



Munich Personal RePEc Archive

**Exogenous Shocks and nonlinearity in
the stock exchange series: Application to
the nonparametric modelling of Orange
stock exchange prices**

CHIKHI, Mohamed

University of Ouargla

2017

Online at <https://mpra.ub.uni-muenchen.de/76815/>
MPRA Paper No. 76815, posted 14 Feb 2017 01:39 UTC

Chocs exogènes et non linéarités dans les séries boursières: Application à la modélisation non paramétrique du cours de l'action Orange

Mohamed CHIKHI¹

Résumé

Cet article vise à analyser le comportement cyclique de la série du cours de l'action Orange du 03/01/2000 à 02/02/2017 par la recherche de la non linéarité à travers d'une classe de modèles non paramétriques hétéroscédastiques, notée NAR-ARCH. L'identification des modèles non paramétriques nécessite une sélection rigoureuse des coefficients de Markov et le choix de la fenêtre qui détermine le degré de lissage de l'estimateur.

Mots-clefs: Erreur de prédiction finale, noyau, fenêtre, processus autorégressif fonctionnel hétéroscédastique, action Orange.

Classification JEL : C14, C22, C58, G17.

Abstract

This paper aims to analyze the cyclical behavior of stock exchange Orange prices from 01/03/2000 to 02/02/2017 by the research of nonlinearities through a class of heteroscedastic non parametric models. The identification of non parametric models requires the selection of the Markov coefficients and the choice of bandwidth, which determines the degree of estimator's smoothing.

Keywords: Final Prediction Error, kernel, bandwidth, heteroscedastic functional autoregressive process, stock exchange Orange.

JEL Classification : C14, C22, C58, G17.

¹ LAQSEF, Université Kasdi Merbah Ouargla. Faculté des Sciences Economiques, des Sciences Commerciales et Sc. de Gestion. Route de Ghardaïa, BP. 511 30000 Ouargla, E-mail: chikhi.m@univ-ouargla.dz

1 Introduction

Les cours boursiers exhibent une structure particulière dont la prise en compte est fondamentale pour aborder le problème de la modélisation. Ainsi, la présence des non linéarités dans les séries financières a des implications importantes particulièrement concernant la propriété d'efficience faible des marchés. En effet, si une série est caractérisée par la dynamique non linéaire, ceci implique des dépendances non linéaires entre des observations.

Les séries financières, notamment les rentabilités boursières, ne sont plus prévisibles par des modèles linéaires puisque ces processus s'appuient sur des hypothèses restrictives, comme par exemple la présence d'une distribution symétrique et la présence d'une structure linéaire. Les modèles linéaires gaussiens ne peuvent générer que des comportements symétriques de la série et ne peuvent pas reproduire les effets de la présence de coûts de transaction variables et d'anticipations hétérogènes des agents sur les marchés boursiers.

Le développement des méthodes de régression non paramétriques ont donné une base théorique solide à l'analyse non linéaire des séries temporelles comme l'indique Robinson (1983). Elles sont considérées comme une alternative au problème de la prise en compte de la non-linéarité.

Les méthodes non paramétriques ont connu un essor important depuis les travaux de Bosq (1979), Collomb (1980) et Robinson (1983). Leur efficacité pratique a été analysée par Bosq et Lecoutre (1992), Härdle (1990), Gannoun (1991), Rosa (1993) et Matzner-Lober (1997). En 1990, Auestad et Tjostheim les ont appliquées pour estimer la moyenne et la variance conditionnelles d'un processus autorégressif non linéaire. Ils choisissent les critères de la variance résiduelle et du *FPE* « *Final Prediction Error* » pour déterminer leurs coefficients de Markov. Cheng et Tong (1992) proposent un critère de validation croisée et Vieu (1995) celui de la validation croisée pénalisée. Cependant, ce n'est qu'en 1998 que Tschernig et Yang suggèrent d'utiliser le critère *FPE* non paramétrique corrigé en introduisant un nouveau terme dans le critère *FPE* retenu par Auestad et Tjostheim (1990). Chikhi et Terraza (2002) appliquent l'estimateur linéaire local à l'action France Télécom pour comparer les résultats en prévision avec ceux obtenus à partir des modèles à mémoires longue. Chikhi et Diebolt (2009) appliquent cette méthode à la série mensuelle de la masse monétaire de la banque centrale de l'Allemagne - La Reichsbank-.

Notre article s'insère dans la lignée de ces travaux et s'intéresse plus particulièrement à l'analyse non paramétrique du comportement cyclique des fluctuations du cours de l'action Orange par les modèles non linéaires hétéroscédastiques.

Nous présentons dans la section 2 l'estimateur linéaire local du processus autorégressif fonctionnel avec erreur ARCH non paramétrique. Afin de détecter une éventuelle non linéarité et la nature du choc exogène, nous présentons une procédure d'estimation non paramétrique et nous l'appliquons à la série du cours de l'action Orange (section 3).

2 Estimateur à noyau d'un processus NAR-ARCH

L'estimateur non paramétrique du processus autorégressif non linéaire $\{Y_t\}_{t \geq 0}$ avec erreur ARCH s'écrit sous la forme (Tschernig and Yang, 1998):

$$Y_t = f(X_t) + \sigma^{1/2}(X_t)\varepsilon_t \quad (1)$$

Où $X_t = (Y_{t-i_1}, \dots, Y_{t-i_k})^T$ est le vecteur de toutes les variables retardées, $i_1 < \dots < i_k$ et $\varepsilon_t \sim IID(0,1)$ et la fonction $f(\cdot)$ représente l'espérance conditionnelle $E(Y_t | X_t)$.

Nous donnons quelques hypothèses sur le processus (1) (Doukhan, 1994) :

(H1) Pour tout entier $M \geq i_k$, le processus $X_{M,t} = (Y_{t-1}, \dots, Y_{t-M})^T$ est strictement stationnaire et β -mélangeant avec $\beta(n) \leq c_0 n^{-(2+\delta)/\delta}$, $\delta > 0, c_0 > 0$:

$$\beta(n) = E \sup \left\{ |P(A | \mathfrak{R}_M^k) - P(A)| : A \in \mathfrak{R}_{n+k}^\infty \right\}$$

Où \mathfrak{R}_t^i est σ -algèbre de $X_{M,t}, X_{M,t+1}, \dots, X_{M,t^i}$.

(H2) La distribution stationnaire du processus $X_{M,t}$ a une densité $\mu_M(x_M), x_M \in \mathcal{R}^M$, continue.

(H3) La fonction $f(\cdot)$ est deux fois continûment différentiable tandis que $\sigma(\cdot)$ est continue et positive.

(H4) Les quatre premiers moments de $\{\varepsilon_t\}_{t \geq i_k}$ sont finis.

Une technique non paramétrique est proposée pour estimer la fonction de la moyenne conditionnelle et celle de la volatilité : l'estimateur linéaire local $\hat{f}(x)$

Notons :

$$X_t = (Y_{t-i_1}, Y_{t-i_2}, \dots, Y_{t-i_k})^T = (X_{t_1}, X_{t_2}, \dots, X_{t_k})^T$$

et

$$Y = (Y_{i_k}, Y_{i_{k+1}}, \dots, Y_n)^T$$

Pour tout $x \in \mathcal{R}^k$, l'estimateur s'écrit donc comme suit :

$$\hat{f}(x) = e^T (Z^T W Z)^{-1} Z^T W Y$$

où :

$$Z = \begin{pmatrix} 1 & \dots & \dots & 1 \\ X_{i_k} - x & \dots & \dots & X_n - x \end{pmatrix}^T$$

$$e = (1 \quad 0_{1 \times k})^T$$

$$W = \text{diag} \left\{ \frac{1}{n} K_h(X_i - x) \right\}_{i=i_k}^n$$

avec : $n' = n - i_k + 1$

(H5) $K : \mathcal{R} \rightarrow \mathcal{R}$ est un noyau positif symétrique et $h = h_n$ est un nombre positif avec $h \rightarrow 0, nh^k \rightarrow \infty$ quand $n \rightarrow \infty$.

Notons : $\|K\|_2^2 = \int K^2(u) du, \sigma_K^2 = \int K(u) u^2 du$. Pour tout $x \in \mathcal{R}^k$, nous avons:

$$K_h(x) = \frac{1}{h^k} \prod_{j=1}^k K\left(\frac{x_j}{h}\right)$$

La convergence et la normalité asymptotiques de l'estimateur linéaire local de ce modèle sont données par :

$$\sqrt{n h^k} \left\{ \hat{f}(x) - f(x) - h^2 b_i(x) \right\} \xrightarrow{D} N(0, V(x)) \quad (2)$$

quand $n \rightarrow \infty$, avec le biais asymptotique :

$$b_c(x) = \frac{\sigma_K^2}{2} \left\{ \frac{2 \nabla f(x)^T \nabla \mu(x)}{\mu(x)} \right\} \quad (3)$$

$$b_i(x) = \frac{\sigma_K^2}{2} \text{Tr}[\nabla^2 f(x)] \quad (4)$$

et la variance asymptotique :

$$V(x) = \frac{\sigma(x)}{\mu(x)} \|K\|_2^{2k} \quad (5)$$

où $\|K\|_2^2 = \int K^2(x_j) dx_j$ and $\sigma_K^2 = \int x_j^2 K(x_j) dx_j$

Le taux de convergence de cet estimateur dépend de la dimension de la matrice des régresseurs. Il nécessite la continuité de la densité pour un taux de convergence optimale (Härdle *et al*, 1998).

Afin d'appliquer l'estimateur linéaire local du modèle *NAR-ARCH*, nous spécifions la fenêtre h et les coefficients de Markov i_1, \dots, i_k . Le choix de la fenêtre est crucial pour la qualité de l'estimation puisqu'il détermine le niveau de lissage de la fonction estimée. Pour estimer une fenêtre optimale, les spécialistes optent régulièrement pour la méthode de validation croisée de (Vieu, 1994). Une approche alternative basée sur le critère (*FPE*) asymptotique est proposée ici, puisque ce dernier donne une meilleure vitesse de convergence que le critère de validation croisée habituel.

Soit $\{\tilde{Y}_t\}$ d'autres séries ayant la même distribution que $\{Y_t\}$, mais indépendantes de $\{Y_t\}$. Le critère *FPE* d'un estimateur \hat{f} de f est défini par :

$$FPE(\hat{f}) = \lim_{t \rightarrow \infty} E \left[\left\{ \tilde{Y}_t - \hat{f}(\tilde{X}_t) \right\}^2 w(\tilde{X}_{M,t}) \right] \quad (6)$$

(H6) Pour la fonction du poids $w: R^M \rightarrow R$, elle est continue positive et son support est compact, nous supposons aussi que $\mu(x_M) > 0$ pour tout $x_M \in \text{supp}(w)$.

Quand le processus $\{Y_t\}$ est asymptotiquement non linéaire stationnaire et \hat{f} est un estimateur non paramétrique, le critère *FPE* est non linéaire.

Pour cet estimateur, nous avons :

$$FPE(h) = AFPE(h) + o(h^4 + \frac{1}{n h^k}) \quad (7)$$

où :

$$AFPE(h) = A + b(h)B + c(h)C$$

avec :

$$A = \int \sigma(x) w(x_M) \mu(x_M) dx_M, B = \int \sigma(x) w(x_M) \mu(x_M) / \mu(x) dx_M,$$

$$C = \int r^2(x)w(x_M)\mu(x_M)dx_M$$

$$r(x) = Tr\{\nabla^2 f(x)\}$$

où :

$$b(n) = \|K\|_2^{2k} n^{-1} h^{-k}, c(h) = \sigma_K^4 h^4 / 4$$

La première partie de la formule (7) présente l'espérance de la variance du bruit blanc des processus $\{Y_t\}$ et $\{\tilde{Y}_t\}$. Le deuxième terme B donne l'espérance de la variance de l'estimateur non paramétrique \hat{f} ; alors que le troisième capture son biais estimé. Cependant, pour un échantillon donné de taille n , nous obtenons une expression qui détermine une fenêtre optimale h_{opt} en minimisant la somme du deuxième et du troisième terme (Yang et Tschernig (1999):

$$h_{opt} = \left\{ k \|K\|_2^{2k} B n^{-1} C^{-1} \sigma_K^{-4} \right\}^{1/(k+4)} \quad (8)$$

Et le critère $AFPE$ minimal est :

$$AFPE_{opt} = A + (k^{-k/(k+4)} + \frac{1}{4} k^{4/(k+4)}) \left\{ \|K\|_2^{8k} B^4 n^{-1} C^k \sigma_K^{4k} \right\}^{1/(k+4)}$$

L'équation (8) est nécessaire pour obtenir l'estimateur de h_{opt} et l'expression (7) contient des valeurs inconnues A , B et C . Il est extrêmement important d'estimer ces trois valeurs afin d'estimer le critère $AFPE$, et ainsi la fenêtre optimale associée :

$$\hat{A} = n^{-1} \sum_{i=i_k}^n \left\{ Y_i - \hat{f}(X_i) \right\}^2 w(X_{M,i}) \quad (9)$$

$$\hat{B} = n^{-1} \sum_{i=i_k}^n \left\{ Y_i - \hat{f}(X_i) \right\}^2 w(X_{M,i}) / \hat{\mu}(X_i) \quad (10)$$

L'estimateur \hat{f} utilise des fenêtres d'ordre $n^{-1/(k+4)}$ et $\hat{\mu}(X_i)$ est un estimateur à noyau de la densité.

Sous les hypothèses (H1)-(H6), quand $n \rightarrow \infty$,

$$\hat{A} = A + \left\{ \|K\|_2^{2k} - 2K(0)^k \right\} n^{-1} h^{-k} B + C \sigma_K^4 h^4 / 4 + o\left\{ h^4 + n^{-1} h^{-k} \right\} + O_p\left\{ n^{-1/2} \right\} \quad (11)$$

Notons qu'avec une fenêtre h de la forme $cn^{-1/(k+4)}$, l'estimateur non paramétrique \hat{A}_a converge vers A et sa convergence est de l'ordre \sqrt{n} si $k \leq 4$. Le deuxième et le troisième terme seront de l'ordre $O\left\{ n^{-1/2} \right\}$.

L'estimateur non paramétrique du FPE asymptotique est :

$$AFPE = \hat{A} + 2K(0)^k n^{-1} h_{opt}^{-k} \hat{B} \quad (12)$$

\hat{A}_a est estimée en utilisant la fenêtre optimale h_{opt} , alors que \hat{B} en utilisant n'importe quelle fenêtre de l'ordre $n^{-1/(k+4)}$.

Yang et Tschernig (1999) montrent que l'estimateur du critère vérifie toutes les propriétés statistiques. Pour pénaliser la surestimation des coefficients de Markov, ils suggèrent d'introduire un facteur de correction :

$$CAFPE = AFPE \left(1 + \frac{k}{\frac{4}{n^{k+4}}} \right) \quad (13)$$

Où la correction dépend du nombre des retards k et le nombre d'observation n . Il s'agit donc de choisir le sous-ensemble $\{\hat{i}_1, \dots, \hat{i}_k\}$ sous les hypothèses (H1)-(H6), telle que Tschernig et Yang (1998, p.7) :

$$P\{\hat{k} = k, \hat{i}_l = i, l = 1, \dots, k\} \xrightarrow{n \rightarrow \infty} 1$$

où

$$e^T = (1 \quad 0_{1 \times k})^T,$$

$$Z_h = \begin{pmatrix} 1 & \dots & 1 \\ X_{k-1} - x & \dots & X_{n-h} - x \end{pmatrix}^T$$

$$W_h = \text{diag}\{K_h(X_{i-1} - x)/n\}_{i=k}^{n-h+1}$$

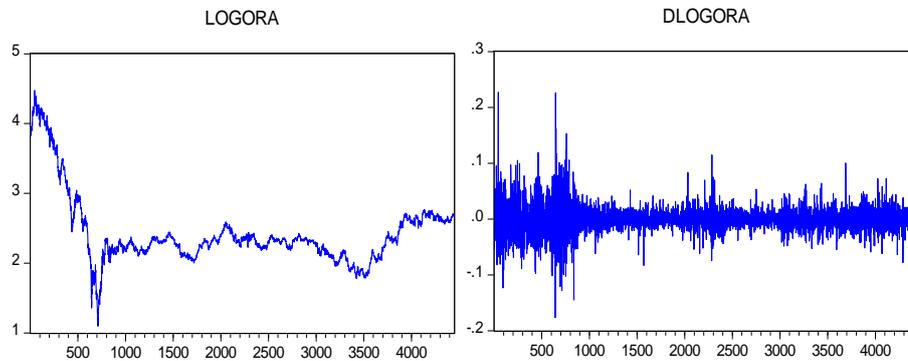
et

$$Y_h = (Y_{k+h-1} \quad \dots \quad Y_n)^T$$

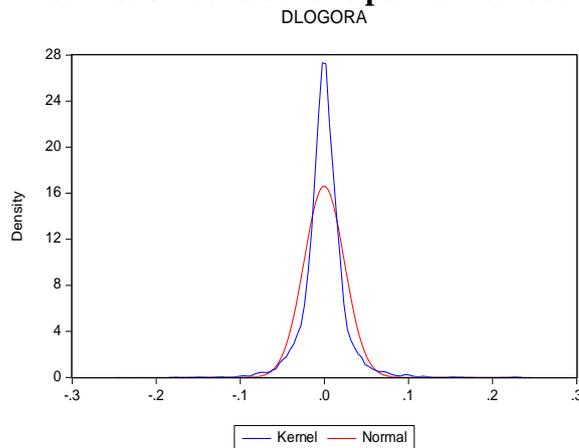
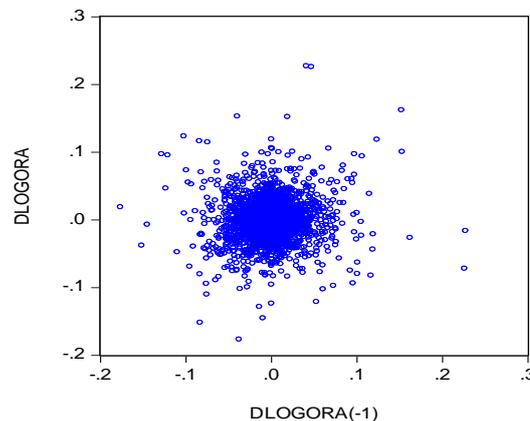
3 Analyse empirique

Nous appliquons la méthode proposée sur la série du cours de l'action Orange en logarithme qui couvre la période du 03/01/2000 à 02/02/2017 (4449 observations). Les résultats des tests de Philips-Perron (Philips et Perron (1988), Elliott-Rothenberg-Stock (Elliott, Rothenberg et Stock (1996)) et Ng-Perron (Ng et Perron (2001)) reportés dans le tableau 1 montrent que cette série est caractérisée par la présence d'une racine unitaire. La série est finalement différenciée pour obtenir les rendements (voir figure 1).

Figure 1 – Evolution de la série journalière de la série du cours de l'action Orange en logarithme et de la rentabilité



Au regard du tableau 2, l'hypothèse de normalité des rendements est clairement rejetée (voir aussi Figure 2). L'asymétrie constatée peut être le signe de la présence de non-linéarités dans le processus d'évolution des rendements. Le diagramme de dispersion de la série (Figure 3) ne se présente pas sous la forme d'un ellipsoïde régulier et confirme la non-linéarité de la série. En outre, la distribution de cette série est leptokurtique. La série des rentabilités est caractérisée par la présence d'un effet ARCH (voir tableau 2).

Figure 2 – Estimation de la densité par la méthode du noyau**Figure 3 – Dispersion des variations successives des rentabilités boursières**

Au regard du tableau 3, le test BDS (Brock, Dechert et Scheinkman (1987)), qui permet de détecter la présence d'une structure de dépendance non linéaire à court terme, rejette l'hypothèse de marche aléatoire. En effet, les statistiques de *BDS*, sont strictement supérieures à la valeur critique au seuil de 5%. Par conséquent, les prix de l'action Orange sont prévisibles à court terme.

Nous analysons le comportement cyclique des rentabilités en travaillant sur des horizons plus longs. Nous estimons alors le coefficient d'intégration fractionnaire par la méthode GPH avec des différentes fenêtres spectrales (Granger et Joyeux (1981) et Hosking (1981)). D'après le tableau 4, il est évident que la série des rentabilités de l'action Orange n'est pas générée par un processus de mémoire longue. Les valeurs de la statistique de Student (avec une puissance de 0.8) sont strictement inférieures à la valeur critique au seuil de 5%. Les rentabilités ne sont donc prévisibles qu'à court terme. L'absence d'une mémoire longue indique que les agents ne peuvent pas anticiper leurs rentabilités à un horizon suffisamment long. En effet, les mouvements observés apparaissent comme le résultat de chocs exogènes transitoires venant affecter le marché c'est à dire que le cours de l'action boursière reviendra vers sa valeur fondamentale, cela signifie également que le processus paramétrique générateur des données est non linéaire mais sa structure n'est pas connue. Devant cet état de fait, nous

pourrions donc utiliser l'approche non paramétrique pour généraliser la classe des modèles paramétriques non linéaires.

L'identification d'un modèle non paramétrique nécessite une sélection des coefficients de Markov et de la fenêtre. Il est donc possible d'appliquer le critère non paramétrique *CAFPE* pour identifier le processus des rendements de l'action Orange. Pour cela, le nombre maximal de retards à considérer est quatre et le retard le plus grand est huit. Il s'agit d'examiner $1 + \sum_{j=1}^4 \binom{8}{j}$ combinaisons

différentes de retards pour la procédure de recherche. Dans cette approche, nous sélectionnons les retards et la fenêtre qui minimisent le critère *CAFPE*.

Nous choisissons donc une fenêtre qui équilibre asymptotiquement le biais et la variance et qui minimise le critère. Pour estimer l'espérance de $b_c(X_t)$, nous estimons les dérivées secondes de (4) en utilisant l'estimateur quadratique local avec la fenêtre :

$$h = 2 \text{var}(Z_t) \left(\frac{4}{k+10} \right)^{\frac{1}{k+12}} n^{-\frac{1}{k+12}}$$

Pour calculer les résidus et la densité qui sont utilisés pour estimer l'espérance de la variance de l'estimateur non paramétrique \hat{f} , nous calculons la fenêtre de Silverman (1986), eq. (4.14), p. 86) :

$$h_s = \text{var}(Z_t) \left(\frac{4}{k+2} \right)^{\frac{1}{k+4}} n^{-\frac{1}{k+4}} \quad (17)$$

De plus, la méthode "*Grid search*" est appliquée pour estimer la fenêtre optimale où la grille couvre l'intervalle $[0, 2h_s, 2h_s]$ dans 24 pas. Nous choisissons alors le noyau gaussien. La fonction du poids sera spécifiée comme suit :

$$w(X_t) = \begin{cases} 1 \\ 0 \end{cases} \text{ if } \begin{cases} \hat{\mu}(X_t) \geq c \\ \hat{\mu}(X_t) < c \end{cases}$$

où la constante c est déterminée telle que n'' observations sont exclues et n' présente le plus grand nombre entier prochain de $0,05n'$.

Le tableau 5 donne les résultats de l'identification du processus de l'espérance conditionnelle. Il contient les retards sélectionnés, les valeurs du critère, de la fenêtre optimale estimées et de la variance résiduelle. Les retards sont sélectionnés en minimisant le critère *CAFPE*. Le critère estimé est égal à 0.00051 et les retards sélectionnés sont : 1, 3, 5 et 8. Notons que l'estimateur utilisé est celui de l'estimateur linéaire local. Nous remarquons que les résidus du modèle NAR (Figure 4) sont caractérisés par une distribution asymétrique et leptokurtique (Figure 5 et tableau 6).

Figure 4 – Résidus du modèle NAR non paramétrique en utilisant les retards 3 et 8

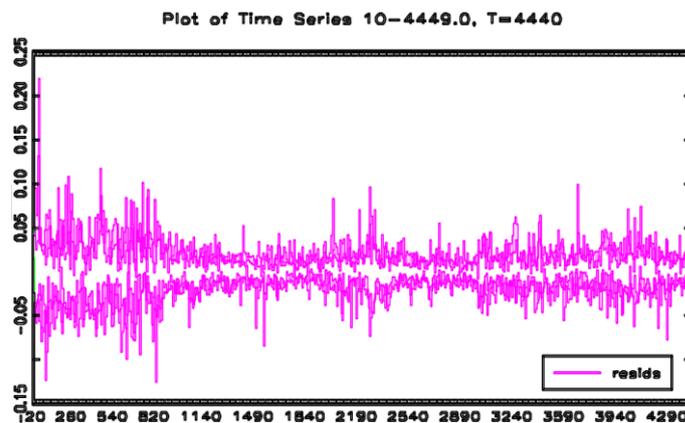
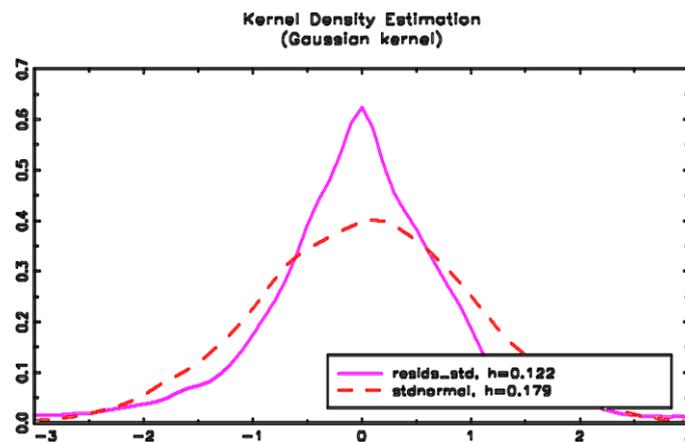


Figure 5 – Estimation de la densité des résidus par la méthode du noyau



Au regard du tableau 7, nous rejetons clairement l'hypothèse de bruit blanc. Ces résidus présentent une structure de dépendance non linéaire car les statistiques de BDS nous ont conduits à rejeter l'hypothèse nulle générateur *iid* pour toutes les dimensions de plongement. Or, les résidus peuvent être modélisés par les modèles *ARCH* (Tableau 6) car les résultats des tests *ARCH LM* sur les résidus du modèle confirment la présence d'un effet *ARCH* ($nR^2 = 13.904 > \chi^2(1)$). Par conséquent, nous estimons la volatilité des rendements de l'action Orange par la méthode du noyau. Le tableau 8 présente les résultats d'estimation de la volatilité conditionnelle. Il contient les retards sélectionnés, la valeur du critère estimée, la fenêtre optimale et la variance résiduelle. Nous avons utilisé aussi l'estimateur linéaire local pour identifier cette volatilité conditionnelle. Les retards sont alors 2,3,7 et 8 et le critère estimé associé est égal à 0.0000014. Nous pouvons remarquer aisément que les résidus du modèle *ARCH* non paramétrique (Figure 6) forment un bruit blanc et ne présentent aucune corrélation sérielle (Tableau 9). En outre, il y'a des signes de non linéarité (Figure 7) où la série a un comportement dynamique asymétrique. Ces dépendances non linéaires dans l'espérance et dans la variance

conditionnelle reproduisent des effets de coûts de transaction variables et d'anticipations hétérogènes des agents sur le marché boursier parisien.

Figure 6 – Résidus du modèle NAR-ARCH non paramétrique en utilisant les retards 3 et 7

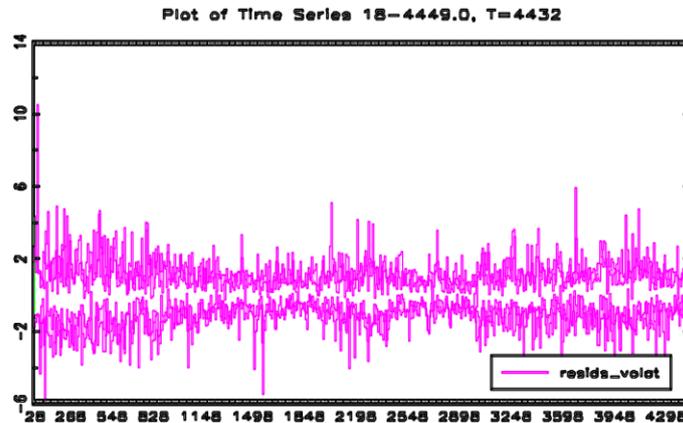
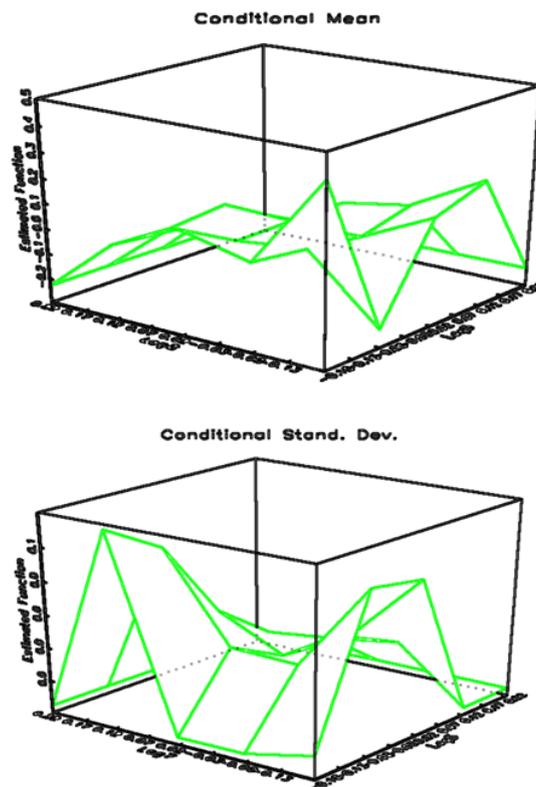


Figure 6 – Fonctions de l'espérance et de la variance conditionnelles



Afin de comparer les performances prévisionnelles du modèle NAR-ARCH proposé et celui de marche aléatoire, nous utilisons le critère de l'erreur moyenne quadratique :

$$QME = H^{-1} \sum_{h=1}^H (\hat{Y}_{n-H+h} - Y_{n-H+h})^2 \quad (17)$$

où h est l'horizon de prévision et H est le nombre total de prévisions correspondant à l'horizon h sur la période prévisionnelle.

Le tableau 10 comporte les résultats de prévisions fournies par notre modèle non paramétrique. Nous remarquons que, quel que soit l'horizon de prévision, la modélisation avec la marche aléatoire est battue par tous le modèle NAR-ARCH non paramétrique. Nous constatons de bonnes capacités prévisionnelles de notre modèle à un horizon de 1, 2, 15, 30, 90 jours.

Étant donné que la série des rentabilités de l'action Orange se caractérise par une présence des dynamiques non linéaires dans les équations de la moyenne et de la variance conditionnelles, et par un phénomène d'hétéroscédasticité, la modélisation non paramétrique permet le calcul des meilleures prévisions à court terme que le modèle de marche aléatoire.

4 Conclusion

Dans cet article, nous avons recherché la présence des non linéarités dans la série des rentabilités du cours de l'action Orange. Dans cette optique, nous avons proposé un modèle non linéaire NAR avec erreur ARCH non paramétrique. Nous avons mis en œuvre l'estimateur linéaire local pour estimer cette classe de modèles non paramétriques en prenant en considération le phénomène de non linéarités au niveau de l'espérance et de la variance conditionnelles. D'après les résultats, les chocs informationnels ont des conséquences transitoires sur la volatilité et notre modèle montre une supériorité évidente sur le modèle de marche aléatoire.

Il existe un problème en estimation fonctionnelle ; il s'agit du choix des coefficients de Markov. La méthode empirique utilisée pour sa détermination ne repose actuellement sur aucun théorème de validation. Les résultats existants admettent en général au moins implicitement que le processus observé est markovien d'ordre k ; hypothèse difficile à vérifier en pratique.

Malgré cette difficulté, les méthodes non paramétriques restent toujours un outil puissant pour étudier les séries temporelles. Si la structure du processus non linéaire n'est pas connue, nous pourrions donc au travers d'une modélisation non paramétrique, tenter de réduire notre ignorance et de minimiser la part du hasard dans les systèmes dynamiques.

Références

- Auestad, B. and Tjostheim, D. (1990), *Identification of nonlinear time series: first order characterization and order determination*, *Biometrika*, 77, 4, 669-687. (1990)
- Bosq, D et Lecoutre, J.P., *Analyse et prévision des séries chronologiques*, Masson, Paris. (1992)
- Bosq, D., *Sur la prédiction non paramétrique de variables aléatoires et mesures aléatoires*, Pub. Interne, UER de Mathématiques, Lille. (1979)
- Bosq, D., *Nonparametric statistics for stochastic processes*, Lecture Notes in statistics, 110, Springer-verlag. (1996)

- Box, G and Pierce, D., *Distribution of residual autocorrelation in autoregressive integrated moving average time series models*, J, Ann., Statist., 6, 461-464. (1970)
- Brockmann, M., *Locally adaptive bandwidth choice for kernel regression estimators*, J. Amer. Statist. Assoc., 88, 1302-1309. (1993)
- Cheng, B. and Tong, H., *On consistent nonparametric order determination and chaos*, Journal of The Royal Statistical Society, Series B, 54, 427-449. (1992)
- Chikhi, M and Diebolt, C., *Nonparametric Analysis of Financial Time Series by the Kernel Methodology*, in: Quality & Quantity. International Journal of Methodology (Springer), 44, pp. 865-880. (2010)
- Chikhi, M and Diebolt, C., *The Reichsbank : A Nonparametric Modelling of Historical Time Series*, in: Applied Economics Letters (Routledge), 16(14), pp. 1409-1414. (2009)
- Chiu, S.T., *Bandwidth selection for kernel estimates with correlated noise*, Statist. Probab. Lett., 8, 347-354. (1989)
- Collomb, G., *Estimation non paramétrique de probabilités conditionnelles*, C.R. Acad. sci. Paris Sér I Math., 291, 427-430. (1980)
- Doukhan, P., *Mixing: Properties and examples*, New York; Springer-Verlag. (1994)
- Elliott, G., Rothenberg, T. J., & Stock, J. H., *Efficient Tests for an Autoregressive Unit Root*. *Econometrica*, 64, 4, 813–836. (1996)
- Engle, R.F., *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*, *Econometrica*, 50(4) 987-1007. (1982)
- Gannoun, A., *Prédiction non paramétrique : médianogramme et méthode du noyau en estimation de la médiane conditionnelle*, *Statistique et Analyse des données*, 16(23), 23-42. (1991)
- Gouriéroux, C., *Modèles ARCH et applications financières*, *Economica*, Paris. (1992)
- Härdle, W and Chen, R., *Nonparametric Time Series Analysis, a selective review with examples*, Proceedings of the 50th session of the ISI, Peking. (1996)
- Härdle, W and Yang, L., *Nonparametric autoregression with Multiplicative Volatility and additive Mean*, Discussion paper 96-62, SFB 373, Humboldt Universität zu Berlin. (1996)
- Härdle, W, Lütkepohl, H and Chen, R., *A review of Nonparametric Time Series Analysis*, Discussion Paper 96-48, SFB 373, Humboldt Universität zu Berlin. (1996)
- Härdle, W, Tsybakov, A. B and Yang, L., *Nonparametric vector autoregression*, *Journal of Statistical Planning and Inference* 68, 221-245. (1998)
- Härdle, W., *Applied nonparametric regression*, Cambridge university press, Cambridge. (1990)
- Lütkepohl, H. and Kräzig, M., *Applied Time Series Econometrics*. (2003)
- Masry, E and Tjostheim, D., *Nonparametric estimation and identification of non-linear ARCH time series: strong convergence and asymptotic normality*, *Thèse de doctorat de l'université de Montpellier II*. (1997)
- Nadaraya, E.A., *On estimating regression*, *Theory probability and their applications*, 9, 134-137. (1964)

- Ng, S and Perron, P., *Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power*, *Econometrica*, 69(6), 1519-1554. (2001)
- Phillips, P.C.B., & Perron, P., *Testing for Unit Roots in Time Series Regression*, *Biometrika*, 75, 335-346. (1988)
- Robinson, P.M., *Nonparametric estimators for time series*, *Journal of Time Series Analysis*, 4, 185-207. (1983)
- Rosa, M. A. C., *Prévision robuste sous une hypothèse ergodique*, Thèse de Doctorat de l'université de Toulouse I. (1993)
- Schwarz, G., *Estimating the dimension of a Model*, *Annals of Statistics*, 6, 461-464. (1978)
- Silverman, B.W., *Density estimation for Statistics and data analysis*, Chapman & Hall. (1986)
- Stute, W., *On almost sure convergence of conditional empirical distribution function*, *Ann. of Prob.*, 14, 891-901. (1986)
- Tjøstheim, D. and Auestad, B., *Nonparametric identification of nonlinear time series: selecting significant lags*, *Journal of American Statistical Association*, 89, 1410-1419. (1994b)
- Tschernig, R and Yang, L., *Nonparametric Lag Selection for Time Series*, *Journal of Time Series Analysis*, forthcoming. (1998)
- Tschernig, R., *Nonlinearities in German Unemployment Rates: A Nonparametric Analysis*, SFB 373 discussion paper 45. (1996)
- Ullah, A., *Nonparametric estimation and hypothesis testing in econometric models*, *Empec*, 13, 223-249. (1988)
- Vieu, P., *order choice in nonlinear autoregressive models*, *statistics, OPA*, 26, 307-328. (1995)
- Watson, G.S., *Smooth regression analysis*, *Sankhyä*, A26, 359-372. (1964)
- Yang, L and Tschernig, R., *Multivariate bandwidth selection for local linear regression*, *Journal of the Royal Statistical Society, Series B*, 61, 793-815. (1999)
- Yang, L and Tschernig, R., *Non- and semiparametric identification of seasonal nonlinear autoregression models*, *Econometric Theory* 18: 1408-1448. (2002)

Annexes statistiques

Tableau 1 – Tests de racine unitaire de Phillips-Perron, Elliott-Rothenberg-Stock and Ng-Perron

	Série	Philips-Perron	Elliott-Rothenberg-Stock	Ng-Perron			
				MZa	MZt	MSB	MPT
H_0 : racine unitaire	Série en logarithmes	-1.407 (-1.94)	0.076 (3.26)	-0.181 (-8.10)	-0.180 (-1.98)	0.994 (0.233)	52.949 (3.17)
	Rentabilités	-62.787 (-1.94)	53.533 (3.26)	-366.018 (-8.10)	-13.52 (-1.98)	0.0369 (0.233)	0.075 (3.17)

(.): les valeurs critiques asymptotiques, Fenêtre optimale : 0.444 pour la série en log et 4.05 pour la série des rentabilités, la méthode d'estimation spectrale utilisée est celle d'Andrews par le noyau de Bartlett. MZa, MZt, MSB, MPT sont les statistiques de Ng-Perron. Le modèle choisi est celui avec constance sans tendance déterministe (à 5%)

Tableau 2 – Caractéristiques statistiques de la série des rentabilités

Skewness	Kurtosis	Jarque-Bera	ARCH(1)
0.457	12.012	15209.64 (0.000)	(0.000)

(.): Probabilité critique.

Tableau 3 – Résultats des tests de BDS sur la série des rentabilités

m	Stat. de BDS
2	20.541
3	26.459
4	30.934
5	34.563
6	37.986
7	41.574
8	45.353
9	49.637
10	54.442

La statistique de BDS est calculée par la méthode de "Fraction of pairs" avec ε égale à 0.7.

Tableau 4 – Estimation semi paramétrique du coefficient d'intégration fractionnaire par la méthode de GPH

La fenêtre	d	Stat. de Student	Prob.
Sans fenêtre	0.035	1.553	0.120
Rectangular	0.006	0.273	0.784
Bartlett	0.012	0.869	0.384
Daniell	0.011	0.673	0.500
Tukey	0.010	0.669	0.503
Parzen	0.015	1.154	0.248
B-priest	0.008	0.418	0.675

d : Le coefficient d'intégration fractionnaire estimé

Tableau 5 – Spécification et estimation non paramétrique de l'espérance conditionnelle en utilisant l'estimateur linéaire local

Sélection non paramétrique des retards		Estimation non paramétrique de l'espérance conditionnelle	
Les retards sélectionnés	1,3,5,8	La fenêtre spécifiée	0.0231
La fenêtre optimale h_n	0.0231	Le critère <i>CAFPE</i> estimé	0.00051
La valeur du critère associée <i>CAFPE</i>	0.00044	La variance résiduelle	0.000463
La variance résiduelle $\hat{\sigma}_\varepsilon^2$	0.000463	La variance conditionnelle des résidus	0.00039

Tableau 6 – Analyse des résidus du modèle NAR

Skewness	Kurtosis	Stat. de J.B	Stat. de ARCH-LM
0.2909	9.5742	8058.386 (0.000)	13.904 (0.000)

(.): Probabilité critique

Tableau 7 – Tests de BDS sur les résidus de l'espérance conditionnelle

m	Stat. de BDS
2	10.243
3	15.574
4	21.547
5	28.279
6	39.412
7	45.254
8	50.247
9	57.791
10	62.432

La statistique de BDS est calculée par la méthode de "Fraction of pairs" avec ε égale à 0.7.

Tableau 8 – Spécification et estimation non paramétrique de la variance conditionnelle en utilisant l'estimateur linéaire local

Sélection non paramétrique des retards		Estimation non paramétrique de la volatilité conditionnelle	
Les retards sélectionnés	2,3,7,8	La fenêtre spécifiée	0.0214
La fenêtre optimale h_n	0.0214	Le critère <i>CAFPE</i> estimé	0.0000014
La valeur du critère associée <i>CAFPE</i>	0.0000012	La variance résiduelle	0.0000012
La variance résiduelle $\hat{\sigma}_\varepsilon^2$	0.0000012	La variance conditionnelle des résidus	0.0000010

Tableau 9 – Tests de BDS sur les résidus standardisés NAR-ARCH

m	Stat. de BDS
2	0.114
3	0.741
4	0.015
5	1.247
6	1.452
7	1.694
8	1.781

La statistique de BDS est calculée par la méthode de "Fraction of pairs" avec ε égale à 0.7.

Tableau 10 – Comparaison des qualités prédictives en utilisant l'erreur moyenne quadratique (prévision hors-échantillon)

Horizon	La fenêtre utilisée pour la prévision	NAR-ARCH non paramétrique	Marche aléatoire
1	0.0231	0.000645	0.000649
2	0.0230	0.000644	0.000648
15	0.0226	0.000628	0.000641
30	0.0221	0.000639	0.000647
90	0.0214	0.000677	0.000689