ÉTUDE COMPARATIVE ENTRE PLUSIEURS TESTS DE DÉTECTION DE LA NON LINÉARITÉ DANS LES MODÈLES AUTORÉGRESSIFS D'ORDRE 1

Nabil AZOUAGH [1] & Said El MELHAOUI [2]

[1] *Université Mohammed Premier, Faculté des Sciences, Département Mathématiques et Informatique.60000 BP 717 Oujda Maroc. E-mail: azouaghnabil@gmail.com*
[2] *Université Mohammed Premier, Faculté de Droit, Département d'Économie, Oujda 60 000 B.P. 724 Maroc. E-mail: elmelhaoui.s@gmail.com*

Résumé. Dans ce travail, on s'intéresse aux tests capables de détecter la non linéarité dans les modèles autorégressifs d'ordre 1 ($AR(1)$). Nous proposons une étude comparative par simulation entre plusieurs tests sélectionnés pour ce problème. Ainsi, nous mettons en compétition les tests de Hinich (1982), Keenan (1985), Tsay (1986), Saikkonen and Luukkonen (1988) avec le test pseudo-Gaussien de Allal and El Melhaoui (2006) basé sur l'approche de Le Cam. De plus, nous appliquons ces tests sur la série temporelle classique de lynx du Canada, qui est bien connue comme une série non linéaire, afin d'examiner leurs pouvoir de détection de la non linéarité pour des données réelles.

Mots-clés. Test de non linéarité, modèle exponentiel autorégressif, test pseudo-gaussien, Méthodologie de Le Cam.

Abstract. In this work, we focus on tests able to detect non linearity in autoregressive models of order 1 ($AR(1)$). We propose a comparative study by simulation between several tests selected for this purpose. Thus, we organize a competition between the tests of Hinich (1982), Keenan (1985), Tsay (1986), Saikkonen and Luukkonen (1988) with the pseudo-Gaussian test of Allal and El Melhaoui (2006) based on Le Cam's approach. In addition, we apply these tests to the classical time series of the Canadian lynx, which is well known as a non-linear series, to examine their ability to detect non-linearity for real data.

Keywords. Nonlinearity test, exponential autoregressive model, pseudo-gaussian test, Le Cam approach.

1 Introduction

Parallèlement à l'apparition des modèles non linéaires, de nombreux tests ont été développés afin de détecter les caractéristiques de la non linéarité et de justifier l'utilisation de tel processus aussi sophistiqués à la place des modèles linéaires traditionnels. Ainsi fut l'extraction de plusieurs tests dans le domaine temporel et dans le domaine fréquentiel pour la détection d'un type spécifique de non linéarité ou pour une dépendance générale non linéaire, voir Tong (1990).

En absence d'un moyen analytique général pour comparer les différents types de tests de non linéarité, spécialement dans le cas réel des échantillons finis, nous ne pouvons qu'effectuer cette comparaison en se basant sur des données simulées pour étudier le comportement des tests et choisir celui qui montre les meilleures performances.

Ainsi, dans cette étude comparative nous avons sélectionné cinq tests, ceux de : Hinich (1982), Keenan (1985), Tsay (1986), Saikkonen and Luukkonen (1988), et le test pseudo-Gaussien développé par Allal and El Melhaoui (2006). L'investigation de ces tests, nous a permis de montrer que la statistique de test de Allal and El Melhaoui (2006), dédiée à détecter la non linéarité de type exponentielle autoregressive d'ordre 1 (EXPAR(1)), dépend d'une nuisance non identifiable sous l'hypothèse nulle notée β. Cette nuisance n'apparait que sous l'hypothèse alternative.

Dans ce travail, nous dérivons un nouveau test pseudo-Gaussien à partir de celui de Allal and El Melhaoui (2006) en contournant le problème de la nuisance β non identifiable sous l'hypothèse nulle. Puis, nous passons à l'étude comparative des performances de test pseudo-Gaussien avec les autres tests de non linéarité sélectionnés dans ce travail. Finalement, nous présentons les résultats de l'application des tests sur les données de lynx Canadien.

2 Principaux résultats

2.1 Test pseudo-gaussien du modèle $AR(1)$ contre un $EXPAR(1)$

Nous rappelons les notations et la statistique de test pseudo-Gaussien construit par Allal and El Melhaoui (2006).

Le problème de test consiste à tester l'hypothèse nulle d'une dépendance linéaire AR(1) contre l'hypothèse alternative d'une dépendance non linéaire de type EXPAR(1). Ce dernier modèle a été introduit par Haggan and Ozaki (1981) comme une tentative de reproduire les caractères non linéaire des vibrations aléatoires. Un modèle exponentiel autorégressif d'ordre 1 est la solution de l'équation aux différences stochastiques de la forme :

$$X_t = \left(\pi + \beta \exp(-\varphi X_{t-1}^2)\right) X_{t-1} + \varepsilon_t, \tag{1}$$

avec $(\pi, \beta, \varphi)'$ est le vecteur des paramètres autorégressifs et de la composante exponentielle $\varphi \geq 0$, et $\{\varepsilon_t;\ t \in \mathbb{Z}\}$ est une suite i.i.d.

Supposons que $X_0^{(n)}$ est observée et soit f la densité de probabilité de ε_t et $I(f)$ l'information de Fisher. Soient les résidus

$$Z_t^{(n)} = X_t^{(n)} - \theta X_{t-1}^{(n)}; \qquad t = 1, \ldots, n, \text{ et } \theta = \pi + \beta\ .$$

Considérons la suite locale $(\pi+\frac{\gamma^{(n)}}{\sqrt{n}},\beta,\frac{\delta^{(n)}}{\sqrt{n}})$, où $\tau^{(n)}=(\gamma^{(n)},\delta^{(n)})'\in\mathbb{R}*\mathbb{R}^{+}$ est une suite de vecteurs vérifiant: $\sup_{n\in\mathbb{N}}(\tau^{(n)'}\tau^{(n)})<\infty$. On signale que le problème de test est unilatéral. On s'intéresse à :

l'hypothèse nulle : $H_f^{(n)}(\pi,\beta,0)$, contre l'alternative: $H_f^{(n)}(\pi+\frac{\gamma^{(n)}}{\sqrt{n}},\beta,\frac{\delta^{(n)}}{\sqrt{n}})$

D'après Allal and El Melhaoui (2006) la statistique de test suit la loi normale centrée réduite et sa forme explicite est :

$$T\left(\theta,\beta\right)=\Delta_f^{(n)'}(\theta,\beta)\frac{\Gamma_f^{-1}(\theta,\beta)e_1}{\sqrt{e_1'\Gamma_f^{-1}(\theta,\beta)e_1}};e_1=(0,1)' \qquad (2)$$

avec,

$$\Delta_f^{(n)}(\theta,\beta)=\frac{1}{\sqrt{n}}\left(\begin{array}{c}\sum_{t=1}^{n}X_{t-1}\phi_f(Z_t)\\ \sum_{t=1}^{n}-\beta X_{t-1}^3\phi_f(Z_t)\end{array}\right),\text{ et }\ \Gamma_f(\theta,\beta)=I(f)\left(\begin{array}{cc}E(X_0^2) & -\beta E(X_0^4)\\ -\beta E(X_0^4) & \beta^2E(X_0^6)\end{array}\right);$$

Le calcul montre que la statistique de test construit par Allal and El Melhaoui (2006) dépend d'une nuisance β non identifiable sous l'hypothèse nulle à travers son signe. En effet, notons $\Gamma^*_{f,21}=\beta^{-1}\Gamma_{f,21}$, $\Gamma^*_{f,22}=\beta^{-2}\Gamma_{f,22}$ and $\Delta^{(n)*}_{f,II}=\beta^{-1}\Delta^{(n)}_{f,II}$, la formule explicite de la statistique de test est :

$$T\left(\theta,\beta\right)=\frac{\beta}{|\beta|}\times\frac{(\Gamma_{f,11}\Delta^{(n)*}_{f,II}-\Gamma^*_{f,21}\Delta^{(n)}_{f,I})}{[(\Gamma_{f,11})^2\Gamma^*_{f,22}-\Gamma_{f,11}(\Gamma^*_{f,21})^2]^{1/2}}=sign(\beta)T\left(\theta\right) \qquad (3)$$

Afin de contourner ce problème nous proposons d'estimer cette nuisance sous l'hypothèse alternative et de la remplacer dans la nouvelle statistique de test (3). En effet, à l'aide de la contiguïté issue de la propriété de la normalité asymptotique local établie par Allal and El Melhaoui (2006), on montre que l'estimation du paramètre nuisible β sous l'hypothèse alternative est asymptotiquement équivalente à son estimation sous l'hypothèse nulle. Puis, on complète par l'estimation de θ, puisque il est identifiable sous l'hypothèse nulle. Finalement nous établissons les caractéristiques asymptotiques du test basé sur la statistique (3) en remplaçant θ et β par leurs estimateurs.

2.2 Extrait des simulations

Nous présentons dans cette section un extrait des résultats de simulation du niveau et de la puissance empirique sous une densité d'innovation gaussienne. Des résultats de simulation supplémentaire seront présentés sous des densités Logistique et Double exponentielle.
Par la suite, notons le test pseudo-Gaussien basé sur la statistique (3) par PS, le test de

Hinich par H, le test de Keenan par K, le test de Tsay par T et celui de Saikkonen and Luukkonen par SL.

Soit le modèle AR(1) avec $|\theta| < 1$ condition de régularité (stationnarité du modèle sous H_0) :

$$X_t = \theta X_{t-1} + \varepsilon_t. \tag{4}$$

Le Tableau 1 donne les niveaux empiriques des tests pour les données générées par le modèle (4) pour différents paramètres θ au niveau $\alpha = 0.05$, on a simulé 1000 pseudo échantillons de taille 80, 100 et 150, la densité d'innovation utilisée est gaussienne centrée réduite.

θ	Sample size	H	K	T	SL	PS
0.3	80	0.066	0.039	0.047	0.043	0.047
	100	0.059	0.036	0.038	0.054	0.043
	150	0.053	0.038	0.03	0.048	0.051
0.6	80	0.071	0.024	0.025	0.035	0.045
	100	0.068	0.042	0.05	0.048	0.048
	150	0.058	0.042	0.036	0.033	0.047
0.9	80	0.078	0.02	0.012	0.079	0.011
	100	0.063	0.02	0.023	0.065	0.012
	150	0.057	0.013	0.019	0.044	0.008

Tableau 1. Niveau empirique des tests de Hinich (H), Keenan (K), Tsay (T), Saikkonen and Luukkonen (SL) et pseudo-Gaussien (PS) sous une dépendance $AR(1)$ au niveau 0.05.

D'après les résultats du Tableau 1, on constate que le test PS ainsi que les tests K et T sont contrôlés par le niveau $\alpha = 0.05$. De plus, on remarque que la validité de test H est compromise pour les échantillons de petite taille. Concernant le test SL, dans le cas $\theta = 0.9$, on observe quelques déviations dues au rapprochement du bord de la condition de stationnarité $|\theta| < 1$.

Considérons les modèles non linéaires suivants: EXPAR (modèle 1 et modèle 2), Bilinéaire, Non Linéaire Moyenne Mobile (NLMA) et le modèle à seuil Autorégressif connu sur le nom Threshold Autorégressif (TAR) :

$$\text{EXPAR(1), modèle 1} : X_t = \left(-0.5 + 1.1\exp(-0.1X_{t-1}^2)\right) X_{t-1} + \varepsilon_t. \tag{5}$$

$$\text{EXPAR(1), modèle 2} : X_t = \left(-1.2 + 0.7\exp(-0.1X_{t-1}^2)\right) X_{t-1} + \varepsilon_t. \tag{6}$$

$$\text{Bilinear} : \quad X_t = 0.5X_{t-1} + 0.5\varepsilon_{t-1}X_{t-1} + \varepsilon_t. \tag{7}$$

$$\text{NLMA} : \ X_t = 0.8\varepsilon_{t-1}\varepsilon_{t-2} + \varepsilon_t. \tag{8}$$

$$\text{TAR} : \begin{cases} X_t = -0.6X_{t-1} + \varepsilon_t, & \text{if } X_{t-1} \leq 1 \\ X_t = 0.5X_{t-1} + \varepsilon_t, & \text{if } X_{t-1} > 1 \end{cases} \tag{9}$$

Le Tableau 2 donne les puissances empiriques des tests pour les données générées par les modèles (5), (6), (7), (8) et (9) au niveau $\alpha = 0.05$, on a simulé 1000 pseudo échantillons de taille 100 et 150, la densité d'innovation utilisée est gaussienne centrée réduite.

Modèles	Sample size	H	K	T	SL	PS
	100	0.066	0.038	0.097	0.461	0.510
EXPAR, Model 1	150	0.076	0.045	0.104	0.672	0.684
	100	0.61	0.034	0.036	0.074	0.721
EXPAR, Model 2	150	0.71	0.027	0.031	0.064	0.887
	100	0.369	0.399	0.343	0.211	0.340
Bilinear	150	0.378	0.466	0.421	0.259	0.423
	100	0.147	0.180	0.168	0.192	0.208
NLMA	150	0.185	0.188	0.204	0.222	0.246
	100	0.100	0.960	0.900	0.384	0.354
TAR	150	0.146	0.990	0.980	0.434	0.434

Tableau 2. Puissance empirique des tests de Hinich (H), Keenan (K), Tsay (T), Saikkonen and Luukkonen (SL) et pseudo-Gaussien (PS) sous une dépendance $EXPAR(1)$ au niveau 0.05.

Concernant le modèle EXPAR (5) et (6), les résultats de la simulation du Tableau 2 montrent la stabilité et la haute performance de la puissance empirique estimée de la procédure pseudo-Gaussienne par rapport aux autres tests. Pour les tests de Keenan et de Tsay, on peut remarquer qu'ils sont biaisés, ces tests ont montré la plus faible puissance empirique sur l'ensemble de cette simulation. Concernant le test de Saikkonen and Luukkonen et le test de Hinich, ils ont montré une amélioration significative par rapport aux tests de Keenan et Tsay. Néanmoins, cette amélioration reste instable puisque leurs puissances empirique varient d'un modèle à l'autre.

Pour la détection du modèle bilinéaire (7) on remarque que les tests pseudo-Gaussiens, Keenan et Tsay sont les statistiques les plus puissantes avec un léger avantage pour le test de Keenan. Le test de Hinich vient ensuite avec une bonne performance et la plus faible puissance empirique est notée pour le test de Saikkonen and Luukkonen.

En ce qui concerne le modèle NLMA (8), on souligne que les performances des tests sélectionnés dans cette étude pour détecter ce type de non linéarité est généralement faible où le test pseudo-Gaussien a montré la meilleure puissance empirique par rapport aux autres.

Pour le modèle TAR (9), bien que les tests de Keenan et de Tsay présentent les puissances empiriques les plus élevées en atteignant le voisinage de 0.99 à la taille d'échantillon

150, les tests pseudo-Gaussien et Saikkonen and Luukkonen sont assez bons pour détecter le modèle TAR (9), alors que le test de Hinich montre la plus faible puissance dans cet ensemble.

2.3 Application des tests sur les données de lynx du Canada

La série chronologique de lynx du Canada est une série classique importante qui présente le nombre de lynx canadiens 'piégés' dans le district du fleuve Mackenzie, dans le nord-ouest du Canada, pour la période 1821-1934. Une modélisation de cette série a été proposée par Haggan and Ozaki (1981) en utilisant un modèle EXPAR. Cependant aucune justification mathématique n'a été fournie par les auteurs pour justifier l'utilisation de ce modèle non linéaire (EXPAR) au lieu d'un modèle linéaire simple.

En appliquant les tests sélectionnés dans ce travail sur la série chronologique de lynx du canada on trouve que les tests de Hinich , Keenan (K), Tsay et Saikkonen and Luukkonen n'ont pas rejeté la linéarité des données. Par contre, le test pseudo-Gaussien a pu révéler la non linéarité dans les données de lynx et justifier ainsi le choix de modèle EXPAR.

Finalement, on constate que le test pseudo-Gaussien possède généralement une bonne performance par rapport aux autres tests en terme de la validité et la puissance empirique, de plus, il a bien détecté la non linéarité des données lynx. Les bons résultats du test pseudo-Gaussien nous motivent dans l'avenir à généraliser ce test pour un ordre quelconque $p \geq 1$.

Bibliographie

[1] Allal, J., & El Melhaoui, S. (2006). Optimal detection of exponential component in autoregressive models. Journal of Time Series Analysis, 27(6), 793-810.

[2] Hinich, M. J. (1982). Testing for Gaussianity and linearity of a stationary time series. Journal of time series analysis, 3(3), 169-176.

[3] L. Le Cam, (1986),*Asymptotic methods in statistical decision theory*, Springer-verlag, New York.

[4] Keenan, D. M. (1985). A Tukey nonadditivity-type test for time series nonlinearity. Biometrika, 72(1), 39-44.

[5] Saikkonen, P., & Luukkonen, R. (1988). Lagrange multiplier tests for testing nonlinearities in time series models. Scandinavian Journal of Statistics, 55-68.

[6] Tong, H. (1990). *Non-linear time series: a dynamical system approach*, Oxford University Press, Oxford.

[7] Tsay, R. S. (1986). Nonlinearity tests for time series. Biometrika, 73(2), 461-466.

[8] V. Haggan, and T. Ozaki, (1981), Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model, Biometrika, Series 68(1), 189-196.