.9941	-0.1964	0.0034	-0.3071
Ecart-type empirique	0.0209	0.0112	0.0215	0.0143

Tableau 1.4

$N = 1000$	θ_0	θ_1	ϕ_0	ϕ_1
Vraie valeur	1	-0.2	0	-0.3
Estimateur moyen (\widehat{b}_{moy})	0.9973	-0.1994	0.0005	-0.2999
Ecart-type empirique	0.0153	0.0827	0.0158	0.0103

Il faut noter que les résultats présentés ci-dessus sont ceux des paramètres du modèle (4.1.2) tel qu'il est présenté par Tong (1990), et cela afin de programmer fidèlement sa méthode d'estimation des paramètres d'un modèle *SETAR*. Par ailleurs, un simple changement de variables permet de retrouver les paramètres du modèle (4.1.1).

En effet, le modèle (4.1.1) est de la forme :

$$y_t = (\phi_0^{(1)} + \phi_1^{(1)} y_{t-1}) I(y_{t-1} \leq r) + (\phi_0^{(2)} + \phi_1^{(2)} y_{t-1}) (1 - I(y_{t-1} \leq r)) + \varepsilon_t$$

et le modèle (4.1.2) est de la forme :

$$y_t = (\theta_0 + \theta_1 y_{t-1}) + I(y_{t-1} \leq r)(\phi_0 + \phi_1 y_{t-1}) + \varepsilon_t$$

on voit bien que :

$$\theta_0 = \phi_0^{(2)}, \quad \theta_1 = \phi_1^{(2)} \quad \text{et} \quad \phi_0 = \phi_0^{(1)} - \phi_0^{(2)}, \quad \phi_1 = \phi_1^{(1)} - \phi_1^{(2)}$$

donc, les paramètres estimés du modèle de base (4.1.1) sont donnés par :

$$\widehat{\phi}_0^{(2)} = \widehat{\theta}_0, \quad \widehat{\phi}_1^{(2)} = \widehat{\theta}_1, \quad \widehat{\phi}_0^{(1)} = \widehat{\phi}_0 + \widehat{\theta}_0, \quad \widehat{\phi}_1^{(1)} = \widehat{\phi}_1 + \widehat{\theta}_1$$

Les résultats de simulation du modèle (4.1.1) sont présentés dans les tableaux (de 5 à 8) suivant :

Tableau 1.5

$N = 50$	$\phi_0^{(1)}$	$\phi_1^{(1)}$	$\phi_0^{(2)}$	$\phi_1^{(2)}$
Vraie valeur	1	-0.5	1	-0.2
Estimateur moyen (\widehat{b}_{moy})	1.0060	-0.5028	0.9487	-0.1781
Ecart-type empirique	0.0194	0.0299	0.0696	0.0367

Tableau 1.6

$N = 200$	$\phi_0^{(1)}$	$\phi_1^{(1)}$	$\phi_0^{(2)}$	$\phi_1^{(2)}$
Vraie valeur	1	-0.5	1	-0.2
Estimateur moyen (\widehat{b}_{moy})	0.9971	-0.5074	0.9996	-0.2007
Ecart-type empirique	0.0944	0.0146	0.0346	0.0184

Tableau 1.7

$N = 500$	$\phi_0^{(1)}$	$\phi_1^{(1)}$	$\phi_0^{(2)}$	$\phi_1^{(2)}$
Vraie valeur	1	-0.5	1	-0.2
Estimateur moyen (\widehat{b}_{moy})	0.9975	-0.5035	0.9941	-0.1964
Ecart-type empirique	0.0584	0.0909	0.1091	0.1124

Tableau 1.8

$N = 1000$	$\phi_0^{(1)}$	$\phi_1^{(1)}$	$\phi_0^{(2)}$	$\phi_1^{(2)}$
Vraie valeur	1	-0.5	1	-0.2
Estimateur moyen (\widehat{b}_{moy})	0.9978	-0.4993	0.9973	-0.1994
Ecart-type empirique	0.0429	0.0644	0.0537	0.0827

• Modèle (2) :

Soit le modèle $SETAR(2, 1, 1)$ donné par :

$$Y_t = \begin{cases} 1 + 0,9y_{t-1} + \varepsilon_t & \text{si } y_{t-1} \leq 1 \\ 1 + 0,7y_{t-1} + \varepsilon_t & \text{si } y_{t-1} > 1 \end{cases}$$

Selon Tong, ce modèle peut être réécrit de façon équivalente comme suit :

$$y_t = (1 + 0,7y_{t-1}) + I(y_{t-1} \leq 1)(0,2y_{t-1}) + \varepsilon_t \quad (4.1.3)$$

Les résultats de simulation du modèle (4.1.3), pour $N = 500$, sont présentés dans le tableau (2.1) suivant :

Tableau 2.1

$N = 500$	θ_0	θ_1	ϕ_0	ϕ_1
Vraie valeur	1	0,7	0	0,2
Estimateur moyen (\widehat{b}_{moy})	1,0212	0,6934	0,0905	0,0852
Ecart-type empirique	0,0344	0,0366	0,0372	0,0514

• Modèle (3) : $SETAR(2, 1, 1)$

$$Y_t = \begin{cases} 1 + 0,9y_{t-1} + \varepsilon_t & \text{si } y_{t-1} \leq 0,1 \\ 1 + 0,5y_{t-1} + \varepsilon_t & \text{si } y_{t-1} > 0,1 \end{cases}$$

Selon Tong, ce modèle peut être réécrit de façon équivalente comme suit :

$$y_t = (1 + 0,5y_{t-1}) + I(y_{t-1} \leq 0,1) (0,4y_{t-1}) + \varepsilon_t \quad (4.1.4)$$

Les résultats de simulation du modèle (4.1.4), pour $N = 500$, sont présentés dans le tableau (3.1) suivant :

Tableau 3.1

$N = 500$	θ_0	θ_1	ϕ_0	ϕ_1
Vraie valeur	1	0,5	0	0,4
Estimateur moyen (\widehat{b}_{moy})	1,0067	0,4964	-0,0195	0,3030
Ecart-type empirique	0,0102	0,0435	0,0285	0,1491

Modèles SETAR d'ordre 2

Pour les modèles *SETAR* d'ordre 2, Nous présentons uniquement les résultats de simulation des paramètres selon le modèle de Tong comme suit.

• **Modèle (1) :**

Soit le modèle *SETAR*(2, 2, 1) donné par :

$$y_t = \begin{cases} 1 - 0.5y_{t-1} + 0.1y_{t-2} + \varepsilon_t & \text{si } y_{t-1} \leq 1 \\ 1 - 0.2y_{t-1} - 0.3y_{t-1} + \varepsilon_t & \text{si } y_{t-1} > 1 \end{cases}$$

Selon Tong, ce modèle peut être réécrit de façon équivalente comme suit :

$$y_t = (1 - 0.2y_{t-1} - 0.3y_{t-2}) + I(y_{t-1} \leq 1) (-0.3y_{t-1} + 0.4y_{t-2}) + \varepsilon_t \quad (4.2.1)$$

Les résultats de simulation du modèle (4.2.1) sont présentés dans les tableaux (de 1 à 4) suivant :

Tableau 4.1

$N = 50$	θ_0	θ_1	θ_2	ϕ_0	ϕ_1	ϕ_2
Vraie valeur	1	-0.2	-0.3	0	-0.3	0.4
Estimateur moyen (\widehat{b}_{moy})	1.0053	-0.2011	-0.3191	0.0315	-0.3077	0.3928
Ecart-type empirique	0.0260	0.0390	0.0203	0.0828	0.0502	0.0285

Tableau 4.2

$N = 200$	θ_0	θ_1	θ_2	ϕ_0	ϕ_1	ϕ_2
Vraie valeur	1	-0.2	-0.3	0	-0.3	0.4
Estimateur moyen (\widehat{b}_{moy})	1.0050	-0.2018	-0.3019	0.0052	-0.2947	0.3934
Ecart-type empirique	0.0348	0.0176	0.0278	0.0372	0.0224	0.0134

Tableau 4.3

$N = 500$	θ_0	θ_1	θ_2	ϕ_0	ϕ_1	ϕ_2
Vraie valeur	1	-0.2	-0.3	0	-0.3	0.4
Estimateur moyen (\widehat{b}_{moy})	0.9936	-0.1962	-0.3019	0.0135	-0.3062	0.3962
Ecart-type empirique	0.0227	0.0111	0.0603	0.0247	0.0146	0.0877

Tableau 4.4

$N = 1000$	θ_0	θ_1	θ_2	ϕ_0	ϕ_1	ϕ_2
Vraie valeur	1	-0.2	-0.3	0	-0.3	0.4
Estimateur moyen (\widehat{b}_{moy})	0.9988	-0.1985	-0.3002	-0.0027	-0.3003	0.4025
Ecart-type empirique	0.0157	0.0791	0.0437	0.0684	0.0108	0.0618

• **Modèle (2) :**

Soit le modèle $SETAR(2, 2, 1)$ donné par :

$$y_t = \begin{cases} 0.2 + 0.1y_{t-1} + 0.3y_{t-2} + \varepsilon_t & \text{si } y_{t-1} \leq 0, 1 \\ 0.2 - 0.1y_{t-1} - 0.2y_{t-2} + \varepsilon_t & \text{si } y_{t-1} > 0, 1 \end{cases}$$

Selon Tong, ce modèle peut être réécrit de façon équivalente comme suit :

$$y_t = (0.2 - 0.1y_{t-1} - 0.2y_{t-2}) + I(y_{t-1} \leq 0.1) (0.2y_{t-1} + 0.5y_{t-2}) + \varepsilon_t \quad (4.2.2)$$

Les résultats de simulation du modèle (4.2.2) sont présentés dans les tableaux (de 5 à 8) suivant :

Tableau 4.5

$N = 50$	θ_0	θ_1	θ_2	ϕ_0	ϕ_1	ϕ_2
Vraie valeur	0.2	-0.1	-0.2	0	0.2	0.5
Estimateur moyen (\widehat{b}_{moy})	0.2154	-0.1203	-0.2354	-0.0023	0.1832	0.4766
Ecart-type empirique	0.0390	0.0356	0.0204	0.0520	0.0499	0.0930

Tableau 4.6

$N = 200$	θ_0	θ_1	θ_2	ϕ_0	ϕ_1	ϕ_2
Vraie valeur	0.2	-0.1	-0.2	0	0.2	0.5
Estimateur moyen (\widehat{b}_{moy})	0.1965	-0.0986	-0.2096	0.0047	0.1895	0.4938
Ecart-type empirique	0.0177	0.0160	0.0983	0.0129	0.0331	0.0385

Tableau 4.7

$N = 500$	θ_0	θ_1	θ_2	ϕ_0	ϕ_1	ϕ_2
Vraie valeur	0.2	-0.1	-0.2	0	0.2	0.5
Estimateur moyen (\widehat{b}_{moy})	0.1932	-0.0914	-0.2009	0.0045	0.1844	0.4960
Ecart-type empirique	0.0113	0.0978	0.0643	0.0268	0.0413	0.0905

Tableau 4.8

$N = 1000$	θ_0	θ_1	θ_2	ϕ_0	ϕ_1	ϕ_2
Vraie valeur	0.2	-0.1	-0.2	0	0.2	0.5
Estimateur moyen (\widehat{b}_{moy})	0.2000	-0.1008	-0.2015	-0.0038	0.1981	0.4976
Ecart-type empirique	0.0810	0.0723	0.0447	0.0107	0.0103	0.0611

Modèles SETAR d'ordre 3

Soit le modèle $SETAR(2, 3, 1)$ donné par :

$$y_t = \begin{cases} 1 + 0.1y_{t-1} + 0, 3y_{t-2} + 0, 1y_{t-3} + \varepsilon_t & \text{si } y_{t-1} \leq 1.5 \\ 1 - 0, 1y_{t-1} - 0, 2y_{t-2} + 0, 1y_{t-3} + \varepsilon_t & \text{si } y_{t-1} > 1.5 \end{cases}$$

Selon Tong, ce modèle peut être réécrit de façon équivalente comme suit :

$$y_t = (1 - 0.1y_{t-1} - 0.2y_{t-2} + 0, 1y_{t-3}) + I(y_{t-1} \leq 1.5) (0.2y_{t-1} + 0.5y_{t-2}) + \varepsilon_t \quad (4.3.1)$$

Les résultats de simulation du modèle (4.3.1) sont présentés dans les tableaux (de 1 à 4) suivant :

Tableau 5.1

$N = 50$	θ_0	θ_1	θ_2	θ_3	ϕ_0	ϕ_1	ϕ_2	ϕ_3
Vraie valeur	1	-0.1	-0.2	0.1	0	0.2	0.5	0
Estimateur moyen (\widehat{b}_{moy})	1.0965	-0.1093	-0.2368	0.0664	0.0904	0.1653	0.4670	-0.0117
Ecart-type empirique	0.0192	0.0466	0.0866	0.0256	0.0273	0.0539	0.0344	0.0235

Tableau 5.2

$N = 200$	θ_0	θ_1	θ_2	θ_3	ϕ_0	ϕ_1	ϕ_2	ϕ_3
Vraie valeur	1	-0.1	-0.2	0.1	0	0.2	0.5	0
Estimateur moyen (\widehat{b}_{moy})	1.0062	-0.0949	-0.2074	0.0879	0.0345	0.1924	0.4903	0.0013
Ecart-type empirique	0.0474	0.0183	0.0119	0.0215	0.0122	0.0221	0.0147	0.0423

Tableau 5.3

$N = 500$	θ_0	θ_1	θ_2	θ_3	ϕ_0	ϕ_1	ϕ_2	ϕ_3
Vraie valeur	1	-0.1	-0.2	0.1	0	0.2	0.5	0
Estimateur moyen (\widehat{b}_{moy})	1.0178	-0.1011	-0.2043	0.0943	0.0045	0.1980	0.4967	-0.0023
Ecart-type empirique	0.0294	0.0137	0.0147	0.0668	0.0323	0.0404	0.0901	0.0885

Tableau 5.4

$N = 1000$	θ_0	θ_1	θ_2	θ_3	ϕ_0	ϕ_1	ϕ_2	ϕ_3
Vraie valeur	1	-0.1	-0.2	0.1	0	0.2	0.5	0
Estimateur moyen (\widehat{b}_{moy})	1.0020	-0.0985	-0.2018	0.0977	0.0032	0.1982	0.5004	0.0008
Ecart-type empirique	0.0129	0.0797	0.0525	0.0478	0.0337	0.0972	0.0652	0.0623

Analyse des résultats d'estimation du vecteur des paramètres du modèle **SE-TAR(2,p,1)**

Les résultats de simulation des différents modèles *SETAR* d'ordre $p = 1, 2, 3$, présentés dans les tableaux ci-dessus, ont permis de mettre en évidence deux aspects, à savoir l'effet de la taille et celui de l'appartenance à la région d'ergodicité.

A partir des résultats de simulation du modèle *SETAR* d'ordre 1 (tables (1.1) jusqu'à (1.4)), on constate que plus que la taille de l'échantillon est importante (de 50 à 1000) plus que les valeurs estimées du vecteur des paramètres se rapprochent des vraies valeurs. cette remarque est valable même pour les modèles *SETAR* d'ordre 2 et 3.

Afin de montrer l'effet de l'ergodicité sur l'estimation des paramètres du modèle *SETAR* d'ordre 1 nous avons comparé la précision des estimateurs de chaque paramètre dans les différents modèles : Modèle (1) (*Tableau 1.3*), Modèle (2) (*Tableau 2.1*) et Modèle (3)

(Tableau 3.1). Comme critère de comparaison, nous avons choisi la valeur absolue du biais. On fait remarquer que le choix des paramètres de ces modèles est tel que si l'on se réfère au théorème de Chan, Petrucelli, Tong et Woolford 1985 (1), les trois modèles sont ergodiques, seulement le modèle (1) est plus loin de la région de la non ergodicité que le sont les modèles (3) et (2). Le tableau suivant montre que les biais sont plus petits dans le modèle (1), et les valeurs les plus importantes apparaissent en modèle (2).

Valeurs absolues des biais de chaque paramètre dans les modèles (1), (2) et (3)

Modèles d'ordre $p = 1$, $N = 500$	θ_0	θ_1	ϕ_0	ϕ_1
Modèle (1) (Tableau 1.3)	0.0059	0.0036	0.0034	0.0071
Modèle (2) (Tableau 2.1)	0.0212	0.0066	0.0905	0.1148
Modèle (3) (Tableau 3.1)	0.0067	0.0036	0.0195	0.0970

4.2 Test du rapport de vraisemblance

La performance du test du rapport de vraisemblance construit est montrée par simulation. Trois processus générateurs de données, de type *Self – Exciting TAR (SETAR)*, donnés par (2.1), d'ordre $p = 1, 2$, et 3 , sont utilisés pour simuler des séries temporelles de taille $N = 30, 50, 80, 100, 150, 200, 300, 500$ et 1000 .

Le test du rapport de vraisemblance développé précédemment dans le chapitre 3, revient à tester l'hypothèse de linéarité (modèle autorégressif) contre l'alternative fournie par le modèle à seuil *SETAR*, c'est à dire tester l'hypothèse nulle suivante :

$$H_0 : \phi_0 = \dots = \phi_q = 0$$

Les valeurs des paramètres sont choisies tels que les modèles sous jacent soit stationnaires, où la condition d'ergodicité est vérifiée. Le paramètre délai d et le paramètre seuil r varient d'un modèle à l'autre. ε_t est un bruit blanc suivant une loi normale de moyenne nulle et de variance 1.

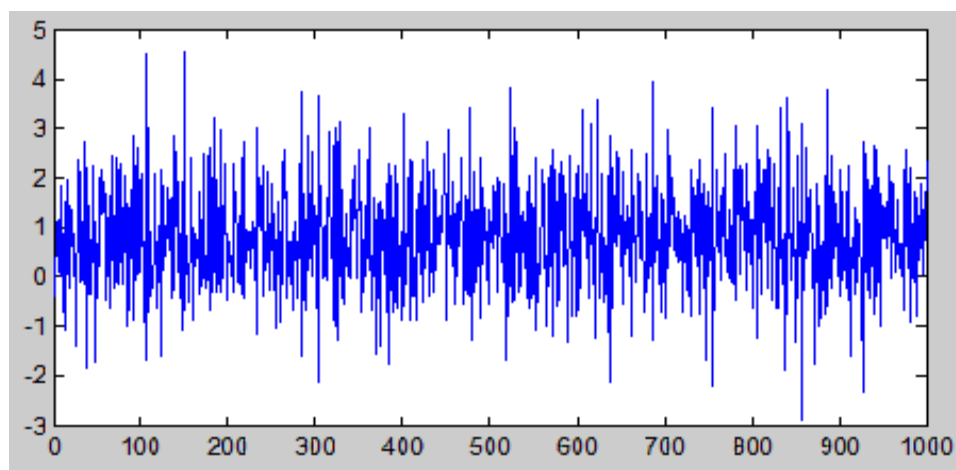
Une caractéristique importante de ces modèles est que le changement de régime est déterminé par des valeurs de la série y_t en fonction de la variable binaire : $I_t = \begin{cases} 1 & \text{si } y_{t-d} \leq r \\ 0 & \text{si } y_{t-d} > r \end{cases}$

En effet, le passage entre les régimes, dans chaque modèle *SETAR*, se fait brusquement en fonction du signe de y_{t-1} ($d = 1$)

Pour chaque processus générateur de données, 1000 réalisations sont effectuées, auxquelles on applique le test du rapport de vraisemblance. Dans les tables 1, 2 et 3 on reporte le nombre de rejets du modèle linéaire.

• **Modèle (1) :** *SETAR*(2, 1, 1)

$$y_t = \begin{cases} 1 - 0,5y_{t-1} + \varepsilon_t & \text{si } y_{t-1} \leq 1 \\ 1 - 0,2y_{t-1} + \varepsilon_t & \text{si } y_{t-1} > 1 \end{cases}$$



Graph 1 : Trajectoire simulée générée à partir du modèle (1)

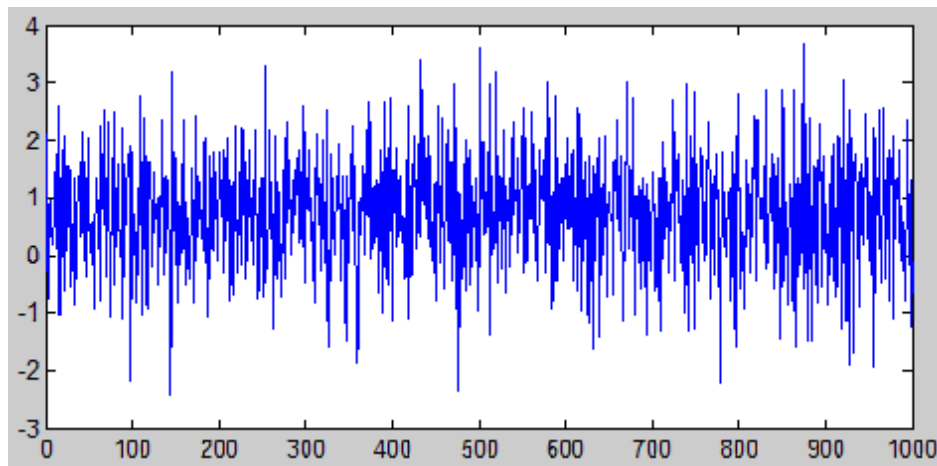
Les résultats sont présentés au tableau suivant :

Table 1 : Fréquence du rejet du modèle linéaire basé sur 1000 réalisations du modèle (1)

N	5%	10%
30	77	144
50	133	181
80	165	283
100	228	332
150	304	409
200	407	512
300	555	670
500	793	871
700	900	952
1000	974	987

• Modèle (2) : SETAR(2, 2, 1)

$$y_t = \begin{cases} 1 + 0,3y_{t-1} - 0,3y_{t-2} + \varepsilon_t & \text{si } y_{t-1} \leq 1 \\ 1 - 0,1y_{t-1} - 0,2y_{t-1} + \varepsilon_t & \text{si } y_{t-1} > 1 \end{cases}$$



Graph 2 : Trajectoire simulée générée à partir du modèle (2)

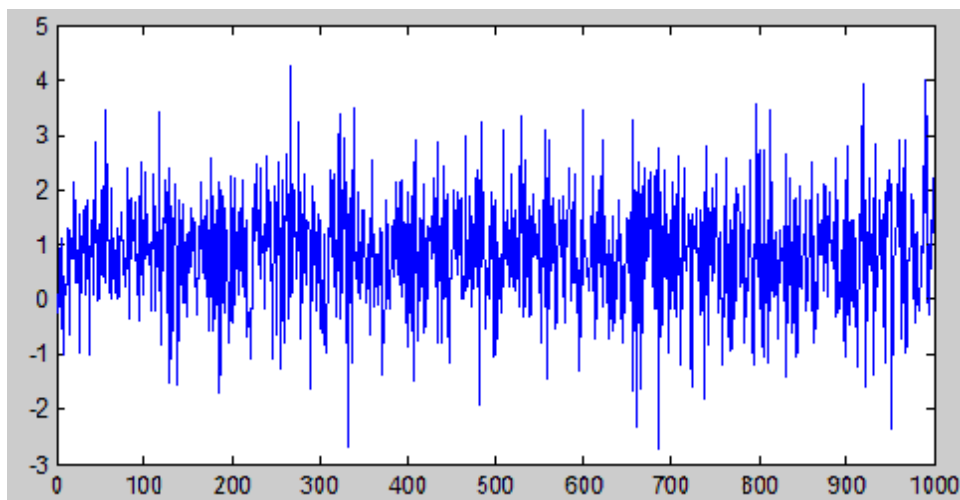
Les résultats sont présentés au tableau suivant :

Table 2 : Fréquence du rejet du modèle linéaire basé sur 1000 réalisations du modèle (2)

N	5%	10%
30	82	168
50	154	239
80	207	312
100	277	382
150	373	518
200	531	639
300	712	830
500	923	955
700	982	990
1000	997	999

• **Modèle (3) : SETAR(2, 3, 1)**

$$y_t = \begin{cases} 1 + 0,1y_{t-1} + 0,1y_{t-2} + 0,1y_{t-3} + \varepsilon_t & \text{si } y_{t-1} \leq 1 \\ 1 - 0,1y_{t-1} - 0,2y_{t-2} + 0,1y_{t-3} + \varepsilon_t & \text{si } y_{t-1} > 1 \end{cases}$$



Graph 3 : Trajectoire simulée générée à partir du modèle (3)

Les résultats sont présentés au tableau suivant :

Table 3 : Fréquence du rejet du modèle linéaire basé sur 1000 réalisations du modèle (3)

N	5%	10%
30	94	172
50	176	238
80	244	358
100	304	467
150	500	623
200	681	760
300	848	911
500	980	995
700	991	998
1000	998	999

Analyse des résultats du test du rapport de vraisemblance

La représentation graphique des trois séries simulées, générées à partir des modèles 1,2, et 3 respectivement, présentent un phénomène périodique caractérisé par des mouvements ascendants suivis par d'autres descendants autour de la moyenne empirique (seuil).

La fréquence du rejet du modèle linéaire (reconnaissance du modèle non linéaire à seuil SETAR) reflète la puissance du test appliqué. En effet, les résultats de simulation des trois modèles $SETAR(2, p, 1)$, $p = \overline{1, 3}$, au seuil 5%, sont acceptables. Ils apparaissent bons pour une taille d'échantillon N comprise entre 150 et 500, et excellentes pour une taille $N \geq 700$.

4.3 Application sur des données réelles

Nous nous proposons maintenant d'appliquer le test du rapport de vraisemblance sur des séries réelles, soit :

- La série du Lynx Canadien (1821-1934)
- La série du Log Lynx Canadien (1821-1934)
- La série des taches solaires Sunspot (1700-1955)
- La série racine carrée des taches solaires Sunspot (1700-1955)
- La série du Nil flow
- La série du Log Nil flow
- et la série des pertes d'électricité algérienne.

Pour chaque série, une description brève sur l'origine des données est proposée. Le graphe de chaque série est ainsi proposé, afin d'illustrer la dynamique d'évolution. Le seuil choisi représente la moyenne empirique des données de la série en question.

A la fin de cette section, nous présentons un tableau récapitulatif des résultats du test appliqué sur ces données réelles pour comparaison avec ceux trouvés par Tong, concernant la non linéarité des séries.

Notons que Tong a étudié la puissance de son test en simulant un modèle à seuil de type $SETAR$ avec $p = 1$ et 2 seulement.

1- Série du Lynx canadien brut

La série de la figure (1) présente les nombres annuels de lynx mis aux enchères à Londres durant les années (1821-1934) par la Hudson Bay Compagny ; soit 114 observations (*source : Time series Library*). Ces nombres reflètent dans une certaine proportion l'évolution de la population des lynx dans la région MacKenzie River au nord-ouest du Canada où ils sont capturés. Cette série historique a fait l'objet de plusieurs études, elle sert de référence pour différentes méthodes de modélisation. Wong et Li (2000) l'ont appliquée sur le modèle Mixture Autoregressif *MAR*, et Tong et Chan (1990) ont proposé le modèle *SETAR*.

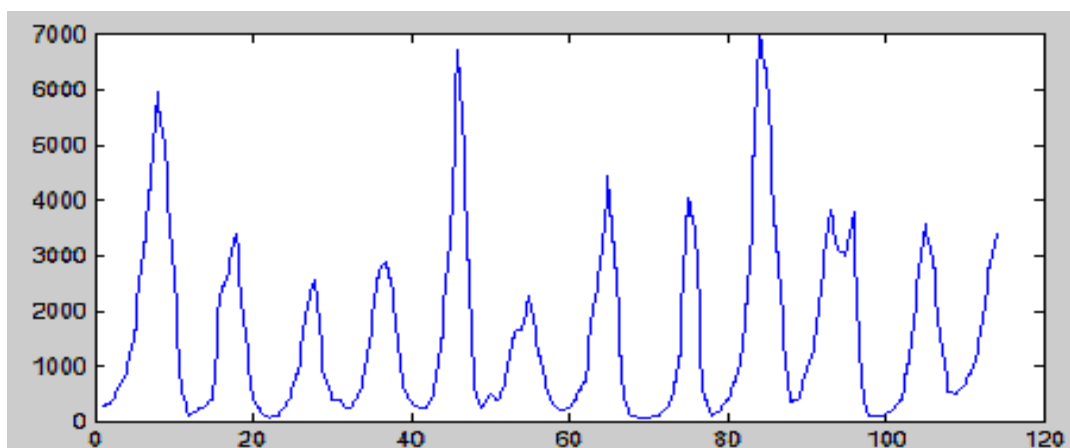


figure 1 : Représentation graphique de la série temporelle du nombre de lynx canadien (1821-1934)

On procède à une application similaire à celle de Tong et on propose le modèle *TAR* (2).

Ainsi, la série des données du lynx est représentée par le modèle suivant :

Modèle d'ordre 1 :

$$y_t = \begin{cases} -0.0045 + 0.0016 y_{t-1} + \varepsilon_t & \text{si } y_{t-1} \leq 1538 \\ 1.1083 + 0.0005 y_{t-1} + \varepsilon_t & \text{si } y_{t-1} > 1538 \end{cases}$$

Modèle d'ordre 2 :

$$y_t = \begin{cases} 0.1389 + 0.002 y_{t-1} - 0.0004 y_{t-2} + \varepsilon_t & \text{si } y_{t-1} \leq 1538 \\ 1.4978 + 0.0010 y_{t-1} - 0.0007 y_{t-2} + \varepsilon_t & \text{si } y_{t-1} > 1538 \end{cases}$$

2- Série du Log Lynx canadien

Pour la série du Log Lynx canadien (*figure 2*), nous avons retenu les modèles suivants :

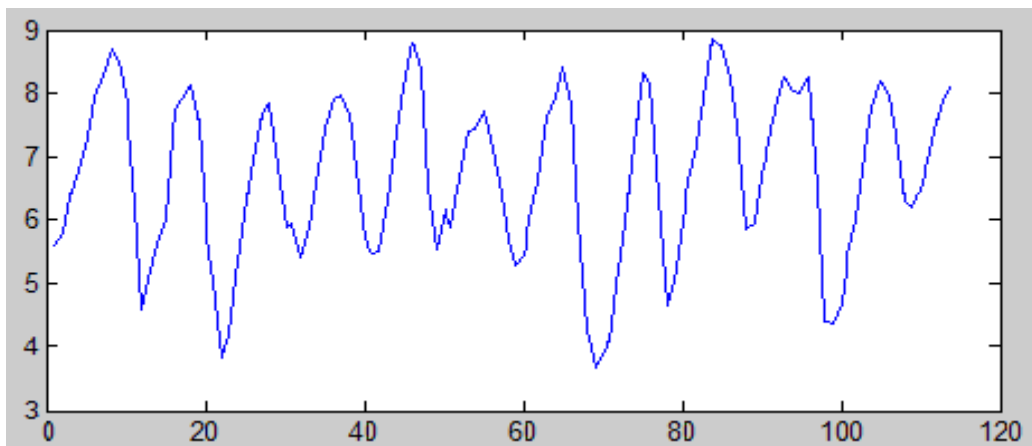


figure 2 : Représentation graphique de la série du \log_{10} du nombre de lynx canadien (1821-1934)

$$y_t = \begin{cases} 0.9163 + 1.3327 y_{t-1} - 0.6432 y_{t-2} + \varepsilon_t & \text{si } y_{t-2} \leq 2,9037 \\ 1.2918 + 1,5818 y_{t-1} - 1.0252 y_{t-2} + \varepsilon_t & \text{si } y_{t-2} > 2,9037 \end{cases}$$

Selon Tong, ce modèle peut être réécrit de façon équivalente comme suit :

$$y_t = (1.2918 + 1,5818 y_{t-1} - 1.0252 y_{t-2}) + I(y_{t-2} \leq 2.9037) (-0.3755 - 0.2491 y_{t-1} + 0.3820 y_{t-2}) + \varepsilon_t$$

Le modèle SETAR d'ordre 1 que l'on propose est comme suit :

$$y_t = \begin{cases} 0.5924 + 0.7934 y_{t-1} + \varepsilon_t & \text{si } y_{t-2} \leq 2.9037 \\ 1.6398 + 0.4935 y_{t-1} + \varepsilon_t & \text{si } y_{t-2} > 2.9037 \end{cases}$$

3- Série du Sunspot brut

La série des taches solaires, "sunspot" en anglais, provient de l'observation du soleil, dont l'activité varie au cours du temps. Plusieurs phénomènes liés à cette activité sont observables, dont des éruptions ou des taches noires, entre autres. Ces taches noires apparaissent quotidiennement selon un cycle approximatif de onze ans.

Le relevé du nombre de taches présentes à la surface du soleil est particulièrement intéressant du fait de l'ancienneté des données disponibles (depuis le début du XVIIIème siècle pour les moyennes annuelles). La série utilisée est composée de 256 observations couvrant les années 1700 à 1955. (*source : Time series Library*).

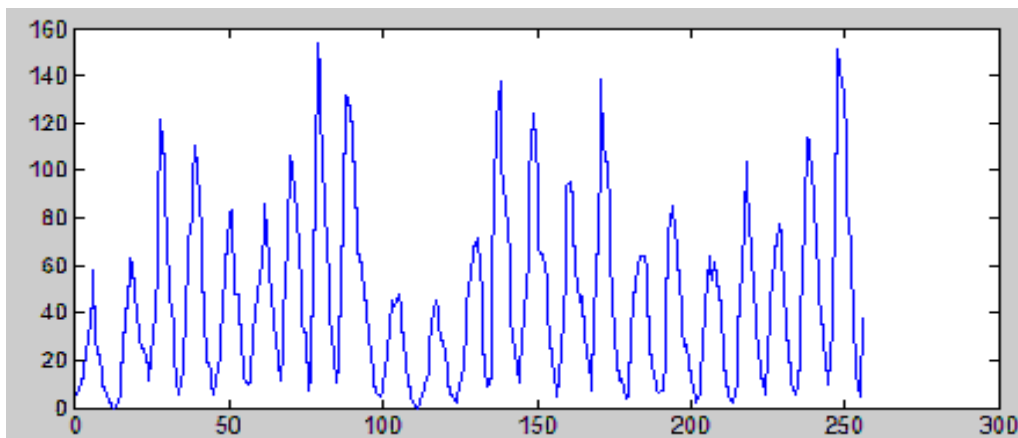


figure 3 : Représentation graphique de la série temporelle du sunspot (1700-1955)

Nous proposons un modèle TAR(2) pour modéliser cette série, donné par :

$$y_t = \begin{cases} 13.1221 + 1.6551 y_{t-1} - 0.9058 y_{t-2} + \varepsilon_t & \text{si } y_{t-1} \leq 44.782 \\ 21.4261 + 1.2021 y_{t-1} - 0.5891 y_{t-2} + \varepsilon_t & \text{si } y_{t-1} > 44.782 \end{cases}$$

4- Série du Nil flow

La série du Nil flow (figure 4) est composée de 100 *observations*, couvrant les années 1871 à 1970 (*source : Time series Library*)

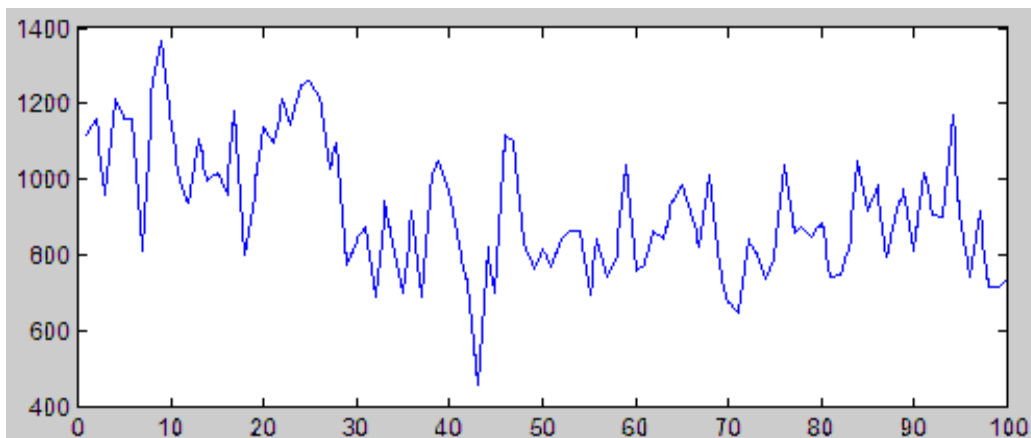


figure 4 : Trajectoire du Nil flow

Nous proposons un modèle TAR(2) pour modéliser cette série, donné par :

$$y_t = \begin{cases} 678.213 - 0.0217 y_{t-1} + 0.2143 y_{t-2} + \varepsilon_t & \text{si } y_{t-2} \leq 919.35 \\ 365.4653 + 0.4494 y_{t-1} + 0.1555 y_{t-2} + \varepsilon_t & \text{si } y_{t-2} > 919.35 \end{cases}$$

5- Série des pertes d'électricité

Les pertes de Distribution électricité (ou gaz) représentent l'écart entre les achats nets d'électricité (ou gaz) pour les besoins de la distribution et les ventes de la clientèle moyenne et basse tension (ou moyenne et basse pression).

Le Taux de pertes de distribution d'électricité (ou gaz) d'une société de Distribution égal au rapport, exprimé en %, des pertes de distribution d'électricité (ou gaz) aux achats nets d'électricité (ou gaz) de cette société de Distribution. Cette série comporte 41 observations, sur une période de 1970 à 2010. (*source : Société Nationale de l'Électricité et du Gaz,*

SONELGAZ).

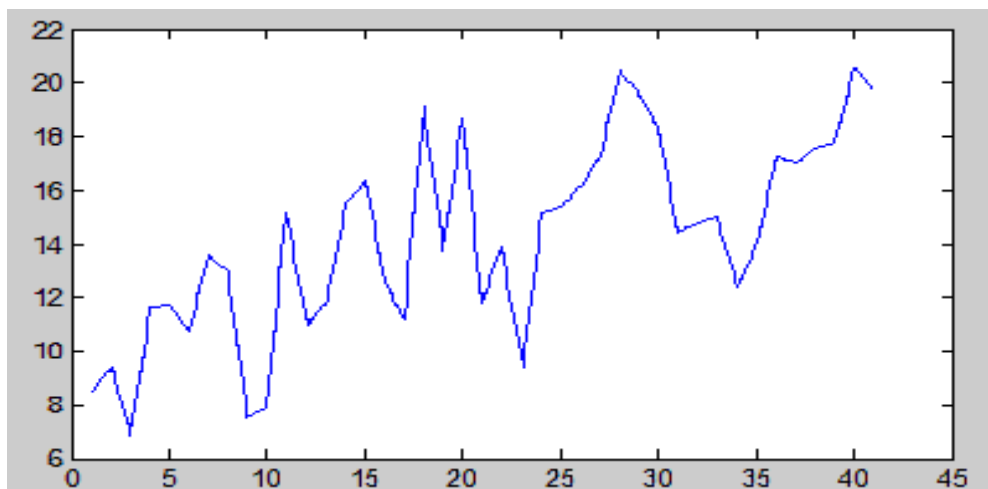


figure 5 : Trajectoire du taux de perte d'électricité

Une fois que l'étape de l'estimation des paramètres du modèle est achevée, nous procédons à l'application du test du rapport de vraisemblance sur ces séries de données, afin de montrer la puissance de ce test.

Le tableau suivant, représente les résultats de simulation du test appliqué au seuil 5%.

	$p=1, d=1$			$p=2, d=1$		
	Test	proba	Tong	Test	proba	Tong
<i>Lynx Canadien (brut)</i>	NL	0.0124	NL	NL	(*)	-
<i>log Lynx Canadien</i>	L	0.2932	-	NL	0.0558	NL
<i>Nil flow</i>	L	0.1090	-	L	0.2781	-
<i>log nil flow</i>	L	0.0972	-	L	0.3481	-
<i>Sunspot (brut)</i>	L	0.8890	-	NL	0.0025	NL
<i>Sqrt Sunspot</i>	L	0.6043	-	L	0.1364	-
<i>Pertes d'électricité</i>	L	0.8909	-	L	0.8802	-

NL : non linéaire ; L : linéaire

(*) : la valeur est de l'ordre de 10^{-6}

Les résultats du test obtenus, sur des séries réelles ont montré sa performance en la reconnaissance du modèle *SETAR* (rejet de l'hypothèse de linéarité ou modèle *AR*), en comparaison avec les résultats de Tong. En effet, les résultats obtenus sont similaires à

ceux trouvés par Tong, malgré que pour certaines séries (lynx canadien et sunspot), les probabilités correspondantes ne sont pas vraiment significatives.

Par ailleurs, nous remarquons que les séries : *log lynx canadien*, et *sunspot brut* passent de la forme linéaire à la forme non linéaire en augmentant l'ordre du modèle p .

Chapitre 5

Modèles à seuil périodiques

5.1 Introduction

Les techniques standards d'analyse de séries chronologiques ont longtemps reposé sur les propriétés fondamentales de linéarité et de stationnarité. Eu égard à leurs coefficients constants et à leur dynamique linéaire, les modèles autorégressifs moyennes mobiles (*ARMA*) ont ainsi fait l'objet d'un intérêt croissant depuis le début des années 1980. Cependant, de nombreuses recherches ont démontré que les hypothèses de linéarité et de stationnarité n'étaient qu'un pis-aller utopique apportant un confort appréciable dans l'étude probabiliste et statistique du modèle (*Guégan 1994*). Le recours à des modèles plus souples est alors vivement apparu comme une nécessité en séries temporelles.

En effet, la structure périodique de nombreuses séries chronologiques et l'insuffisance pragmatique des processus périodiquement corrélés (modèles saisonniers *ARMA* à coefficients constants) ont souligné dans un premier temps la nécessité d'introduire des coefficients dépendant du temps dans l'écriture du modèle. Par ailleurs, de nombreuses séries chronologiques rencontrées en pratique possèdent des caractéristiques non-linéaires, ce qui exclut toute éventualité de modélisation *ARMA*.

Les divers avantages et propriétés intéressantes satisfaites par les modèles à seuil classiques ou à coefficients constants ont motivé l'extension de ces modèles aux modèles à seuil à coefficients périodiques pour capturer, en plus, le phénomène de périodicité. Le premier travail accompli, à notre connaissance, dans cette direction est celui de Lewis et Ray (2002), dans lequel, ils ont introduit un modèle autorégressif à seuil périodique avec l'estimation des

paramètres en appliquant l'algorithme *TSMARS* pour la modélisation des données de la rivière Fraser en Bretagne. Puisque un modèle à seuil périodique compte plus de paramètres que son analogue traditionnel et donc il est plus compliqué.

Ce chapitre est consacré à la présentation du test d'un modèle autorégressif à seuil classique contre un modèle autorégressif à seuil périodique développé par (Bentarzi et Merzougui (2009)). Ce chapitre est organisé comme suit :

Dans la première section, les notations et définitions de base sur les modèles autorégressifs à coefficients classiques et périodiques sont données. En suite, la propriété de Normalité Asymptotique Locale (LAN) est établie, en adaptant les conditions de Swensen (1985) au modèle autorégressif à seuil. Ainsi que la Quadratique Asymptotique Locale (LAQ), la Normalité Asymptotique Locale (LAN) et la linéarité Asymptotique Locale.

Enfin, on verra le test localement asymptotiquement optimal pour tester un modèle autorégressif à seuil classique contre un autorégressif à seuil périodique, pour le cas où la densité des innovations est supposée être connue.

5.2 Notations et Définitions principales

5.2.1 Processus Autorégressif à seuil Self-Exciting Classique (*SETAR*)

Définition 2 *Un processus univarié, du second ordre, $\{Y_t; t \in \mathbb{Z}\}$ est dit suivre le modèle particulier Autorégressif à seuil Self-Exciting $SETAR(2, 1, 1)$, s'il est solution de l'équation aux différences non linéaire stochastique de la forme :*

$$Y_t = \phi_1 y_{t-1} I(y_{t-1} \leq r) + \phi_2 y_{t-1} I(y_{t-1} > r) + \varepsilon_t, \quad t \in \mathbb{Z}$$

où $\{\varepsilon_t; t \in \mathbb{Z}\}$ est un processus bruit blanc, i.e., une suite de variables indépendantes identiquement distribuées, de moyenne 0 et de variance σ^2 , avec une densité de probabilité $f(\cdot)$, pas nécessairement gaussienne et où r est le paramètre de seuil. Ce paramètre est, dans le présent travail, considéré connu; alors on peut, sans perte de généralité, le prendre égale à 0, désormais le modèle sera réécrit sous la forme suivante :

$$Y_t = \phi_1 Y_{t-1}^- + \phi_2 Y_{t-1}^+ + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (5.2.1)$$

où

$$Y_{t-1}^- = y_{t-1}I(y_{t-1} \leq 0), \text{ et } Y_{t-1}^+ = y_{t-1}I(y_{t-1} > 0).$$

5.2.2 Processus Autorégressif à seuil Self-Exciting Périodique (PSETAR)

Définition 3 *Un processus du second ordre $\{Y_t; t \in \mathbb{Z}\}$ est dit suivre le modèle Autorégressif à seuil Self-Exciting Périodique PSETAR(2,1,1), de période S ($S \geq 2$), s'il est solution de l'équation aux différences non linéaire stochastique périodique de la forme :*

$$Y_t = \phi_{t,1} y_{t-1} I(y_{t-1} \leq r_t) + \phi_{t,2} y_{t-1} I(y_{t-1} > r_t) + \varepsilon_t, \quad t \in \mathbb{Z}$$

où $\{\varepsilon_t; t \in \mathbb{Z}\}$ est un processus bruit blanc, i.e., une suite de variables indépendantes identiquement distribuées, de moyenne 0 et de variance finie σ_t^2 , avec une densité de probabilité $f(\cdot)$, pas nécessairement Gaussienne et où r_t sont les paramètres de seuil. Les paramètres Autorégressifs $\phi_{t,1}$ et $\phi_{t,2}$, la variance des innovations σ_t^2 et les paramètres de seuil r_t sont périodiques, dans le temps, de période S , i.e.,

$$\begin{aligned} \phi_{t+kS,1} &= \phi_{t,1} \quad \text{et} \quad \phi_{t+kS,2} = \phi_{t,2}, \quad \forall k, t \in \mathbb{Z}, \\ \sigma_{t+kS}^2 &= \sigma_t^2 \quad \text{et} \quad r_{t+kS} = r_t, \quad \forall k, t \in \mathbb{Z}. \end{aligned}$$

Dans ce travail, la variance des innovations est considérée être un paramètre invariant dans le temps, i.e., $\sigma_t^2 = \sigma^2$. En plus, les paramètres r_s , $s = 1, 2, \dots, S$, sont supposés être connus, alors on peut, sans perte de généralité, les prendre égales à 0, d'où le modèle sera réécrit sous la forme suivante :

$$Y_t = \phi_{t,1} Y_{t-1}^- + \phi_{t,2} Y_{t-1}^+ + \varepsilon_t, \quad t \in \mathbb{Z}, \tag{5.2.2}$$

où $Y_{t-1}^- = y_{t-1}I(y_{t-1} \leq 0)$, et $Y_{t-1}^+ = y_{t-1}I(y_{t-1} > 0)$

Notons $H_f^{(n)}(\underline{\phi})$ une suite des hypothèses nulles sous laquelle $\{Y_t^{(n)}, t \in \mathbb{Z}\}$ est une suite de réalisations d'un processus satisfaisant le modèle (5.2.1), où $\underline{\phi} = (\phi_1, \phi_2)'$ et $H_f^{(n)}(\underline{\phi}^{(n)})$ la suite des hypothèses alternatives sous laquelle la suite $\{Y_t^{(n)}, t \in \mathbb{Z}\}$ est une suite de

suffisante de causalité du modèle (2.1), i.e., $\phi_1 < 1$, $\phi_2 < 1$ et $\phi_1 \phi_2 < 1$,

H2. : La densité des innovations $f(\cdot)$ est supposée satisfaire les conditions suivantes :

(a) $f(x) > 0, \forall x \in \mathbb{R}$,

(b) $f(\cdot)$ est absolument continue par rapport à la mesure de Lebesgue μ : il existe une

fonction $\dot{f}(\cdot)$ tel que, pour $-\infty < a < b < \infty$, nous avons $f(b) - f(a) = \int_a^b \dot{f}_\sigma(x) d\mu(x)$,

(c) L'information de Fisher $I(f) = \int (\Psi_f(x))^2 f_\sigma(x) dx$ est finie, où $\Psi_f = -\frac{\dot{f}_\sigma(\cdot)}{f_\sigma(\cdot)}$,

(d) $\int xf(x) dx = 0$ et la variance est finie, i.e., $\sigma^2 = E(X_t^2) < \infty$.

On note que les deux conditions (b) et (c) impliquent la différentiabilité en moyenne quadratique de la fonction $f(\cdot)^{\frac{1}{2}}$. i.e.,

$$\lim_{\lambda \rightarrow 0} \lambda^{-2} \int \left(f_{\sigma}^{\frac{1}{2}}(x + \lambda) - f_{\sigma}^{\frac{1}{2}}(x) - \lambda \frac{\dot{f}_{\sigma}(x)}{f_{\sigma}^{\frac{1}{2}}(x)} \right)^2 dx = 0.$$

(voir, Lemme 3, page 191, Hájek (1972) ou Hájek et Šidák (1967)).

5.3 Normalité Asymptotique Locale

5.3.1 Suite de Rapports de Vraisemblance

Notons $\underline{Y}^{(n)} = (Y_1^{(n)}, \dots, Y_n^{(n)})$ une réalisation de taille finie n d'un processus autorégressif périodiquement corrélé $\{Y_t^{(n)}; t \in \mathbb{Z}\}$ satisfaisant le modèle Autorégressif à seuil causale (2.1) et soit $Y_0^{(n)}$ la valeur initiale dont les densités sont $g_0^{(n)}(Y_0^{(n)}; \underline{\phi}, \sigma)$ et $g_0^{(n)}(Y_0^{(n)}; \underline{\phi}^{(n)}, \sigma)$ sous $H_f^{(n)}(\underline{\phi})$ et $H_f^{(n)}(\underline{\phi}^{(n)})$, respectivement. En plus, $g_0^{(n)}(Y_0^{(n)}; \underline{\phi}^{(n)}, \sigma) - g_0^{(n)}(Y_0^{(n)}; \underline{\phi}, \sigma)$ est supposée convergente en probabilité vers 0, quand $\underline{\phi}^{(n)} \rightarrow \underline{\Phi}$ avec $n \rightarrow \infty$, et pour simplifier les notations, la taille de la série temporelle observée n est supposée un multiple de S , i.e., $n = mS$, $m \in \mathbb{N}^*$; $t = s + S\tau$, $s = 1, \dots, S$ et $\tau = 0, 1, \dots, m - 1$. $Z_t^{(n)}(\underline{\phi})$ et $Z_t^{(n)}(\underline{\phi}^{(n)})$,

$t \in \mathbb{Z}$, représentent les résidus calculés sous $H_f^{(n)}(\underline{\phi})$ et $H_f^{(n)}(\underline{\phi}^{(n)})$, respectivement. D'où :

$$Z_{s,\tau}^{(n)}(\underline{\phi}) = Y_{s+S\tau}^{(n)} - \phi_1 Y_{s-1+S\tau}^{(n)-} - \phi_2 Y_{s-1+S\tau}^{(n)+},$$

et

$$\begin{aligned} Z_{s,\tau}^{(n)}(\underline{\phi}^{(n)}) &= \\ & Y_{s+S\tau}^{(n)} - \left(\phi_1 + \frac{1}{\sqrt{n}} (\lambda_1^{(n)} + h_{s,1}^{(n)}) \right) Y_{s-1+S\tau}^{(n)-} - \left(\phi_2 + \frac{1}{\sqrt{n}} (\lambda_2^{(n)} + h_{s,2}^{(n)}) \right) Y_{s-1+S\tau}^{(n)+}, \\ &= Z_{s,\tau}^{(n)}(\underline{\phi}) - \frac{1}{\sqrt{n}} (\lambda_1^{(n)} + h_{s,1}^{(n)}) Y_{s-1+S\tau}^{(n)-} - \frac{1}{\sqrt{n}} (\lambda_2^{(n)} + h_{s,2}^{(n)}) Y_{s-1+S\tau}^{(n)+}, \\ &= Z_{s,\tau}^{(n)}(\underline{\phi}) - \frac{1}{\sqrt{n}} \underline{\mathcal{I}}^{*(n)'} Y_{s-1+S\tau}^{(n)} = Z_{s,\tau}^{(n)}(\underline{\phi}) - \gamma_{s,\tau}^{(n)}, \end{aligned}$$

où $\underline{\mathcal{I}}^{*(n)} = \left(\underline{\mathcal{I}}_1^{*(n)'}, \underline{\mathcal{I}}_2^{*(n)'}, \dots, \underline{\mathcal{I}}_S^{*(n)'} \right)'$ avec $\underline{\mathcal{I}}_s^{*(n)} = (\lambda_1^{(n)} + h_{s,1}^{(n)}, \lambda_2^{(n)} + h_{s,2}^{(n)})$, et où $\underline{Y}_{s-1+S\tau}^{(n)} = \left(Y_{s-1+S\tau}^{(n)-}, Y_{s-1+S\tau}^{(n)+} \right)'$, $s = 1, \dots, S$ et $S \in \mathbb{Z}$. La variance empirique correspondante est donnée par $\hat{\sigma}^2 = n^{-1} \sum_{t=1}^n \left(Z_t^{(n)}(\underline{\phi}) \right)^2$.

Alors, le logarithme du rapports de vraisemblance, $\Lambda_f^{(n)}(\underline{\phi}^{(n)}) = \Lambda_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)} \underline{\mathcal{I}}^{(n)})$ pour $H_f^{(n)}(\underline{\phi})$ contre $H_f^{(n)}(\underline{\phi}^{(n)})$, est alors donné, pour $n = mS$, par :

$$\Lambda_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)} \underline{\tau}^{(n)}) = \sum_{s=1}^S \sum_{\tau=0}^{m-1} \log \frac{f_\sigma \left(Z_{s,\tau}^{(n)} \left(\underline{\phi}^{(n)} \right) \right)}{f_\sigma \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) \right)} + \log \frac{g_0^{(n)} \left(Y_0^{(n)}; \underline{\phi}^{(n)}, \sigma \right)}{g_0^{(n)} \left(Y_0^{(n)}; \underline{\phi}, \sigma \right)},$$

par conséquent, nous avons, sous $H_f^{(n)}(\underline{\phi})$, l'approximation asymptotique locale :

$$\begin{aligned} \Lambda_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)} \underline{\tau}^{(n)}) &= \sum_{s=1}^S \sum_{\tau=0}^{m-1} \log \frac{f_\sigma \left(Z_{s,\tau}^{(n)} \left(\underline{\phi}^{(n)} \right) \right)}{f_\sigma \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) \right)} + 0_P(1), \\ &= \sum_{s=1}^S \sum_{\tau=0}^{m-1} \left[\log \left(f_\sigma \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) - \gamma_{s,\tau}^{(n)} \right) \right) - \log \left(f_\sigma \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) \right) \right) \right] + 0_P(1). \end{aligned}$$

où le terme $0_P(1)$ correspond à l'influence de la valeur initiale $Y_0^{(n)}$.

5.3.2 Normalité Asymptotique Locale

Afin d'adapter les conditions de Swensen pour traiter le modèle autorégressif à seuil périodique $PSETAR(2, 1, 1)$, un ensemble de variables est introduit :

$$\begin{aligned} \xi_{s+S\tau}^{(n)}(\underline{\phi}) &= \frac{f^{\frac{1}{2}} \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) - \gamma_{s,\tau}^{(n)} \right)}{f^{\frac{1}{2}} \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) \right)} - 1, \\ \zeta_{s+S\tau}^{(n)}(\underline{\phi}) &= \frac{1}{2} \Psi \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) \right) \gamma_{s,\tau}^{(n)} = \frac{1}{2} \delta_{s,\tau}^{(n)'} \underline{\tau}_s^{*(n)} \text{ où } \delta_{s,\tau}^{(n)} = \frac{1}{\sqrt{n}} \Psi \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) \right) Y_{s-1+S\tau}^{(n)}, \\ \underline{\delta}^{(n)}(\underline{\phi}) &= (\delta_1^{(n)'}, \dots, \delta_S^{(n)'})' \in \mathbb{R}^{2S} \text{ où } \delta_s^{(n)} = \frac{1}{\sqrt{n}} \sum_{\tau=0}^{m-1} \Psi \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) \right) Y_{s-1+S\tau}^{(n)}, \quad s = 1, \dots, S, \\ \delta_{s,1}^{(n)} &= \frac{1}{\sqrt{n}} \sum_{\tau=0}^{m-1} \Psi \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) \right) Y_{s-1+S\tau}^{(n)-} \text{ et } \delta_{s,2}^{(n)} = \frac{1}{\sqrt{n}} \sum_{\tau=0}^{m-1} \Psi \left(Z_{s,\tau}^{(n)} \left(\underline{\phi} \right) \right) Y_{s-1+S\tau}^{(n)+}. \end{aligned}$$

(5.3.1)

Considérons la matrice bloc diagonale de dimension $2S \times 2S$:

$$\underline{\Gamma}(\underline{\phi}, \sigma) = \frac{1}{\sigma^2} \begin{pmatrix} \underline{\Gamma}(\underline{\phi}, \sigma) & 0 & \dots & 0 \\ 0 & \underline{\Gamma}(\underline{\phi}, \sigma) & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \ddots & \underline{\Gamma}(\underline{\phi}, \sigma) \end{pmatrix} = I_{S \times S} \otimes \left(\frac{1}{\sigma^2} \underline{\Gamma}(\underline{\phi}, \sigma) \right), \quad (5.3.2)$$

où $\Gamma(\underline{\phi})$ est la matrice de variance du vecteur $\underline{Y}_{s-1+S\tau}^{(n)}$,

$$\underline{\Gamma}(\underline{\phi}, \sigma) = \begin{pmatrix} E(Y_0^{-2}) & 0 \\ 0 & E(Y_0^{+2}) \end{pmatrix}.$$

En utilisant les définitions et notations précédentes, on peut formuler les conditions de Swensen adaptées, qui impliquent la propriété LAN .

Proposition 4 (*Conditions de Swensen Adaptées (1985), Bentarzi et Merzougui (2009a)*)

Les conditions adaptées suivantes sont, sous les hypothèses (H1) et (H2), satisfaites :

- (1) $\lim_{m \rightarrow \infty} E \sum_{s=1}^S \sum_{\tau=0}^{m-1} \left(\zeta_{s+S\tau}^{(n)}(\underline{\phi}) - \zeta_{s+S\tau}^{(n)}(\underline{\phi}) \right)^2 = 0,$
- (2) $\sup_m E \sum_{s=1}^S \sum_{\tau=0}^{m-1} \left(\zeta_{s+S\tau}^{(n)2}(\underline{\phi}) \right) < \infty,$
- (3) $\max_s \max_{\tau} \left| \zeta_{s+S\tau}^{(n)}(\underline{\phi}) \right| = 0_p(1),$
- (4) $\sum_{s=1}^S \sum_{\tau=0}^{m-1} \left(\zeta_{s+S\tau}^{(n)2}(\underline{\phi}) \right) - \frac{1}{4S} I(f_1) \underline{I}^{*(n)'} \underline{\Gamma}(\underline{\phi}, \sigma) \underline{I}^{*(n)} = 0_p(1),$
- (5) $\sum_{s=1}^S \sum_{\tau=0}^{m-1} E \left[\zeta_{s+S\tau}^{(n)2}(\underline{\phi}) I_{\left[\left| \zeta_{s+S\tau}^{(n)}(\underline{\phi}) \right| > \frac{1}{2} \right]} / \mathcal{B}_{n,s-1+S\tau} \right] = 0_p(1),$
 $\mathcal{B}_{n,s-1+S\tau}$ est la σ -algèbre engendrée par le passé du processus jusqu'au temps $s-1+S\tau$,
- (6) $E \left(\zeta_{s+S\tau}^{(n)}(\underline{\phi}) / \mathcal{B}_{n,s-1+S\tau} \right) = 0_p(1),$

Proposition 5 (*Bentarzi et Merzougui (2009a)*)

Supposons que les hypothèses H1 et H2 sont satisfaites, alors nous avons, sous $H_f^{(n)}(\underline{\phi})$, quand $n \rightarrow \infty$:

(i) *Quadratique Asymptotique Locale (LAQ)*

$$\Lambda_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau}^{(n)}) = \underline{\tau}^{(n)'} \underline{\Delta}_f^{(n)}(\underline{\phi}) - \frac{1}{2} \underline{\tau}^{(n)'} \Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \sigma) \underline{\tau}^{(n)} + 0_p(1),$$

où la matrice carrée $\Gamma^{\Delta_f^{(n)}}(\underline{\phi}) = \frac{I(f_1)}{S} K' \Gamma(\underline{\phi}, \sigma) K$ et la suite centrale $\underline{\Delta}_f^{(n)}(\underline{\phi})$ est donnée par

$$\begin{aligned} \underline{\Delta}_f^{(n)}(\underline{\phi}) &= K' \underline{\delta}^{(n)}(\underline{\phi}), \\ \underline{\Delta}_f^{(n)}(\underline{\phi}) &= \left(\sum_{s=1}^S \delta_{s,1}^{(n)}, \sum_{s=1}^S \delta_{s,2}^{(n)}; \delta_{1,1}^{(n)} - \delta_{S,1}^{(n)}, \delta_{1,2}^{(n)} - \delta_{S,2}^{(n)}; \dots; \delta_{S-1,1}^{(n)} - \delta_{S,1}^{(n)}, \delta_{S-1,2}^{(n)} - \delta_{S,2}^{(n)} \right)'. \end{aligned}$$

(ii) *Normalité Asymptotique Locale de la suite centrale $\underline{\Delta}_f^{(n)}(\underline{\phi})$*

$$\underline{\Delta}_f^{(n)}(\underline{\phi}) \rightarrow N_{2S} \left(\underline{0}, \Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \sigma) \right).$$

Comme conséquences de la propriétés LAN :

i) $H_f^{(n)}(\underline{\phi})$ et $H_f^{(n)}(\underline{\phi}^{(n)})$ sont contiguës.

ii) La suite centrale $\underline{\Delta}_f^{(n)}(\underline{\phi})$ satisfait la Linéarité Asymptotique Locale suivante :

$$\underline{\Delta}_f^{(n)}(\underline{\phi}^{(n)}) - \underline{\Delta}_f^{(n)}(\underline{\phi}) = -\Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \sigma) \underline{\tau}^{(n)} + 0_p(1)$$

Le corollaire suivant établit les distributions asymptotiques de la suite centrale $\underline{\Delta}_f^{(n)}(\underline{\phi})$ et du logarithme du rapport de vraisemblance $\Lambda_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau}^{(n)})$ sous l'hypothèse alternative $H_f^{(n)}(\underline{\phi}^{(n)})$.

Corollaire 6 Si les hypothèses H1 et H2 sont satisfaites, alors :

$$1) \underline{\Delta}_f^{(n)}(\underline{\phi}) \Longrightarrow N \left(\Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \sigma) \underline{\tau}^{(n)}, \Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \sigma) \right) \text{ sous } H_f^{(n)}(\underline{\phi}^{(n)}),$$

$$2) \Lambda_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau}^{(n)}) \Longrightarrow N \left(\frac{1}{2} \underline{\tau}^{(n)'} \Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \sigma) \underline{\tau}^{(n)}, \underline{\tau}^{(n)'} \Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \sigma) \underline{\tau}^{(n)} \right) \text{ sous } H_f^{(n)}(\underline{\phi}^{(n)}).$$

Ces résultats sont des conséquences de la propriété LAN.

5.4 Test Asymptotique Locale (le plus Stringent)

L'une des conséquences essentielles de la propriété LAN est que la suite des expériences locales :

$$\xi_f^{(n)}(\underline{\phi}) = \left\{ P_{\underline{\Phi} + \underline{\nu}^{(n)} \underline{\tau}^{(n)}}; \underline{\tau}^{(n)} \in \mathbb{R}^{2S} / \sup_n \underline{\tau}^{(n)'} \underline{\tau}^{(n)} < \infty, \underline{\tau}^{(n)} \rightarrow \underline{\tau} \text{ quand } n \rightarrow \infty \right\},$$

converge faiblement vers l'expérience de position Gaussienne de dimension $2S$.

$$E_f = \{ N(\Gamma^\Delta(\underline{\phi}, \sigma) \underline{\tau}, \Gamma^\Delta(\underline{\phi}, \sigma)), \quad \underline{\tau} \in \mathbb{R}^{2S} \}.$$

Ainsi, dans le cas des tests d'hypothèse, si un test $\varphi(\Delta)$ est exactement optimal au niveau α dans E_f , alors, la suite de tests $\varphi(\Delta^{(n)}(\underline{\phi}))$ sera localement et asymptotiquement optimale au niveau asymptotique α dans la suite d'expériences originales $\xi_f^{(n)}(\underline{\phi})$. On rappelle que :

- 1) $\underline{\Delta}_f^{(n)}(\underline{\phi}) \implies N(\underline{0}, \Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \sigma)), \quad \text{sous } H_f^{(n)}(\underline{\phi}),$
- 2) $\underline{\Delta}_f^{(n)}(\underline{\phi}) \implies N(\Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \sigma) \underline{\tau}^{(n)}, \Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \sigma)), \quad \text{sous } H_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)} \underline{\tau}).$

Soit la décomposition suivante de la suite centrale $\underline{\Delta}_f^{(n)}(\underline{\theta}) = \begin{pmatrix} \underline{\Delta}_{I,f}^{(n)'} & \underline{\Delta}_{II,f}^{(n)'} \end{pmatrix}'$ où les vecteurs colonnes $\underline{\Delta}_{I,f}^{(n)}$ et $\underline{\Delta}_{II,f}^{(n)}$ de dimensions 2 et $2(S-1)$, respectivement, sont donnés par :

$$\begin{aligned} \underline{\Delta}_{I,f}^{(n)} &= \begin{pmatrix} \sum_{s=1}^S \delta_{s,1}^{(n)} & \sum_{s=1}^S \delta_{s,2}^{(n)} \end{pmatrix}', \\ \underline{\Delta}_{II,f}^{(n)} &= \left(\delta_{1,1}^{(n)} - \delta_{S,1}^{(n)}, \delta_{1,2}^{(n)} - \delta_{S,2}^{(n)}, \dots, \delta_{S-1,1}^{(n)} - \delta_{S,1}^{(n)}, \delta_{S-1,2}^{(n)} - \delta_{S,2}^{(n)} \right)', \end{aligned}$$

et soit la partition en bloc correspondante de la matrice de variance :

$$\Gamma^\Delta(\underline{\phi}, \sigma) = \begin{pmatrix} 2\Gamma_{11}^\Delta(\underline{\phi}, \sigma) & 0 \\ 0 & \Gamma_{22}^\Delta(\underline{\phi}, \sigma) \end{pmatrix},$$

où la matrice, de dimension 2×2 , $\mathbf{\Gamma}_{11}^{\Delta}(\underline{\phi}, \sigma) = I(f) \Gamma(\underline{\phi}, \sigma)$ et la matrice, de dimension $2(S-1) \times 2(S-1)$,

$$\mathbf{\Gamma}_{22}^{\Delta}(\underline{\varphi}, \sigma) = \frac{I(f)}{S} \begin{pmatrix} 2\Gamma(\underline{\phi}, \sigma) & \Gamma(\underline{\phi}, \sigma) & \dots & \Gamma(\underline{\phi}, \sigma) \\ & 2\Gamma(\underline{\phi}, \sigma) & \dots & \Gamma(\underline{\phi}, \sigma) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma(\underline{\phi}, \sigma) & \Gamma(\underline{\phi}, \sigma) & \dots & 2\Gamma(\underline{\phi}, \sigma) \end{pmatrix}.$$

Soit $\eta = \mathbf{\Gamma}^{\Delta}(\underline{\phi}, \sigma) \begin{pmatrix} \lambda \\ \underline{h} \end{pmatrix}$, alors le problème de tester l'hypothèse nulle $H_f^{(n)}(\underline{\phi})$ contre l'alternative locale $H_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)} \underline{\tau}^{(n)})$, i.e., tester un modèle $SETAR(2, 1, 1)$ à temps-invariant, donné par (5.2.1), contre, un modèle $SETAR(2, 1, 1)$ localement S -périodique, donné par (5.2.2), revient, tout simplement, au problème de test lié à l'expérience de position Gaussienne. Plus précisément : tester l'hypothèse nulle

$$H_{0,f} : N(\eta_0, \mathbf{\Gamma}^{\Delta}(\underline{\phi}, \sigma)), \quad \left(\eta_0 = \mathbf{\Gamma}^{\Delta}(\underline{\phi}, \sigma) \begin{pmatrix} \lambda \\ \underline{0} \end{pmatrix} \right),$$

contre l'hypothèse alternative

$$H_{1,f} : N(\eta, \mathbf{\Gamma}^{\Delta}(\underline{\phi}, \sigma)), \quad \left(\eta = \mathbf{\Gamma}^{\Delta}(\underline{\phi}, \sigma) \begin{pmatrix} \lambda \\ \underline{h} \end{pmatrix}, \underline{h} \neq \underline{0} \right).$$

La proposition suivante établit le test localement asymptotiquement optimal pour tester $H_{0,f}$ contre $H_{1,f}$.

Proposition 7 (Bentarzi et Merzougui (2009a))

Sous les conditions H1 et H2, le test qui rejette l'hypothèse nulle $H_f^{(n)}(\underline{\phi})$ si :

$$\widehat{Q}_f^{(n)}(\widehat{\underline{\phi}}^{(n)}) = \widehat{\Delta}_{II,f}^{(n)'}(\widehat{\underline{\phi}}^{(n)}) \left(\mathbf{\Gamma}_{22}^{\Delta}(\widehat{\underline{\phi}}^{(n)}) \right)^{-1} \widehat{\Delta}_{II,f}^{(n)}(\widehat{\underline{\phi}}^{(n)}) > \chi_{2(S-1), 1-\alpha}^2,$$

i) est de niveau asymptotique α (sous $H_f^{(n)}(\underline{\phi})$),

ii) est de puissance asymptotique :

$$1 - \mathcal{F} \left(\chi_{1-\alpha}^2; 2(S-1), \underline{h}' \mathbf{\Gamma}_{22}^{\Delta}(\widehat{\underline{\phi}}^{(n)}) \underline{h} \right), \text{ sous } H_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)} \underline{\tau}),$$

où $\mathcal{F}(\chi_{1-\alpha}^2; r, \nu)$ désigne la fonction de répartition de la loi khi deux non centré à r degré de liberté et de paramètre de non centralité ν ,

iii) est un test localement asymptotiquement "most stringent" contre $H_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{I})$.

5.5 Test Adaptatif

Le test décrit dans la proposition précédente est valide et optimal quand la densité des innovations f , qui intervient dans la statistique du test à travers la fonction $\Psi_f(\cdot) = -\frac{\dot{f}_\sigma(\cdot)}{f_\sigma(\cdot)}$, est spécifiée. Cependant, en pratique cette densité est fréquemment inconnue, par conséquent, elle est considérée comme un paramètre de nuisance.

Soit le modèle semi-paramétrique dont le vecteur des paramètres est $(\underline{\phi}, f)$ où f est le paramètre de nuisance qui est supposé appartenir à la classe des densités symétriques, \mathcal{F}^+ . L'hypothèse supplémentaire est donnée par :

Hypothèse (H3). La fonction de densité d'innovation f est symétrique avec des moments d'ordres quatre finis.

Un test est dit **adaptatif** s'il a, quand f est inconnue, la même puissance asymptotique que le test optimal dans le modèle où cette densité d'innovation est spécifiée. En suivant les étapes de Kreiss (1987), lorsqu'il a appliqué la méthode de la fonction du noyau, l'estimation de la densité d'innovation f est établi.

Soit :

$$(i) \quad g(x; \eta) = \frac{1}{\sqrt{2\pi\eta^2}} \exp\left(-\frac{x^2}{2\eta^2}\right), x \in \mathbb{R},$$

$$(ii) \quad f_\eta(x) = \int_{-\infty}^{+\infty} g(x-y; \eta) f(y) dy,$$

$$(iii) \quad \hat{f}_{\eta, \tau}(x, \underline{\phi}) = \frac{1}{2(m-1)} \sum_{\substack{\tau_0=0 \\ \tau_0 \neq \tau}}^{m-1} \{g(x+z_{s, \tau_0}, \eta) + g(x-z_{s, \tau_0}, \eta)\}; \tau = 0, \dots, m-1,$$

et soit $\hat{q}_{n, \tau}(x, \underline{\phi})$, pour un $\underline{\phi}$ donné, l'estimateur de $\Psi_f(\cdot)$ donné par

$$\hat{q}_{n,\tau}(x, \underline{\phi}) = \begin{cases} -\frac{1}{2} \frac{\hat{f}'_{\eta(n),\tau}(x, \underline{\phi})}{\hat{f}_{\eta(n),\tau}(x, \underline{\phi})} & \text{si } \begin{cases} \hat{f}_{\eta(n),\tau}(x, \underline{\phi}) \geq d_n, |x| \leq g_n, \\ |\hat{f}'_{\eta(n),\tau}(x, \underline{\phi})| \leq r_n \hat{f}_{\eta(n),\tau}(x, \underline{\phi}) \end{cases} \\ 0, & \text{sinon,} \end{cases}$$

avec $r_n \rightarrow \infty, g_n \rightarrow \infty, \eta(n) \rightarrow 0, d_n \rightarrow 0$. En remplaçant, dans l'expression de $\delta_s^{(n)}$, $\Psi_f(\cdot)$ par $\hat{q}_{n,\tau}(x, \underline{\phi})$, on obtient

$$\tilde{\delta}_s^{(n)} = \frac{1}{\sqrt{n}} \sum_{\tau=0}^{m-1} \hat{q}_{n,\tau}(x, \underline{\phi}) Y(s-1+d\tau), \quad s = 1, 2, \dots, S.$$

Lemme 8 Soit $(\bar{\phi}_n)$ une suite discrète d'estimateurs \sqrt{n} -consistents pour $\underline{\phi}$ et soit $\underline{\Delta}_f^{(n)}(\theta) = K' \tilde{\delta}^{(n)}$. Alors, nous avons, sous les hypothèses (H1) – (H5) :

$$\underline{\Delta}_f^{(n)}(\bar{\phi}_n) - \underline{\Delta}_f^{(n)}(\underline{\phi}_n) = 0_p(1),$$

pour $r_n \rightarrow \infty, g_n \rightarrow \infty, \eta(n) \rightarrow 0, d_n \rightarrow 0, \eta(n)r_n \rightarrow 0, g_n\eta(n)^{-4}/n \rightarrow 0$ et $n\eta(n)^9$ reste borné.

Lemme 9 $\hat{I}_n(\tilde{\phi}_n) = \frac{1}{S} \sum_{s=1}^S \left(\frac{1}{m} \sum_{\tau=0}^{m-1} \hat{q}_{n,\tau}^2(Z_{s,\tau}^{(n)}(\tilde{\phi}_n), \tilde{\phi}_n) \right)$ est un estimateur consistant de l'information de Fisher $I(f_1)$, i.e.,

$$\hat{I}_n(\tilde{\phi}_n) = I(f_1) + 0_p(1).$$

La proposition suivante établit le test adaptatif du modèle $SETAR(2, 1, 1)$ classique contre un autre périodique.

Proposition 10 (Bentarzi et Merzougui (2009a))

Sous les hypothèses (H1) – (H5), nous avons, sous $H_f^{(n)}(\underline{\phi})$, la statistique du test

$$\tilde{Q}_f^{(n)}(\hat{\underline{\phi}}^{(n)}) = \tilde{\Delta}_{II,f}^{(n)'}(\hat{\underline{\phi}}^{(n)}) \left(\hat{\Gamma}_{22}^{\Delta}(\hat{\underline{\phi}}^{(n)}) \right)^{-1} \tilde{\Delta}_{II,f}^{(n)}(\hat{\underline{\phi}}^{(n)}) > \chi_{2(S-1), 1-\alpha}^2,$$

est telle que :

i) le niveau asymptotique est α sous $H_f^{(n)}(\underline{\phi})$,

ii) est un test adaptatif locale de $H_f^{(n)}(\underline{\phi})$ contre $H_f^{(n)}(\underline{\Phi} + \nu^{(n)}\underline{\tau})$, i.e., pour tester $\underline{h} = \underline{0}$ contre $\underline{h} \neq \underline{0}$, avec un niveau asymptotique α et $f \in F^+$,

iii) La puissance asymptotique de ce test est la même que celle obtenue dans le cas paramétrique.

Conclusion et Perspectives

La présence de fluctuations cycliques asymétriques et de diverses ruptures dans la dynamique des variables économétriques est considérée comme l'un des problèmes majeurs dans l'analyse des séries temporelles. Pour la modélisation de ces différentes caractéristiques de cycles économiques, une large variété de techniques de séries temporelles linéaires et non linéaires se sont employées. Comme les modèles linéaires sont incapables de capturer de telles caractéristiques, un très grand intérêt est accordé récemment aux spécifications non linéaires, dont on introduit une distinction significative entre les phases d'expansion et les phases de récession. Ces modèles sont suffisamment flexibles pour qu'on puisse tenir compte des différentes spécifications et des relations correspondantes à chaque phase. Parmi ces modèles non linéaires, on peut citer les modèles autorégressifs à seuil (*TAR*, Tiao et Say 1994), les modèles *SETAR* (Teräsvirta et Anderson, 1992), et les modèles à changements de régimes (Markov Switching, Hamilton 1989).

Dans ce mémoire nous avons abordé la classe des modèles à seuil de type Self-Exciting Threshold Autoregressive (*SETAR*) classique, et la classe des modèles *SETAR* périodique.

Pour le cas classique, nous avons présenté les différentes étapes d'identification, spécification et estimation de ce type de modèles. Nous avons proposé, par la suite, un test basé sur le test du rapport de vraisemblance pour tester l'hypothèse linéaire (*AR*) contre l'hypothèse alternative de non linéarité (*SETAR*). La performance de ce test est montrée par simulation. Ainsi que la performance de la méthode d'estimation des paramètres du modèles *SETAR*. Pour cette dernière, nous avons retenu, l'estimateur moyen pour chaque paramètre ainsi que la variance empirique. Les paramètres estimés sont vraiment proche des paramètres réels du modèle simulé. Pour le cas périodique, nous avons présenté le test adaptatif développé par Bentarzi et Merzougui (2009), du modèle *SETAR* classique contre le modèle *SETAR* périodique

Pour les modèles *PAR* périodiques il y a une abondance de travaux par plusieurs chercheurs alors que pour les modèles *SETAR* périodiques il reste beaucoup de choses à faire à commencer par l'ergodicité, l'existence de solution stationnaire unique et une théorie d'iden-

tification avec l'estimation des paramètres de seuil supposé connu et fixé à 0, les critères de choix de l'ordre et les tests de non linéarités et l'estimation des paramètres. Ces modèles sont en effet très complexes et la théorie dans ce domaine est encore fort incomplète.

Bibliographie

- [1] Bentarzi, M. and Merzougui, M. (2009). Adaptive Test for Periodicity in Self-Exciting Threshold Autoregressive Models. *Comm. Simulation Comput.* 38, 1-19.
- [2] Bentarzi, M. and Merzougui, M. (2010). Adaptive Test for Periodicity in Autoregressive Conditions Heteroskedastic Processes. (en cours). *Communications in Statistic- Simulation and Computation.*
- [3] Bensalem. M. , Perraudin. C. Tests de linéarité, Spécification et Estimation de modèle à seuil : une analyse comparée des méthodes de Tsay et de Hansen. *La documentation française, Economie et Prévision, 2001/2- n° 148, ISSN 0249- 4744, pages 157 à 176.*
- [4] Chan, K. S. (1990). Testing for Threshold Autoregression. *The Annals of Statistics.* Vol. 18, No. 4, 1886-1894.
- [5] Chan, K. S. J. D. Petrucci ; H. Tong ; Samuel W. Woolford. A Multiple-Threshold AR(1) Model. *Journal of Applied Probability, Vol. 22, No. 2. (Jun., 1985), pp. 267-279.*
- [6] Chan, K. S. Ruey S. Tsay. Limiting Properties of the Least Squares Estimator of a Continuous Threshold Autoregressive Model. *Biometrika, Vol. 85, No. 2. (Jun., 1998), pp. 413-426.*
- [7] Chan, K. S. Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model. *The Annals of Statistics, Vol. 21, No. 1. (Mar., 1993), pp. 520-533.*
- [8] Chan, K. S. H. Tong. On the Use of the Deterministic Lyapunov Function for the Ergodicity of Stochastic Difference Equations. *Advances in Applied Probability, Vol. 17, No. 3. (Sep., 1985), pp. 666-678.*

- [9] Charpentier. A. Cours de Séries Temporelles- DESS Mathématiques de la décision & DESS Actuariat. Volume 2, p 225.
- [10] Colletaz. G. Hurlin. C. Modèles Non Linéaires et Prévisions. *LEO – Laboratoire d’Economie d’Orléans, UMR CNRS 6221, Novembre 2006.*
- [11] Colletaz. G. Hurlin. C. Threshold Effects of the Public Capital Productivity : an International Panel Smooth Transition Approach. *Laboratoire d’Economie d’Orléans, N° 2006-01.*
- [12] Ghaddar D. K. H. Tong. Data Transformation and Self-Exciting Threshold Autoregression. *J. Applied Statistics, Vol. 30, No. 3. (1981), pp. 238-248.*
- [13] Jeroen V. K. Rombouts Techniques d’analyse et de prévision de la conjoncture (6-837-77). *Hiver 2007, 23 Novembre 2006.*
- [14] Lardic.S. Mignon.V. Econométrie des Séries Temporelles Macroéconomiques et Financières. *Economica. 2002.*
- [15] Mohamed Ali Houfi et Moukhtar Kouki. Modèles non linéaires et tests d’efficience informationnelle. *Correspondance : LEGI-Ecole Polytechnique de Tunisie, Septembre 2006.*
- [16] Michael P. Clements, Jeremy Srnith. The performance of alternative forecasting methods for SETAR models.
- [17] Mélard. G. Roy. R. Modèles de séries chronologiques avec seuil. *Revue statistique appliquée, tome36, n° 4 (1988), p.5-23.*
- [18] Oyvind Eitrheim , Timo Teraisvirta. Testing the adequacy of smooth transition autoregressive models. *Journal of Econometrics 74 (1996) 59-75.*
- [19] Petrucci. J. D. and N Davies (1986). A Portmanteau test for self-exciting threshold autoregressive-type nonlinearity in time series. *Biometrika, 73, 687-694.*
- [20] Petrucci. J. D. and S. W. Woolford. (1984). A threshold AR(1) model. *J.Appl. Probab., 21, 270-286.*
- [21] Remzi Uctum. Econométrie des modèles à changement de régimes : un essai de synthèse. *L’Actualité économique, vol.83, n° 4, 2007, p.447-482.*
- [22] Sophie Lamouroux. Estimation des modèles à seuil STAR et SETAR par deux méthodes d’estimation indirecte. Rapport de recherche, Septembre 1998.

- [23] Shiqing Ling. Howell Tong. Dong Li. Ergodicity and Invertibility of Threshold MA Models. *AMS 1980 subject classifications : Primary 62F05, 62M10; secondary 60G10.*
- [24] Shiqing Ling. Howell Tong. Testing for a linear MA model against Threshold MA models. *The Annals of Statistics. 2005, Vol. 33, No. 6, 2529–2552*
- [25] Tong, H (1978). On threshold models. In *Pattern Recognition and Signal Processing*, (ed. C. H. Chen). Amsterdam : Sijhoff & Noordhoff.
- [26] Tong, H. (1983). Threshold models in nonlinear time series. *New York. Springer.*
- [27] Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *J. Royal. Soc. Ser. B, 4, 245-292.*
- [28] Tsay, R. S. (1989). Testing and Modeling Threshold Autoregressive Processes. *Journal of the American Statistical Association. Vol. 84, 231-240.*
- [29] Tsay, R. S. (1998). Testing and Modelling Multivariate Threshold Models. *Journal of the American Statistical Association. 93 : 1188-202.*

- ANNEXE -

Théorème (1)

Soit la matrice symétrique nonsingulière (régulière)

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{pmatrix},$$

et notons son inverse

$$\mathbf{V}^{-1} = \begin{pmatrix} \mathbf{V}^{11} & \mathbf{V}^{12} \\ \mathbf{V}^{21} & \mathbf{V}^{22} \end{pmatrix},$$

alors

1) Si V_{11} est régulière, on a :

$$|\mathbf{V}| = |\mathbf{V}_{11}| \left| \mathbf{V}_{22} - \mathbf{V}_{21}^{-1} \mathbf{V}_{11} \mathbf{V}_{12} \right|, \quad (1.1)$$

et

$$\mathbf{V}^{-1} = \begin{pmatrix} \mathbf{V}_{11}^{-1} + \mathbf{V}_{11}^{-1} \mathbf{V}_{12} \mathbf{V}^{22} \mathbf{V}_{21} \mathbf{V}_{11}^{-1} & -\mathbf{V}_{11}^{-1} \mathbf{V}_{12} \mathbf{V}^{22} \\ -\mathbf{V}^{22} \mathbf{V}_{21} \mathbf{V}_{11}^{-1} & \mathbf{V}^{22} \end{pmatrix}, \quad (1.2)$$

où

$$\mathbf{V}^{22} = (\mathbf{V}_{22} - \mathbf{V}_{21} \mathbf{V}_{11}^{-1} \mathbf{V}_{12})^{-1}, \quad (1.3)$$

2) Si V_{22} est régulière on a :

$$|\mathbf{V}| = |\mathbf{V}_{22}| \left| \mathbf{V}_{11} - \mathbf{V}_{12} \mathbf{V}_{22}^{-1} \mathbf{V}_{21} \right|, \quad (2.1)$$

et

$$\mathbf{V}^{-1} = \begin{pmatrix} \mathbf{V}^{11} & -\mathbf{V}^{11} \mathbf{V}_{12} \mathbf{V}_{22}^{-1} \\ -\mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{V}^{11} & \mathbf{V}_{22}^{-1} + \mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{V}^{11} \mathbf{V}_{12} \mathbf{V}_{22}^{-1} \end{pmatrix}, \quad (2.2)$$

où

$$\mathbf{V}^{11} = (\mathbf{V}_{11} - \mathbf{V}_{12} \mathbf{V}_{22}^{-1} \mathbf{V}_{21})^{-1}. \quad (2.3)$$

Remarque : les formules (1.1) et (2.1) sont dites formules de Dodge.

Théorème (2)

Si X est une variable aléatoire suivant une loi normale de moyenne μ et de variance v . Les deux quadratiques forme $X'AX$ et $X'BX$ sont indépendantes, si et seulement si $AVB = 0$ ($BVA = 0$).

Théorème (3)

Si X est une variable aléatoire suivant une loi normale de moyenne μ et de variance v , alors la quadratique forme $X'AX$ suit χ^2 $\left[r(A), \frac{1}{2}\mu'AX\mu \right]$ (khi deux non centré) si et seulement si AV est idempotente.

$r(A)$: rang de A , et $\frac{1}{2}\mu'AX\mu$ appelé paramètre de non centralité.

*** Loi de probabilité de Q_H**

Démonstration. a) **Indépendance de SSE et $Q\left(\underline{b}/Y, \underline{X}\right)_H$.**

Afin d'appliquer les théorèmes 2 et 3, fournissant respectivement une condition nécessaire et suffisante pour que deux formes quadratiques en une même variable aléatoire normalement distribuée soient indépendantes et une condition nécessaire et suffisante pour qu'une quadratique forme en une variable aléatoire normale suit une loi de khi-deux, on doit, d'abord, exprimer les deux variables aléatoires des deux quadratiques en une même variable aléatoire normale en fonction du vecteur aléatoire expliqué Y . A cet effet, on peut aisément vérifier que la variable aléatoire vectorielle (plus précisément la statistique vectorielle) $C'\hat{\underline{b}} - \beta$ peut s'écrire sous la forme suivante :

$$\left(Y - \underline{X}\hat{\underline{b}} \right)' \underline{X} = Y' \left(I - \underline{X} (\underline{X}'\underline{X})^{-1} \underline{X}' \right) \underline{X} = 0, \quad (1).$$

$$\hat{\underline{b}}_H = \hat{\underline{b}} - C (C' C)^{-1} (C' \hat{\underline{b}} - \beta), \quad (2)$$

$$C' \hat{\underline{b}} - \beta = C' (\underline{X}'\underline{X})^{-1} \underline{X}' (Y - \underline{X} C (C' C)^{-1} \beta), \quad (3)$$

ce qui permet de réécrire la statistique quadratique $Q\left(\underline{b}/Y, \underline{X}\right)_H$ sous la forme désirée suivante :

$$Q\left(\underline{b}/Y, \underline{X}\right)_H = (Y - \underline{X} C (C' C)^{-1} \beta)' B (Y - \underline{X} C (C' C)^{-1} \beta), \quad (4)$$

$$Q\left(\underline{b}/Y, \underline{X}\right)_H = (Y - X - C(C'C)^{-1} b)' B (Y - X - C(C'C)^{-1} b)$$

où la matrice carrée B , de dimension $n \times n$, est donnée par

$$\underline{X} (\underline{X}' \underline{X})^{-1} C \left(C' (\underline{X}' \underline{X})^{-1} C \right)^{-1} C' (\underline{X}' \underline{X})^{-1} \underline{X}'$$

D'autre part, la statistique SSE peut elle aussi s'écrire sous forme quadratique en utilisant la même variable normale $Y - \underline{X} C (C' C)^{-1} \beta$. En effet, en tenant compte que :

$$\underline{X} \left(I - \underline{X} (\underline{X}' \underline{X})^{-1} \underline{X}' \right) \underline{X}' = 0;$$

on peut facilement voir que :

$$\begin{aligned} SSE &= Y' \left(I - \underline{X} (\underline{X}' \underline{X})^{-1} \underline{X}' \right) Y, \quad (5) \\ &= (Y - \underline{X} C (C' C)^{-1} \beta)' \left(I - \underline{X} (\underline{X}' \underline{X})^{-1} \underline{X}' \right) (Y - \underline{X} C (C' C)^{-1} \beta) \end{aligned}$$

Comme le vecteur $Y - \underline{X} C (C' C)^{-1} \beta$ suit, sous l'hypothèse de normalité SH_2^* , une loi normale de variances-covariances $I \sigma^2$, alors on peut aisément voir que la condition nécessaire et suffisante donnée par le théorème (2), pour que les formes quadratiques $Q\left(\underline{b}/Y, \underline{X}\right)_H$ et SSE données par les expressions (4) et (5), respectivement, soient indépendantes est remplie d'où découle l'indépendance de ces deux formes quadratiques.

b) **Distribution de** $\frac{Q(\underline{b}/Y, \underline{X})_H}{\sigma^2}$

On peut facilement voir que la condition nécessaire et suffisante pour qu'une forme quadratique, en une variable aléatoire normale, donnée par le théorème (3) est satisfaite. En effet, la matrice $\frac{1}{\sigma^2} B$ associée à la forme quadratique $\frac{Q(\underline{b}/Y, \underline{X})_H}{\sigma^2}$ et la matrice de variances-covariances $I \sigma^2$ du vecteur aléatoire normale satisfont la condition du théorème précité :

$$\frac{Q\left(\underline{b}/Y, \underline{X}\right)_H}{\sigma^2} B I \sigma^2 = B, \quad \text{car la matrice } B \text{ est idempotente}$$

Alors, $\frac{Q\left(\underline{b}/Y, \underline{X}\right)_H}{\sigma^2}$ suit, sous le système d'hypothèses, une loi de khi-deux centrée à $(q+1)$ degrés de liberté. On remarque que le paramètre de non centralité et le degrés de liberté découlent immédiatement de de la matrice associée à la forme quadratique en question à savoir

$\frac{Q(\underline{b}/Y, \underline{X})}{\sigma^2}$ qui est de rang $(q+1)$. Le paramètre de non centralité est donné par :

$$z = \frac{1}{\sigma^2} (C'\underline{b} - \beta)' \left(C' (\underline{X}' \underline{X})^{-1} C \right)^{-1} (C'\underline{b} - \beta), \quad (6)$$

et qui est clairement nul sous la contrainte linéaire (3.3.1).

■