

# DÉTECTION DE SYSTÈMES DYNAMIQUES DANS DES DONNÉES BOURSIÈRES : DÉBRUITAGE PAR ONDELETTES

Matthieu Garcin <sup>1</sup> & Dominique Guégan <sup>2</sup>

<sup>1</sup> *MSE - CES, Université Paris1 Panthéon-Sorbonne, 106 boulevard de l'hôpital, 75013 Paris, France, et Natixis Asset Management, prénom.nom [ at ] polytechnique.edu*

<sup>2</sup> *MSE - CES, Université Paris1 Panthéon-Sorbonne, 106 boulevard de l'hôpital, 75013 Paris, France, prénom.nom [ at ] univ-paris1.fr*

**Résumé.** Nous nous intéressons à des signaux financiers que nous tentons de décrire par un système dynamique perturbé par un bruit non-linéaire. En utilisant des ondelettes et un filtrage par seuil, nous construisons un estimateur du système non bruité, qui nous permet ensuite de faire des prévisions sur les séries financières. Le choix du seuil dans le filtrage se fait en minimisant un estimateur de l'erreur de reconstruction.

**Mots-clés.** Système dynamique, chaos, ondelettes, SURE, prévision, non-linéaire.

**Abstract.** We are interested in financial signals that we try to describe by a dynamical system disrupted by a nonlinear noise. We use wavelets and a threshold-filter in order to construct an estimator of the pure dynamical system. Then, such an estimator allows us to make forecasts for the financial series. The choice of the threshold in the filtering is done by minimizing an estimate of the reconstruction error.

**Keywords.** Dynamical system, chaos, wavelets, SURE, forecast, nonlinear.

## 1 Introduction

Nous nous intéressons à la modélisation de cours boursiers par des systèmes dynamiques ou chaotiques à temps discret, perturbés par du bruit, comme dans Guégan (2003). Ainsi, à partir d'un signal financier, nous voulons estimer un système dynamique afin de pouvoir faire des prévisions et construire des mesures de risque.

Notre méthode d'estimation utilise les ondelettes, dont la théorie est déjà très développée mais s'intéresse principalement à des cas où l'impact du bruit sur un signal est linéaire. Dans notre cadre de travail, l'impact du bruit est non-linéaire et il faut imaginer de nouveaux outils.

## 2 Le modèle

Nous écrivons  $(x_t)_{t \in \mathbb{Z}}$  le système dynamique non perturbé par le bruit. Deux états successifs sont liés par la fonction  $z$  (nous pouvons étendre les résultats qui suivent à

plusieurs dimensions, ce qui signifie, pour une dimension  $d$ , que  $x_t$  n'est pas seulement fonction de  $x_{t-1}$  mais aussi de  $x_{t-2}, \dots, x_{t-d}$  :

$$x_t = z(x_{t-1}).$$

Le système dynamique bruité  $(u_t)$  a, lui, pour fonction de chaos  $z^\varepsilon$  :

$$u_t = z^\varepsilon(u_{t-1}).$$

Nous considérons un bruit de mesure, tel que :

$$\begin{cases} u_t = x_t + \varepsilon_t \\ x_t = z(x_{t-1}), \end{cases}$$

où les  $\varepsilon_t$  sont des variables aléatoires gaussiennes iid.

Nous observons des états successifs du système bruité,  $(u_t)_{t \in \mathbb{Z}}$ , qui sont, dans le cas que nous étudions, des rendements. Notre objectif est de reconstruire  $z$ .

### 3 Méthode de débruitage

Nous ne voulons pas imposer une forme particulière à la fonction de chaos  $z$ , donc, pour son estimation, nous privilégions une approche non-paramétrique. Pour ce faire, nous décomposons la fonction bruitée  $z^\varepsilon$  dans une base d'ondelettes. Ainsi, nous calculons les coefficients d'ondelettes du système bruité, conditionnellement à la grille d'observations  $u_1, \dots, u_N$  :

$$\hat{z}_{j,k}^\varepsilon = \langle z^\varepsilon, \psi_{j,k} \rangle.$$

Ensuite, nous pouvons appliquer à ces coefficients bruités un filtre, qui est un opérateur  $F$ , et on obtient un estimateur de la fonction de chaos pure,  $z$  :

$$\tilde{z}(u) = \sum_{j \in J} \sum_{k \in K} F \hat{z}_{j,k}^\varepsilon \psi_{j,k}(u),$$

où  $J$  et  $K$  sont des parties de  $\mathbb{Z}$ . Plusieurs types de filtres existent et sont adaptés selon la connaissance *a priori* que l'on a de  $z$  (Donoho et Johnstone (1994), Donoho (1995a), Donoho, Johnstone, Kerkyacharian et Picard (1995), Mallat (2000)). Dans notre cas, nous utilisons plus particulièrement des filtres non-linéaires : les filtres à seuil. Pour un seuillage dur, de seuil  $\lambda \geq 0$ , on prend :

$$F \hat{z}_{j,k}^\varepsilon = \hat{z}_{j,k}^\varepsilon \mathbf{1}_{\{|\hat{z}_{j,k}^\varepsilon| \geq \lambda\}}$$

et pour un seuillage doux, de seuil  $\lambda \geq 0$ , on prend :

$$F \hat{z}_{j,k}^\varepsilon = (\hat{z}_{j,k}^\varepsilon - \lambda) \mathbf{1}_{\{\hat{z}_{j,k}^\varepsilon \geq \lambda\}} + (\hat{z}_{j,k}^\varepsilon + \lambda) \mathbf{1}_{\{\hat{z}_{j,k}^\varepsilon \leq -\lambda\}}.$$

Le choix du seuil  $\lambda$  est primordial pour obtenir le meilleur estimateur de  $z$ . Pour un seuillage doux, un seuil particulier (le seuil SURE), permet, dans le cas classique, de minimiser un estimateur de l'erreur de reconstruction (Stein (1981), Donoho et Johnstone (1995), Mallat (2000)). En effet, nous n'avons connaissance que du système bruité et pas du système pur ; il nous est donc impossible de calculer directement l'erreur de reconstruction, définie comme l'écart quadratique entre les fonctions  $z$  et  $\tilde{z}$ . Nous devons donc avoir recours à un estimateur. Dans le cadre classique, où l'impact du bruit est linéaire, cet estimateur est sans biais. Dans notre cadre de travail, nous avons construit un tel estimateur et nous sommes capables de majorer son biais. Notre estimateur de l'erreur de reconstruction est

$$\sum_{l=1}^N \mathcal{S}_{j,k}(\tilde{z}_{j,k}^\varepsilon),$$

où

$$\mathcal{S}_{j,k} : x \mapsto \begin{cases} \lambda^2 + \sigma^2(1 + \Gamma_{j,k}) & \text{si } |x| \geq \lambda \\ x^2 - \sigma^2(1 + \Gamma_{j,k}) & \text{sinon} \end{cases}$$

et où  $\Gamma_{j,k} = \langle z'^2, \psi_{j,k}^2 \rangle$ .

Dans la Figure 1, nous observons le bon comportement de l'estimateur de l'erreur sur un exemple.

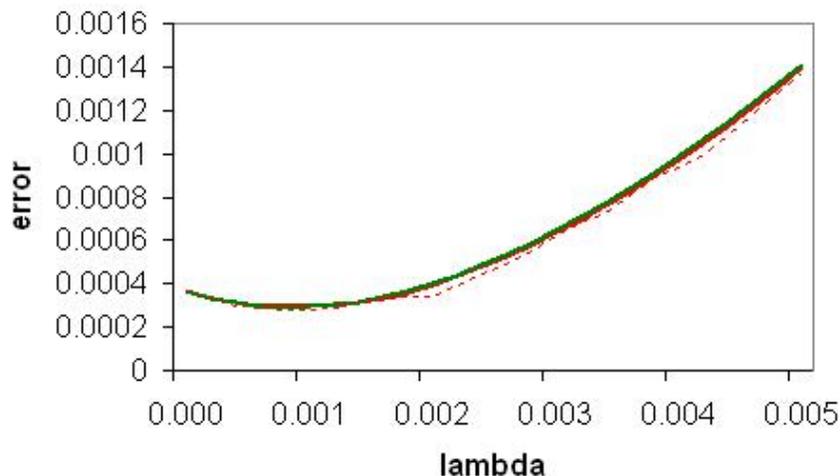


FIGURE 1 – Erreur de reconstruction d'un filtre à seuillage doux, pour différentes valeurs de seuil, pour une carte logistique de paramètre 4 et un bruit gaussien d'écart-type 1%. La courbe verte correspond à l'erreur calculée analytiquement. La courbe rouge pleine, superposée à la courbe verte, correspond à notre estimateur de l'erreur (calculé en moyenne sur 100 tirages). La courbe rouge pointillée correspond à notre estimateur de l'erreur pour un unique tirage.

## 4 Test sur données financières

Nous tentons de détecter, par la méthode décrite ci-avant, un système dynamique dans des données boursières. Dans l'exemple qui suit, nous observons un indice d'obligations européennes de 2003 à 2012 et reconstruisons la fonction  $z$  liant deux rendements quotidiens successifs. Nous obtenons la Figure 2. En augmentant la dimension (c'est-à-dire que nous expliquons le rendement en  $t$  par une fonction  $z$  prenant en argument les rendements en  $t - 1$  et  $t - 2$ ), nous obtenons une autre estimation de  $z$ , Figure 3. On peut itérer et augmenter encore la dimension.

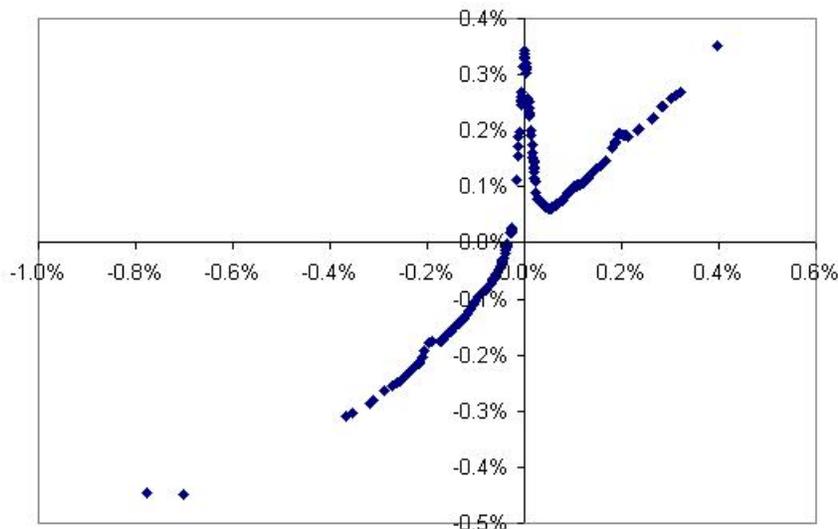


FIGURE 2 – Fonction de chaos en dimension 1.

En plus de notre estimation de l'erreur de reconstruction, nous pouvons nous intéresser à la performance de prévision à partir de ces deux fonctions  $z$  de dimensions différentes, Figure 4. Le test est établi sur un échantillon d'un an (2013), postérieur à l'échantillon qui nous a permis d'estimer  $z$  (2003-2012).

## 5 Conclusion

Notre travail permet d'estimer des systèmes déterministes sur des données boursières à partir d'un algorithme dont la justification théorique fait appel à une extension de résultats classiques de la théorie des ondelettes au cas d'un bruit non-linéaire. Cette méthode, facile à mettre en place, permet de faire des prévisions sur des cours d'actifs financiers.

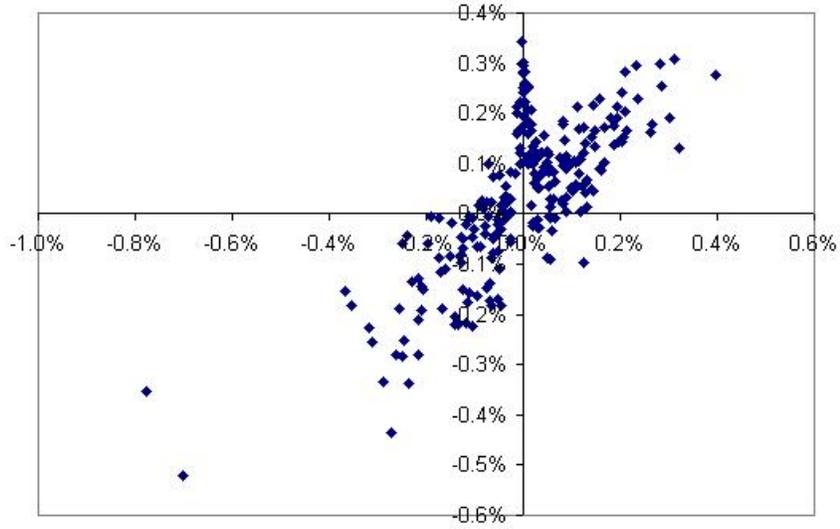


FIGURE 3 – Fonction de chaos en dimension 2.

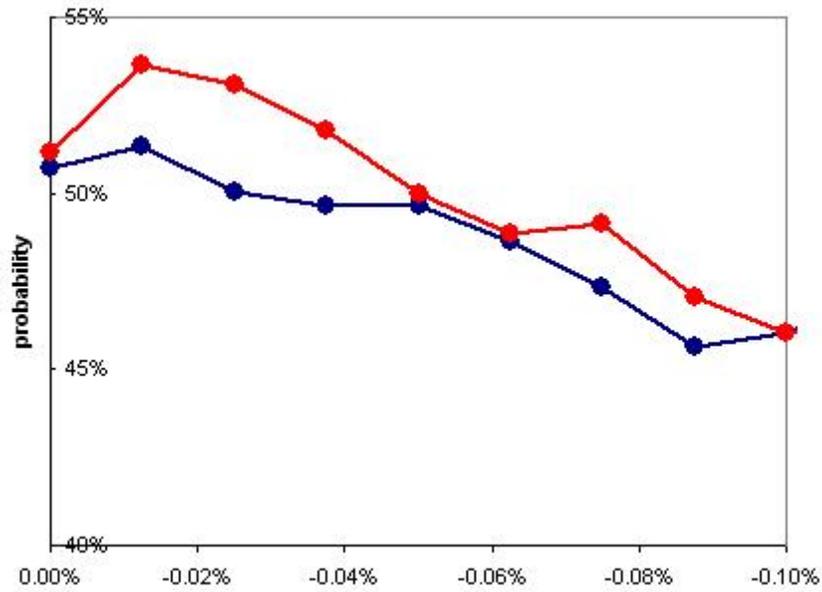


FIGURE 4 – Moyenne du risque de première espèce et du risque de deuxième espèce pour le test de l'hypothèse  $H_0$  d'un rendement inférieur au seuil indiqué en abscisse, en dimension 1 (rouge) et en dimension 2 (bleu).

## Bibliographie

- [1] Donoho, D. (1995a), Denoising via soft thresholding, *IEEE transactions on information theory*, 41, 613–627
- [2] Donoho, D. et Johnstone, I. (1994), Ideal spatial adaptation via wavelet shrinkage, *Biometrika*, 81, 425–455
- [3] Donoho, D. et Johnstone, I. (1995), Adapting to unknown smoothness via wavelet shrinkage, *Journal of the American statistical association*, 90, 432, 1200–1244
- [4] Donoho, D., Johnstone, I., Kerkyacharian, G. et Picard, D. (1995), Wavelet shrinkage : Asymptopia?, *Journal of the Royal Statistical Society, Series B*, 57, 2, 301–369
- [5] Guégan, D. (2003), *Les Chaos en finance : approche statistique*, Economica, Paris
- [6] Mallat, S. (2000), *Une exploration des signaux en ondelettes*, Ellipses, Éditions de l'École Polytechnique, Paris
- [7] Stein, C. (1981), Estimation of the mean of a multivariate normal distribution, *The annals of statistics*, 9, 6, 1135–1151