

Etude de la volatilité des rendements boursiers européens par l'approche hybride RCP-Ondelettes.

Salma TARCHELLA¹ & Zied KACEM² & Saloua BENAMMOU³

1 Faculté des sciences économiques et de gestion de Sousse, email : salmatarchella@yahoo.fr

2 Ecole polytechnique privée de Sousse, email : zied.kacem2013@gmail.com

3 Computational Mathematics Laboratory, email: saloua.benam mou@yahoo.fr

Résumé. La présence de bruit dans les domaines de l'analyse des données financières est toujours un problème où la qualité des estimations et d'ajustement du modèle de régression économétriques en souffre. Dans ce travail nous présentons une approche d'analyse des données hybride basée sur la combinaison des méthodes d'analyse classiques, telles que la régression sur composantes principales (RCP) et la théorie des ondelettes, c'est l'approche RCP-Ondelettes. Cette approche donne une nouvelle perspicacité à l'analyse des séries chronologiques et offre aux investisseurs un meilleur arrangement des mécanismes qui expliquent la dynamique des rendements courants en se basant sur les modèles d'évaluation par arbitrage (MEA), pendant qu'elle enlève le bruit qui affecte les données financières par un seuillage adéquat. Le but de cette étude est d'étudier la dynamique des rendements courants au sein du marché boursier français et allemand. Les résultats empiriques prouvent que la méthode suggérée permet d'améliorer les résultats de la régression sur ces marchés et de progresser la qualité d'ajustement du modèle.

Mots-clés. ACP, RCP, Seuillage, débruitage, Ondelettes, MEA, CAC40, DAX.

Abstract. The presence of noise in financial data analysis is still a major issue that raises several problems given that it can cause significant estimation biases and can deteriorate the adjusting quality of many econometric models such as Arbitrage Pricing Model (APT). In this work we present a hybrid data analysis approach based on the combination principal component regression (PCR) and wavelet theory, this is the Wavelet-PCR approach recently initiated. This approach give e new insight in time series analysis and offers to investors a better understanding of the mechanisms explaining the dynamics of the stock returns within the APT framework in order. It allows suppressing the noise that affects the financial data by a wavelet-based thresholding strategy. The aim of this study is to investigate the dynamics of the equity returns in the French and the German stock markets. The empirical results show that the suggested method can improve the results of the regression on these markets and ameliorate the quality of the model fit allowing.

Keywords. PCA, PCR, thresholding, denoising, Wavelets, CAPM, CAC40, DAX.

1. Introduction

Plusieurs méthodes d'analyse des données étudient le rapport entre un groupe de variables explicatives et un groupe différent de variables expliquées. En se basant sur un arrangement de régression, la régression linéaire multiple nous permet d'expliquer l'évolution de rentabilités des actions en fonction de différentes variables, Bowersman, L. et al. (1990). Cependant, il semble difficile d'effectuer une linéarisation sur un nombre infini de variables explicatives, dans ce cas on opte pour une méthode de réduction de dimension, on parle de la méthode ACP, d'où on aura la régression sur composantes principales (RCP).

La qualité des estimations et d'ajustement du modèle de régression économétriques en souffre à cause des bruits qui affectent les données initiales, Mallat, S.G. (1989). Une combinaison entre la RCP et la théorie des ondelettes peut être une solution à ce problème, Saloua, B. et al. (2010). En se comparant aux autres méthodes d'analyse des données classiques, le recours à la méthode hybride RCP-Ondelettes nous a donné des résultats indéniables dans la qualité d'estimation et de prévision de la volatilité boursière sur les marchés français et allemand.

Dans ce papier nous définissons la volatilité boursière et le principe du modèle d'évaluation par arbitrage (MEA), ensuite nous présentons les étapes de la RCP-Ondelettes, puis nous décrivons les données et la méthodologie utilisées dans cette étude. Finalement, nous présentons les principaux résultats obtenus et les principales conclusions.

2. La volatilité boursière et le MEA

La volatilité boursière peut être définie comme étant une mesure statistique des fluctuations connues des rendements pour un titre ou un indice boursier durant une période donnée, Daly, K. (2011). C'est un indicateur de risque le plus utilisé sur les marchés financiers qui peut aider les agents économiques et financiers à éviter les situations d'incertitude de prendre des décisions. La théorie financière a mis en avant deux modèles pour déterminer la volatilité, le modèle d'évaluation des actifs financiers (MEDAF) et le modèle d'évaluation par arbitrage (MEA), un modèle potentiel qui permet de surmonter les faiblesses reconnus par le précédent, Willam, S.G. (1990). Le MEA prédit que le rendement des actions est une fonction linéaire d'un certain nombre de facteurs économiques, Yli-Olli, P. et Virtanen, I. (1992) :

$$R_{it} = E(R_{it}) + b_{i1} \delta_{1t} + b_{i2} \delta_{2t} + \dots + b_{ik} \delta_{kt} + \varepsilon_{it}$$

Avec :

$E(R_{it})$, $i = 1, 2, \dots, n$, est le rendement attendu de l'action i ;

δ_j , $j = 1, 2, \dots, k$, se sont des facteurs économiques inobservés ;

b_{ij} : est la sensibilité de l'action i pour le facteur économique j ;

ε_{it} : présente le terme d'erreur.

L'étude de cette volatilité a mis en avant certains facteurs qui se différencient d'un modèle à un autre et d'un auteur à un autre, Rasiah, D. et al. (2011), où on est toujours confronté aux bruits qui affectent les données initiales.

3. La RCP-Ondelettes

L'approche RCP-Ondelettes se déroule en différentes étapes, comme le montre le schéma suivant, Saloua, B. et al. (2010) :

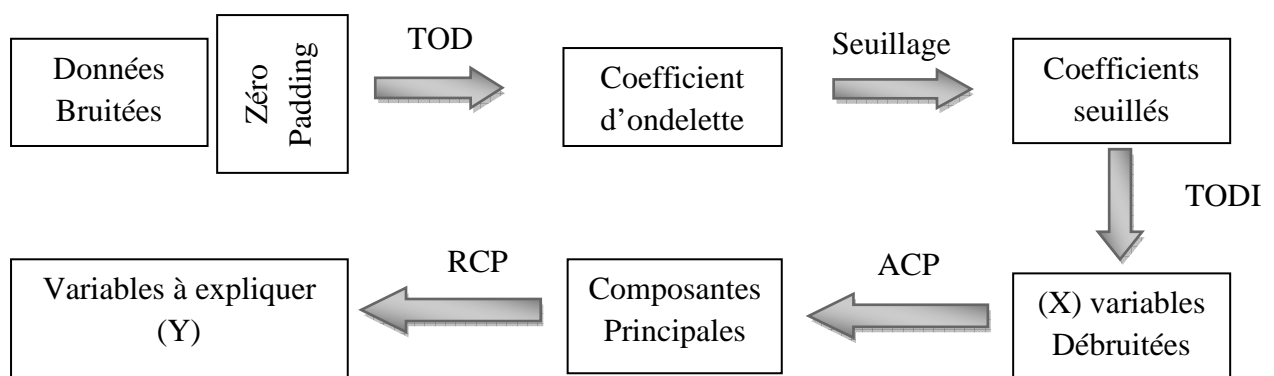


Figure 1: Les étapes de la RCP-Ondelettes

3.1 La transformée en ondelettes

C'est un outil de traitement de signal récemment mis au point, permettant l'analyse des propriétés locales des signaux chronologiques complexes sur plusieurs échelles de temps. Cette procédure ne peut être appliquée que si les vecteurs des données explicatifs soient de longueur dyadique, donc une obligation au recours de la méthode "Zéro Padding" qui consiste à ajouter des zéros aux colonnes manquantes.

L'analyse par ondelette est une représentation "temps-fréquence" qui permet de décomposer le signal à la fois en temps et en fréquence, Meyer, Y. et Rioul, O. (1988), alors que l'analyse multirésolution est une suite de sous-espaces emboîtés qui permet de décomposer les signaux à différents niveaux de résolution, de telles sortes que le passage d'un espace à un autre, soit la conséquence d'un changement d'échelle.

Contrairement à l'analyse de Fourier qui est basée sur des fonctions de sinus et de cosinus, la transformée par ondelette utilise un ensemble des fonctions $\psi(x)$ appelées "Ondelettes" où on distingue une transformée continue (TOC) et une autre discrète (TOD), Daubechies, I. (1992).

Les ondelettes se sont des fonctions élémentaires de ψ_{jk} qui représentent les vibrations les plus courtes et les plus élémentaires que l'on puisse envisager, Mallat, S.G. (1989). Elles sont dilatées (par un facteur 2^j) et/ou translatées (par un facteur k) telle que la fonction ψ forme une base orthonormée dans l'espace $L^2(\mathbb{R})$ (l'ensemble des fonctions de carrée intégrable) :

$$\{ \psi_{j,n} = 2^{-j/2} \psi(2^{-j}x - n), \text{ tel que } j,n \in \mathbb{Z} \}.$$

Ce qui permet d'écrire une fonction d'échelle $f \in L^2(\mathbb{R})$ de la façon suivante :

$$f(x) = \sum_{k \in \mathbb{Z}} \alpha_k \varphi_{0k}(x) + \sum_{j=0}^{\infty} \sum_{k \in \mathbb{Z}} d_{jk} \psi_{jk}(x), \quad (j, k \in \mathbb{Z})$$

Où α_k et d_{jk} présentent respectivement les coefficients d'approximation et de détail.

3.2 Le débruitage et les techniques de seuillage

Le débruitage est imposé pour améliorer la qualité des signaux et le pouvoir de les user au mieux postérieurement, Gencay, R. et al. (2002). Dans la littérature récente, beaucoup de méthodes ont été développées dans le but d'avoir réalisé les objectifs de débruitage dans le domaine des ondelettes, tel que les techniques de seuillage "dur" et "doux", qui ont été présentées par Donoho, D. et Johnstone, I. (1994,1998). Dans les études de traitement des signaux chronologiques, le seuillage doux (Soft Thresholding) présente un biais plus important que le seuillage dur, Donoho, D. (1998). Il conduit à mettre à zéro les valeurs de coefficients qui sont inférieures à un seuil (λ) et à ne conserver que ce qui dépasse le seuil pour les autres coefficients. Le signal après seuillage est écrit et illustré selon la figure suivante :

$$\theta(x) = \begin{cases} x + \lambda & \text{si } x \leq -\lambda \\ x - \lambda & \text{si } x \geq \lambda \\ 0 & \text{si } |x| \leq \lambda \end{cases}$$

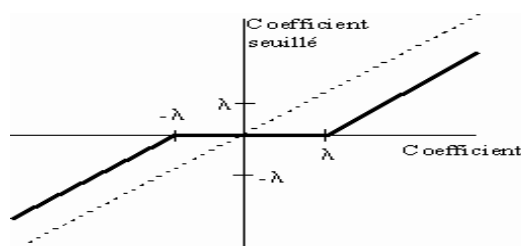


Figure 2 : Seuillage doux de DONOHO (au seuil λ).

3.3 La régression sur composante principale (RCP)

C'est une extension simple de la régression linéaire multiple (RLM) et l'analyse en composantes principales (ACP), Lebart, L. et al. (2004).

La méthode ACP permet de réduire la dimensionnalité de l'ensemble de variables et d'extraire l'information principale des données de grande dimension par projection dans un espace de dimension plus inférieure. La RCP permet de remplacer les variables explicatives en un nouvel ensemble de variables orthogonales, appelées composantes principales (CP). Cette transformation classe les nouvelles variables orthogonales dans l'ordre de leur importance, en éliminant certaines des (CP), les moins importants, pour effectuer une réduction de la variance. Ainsi une analyse de régression multiple de la variable de réponse à la base de l'ensemble réduit de composantes principales est réalisée à l'aide des estimations faites par les MCO, Lebart, L. et al. (2004).

4. Données et méthodologie

Pour illustrer notre étude, nous avons utilisé des données boursières réelles, il s'agit des rentabilités journalières des indices boursiers CAC40 et DAX, relatives respectivement aux marchés français et allemand sur la période allant du début 2011 jusqu'au fin 2014. Pour avoir une description simultanée des liaisons entre ces rendements et les similitudes entre les années, la méthode ACP nous a permis de réduire la dimensionnalité de l'ensemble de variables en un nombre plus réduit et centré et d'avoir des représentations graphiques approchées.

Comme étape préliminaire pour séparer le signal du bruit, nous avons fait recours à une TOD à cinq niveaux de résolution ($n = 5$) pour les 36¹ actions du CAC40 et les 30 actions du DAX. Nous avons fait recours à la famille des ondelettes de Daubechies "db5", qui nous a fournit par la suite cinq niveaux de coefficients de détail et un niveau unique de coefficient d'approximation.

Dans notre étude nous avons appliqué un seuillage doux aux coefficients de détails obtenus antérieurement par la TOD et en se basant sur un principe de « Minimax », où le seuil suggéré est calculé par approximation numérique et optimiser par minimisation d'un critère du risque R dans l'espace $L^2(R)$. L'idée est de choisir un estimateur \hat{t} atteignant le risque de « Minimax » pour une classe particulière des fonctions lisses (F), Donoho, D. (1994) :

$$\tilde{R}(F) = \inf_{\hat{t}} \sup_t R(\hat{t}, t)$$

L'utilisation de la procédure « Minimax », nous a permis de fixer un seuil qui consiste à choisir un estimateur \hat{t} de t qui atteint le risque défini par la relation ci-dessus. Une application inverse de la transformée en ondelettes discrète (TODI) doit être mise en jeu pour la reconstitution du signal de séries financières débruitées, où nous allons appliquer au nouveau une ACP et une RCP.

Nous supposons que les rendements boursiers sont en partie expliqués par des facteurs communs indéterminés, Saloua, B. et al. (2010) :

$$R_{it} = \sum_{j=1}^k \beta_j^i \times F_{jt} + \varepsilon_{it}$$

Avec :

R_{it} : Le rendement de l'action i à la date t ;

β_j^i : La sensibilité de l'action i au j ième facteur ;

F_{jt} : Le j ième facteur, sachant que $j=1 \dots k$ (Le nombre de facteurs retenus par l'ACP).

ε_{it} : Le risque spécifique de l'action i à la date t ;

¹ Les 4 sociétés manquantes de l'indice CAC40 sont exclues à cause du non disponibilité des données.

5. Résultats empiriques

En comparant les résultats empiriques obtenus avant et après débruitage, nous avons remarqué une diminution de coefficients de corrélation entre les variables explicatives en débruitant les données qui conduit à une augmentation de nombre des facteurs retenus par l'ACP, où il a dépassé de 4 à 7 facteurs, ceci est prévu pour les deux indices CAC40 et DAX.

L'approche RCP-Ondelettes surpasse la RCP classique, elle semble être la plus meilleure, suite d'une augmentation de R^2 ajusté et une diminution importante de la moyenne des carrés des erreurs (MCE) pour toutes les actions du CAC40 et DAX, ce qui coïncide avec les résultats obtenus par les travaux menés antérieurement par Saloua, B. et al. (2010) sur le marché français dans la période allant de janvier 2004 à janvier 2008. Ceci peut être vu sur les schémas ci-dessous où les bâtons représentant le MCE après débruitage se trouvent au-dessous de ceux du MCE avant débruitage :

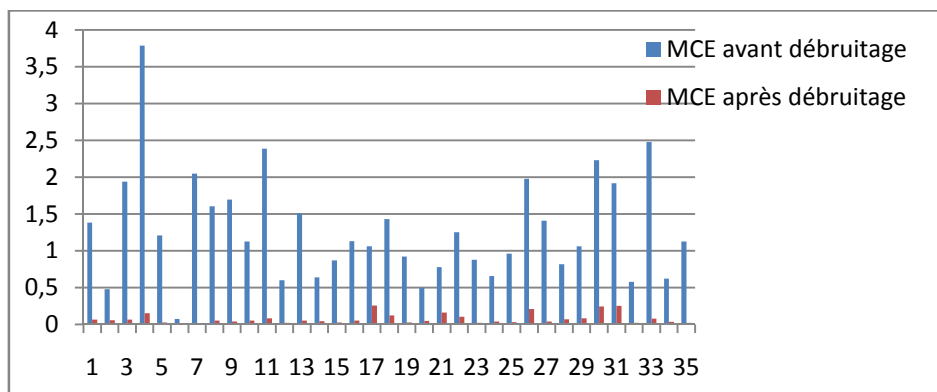


Figure 3 : MCE avant et après débruitage pour l'indice CAC40.

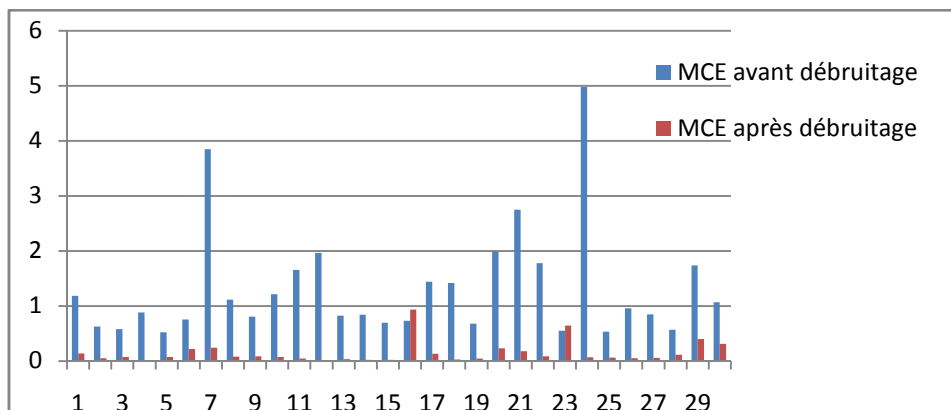


Figure 4 : MCE avant et après débruitage pour l'indice DAX.

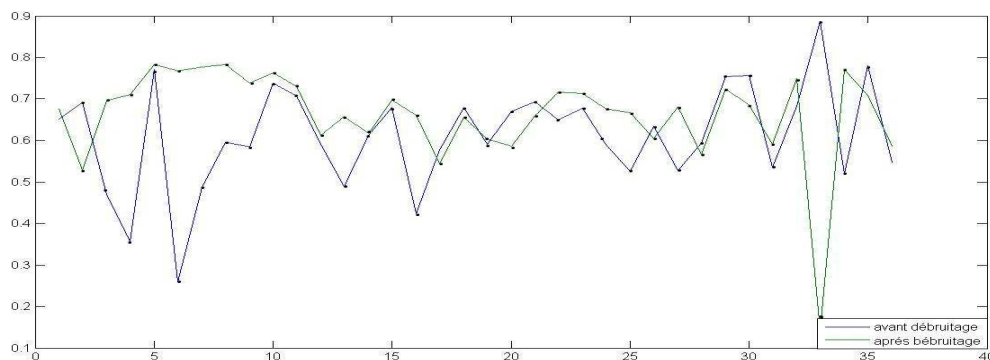


Figure 5 : R^2 ajusté avant (bleu) et après débruitage (vert) pour l'indice CAC40.

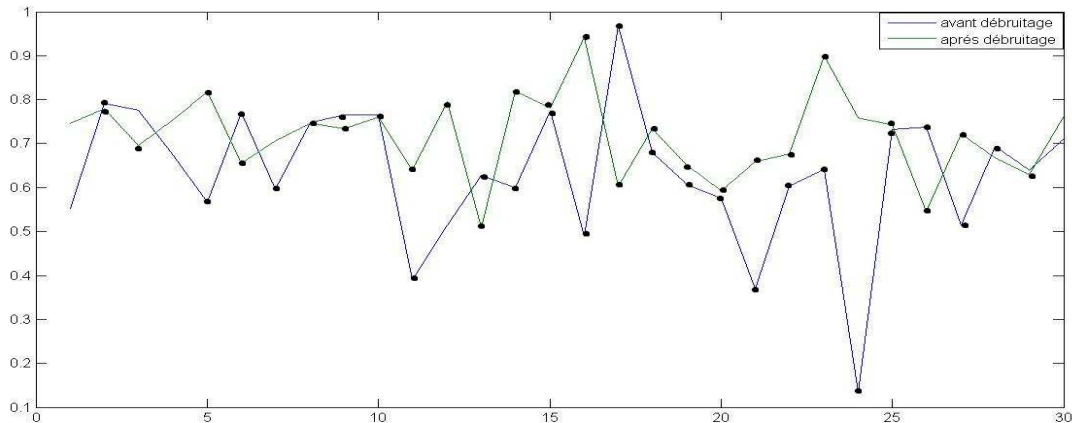


Figure 6: R^2 ajusté avant (bleu) et après débruitage (vert) pour l'indice DAX.

6. Conclusion

L'application de cette approche hybride sur les rendements des actions de deux marchés européens, à savoir le marché français et le marché allemand nous a permis une amélioration significative en termes d'exactitude de prévision et de qualité d'ajustement du modèle de régression. La RCP-Ondelettes surpasse la RCP classique, en donnant les meilleurs résultats d'estimation, en termes de diminution du résidu et du MCE. Une meilleure estimation des facteurs explicatifs de la volatilité boursière a été eu lieu suite de l'utilisation de cette approche, où le nombre des facteurs a augmenté de 4 à 7 facteurs, n'empêche que nous pouvons identifier ces facteurs et essayer de trouver les facteurs en communs entre ces deux marchés afin de donner un nouveau modèle standard pour tous les marchés européens.

Bibliographie

- [1] Bowersman, L. et al. (1990), Linear Statistical Models: An Applied Approach, 2^{eme} édition, Duxbury Press, Belmont, California.
- [2] Daly, K. (2011), An Overview of the Determinants of Financial Volatility, Modern Applied Science, 5, 46–63.
- [3] Daubechies, I. (1992), Ten Lectures on Wavelets, SIAM, Philadelphia.
- [4] Donoho, D. et Johnstone, I. (1994), Ideal spatial adaptation by wavelet shrinkage, Biometrika, 81, 425–55.
- [5] Donoho, D. et Johnstone, I. (1998), Minimax estimation via wavelet shrinkage, Ann. Statist, 26, 879–921.
- [6] Gencay, R. et al. (2002), An Introduction to Wavelets and Other Filtering Methods in Finance and Economics, Academic Press, New York.
- [7] Lebart, L. et al. (2004), Statistique exploratoire multidimensionnelle, Dunod Paris.
- [8] Mallat, S.G. (1989), A theory for multiresolution signal decomposition: the wavelet representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 11, 674–93.
- [9] Meyer, Y. et Rioul, O. (1988), L'Analyse par Ondelettes, Pour La Science, 119, 28–37.
- [10] Rasiah, D. et al. (2011), The Effectiveness of Arbitrage Pricing Model in Modern Financial Theory, International Research Journal, 3, 125–135
- [11] Saloua, B. et al. (2010), Risk reduction using wavelets for denoising principal-components regression models, Emerald, The Journal of Risk Finance, 2, 180–203.
- [12] Saloua, B. et al. (2010), Sur l'utilisation conjointe de la régression sur composantes principales et des ondelettes, Revue MODULAD, 41, 38–58.
- [13] Willam, S.G. (1990), Stock Market Volatility, Financial Analyst Journal, 46, 23–34.
- [14] Yli-Olli, P. et Virtanen, I. (1992), Some empirical tests of the arbitrage pricing theory using transformation analysis, Empirical Economics, 4, 507–522