

Identification des modèles *SETAR* par le critère de la densité prédictive

Fayçal HAMDI et Abderaouf KHALFI

Laboratoire RECITS, Faculté de Mathématiques
USTHB, BP 32 El-Alia, Bab-Ezzouar, Alger, Algérie.

fhamdi@usthb.dz, khalfi.doc@gmail.com

Abstract: In this note, we consider the class of Self-Exciting Threshold Autoregressive (*SETAR*) models where it is question to study the problem of selecting the order of each autoregressive component of the model. These models have proved useful for taking into account some stylized features exhibited by many time series such as asymmetric cycles, high-amplitude changes in the mean and the leptokurtic distributions. We propose a criterion based on the predictive density function which generalizes the one proposed by Djuric and Kay (1992). Note that this important problem in the modeling of real data has not received the interest that it merits.

Keywords: Time series with changes in regime, Self-Exciting Threshold Autoregressive models, order selection, predictive density criterion.

Résumé : Dans cette note, on considère la classe des modèles Self-Exciting autorégressifs à seuils (*SETAR*) où il est question d'étudier le problème de sélection de l'ordre de chaque composante autorégressive du modèle. Ces modèles se sont avérés utiles pour prendre en charge certaines caractéristiques rencontrés souvent lors de la modélisation de beaucoup de séries chronologiques telles que les cycles asymétriques, des ruptures de forte amplitude dans la moyenne et les distributions leptokurtiques. Nous proposons un critère fondé sur la fonction de densité prédictive et qui généralise celui proposé par Djuric et Kay (1992). Notons que ce problème important dans la modélisation des données réelles n'a pas reçu l'intérêt qu'il mérite.

Mots clés : Séries chronologiques avec changement de régimes, modèles Self-Exciting autorégressifs à seuils, sélection des ordres, critère de la densité prédictive.

1 Introduction

Depuis son introduction par Tong (1978), la famille des modèles à seuil, en particulier le modèle autorégressif à seuil (*TAR*) et ses extensions, s'est avérée utile dans de nombreuses applications (e.g. Tong et Lim, 1980, Tong, 1990; Tsay, 1998; Clements et Smith, 1997 et 2000 et Hansen, 2011). Dès lors, elle est devenue l'une des classes de modèles de séries chronologiques non linéaires les plus utilisées. Les modèles *TAR* reposent sur l'idée qu'une relation peut être non linéaire sur la période globale considérée, mais linéaire par sous périodes. Une représentation *TAR* est ainsi formée de modèles linéaires par morceaux. En d'autres termes elle est formée d'un mélange de modèles autorégressifs où le changement de régime est régi par des seuils, qui forment une partition de l'espace des réels. Les modèles les plus couramment utilisés dans cette catégorie sont les modèles Self-Exciting *TAR* (*SETAR*) de Tong et Lim (1980), où la variable de seuil est considérée comme étant une valeur retardée de la série chronologique elle-même. Tong et Lim (1980) ont montré que tout modèle non-linéaire autorégressif d'ordre p (*NLAR*) exprimé à travers la formulation $y_t = f(y_{t-1}, \dots, y_{t-p}) + \varepsilon_t$, peut être approché, sous certaines conditions générales sur la fonction f , par un modèle *TAR*. Ils ont fait également ressortir la capacité des modèles *SETAR* à reproduire des cycles limites. Dans le même papier, Tong et Lim (1980) ont appliqué avec succès la modélisation *SETAR* à la célèbre série du nombre annuel de lynx canadien, entre 1821 et 1934, où ils ont détecté la présence d'un cycle limite d'une période de 9 ans.

Le problème de l'identification ou la recherche du meilleur modèle dans une classe de modèles candidats est une préoccupation majeure dans l'analyse des séries chronologiques. En particulier, dans la classe des modèles *SETAR*. Notons, qu'outre la sélection de l'ordre de chaque modèle autorégressif du mélange, l'identification est intimement liée aux nombre de régimes, les seuils et la variable délai. Ceci présente une difficulté majeure pour cette étape cruciale dans la modélisation des séries chronologiques par les modèles *SETAR*.

Le problème de la sélection du meilleur modèle *SETAR* a été étudié par plusieurs chercheurs. Tong (1990) a suggéré d'utiliser le critère d'information d'Akaike (*AIC*) (Akaike, 1973), mais aucune justification théorique n'a été donnée. Wong et Li (1998) ont proposé une version corrigée du critère *AIC* (*AICc*) pour la sélection du modèle *SETAR*. Ils ont comparé la performance de ce critère *AICc* par rapport aux critères *AIC* et *BIC* classiques en se basant sur une étude de simulation. Dans cette étude, ils ont montré que la performance du *AICc* est meilleure que celle du *AIC* et *BIC* classique, en particulier lorsque la taille de la série est petite. De Gooijer (2001) a proposé trois critères de validation croisée où le retard et la valeur de seuil sont considérés comme inconnus. Kapetanios (2001) a étendu certains critères connus dans la littérature des modèles linéaires au cas des modèles à seuils. Galeano et Peña (2007) ont généralisé les critères de sélection proposés par Hurvich et al (1990) dans le cadre des modèles *AR* au cas des modèles *SETAR*. Öhrvik et Schoier (2005) et Fenga et Politis (2015) ont étudié la performance de plusieurs critères de sélection basés sur la technique du bootstrap.

Nous nous intéressons dans cette note à l'extension du Critère de la Densité Prédicative (*PDC*) proposé par Djuric and Kay (1992) au cas des modèles *SETAR*. Le reste de ce papier est organisé comme suit. Dans la section 2, nous rappelons la définition du modèle

SETAR. Dans la section 3, nous dérivons l'expression de la densité prédictive dont nous avons besoin pour définir notre critère *PDC*. Par la suite, nous proposons une version étendue à la classe des modèles *SETAR*, où une procédure de calcul récursif de ce critère est établie dans la dernière section.

2 Modèles *SETAR*

Les modèles autorégressifs à seuils (*TAR*) introduits par Tong (1978) constituent l'une des formulations possibles de la grande famille des modèles non-linéaires, à savoir la famille des modèles à changement de régime. En effet, une façon d'introduire la non-linéarité est de considérer une dynamique linéaire par morceaux en introduisant $K + 1$ seuils, $-\infty = r_0 < r_1 < \dots < r_K = +\infty$, qui forment une partition de l'espace des réels. En d'autres termes, il existe plusieurs dynamiques (K régimes), pour une même série, chacune s'appliquant à un sous-échantillon. Ces modèles ont été étudiés par plusieurs chercheurs (Tong et Lim, 1980; Tong, 1983; Chan et Tong, 1986; Tsay, 1989; Tong, 1990; Hili, 1992; Chan, 1993 et bien d'autres). Ils offrent la possibilité d'une modélisation endogène du mécanisme de changement de régime. Un cas particulier survient lorsque la variable de seuil est considérée comme étant une valeur retardée de la série chronologique elle-même. Dans ce cas et comme le régime est déterminé par cette même série chronologique, le modèle résultant est appelé un modèle Self-Exciting *TAR* (*SETAR*).

Rappelons qu'une série chronologique (y_1, y_2, \dots, y_n) est issue d'un modèle *SETAR* à K régimes et d'ordres (p_1, \dots, p_K) , noté *SETAR* $(K; p_1, \dots, p_K)$, si

$$y_t = \sum_{k=1}^K \left(\sigma_k \varepsilon_t - \sum_{j=1}^{p_k} \phi_{k,j} y_{t-j} \right) \mathbf{1}_{(r_{k-1} \leq y_{t-d} < r_k)}, \quad (1)$$

où l'ensemble des coefficients $\phi_{k,j}$ ($k = 1, \dots, K$ et $j = 1, \dots, p_k$) sont des réels, d est un entier positif appelé paramètre de délai (ou de retard) et les coefficients r_k , $k = 0, \dots, K$, sont les paramètres de seuil pour lesquels le système passe d'un régime à l'autre. $\mathbf{1}_A$ désigne la fonction indicatrice de l'ensemble A et $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un processus bruit blanc de variance unitaire. Enfin, les constantes σ_k , $k = 1, 2, \dots, K$, sont des réelles strictement positives.

3 Calcul de la densité prédictive

Pour procéder à l'identification des ordres d'un modèle *SETAR* en utilisant le critère *PDC*, il est nécessaire d'avoir l'expression de la densité prédictive. Pour cela, considérons une réalisation donnée $y = (y_1, y_2, \dots, y_n)$ du modèle *SETAR* défini dans (1), que nous supposons stationnaire. Sous l'hypothèse de normalité de ε_t , la densité prédictive de y_t sachant l'information disponible à l'instant $t - 1$, récapitulée par le vecteur $y_{1,t-1} = (y_1, y_2, \dots, y_{t-1})$, et les ordres du modèle (p_1, \dots, p_K) , peut être exprimée comme suit

$$f(y_t | y_{1,t-1}, p_1, \dots, p_K) = \int_{\sigma} \int_{\Phi} f(y_t | y_{1,t-1}, p_1, \dots, p_K, \sigma, \Phi) f(\sigma, \Phi | y_{1,t-1}, p_1, \dots, p_K) d\Phi d\sigma,$$

où $\sigma = (\sigma_1, \dots, \sigma_K)'$ et $\Phi = (\Phi'_{1,p_1}, \dots, \Phi'_{K,p_K})'$, où $\Phi_{k,p_k} = (\phi_{k,1}, \dots, \phi_{k,p_k})'$.

Soit $m = \max\{p_1, \dots, p_K, d\}$. Selon le théorème de Bayes, la densité a posteriori des paramètres σ et Φ après avoir observé le vecteur $y_{1,t-1}$, est donnée par

$$f(\sigma, \Phi | y_{1,t-1}, p_1, \dots, p_K) \propto f(y_{m+1,t-1} | y_{1,m}, p_1, \dots, p_K, \sigma, \Phi) f(\sigma, \Phi | p_1, \dots, p_K),$$

où

$$f(y_{m+1,t-1} | y_{1,m}, p_1 > 0, \dots, p_K > 0, \sigma, \Phi) = \prod_{i=m+1}^{t-1} \left(\frac{1}{\sqrt{2\pi\sigma_{R_i}^2}} \exp \left\{ -\frac{1}{2\sigma_{R_i}^2} \left(y_i + \sum_{j=1}^{p_{R_i}} \phi_{R_j,j} y_{i-j} \right)^2 \right\} \right),$$

$y_{m+1,t-1} = (y_{m+1}, y_2, \dots, y_{t-1})$ et $\{R_t, t \in Z\}$ est une suite de variables aléatoires indicatrices prenant des valeurs dans $\{1, 2, \dots, K\}$ telle que $R_t = k$, si $r_{k-1} \leq y_{t-d} < r_k$.

Du fait que nous ne disposons d'aucune information a priori sur les paramètres σ et Φ , la loi souvent utilisée dans ce cas et pour ces deux paramètres est une loi a priori non informative. Celle-ci peut être construite à partir de la distribution d'échantillonnage, puisque c'est la seule information disponible (e.g. Robert, 2007). Dans notre cas, elle est choisie comme suit

$$f(\sigma, \Phi | p_1, \dots, p_K) \propto \prod_{k=1}^K \frac{1}{\sigma_k}, \quad \Phi_{k,p_k} \in \mathbb{R}^{p_k} \text{ et } \sigma_k \in \mathbb{R}_+^*, \text{ pour } k = 1, \dots, K.$$

Par conséquent, la densité prédictive est donnée par

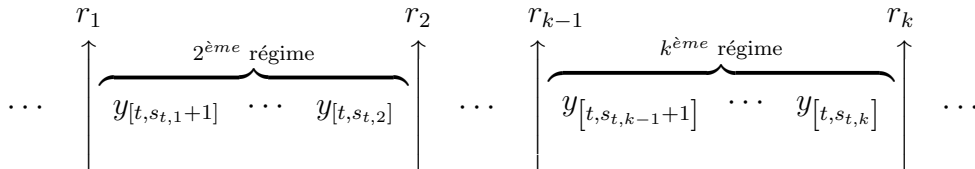
$$\begin{aligned} & f(y_t | y_{1,t-1}, p_1 > 0, \dots, p_K > 0) \\ & \propto \int_{\sigma} \int_{\Phi} f(y_{m+1,t} | y_{1,m}, p_1, \dots, p_K, \sigma, \Phi) f(\sigma, \Phi | p_1, \dots, p_K) d\Phi d\sigma \\ & \propto \int_{\sigma} \int_{\Phi} \left(\prod_{k=1}^K \frac{1}{\sigma_k} \right) \left(\prod_{i=m+1}^t \frac{1}{\sqrt{2\pi\sigma_{R_i}^2}} \right) \exp \left\{ -\sum_{i=m+1}^t \frac{1}{2\sigma_{R_i}^2} \left(y_i + \sum_{j=1}^{p_{R_i}} \phi_{R_j,j} y_{i-j} \right)^2 \right\} d\Phi d\sigma. \end{aligned}$$

En se basant sur la représentation donnée par Tsay (1989), nous allons réécrire cette intégrale d'une manière qui peut simplifier son évaluation (voir la figure 1). Pour cela, soient $E_t = \{y_{m+1-d}, \dots, y_{t-d}\}$ l'ensemble des valeurs possibles de la variable seuil et $[t, i]$ l'indice de la $i^{\text{ème}}$ petite valeur de cet ensemble E_t . Ainsi,

$$\varepsilon_{[t,i]+d} = y_{[t,i]+d} + \sum_{j=1}^{p_k} \phi_{k,j} y_{[t,i]+d-j}, \text{ si } s_{t,k-1} < i \leq s_{t,k},$$

où $s_{t,0} = 0$ et $s_{t,k} = \min_{1 \leq i \leq t-d} \{i, \text{ tel que } r_k \leq y_{[t,i+1]}\}$, pour $k = 1, \dots, K$.

Figure 1



En utilisant ces notations, la densité prédictive peut être réécrite comme suit

$$f(y_t \mid y_{1,t-1}, p_1 > 0, \dots, p_K > 0) = \frac{C}{(2\pi)^{\frac{t-m}{2}}} \times \int_{\sigma} \int_{\Phi} \prod_{k=1}^K \left[\frac{1}{\sigma_k^{s_{t,k}-s_{t,k-1}+1}} \exp \left\{ - \sum_{i=s_{t,k-1}+1}^{s_{t,k}} \frac{1}{2\sigma_k^2} \left(y_{[t,i]+d} + \sum_{j=1}^{p_k} \phi_{k,j} y_{[t,i]+d-j} \right)^2 \right\} \right] d\Phi d\sigma,$$

où C est la constante de normalisation à déterminer à la fin du calcul. Soit maintenant $n_{t,k} = s_{l,k} - s_{l,k-1}$. Considérons la matrice $H_p(l, k)$ et le vecteur $Y(l, k)$ de dimensions respectives $n_{l,k} \times p$ et $n_{l,k} \times 1$, définis par

$$H_{p_k}(t, k) = \begin{bmatrix} y_{[t,s_{t,k-1}+1]+d-1} & y_{[t,s_{t,k-1}+1]+d-2} & \cdots & y_{[t,s_{t,k-1}+1]+d-p_k} \\ y_{[t,s_{t,k-1}+2]+d-1} & y_{[t,s_{t,k-1}+2]+d-2} & \cdots & y_{[t,s_{t,k-1}+2]+d-p_k} \\ \vdots & \vdots & & \vdots \\ y_{[t,s_{t,k}]+d-1} & y_{[t,s_{t,k}]+d-2} & \cdots & y_{[t,s_{t,k}]+d-p_k} \end{bmatrix},$$

et

$$Y(t, k) = (y_{[t,s_{t,k-1}+1]+d}, y_{[t,s_{t,k-1}+2]+d}, \dots, y_{[t,s_{t,k}]+d})',$$

à partir de laquelle, nous pouvons voir que

$$\begin{aligned} & \exp \left\{ - \frac{1}{2\sigma_k^2} \sum_{i=s_{t,k-1}+1}^{s_{t,k}} \left(y_{[t,i]+d} + \sum_{j=1}^{p_k} \phi_{k,j} y_{[t,i]+d-j} \right)^2 \right\} \\ = & \exp \left\{ - \frac{Y'(t, k) P_{p_k}(t, k) Y(t, k) - [\Phi_{k,p_k} + A_{p_k}(t, k)]' \Sigma_{p_k}(t, k) [\Phi_{k,p_k} + A_{p_k}(t, k)]}{2\sigma_k^2} \right\}, \end{aligned}$$

où

$$\begin{aligned} \Sigma_{p_k}(t, k) &= H_{p_k}'(t, k) H_{p_k}(t, k), \\ A_{p_k}(t, k) &= \Sigma_{p_k}^{-1}(t, k) H_{p_k}'(t, k) Y(t, k), \\ P_{p_k}(t, k) &= I_{n_{t,k}} - H_{p_k}(t, k) [H_{p_k}'(t, k) H_{p_k}(t, k)]^{-1} H_{p_k}'(t, k), \end{aligned}$$

et $I_{n_{t,k}}$ est la matrice d'identité de dimension $n_{t,k}$. Il s'en suit que

$$\begin{aligned} & f(y_t \mid y_{1,t-1}, p_1 > 0, \dots, p_K > 0) \\ = & \frac{C}{(2\pi)^{\frac{t-m}{2}}} \prod_{k=1}^K \left[\int_{\sigma} \left(\frac{1}{\sigma_k^{n_{t,k}+1}} \exp \left\{ - \frac{Y'(t, k) P_{p_k}(t, k) Y(t, k)}{2\sigma_k^2} \right\} \right. \right. \\ & \left. \left. \times \int_{\Phi_{k,p_k}} \exp \left\{ - \frac{[\Phi_{k,p_k} + A_{p_k}(t, k)]' \Sigma_{p_k}(t, k) [\Phi_{k,p_k} + A_{p_k}(t, k)]}{2\sigma_k^2} \right\} d\Phi_{k,p_k} \right) d\sigma_k \right]. \end{aligned}$$

Or

$$\begin{aligned} & \int_{\mathbb{R}^{p_k}} \exp \left\{ - \frac{[\Phi_{k,p_k} + A_{p_k}(t, k)]' \Sigma_{p_k}(t, k) [\Phi_{k,p_k} + A_{p_k}(t, k)]}{2\sigma_k^2} \right\} d\Phi_{k,p_k} \\ = & (2\pi)^{\frac{p_k}{2}} \sigma^{p_k} |\Sigma_{p_k}^{-1}(t, k)|^{\frac{1}{2}} = \frac{(2\pi)^{\frac{p_k}{2}} \sigma^{p_k}}{|\Sigma_{p_k}(t, k)|^{\frac{1}{2}}}. \end{aligned}$$

Par conséquent,

$$\begin{aligned} f(y_t \mid y_{1,t-1}, p_1 > 0, \dots, p_K > 0) \\ &= \frac{C}{(2\pi)^{\frac{t-m}{2}}} \prod_{k=1}^K \frac{(2\pi)^{\frac{p_k}{2}}}{|\Sigma_{p_k}(t,k)|^{\frac{1}{2}}} \left(\int_0^{+\infty} \left(\sigma_k^{-n_{t,k}+p_k-1} \exp \left\{ -\frac{Y'(t,k)P_{p_k}(t,k)Y(t,k)}{2\sigma_k^2} \right\} \right) d\sigma_k \right) \\ &= \frac{C}{(2\pi)^{\frac{t-m}{2}}} \prod_{k=1}^K \frac{(2\pi)^{\frac{p_k}{2}}}{2|\Sigma_{p_k}(t,k)|^{\frac{1}{2}}} \frac{\Gamma\left(\frac{n_{t,k}-p_k}{2}\right)}{\left(\frac{Y'(t,k)P_{p_k}(t,k)Y(t,k)}{2}\right)^{\frac{n_{t,k}-p_k}{2}}}. \end{aligned}$$

Finalement, nous pouvons facilement montrer que cette densité est bien définie et que la constante de normalisation C est donnée par

$$C = \frac{1}{(2\pi)^{\frac{t-1-m}{2}}} \prod_{k=1}^K \frac{(2\pi)^{\frac{p_k}{2}}}{2|\Sigma_{p_k}(t-1,k)|^{\frac{1}{2}}} \frac{\Gamma\left(\frac{n_{t-1,k}-p_k}{2}\right)}{\left(\frac{Y'(t-1,k)P_{p_k}(t-1,k)Y(t-1,k)}{2}\right)^{\frac{n_{t-1,k}-p_k}{2}}}.$$

Proposition 1 *La densité prédictive du modèle gaussien (1) est donnée par*

$$f(y_t \mid y_{1,t-1}, p_1, \dots, p_K) = \frac{1}{\sqrt{2\pi}} \prod_{k=1}^K g_{p_k}(t, k), \quad (2)$$

où pour tout $n_{t-1,k} > p_k$,

$$g_{p_k}(t, k) = \begin{cases} \frac{|\Sigma_{p_k}(t-1,k)|^{\frac{1}{2}}}{|\Sigma_{p_k}(t,k)|^{\frac{1}{2}}} \frac{\Gamma\left(\frac{n_{t,k}-p_k}{2}\right)}{\Gamma\left(\frac{n_{t-1,k}-p_k}{2}\right)} \frac{\left(\frac{Y'(t-1,k)P_{p_k}(t-1,k)Y(t-1,k)}{2}\right)^{\frac{n_{t-1,k}-p_k}{2}}}{\left(\frac{Y'(t,k)P_{p_k}(t,k)Y(t,k)}{2}\right)^{\frac{n_{t,k}-p_k}{2}}} & \text{si } p_k > 0, \\ \frac{\Gamma\left(\frac{n_{t,k}}{2}\right)}{\Gamma\left(\frac{n_{t-1,k}}{2}\right)} \frac{\left(\frac{Y'(t-1,k)Y(t-1,k)}{2}\right)^{\frac{n_{t-1,k}}{2}}}{\left(\frac{Y'(t,k)Y(t,k)}{2}\right)^{\frac{n_{t,k}}{2}}} & \text{si } p_k = 0, \end{cases}$$

avec

$$\begin{aligned} \Sigma_{p_k}(t, k) &= H'_{p_k}(t, k) H_{p_k}(t, k), \\ A_{p_k}(t, k) &= \Sigma_{p_k}^{-1}(t, k) H'_{p_k}(t, k) Y(t, k), \\ P_{p_k}(t, k) &= I_{n_{t,k}} - H_{p_k}(t, k) [H'_{p_k}(t, k) H_{p_k}(t, k)]^{-1} H'_{p_k}(t, k), \end{aligned}$$

et $I_{n_{t,k}}$ est la matrice d'identité de dimension $n_{t,k}$.

4 Critère de la densité prédictive

Djuric et Kay (1992) ont proposé un critère pour choisir l'ordre d'un modèle autorégressif stationnaire. Celui-ci est nommé *PDC* et il est fondé sur la densité prédictive qui est utilisée dans le but d'élaborer un critère qui minimise la probabilité de l'erreur globale, afin d'obtenir l'ordre autorégressif optimal. Cette même approche bayésienne a été généralisée

par Bentarzi et al. (2008) au cas des modèles autorégressifs à coefficients périodiques (*PAR*). Cette section portera sur le problème de l'identification d'un modèle *SETAR* où nous visons à généraliser le critère de sélection *PDC* proposé par Djuric et Kay (1992) pour identifier un modèle *AR*, au cas d'un modèle *SETAR*.

Rappelons que dans le cas autorégressif linéaire, le critère *PDC* est défini à partir de la densité prédictive qui est donné par (voir Djuric et Kay, 1992)

$$PDC(p) = \sum_{t=2}^n -\log f(y_t | y_1, \dots, y_{t-1}, p)$$

où $f(y_t | y_1, \dots, y_{t-1}, p)$ est la densité prédictive de y_t sachant le passé y_1, \dots, y_{t-1} et l'ordre autorégressif p du modèle candidat. Pour le cas des modèles *SETAR*, l'expression du critère *PDC* défini à partir de la densité prédictive (2) est donnée par

$$PDC(p_1, \dots, p_K) = \sum_{t=t_0}^n -\log f(y_t | y_{1,t-1}, p_1, \dots, p_K),$$

où $t_0 = \max_{k=1, \dots, K} t_{k,0}$, avec $t_{k,0} = \min_{t=2, \dots, n} \{t, \text{tel que } n_{t-1,k} > 0\}$. Nous définissons les ordres optimaux par

$$(\hat{p}_1, \dots, \hat{p}_K) = \arg \left[\min_{(p_1, \dots, p_K) \in \{0,1, \dots, P\}^K} \left\{ -\sum_{k=1}^K \sum_{t=t_0}^n \log(g_{p_k}(t, k)) \right\} \right],$$

où P est un entier positif. Pour des raisons de calcul, nous ne pourrions pas écrire la densité prédictive lorsque nous n'avons pas le nombre minimum d'observations $n_{t-1,k} > p_k$, pour $k = 1, \dots, K$. Pour cette raison, nous proposons d'utiliser les mêmes arguments donnés par Djuric et Kay (1992), pour le calcul du critère *PDC* dans le cas d'un modèle *SETAR*.

Pour $p_1 = \dots = p_K = 0$, nous trouvons que

$$J_{(0,0, \dots, 0)} = -\sum_{k=1}^K \sum_{t=t_0}^n \log(g_0(t, k)),$$

Le nombre d'observations utilisé pour le calcul du *PDC*, dans ce cas, est égal à $n - t_0 + 1$. Si maintenant $(p_1, p_2, \dots, p_K) = (p_1, 0, \dots, 0)$, et afin d'exploiter le même nombre d'observations $n - t_0 + 1$, nous proposons d'utiliser

$$J_{(p_1, 0, \dots, 0)} = -\sum_{i=0}^{p_1-1} \sum_{t=t_{1,i}}^{t_{1,i+1}-1} \log(g_i(t, 1)) - \sum_{t=t_{1,p_1}}^n \log(g_{p_1}(t, 1)) - \sum_{k=2}^K \sum_{t=t_{k,0}}^n \log(g_0(t, k)).$$

Nous procédons de la même manière avec les modèles *SETAR* d'ordres supérieurs. Ainsi, nous pouvons calculer l'estimateur $(\hat{p}_1, \dots, \hat{p}_K)$ comme étant la valeur qui minimise $J_{(p_1, \dots, p_K)}$. D'où

$$(\hat{p}_1, \dots, \hat{p}_K) = \arg \left\{ \min_{(p_1, \dots, p_K) \in \{0,1, \dots, P\}^K} J_{(p_1, \dots, p_K)} \right\},$$

où

$$J_{(p_1, \dots, p_K)} = -\sum_{k=1}^K \sum_{i=0}^{p_k-1} \sum_{t=t_{k,i}}^{t_{k,i+1}-1} \log(g_i(t, k)) - \sum_{k=1}^K \sum_{t=t_{k,p_k}}^n \log(g_{p_k}(t, k)),$$

et $t_{k,i} = \min_{t=2, \dots, n} \{t, \text{tel que } n_{t-1,k} > i\}$.

Références

- [1] Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In : Proceedings of the Second International Symposium on Information Theory. Akademiai Kiadó, Budapest, 267-281.
- [2] Bentarzi, M., Guerbyenne, H., Hemis, R., 2008. Predictive Density Order Selection of Periodic *AR* Models. *Communications in Statistics - Simulation and Computation*, **37**, 1167-1182.
- [3] Chan, K.S., 1993. Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *The annals of statistics*, **21**, 520-533.
- [4] Chan, K. S., Tong, H., 1986. On estimating thresholds in autoregressive models. *Journal of time series analysis*, **7**, 179-190.
- [5] Clements, M. P., Smith, J., 1997. The performance of alternative forecasting methods for *SETAR* models. *International Journal of Forecasting*, **13**, 463-475.
- [6] Clements, M. P., Smith, J., 2000. Evaluating the forecast densities of linear and non-linear models : applications to output growth and unemployment. *Journal of Forecasting*, **19**, 255-276.
- [7] De Gooijer, J.G., 2001. Cross-validation criteria for *SETAR* model selection. *Journal of time series analysis*, **22**, 267-281.
- [8] Djuric, P. M., Kay, S. M., 1992. Order Selection of Autoregressive Models, *IEEE Tranaction on Signal Processing*, **40**, 2829-2833.
- [9] Fenga, L., Politis, D. N., 2015. Bootstrap order selection for *SETAR* models. *Journal of Statistical Computation and Simulation*, **85**, 235-250.
- [10] Galeano, P., Peña, D., 2007. Improved model selection criteria for *SETAR* time series models. *Journal of statistical planning and inference*, **137**, 2802-2814.
- [11] Hansen, B. E., 2011, Threshold Autoregressions in Economics, *Statistics and Its Interface*, **4**, 123-127.
- [12] Hili, O., 1992. Sur les modèles autorégressifs à seuils. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, **314**, 573-576.
- [13] Hurvich, C.M., Shumway, R., Tsai, C., 1990. Improved estimators of Kullback-Leibler information for autoregressive model selection in small samples. *Biometrika*, **77**, 709-719.
- [14] Kapetanios, G., 2001. Model selection in threshold models. *Journal of time series analysis*, **22**, 733-754.
- [15] Öhrvik, J., Schoier, G., 2005. *SETAR* model selection-A bootstrap approach. *Computational Statistics*, **20**, 559-573.
- [16] Robert, C., 2007. *The Bayesian choice : from decision-theoretic foundations to computational implementation*. Springer Science & Business Media.
- [17] Tong, H., 1978. *On a threshold model*. In Pattern Recognition and Signal Processing (ed. C. H. Chen). Amsterdam : Sijthoff and Noordhoff, 575-586.
- [18] Tong, H., Lim, K.S., 1980. Threshold autoregression, limit cycles and cyclical data (with discussion). *Journal of the Royal Statistical Society. Series B (Methodological)*, 245-292.

-
- [19] Tong, H., 1983. *Threshold models in non-linear time series analysis*. Springer Lecture Notes in Statistics, Vol. 21. New York : Springer.
- [20] Tong, H., 1990. *Non-linear Time Series : A Dynamical System Approach*. Oxford University Press, Oxford.
- [21] Tsay, R. S., 1989. Testing and Modeling Threshold Autoregressive Processes. *Journal of American Statistical Association*, **84**, 231-240.
- [22] Tsay, R. S., 1998. Testing and modeling multivariate threshold models. *Journal of the american statistical association*, **93**, 1188-1202.
- [23] Wong, C.S., Li, W.K., 1998. A note on the corrected Akaike information criterion for the threshold autoregressive models. *Journal of time series analysis*, **19**, 113-124.