



Munich Personal RePEc Archive

Study of the returns nonlinear dynamics of the Casablanca stock exchange

RIANE, Nizare

Faculté de Droit et d'Economie, Université Mohamed V, Agdal,
Rabat, Morocco.

10 September 2014

Online at <https://mpra.ub.uni-muenchen.de/61957/>

MPRA Paper No. 61957, posted 07 Feb 2015 05:21 UTC

Etude de la dynamique non-linéaire des rentabilités de la bourse de Casablanca

Résumé. La prépondérance de l'approche linéaire dans la modélisation du marché boursier est le fruit du paradigme de Frisch-Slutsky selon lequel le marché ne peut que converger vers un point d'équilibre ou diverger, selon une trajectoire monotone ou oscillatoire. Par ailleurs, cette description de la réalité est insuffisante, d'abord par son incapacité à décrire les fluctuations qui ont tendance à persister ainsi que les anomalies du marché, ensuite par la faiblesse des tests statistiques linéaires face à des processus plus complexes. Dans ce papier, nous examinons l'existence d'une dynamique non-linéaire qui gouverne l'évolution de l'indice MASI. L'analyse utilise les concepts d'exposants de Lyapunov, dimension de corrélation et d'autres outils pour déterminer la nature du processus sous-jacent. Les résultats démontrent l'existence d'un processus non-linéaire, mais le caractère déterministe reste contesté.

Classification JEL : C12 C19 C22 G12

Mots clés : Chaos, attracteur, non linéarité, déterminisme, exposant de Lyapunov, dimension de corrélation.

Study of the returns nonlinear dynamic of the Casablanca stock exchange

Abstract. The preponderance of the linear approach in the stock market modeling is the result of the Frisch-Slutsky paradigm which implies that the market can only converge to an equilibrium point or diverge, according to a monotonic or oscillatory trajectory. Moreover, this description of reality is insufficient, first by his inability to describe the fluctuations that tend to persist and market anomalies, second by the weakness of the linear statistical tests facing more complex processes. In this paper, we examine the existence of a non-linear dynamics that govern the evolution of the MASI Index. The analysis uses the concepts of Lyapunov exponents, correlation dimension and other tools to determine the nature of the underlying process. The results provide evidence of a non-linear process, but the determinism remains contested.

JEL Classification : C12 C19 C22 G12

Key words: Chaos, attractor, nonlinearity, determinism, Lyapunov exponent, correlation dimension.

Introduction :

L'approche linéaire en économie s'inspire du paradigme dominant selon lequel les variables économiques connaissent deux types de trajectoire, elles convergent vers un point d'équilibre ou elles divergent, selon des trajectoires monotones ou oscillatoires. Néanmoins, cette approche rencontre des défis majeurs en finance qu'elle n'arrive pas à surmonter d'une manière convaincante : les difficultés de modéliser les anomalies du marché et les éléments du monde réel qui écartent les marchés financiers de l'équilibre recherché.

Une question qui se pose naturellement est de chercher à connaître les sources de défaillance de cette approche pour proposer une alternative de modélisation, et la réponse qui s'est imposée à nous est que la linéarité des modèles est la cause de ces insuffisances et qu'il faut adopter un nouveau paradigme dans lequel la description de la dynamique financière est non-linéaire.

L'analyse non-linéaire apparaît alors comme une solution pour dépasser ces difficultés. Elle vient couronner le développement mathématique de l'étude des systèmes dynamiques. Il s'agit d'un développement qui a eu des répercussions d'abord dans les sciences naturelles à commencer par la physique.

Le progrès de l'outil informatique a accompagné l'analyse numérique. Cette branche des mathématiques - qui a parmi ces objectifs d'approcher les systèmes dynamiques à un degré de tolérance et à une marge d'erreur maîtrisables - a réussi à élucider le mystère des équations complexes et à en trouver la solution (comme la suite logistique qui a été développée par Verhulst en 1845 et qui a occupé les mathématiciens jusqu'aux années quatre-vingt du 20ème siècle), et a encouragé le déploiement et la mise en œuvre des

algorithmes de tests empiriques à partir de données réelles, et l'approche non-linéaire a pris un grand essor :

F. Takens donne un nouvel essor à l'étude empirique des processus non-linéaires, dans son article majeur « Detecting strange attractors in turbulence ». Il donne une approche pratique de reconstruction de la dynamique d'un système en se basant sur la seule information contenue dans son signal.

En 1983, P. Grassberger et I. Procaccia donnent naissance à la dimension de corrélation. Il s'agit d'une mesure de la dimension de l'espace occupé par un ensemble de points.

L'équipe d'A. Wolf publie en 1984 un algorithme de calcul de l'exposant de Lyapunov à partir de données empiriques. Cet exposant doit son nom à Alexandre Lyapunov qui l'a construit pour la première fois pour quantifier la stabilité ou l'instabilité d'un système dynamique.

Une révolution survient dans le domaine des réseaux de neurones artificiels avec l'introduction en 1986 du Perceptron Multi-Couche par Rumelhart. Les premiers travaux sur les réseaux de neurones remontent à la fin des années 1950. Il s'agit des neurologues Warren McCulloch et Walter Pitts qui ont posé le premier pilier avec leur article « What the frog's eye tells to the frog's brain ? ».

W. Brock, W. D. Dechert, J. Scheinkman mettent en place en 1987 un test de linéarité basé sur la dimension de corrélation qui portera leur nom. Ils exploitent le fait que cette statistique croît exponentiellement avec la dimension de plongement pour un processus identiquement et indépendamment distribué.

En 1989, Tong introduit le modèle TAR (Threshold Autoregressive) linéaire par morceaux, un modèle qui prend en considération les changements de régime.

En 1994, T. Kaplan construit un test de déterminisme à partir d'une série chronologique. Il est parti de la propriété de la sensibilité aux conditions initiales qui caractérise les dynamiques non linéaires pour tester le déterminisme.

Plusieurs travaux ont accompagné cette lancée, soit pour proposer des algorithmes de calcul plus efficaces ou bien pour explorer d'autres caractéristiques des systèmes dynamiques. Et la théorie a vite débordé le domaine des sciences exactes. En météorologie, on rencontre l'étude du changement climatique (phénomène El Niño) ; en géologie, on a l'étude des tremblements de terres ; en médecine, il y a l'étude du rythme cardiaque...etc.

L'économie n'a pas pu échapper à cette vague et elle s'est trouvée tout de suite concernée par la théorie suite aux premiers essais des économistes qui ont manifesté un enthousiasme rapide pour cette nouvelle approche, surtout en finance qui procure un nombre assez élevé de données, et où le marché peut être observé en temps réel.

Des situations instables et des comportements complexes caractérisent les séries financières et sont difficiles à expliquer avec les modèles traditionnels, et leur modélisation repose toujours sur l'hypothèse de l'équilibre général selon laquelle le marché est efficace, efficient et permet une allocation optimale des ressources et des informations. Ce résultat implique un autre, celui du marché aléatoire imprévisible, réduisant ainsi les situations d'incertitude à des situations de risque, en fermant ainsi le futur afin d'obtenir des solutions stables aux modèles économiques.

En effet, l'approche linéaire dans l'étude des cours boursiers se heurte à plusieurs difficultés d'ordre théorique et empirique, et elle est très controversée pour plusieurs raisons :

1. En l'absence de chocs, les fluctuations représentées par un modèle linéaire ne peuvent que s'estomper ou exploser. Or dans le monde réel, d'une part, le modèle linéaire ne peut expliquer le caractère auto-entretenu des fluctuations, et d'autre part, il y a des limites institutionnelles et techniques qui empêchent les variables de dépasser un certain seuil et d'exploser.
2. L'inadaptation de l'approximation linéaire des phénomènes non-linéaires, surtout lorsque la dynamique en question est assez complexe : dans le cas d'un modèle chaotique, l'erreur commise par l'approximation linéaire augmente d'une manière exponentielle avec le temps.
3. Cette modélisation reste incapable de prévoir les ruptures de la dynamique des phénomènes suite à un changement structurel : la modification d'un paramètre critique du système modifie la dynamique en question d'une façon catastrophique, et on peut passer d'un modèle stable autour d'un point fixe à un modèle complètement chaotique. Il s'agit de la théorie des catastrophes.
4. L'apparition et le développement des modèles non-linéaires soutenus par le progrès de l'outil informatique comme alternative de modélisation.

Dans son article « A Chaotic Attractor for the S&P 500 », Edgar E. Peters conclut sur la nature déterministe non-linéaire de l'indice S&P500. Les statistiques de l'étude montrent de façon très claire le caractère chaotique de l'indice, alors que dans le travail de Girerd-Potin I.

et Taramasco O. sur les rentabilités boursières de la bourse de Paris, les auteurs n'arrivent pas à rejeter l'hypothèse de la nature stochastique des cours même s'ils ont pu rejeter l'hypothèse de la linéarité. Devant une telle divergence de résultat, une conclusion intuitive sera de considérer que la dynamique du marché boursier est relative et dépend du stade de l'évolution de ce marché ainsi que des caractéristiques de l'économie en question. Il sera donc incontournable de répondre à certaines questions qui représentent l'essence de cette étude : le marché boursier marocain peut-il être décrit par un processus non-linéaire ?

I. Le schéma de l'étude :

Les difficultés majeures rencontrées par les chercheurs dans l'analyse des données économiques par les méthodes non linéaires sont la non stationnarité et le bruit dans les données. Si on prend par exemple deux articles qui ont porté sur les données boursières : l'article de Peters sur l'indice boursier S&P 500, et celui de Girerd-Potin et Taramasco sur les rentabilités. Le problème qui se pose avec l'article de Peters est la non-stationnarité. Il propose comme solution la régression de l'indice S&P 500 sur l'indice générale des prix, mais rien ne garantit que sa méthode soit efficace. Avant d'entamer n'importe quelle étude sur les phénomènes chaotiques à partir des séries chronologiques il faut s'assurer que la série ne présente aucune tendance, et si c'est le cas, il faut absolument la stationnariser. L'instationnarité de la série affecte les mesures invariantes caractérisant la dynamique en question et donne des résultats biaisés : une série non stationnaire voit sa dimension de corrélation croître d'une manière infinie, des effets similaires affectent les autres mesures invariantes et leurs valeurs deviennent indéterminables.

La transformation courante introduite sur la série pour éliminer la non-stationnarité est la différenciation :

$$\Delta X_t = X_t - X_{t-1} \quad (1)$$

Selon Schreiber et Kantz (2004), cette technique même si elle permet de stationnariser la série, elle tend à augmenter le bruit.

Le problème qui se présente avec le 2^{ème} article est celui du bruit, puisqu'il influence l'ensemble des mesures invariantes, les données économiques sont connues par la forte présence du bruit.

Plusieurs techniques existent pour éliminer cette composante, les techniques de filtrage traditionnelles (filtre passe-bande, filtre de Kalman ...) ne sont pas adaptées pour ce type de problèmes car elles détruisent la structure sous-jacente de la série. Des méthodes plus adaptées à ce genre de problème sont utilisées, la technique de projection locale ou le filtrage par les ondelettes. C'est cette dernière technique qui sera utilisé dans ce travail.

Notre étude va porter sur les rentabilités journalières de l'indice MASI sur période [12/04/1994 - 15/09/2011], ce qui représente un total de 4096 observations, les rentabilités sont exprimées sous forme continue

$$r_t = \log_{10}(P_t) - \log_{10}(P_{t-1}) \quad (2)$$

Le schéma de notre travail est le suivant

- 1) Tester les conditions d'existence d'une dynamique chaotique dans la série : Stationnarité, linéarité, déterminisme.
- 2) Déterminer les caractéristiques du processus sous-jacent.
- 3) Modéliser le comportement de la série.

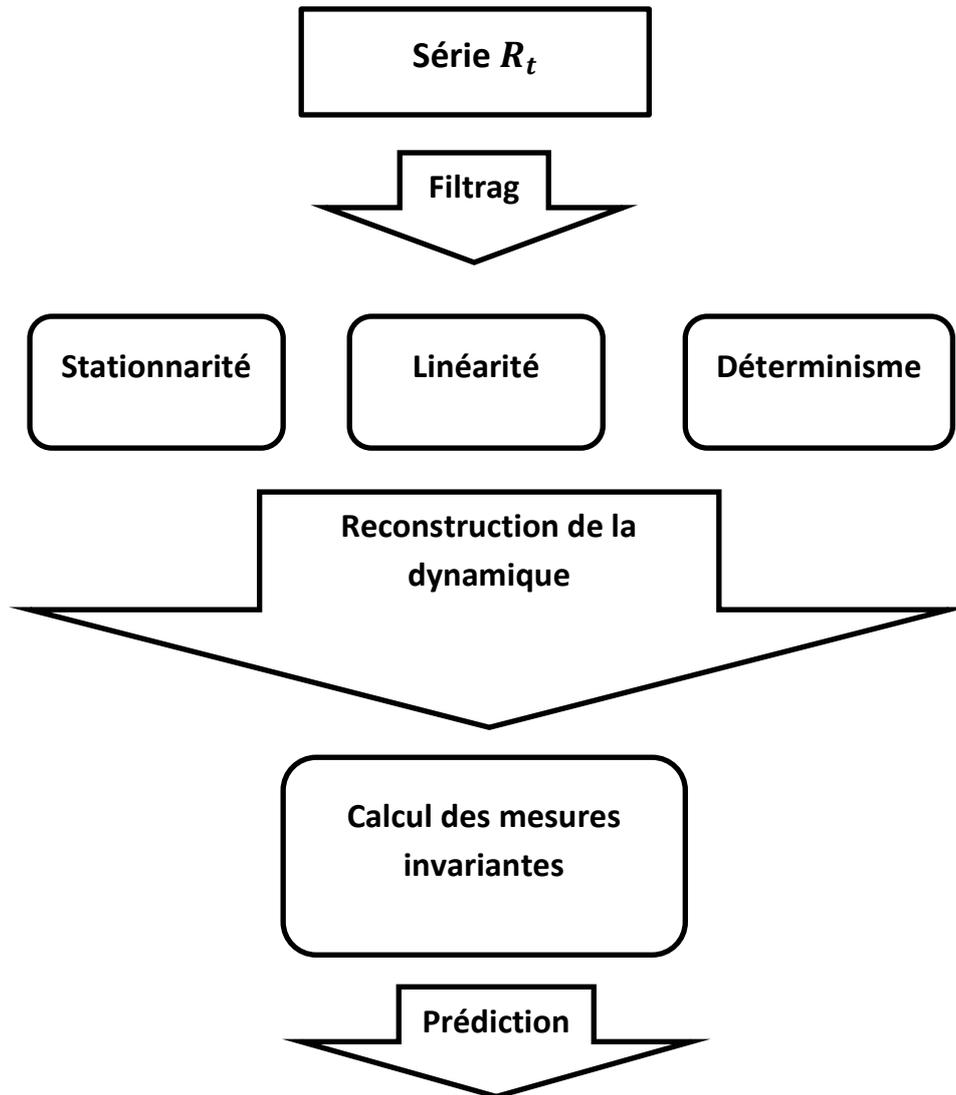


Figure 1. Schéma de l'étude

II. Analyse préliminaire des rentabilités :

Un premier regard sur le graphique de l'évolution des rentabilités nous permet de tirer les conclusions suivantes :

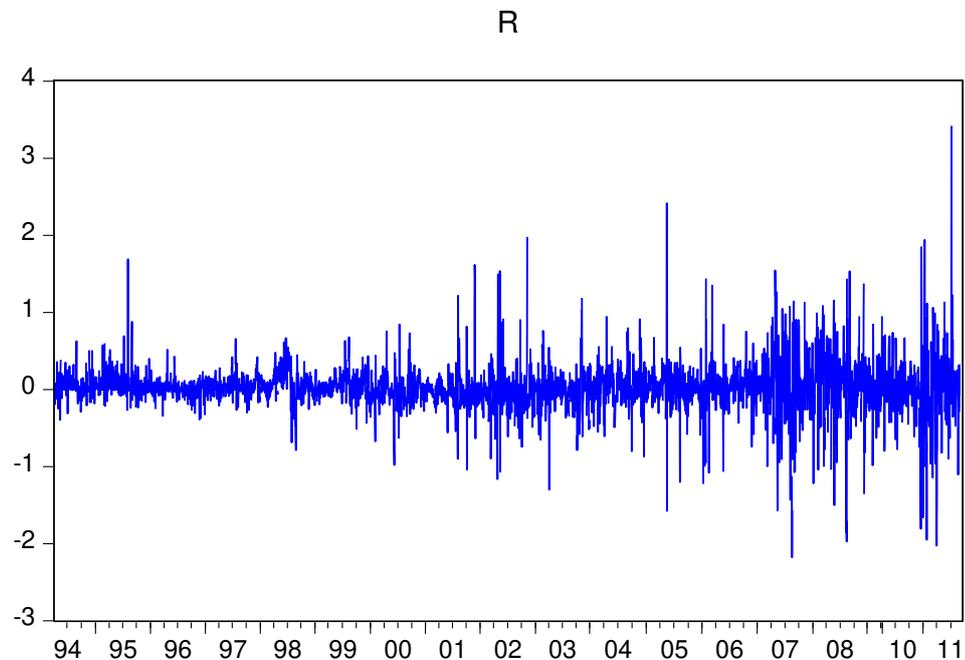


Figure 2. Série des rentabilités journalières de l'indice MASI

- Les fluctuations s'accroissent avec le temps, ceci émane de l'expansion du marché boursier.
- Les grandes fluctuations se concentrent dans la période de formation de la bulle de 2008.
- La volatilité semble suivre un processus dépendant du temps.

D'autres informations sont à tirer de la distribution de la série

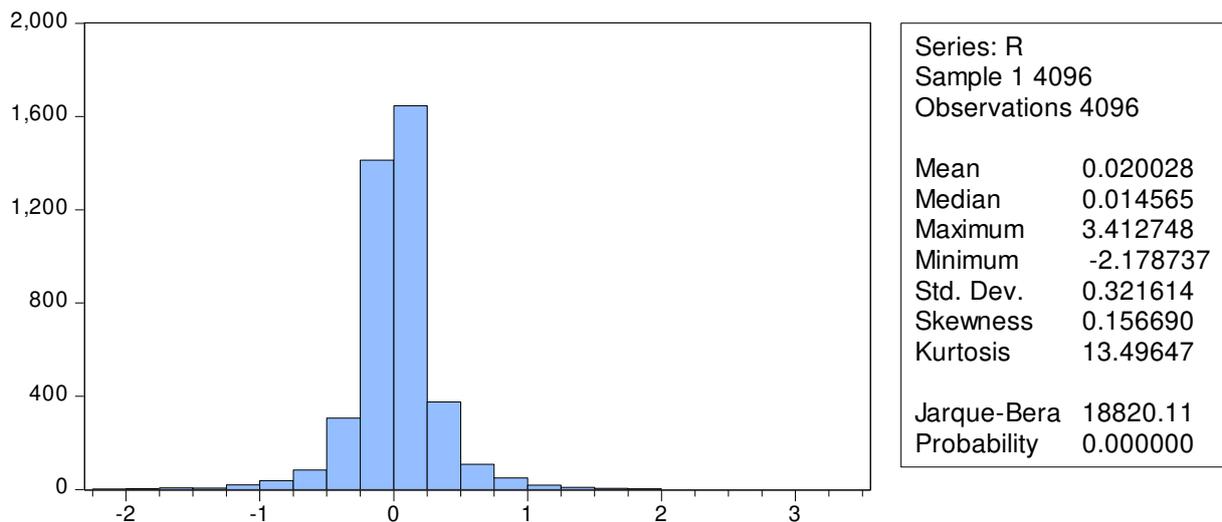


Figure 3. Histogramme des rentabilités de l'indice MASI

La distribution est clairement Leptokurtique, ceci implique que la majorité des rendements se situe autour de la moyenne 0.02, elle est également étalée à droite ce qui indique que la volatilité a tendance à augmenter plus après un choc positif qu'après un choc négatif. Ces résultats sont confirmés par les statistique Kurtosis et Skewness, la normalité est rejetée complètement par le test de Jarque-Bera. Il s'agit-la du premier signe de non linéarité.

1. Le problème du bruit :

Nous avons déjà évoqué le problème qui se pose avec l'existence du bruit dans les données et son influence sur les caractéristiques invariantes de la dynamique sous-jacente de la série.

Afin de dépasser cet obstacle, plusieurs techniques de débruitage sont mises en œuvre, mais comme nous l'avons déjà signalé, les techniques traditionnelles risquent de détruire la structure de l'attracteur. Le recours à des techniques plus adaptées est nécessaire.

La méthode qu'on a choisi d'utiliser est le débruitage par les ondelettes. Cette méthode se base sur la transformée de Fourier en fenêtres glissantes. Les bases d'ondelettes sont construites à partir de deux fonctions : fonction d'échelle et ondelette mère, l'objectif est d'approximer la « vrai » fonction mathématique qui régit la dynamique de la série. Cette construction met en évidence l'importance du bruit dans le signal et nous permet d'extraire le signal pur, la famille d'ondelette utilisée est celle de *Daubechies*, la série filtrée présente évidemment moins de turbulence que la série originale dont la cause principale est le bruit.

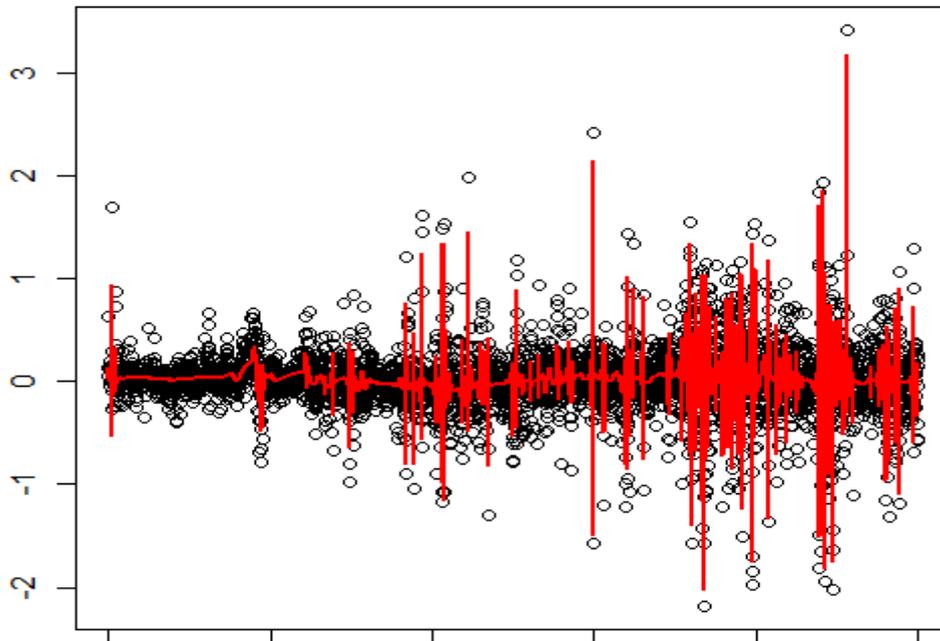


Figure 4. Graphe de la série brute (noir) et filtrée (rouge)

2. Etude de la stationnarité :

La stationnarité de la série est une condition nécessaire pour la pertinence de notre analyse. Le non-respect de cette condition affecte l'ensemble des mesures invariantes et notamment la dimension de corrélation.

Les tests de stationnarité classique basé sur la détection des racines unitaires (Test Dickey-Fuller, Phillips-Perron...) dans les séries sont incompatibles avec la philosophie de l'approche non-linéaire (évident puisque ces test se base sur des modèles linéaires). Par exemple, des travaux récents ont prouvé l'incapacité de ces tests à détecter la non-stationnarité des modèles à changement de régime (Nelson et al (2001)).

Plusieurs techniques sont proposées pour tester cette hypothèse sans avoir besoin de supposer une structure particulière de la dynamique sous-jacente de la série : l'analyse spectrale de la série, le diagramme de récurrence ou le digramme de séparation spatiotemporelle. On propose d'utiliser la méthode de Provenzale et al. conçu pour éviter l'effet de la corrélation temporelle sur l'estimation de la dimension de corrélation.

Cet outil a été élaboré par Provenzale et al en 1992. Dans la présence de corrélation temporelle, la probabilité pour qu'un ensemble de paires de points d'état $\{x_{t_i}, x_{t_i-\Delta t}, x_{t_i-2\Delta t} \dots\}$ dans l'espace d'état aient une distance inférieure à ϵ ne dépend pas seulement de ϵ mais aussi du temps écoulé entre les deux mesures. Cette dépendance peut être détectée en représentant cet ensemble en fonction de deux variables : la séparation temporelle Δt et la distance spatiale ϵ , cela nous permet de définir la fenêtre de Theiler qui indique le stade où la corrélation temporelle disparaît.

Dans le cas d'une série stationnaire, les courbes représentant l'évolution de chaque vecteur atteignent la saturation à un niveau donné, pour le cas inverse (non-stationnaire), les courbes ont une tendance à la hausse de manière continue.

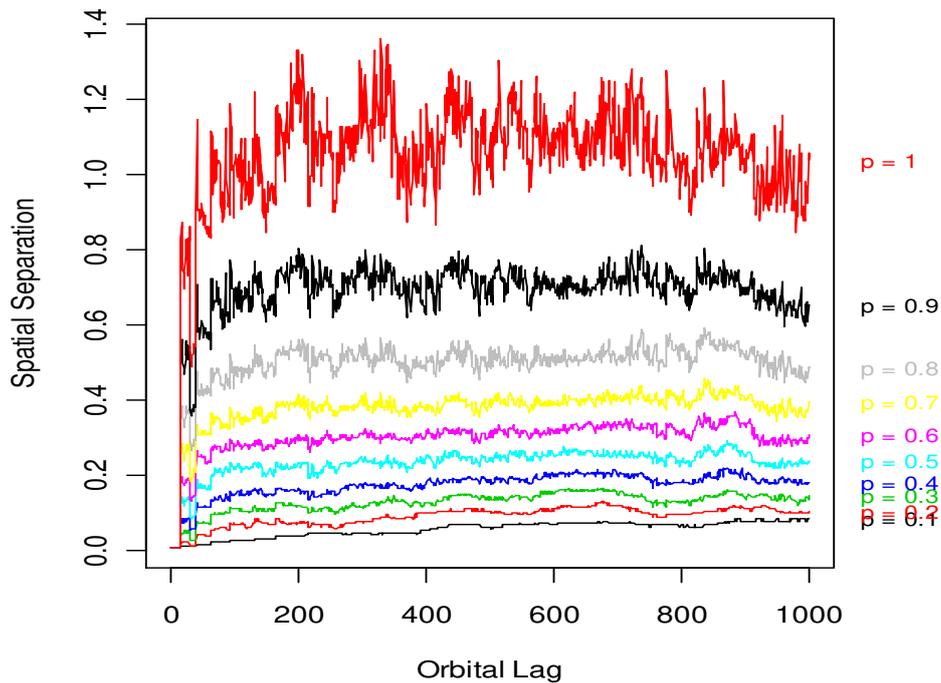


Figure 5. Diagramme de séparation spatio-temporelle

Les courbes atteignent la saturation à une séparation de $\Delta t \approx 400$. On fixe alors la fenêtre de Theiler à $t_{\text{Theiler}} = 500$ (on risque rien si on fixe une fenêtre plus large), il s'agit de l'ordre à partir duquel l'effet temporelle disparaît. Il faut ajouter aussi que l'existence de fluctuation peut être un signe de l'existence d'une périodicité dans le signal. La série est donc stationnaire.

3. Test de non-linéarité :

Une condition nécessaire de l'existence d'une dynamique chaotique est la non-linéarité. Toute la nature de l'analyse dépend de ce test. Cette propriété nous permettra de juger s'il est plutôt judicieux de procéder par des méthodes linéaires ou de passer plutôt par les outils non linéaires.

Le test BDS (Brock, dechert, Scheinkman) de linéarité se base sur l'intégrale de corrélation pour vérifier l'hypothèse d'une série iid contre une hypothèse non spécifiée, cette quantité est essentielle pour le calcul de la dimension de corrélation, selon Grassberger et Procaccia (1983) cette mesure suit une loi d'échelle de type $C_m(\epsilon) \propto \epsilon^D$, où C est l'intégrale de corrélation, D la dimension de corrélation, m la dimension de plongement.

Pour un processus iid, la mesure D croît parallèlement à la dimension de plongement m, alors qu'un processus déterministe a une dimension constante.

Brock, dechert, Scheinkman et le Baron ont exploité cette propriété pour un processus iid, pour ce type de processus l'intégrale de corrélation suit une loi exponentielle de la forme $C_m(\epsilon) = C_1(\epsilon)^m$ avec

$$C_m(\epsilon) = \frac{2}{(N - n_{min})(N - n_{min} - 1)} \sum_{i=1}^N \sum_{j=i+1+n_{min}}^N \Theta(\epsilon - \|x_i - x_j\|) \quad (1)$$

n_{min} Découle de la notion de la fenêtre de Theiler.

Θ est la fonction de Heavyside qui prend 1 si le couple de point est dans le même boule de centre ϵ et zéro sinon.

La statistique BDS est définie par

$$BDS_{m,T}(\epsilon) = \sqrt{T} \frac{C_{m,T}(\epsilon) - C_{1,T}(\epsilon)^m}{\sigma_{m,T}(\epsilon)} \quad (2)$$

Cette statistique suit asymptotiquement la loi normale. L'inconvénient majeur de ce test est qu'il teste l'hypothèse Nulle : processus iid contre une hypothèse alternative inconnue.

On applique le test sur la série brute, puis la série après une transformation linéaire. Dans notre cas, nous allons tester la série brute et la série après une transformation non linéaire en variance de type hétéroscédastique, cela nous permettra d'écarter aussi l'hypothèse de non linéarité en variance, le modèle optimale à choisir est un AR(1)-EGARCH(2,2).

Le test BDS sur la série brute puis sur les résidus de la transformation EGARCH est donné dans les tableaux suivants :

ϵ/σ	<i>M=2</i>	<i>M=3</i>	<i>M=4</i>	<i>M=5</i>	<i>M=6</i>	<i>M=7</i>	<i>M=8</i>	<i>M=9</i>
0.25	27.62506	38.16367	50.93245	70.55318	100.1627	147.0050	225.2580	351.9228
0.5	29.12538	37.43357	45.96383	56.87476	71.55675	92.53239	124.4465	170.5914
0.75	30.19899	36.85043	42.48387	49.03856	56.78185	66.38387	79.48362	96.38013
1	30.07890	35.70860	39.41795	43.35244	47.70721	52.57056	58.71011	66.08706
2	25.73609	28.69891	29.40294	29.87748	30.42391	30.87232	31.21131	31.59325

Tableau 1. Test BDS sur la série

ϵ/σ	$M=2$	$M=3$	$M=4$	$M=5$	$M=6$	$M=7$	$M=8$	$M=9$
0.25	23.82784	33.97171	45.44979	61.68157	85.67230	124.0855	187.1358	291.7150
0.5	24.07238	32.51738	40.45408	50.12526	62.36150	80.01527	105.8599	143.2680
0.75	24.37099	31.47867	37.04917	43.10708	49.80155	58.35313	69.47719	83.68263
1	24.71040	30.70649	34.70249	38.64279	42.65186	47.18641	52.76645	59.28446
2	22.69408	26.73297	28.18069	28.99305	29.70624	30.25708	30.72949	31.18318

Tableau 2. Test BDS sur les résidus

Pour l'ensemble des couples (ϵ, M) , la statistique BDS est supérieure à la valeur critique de la loi normale, ceci nous amène à rejeter l'hypothèse nulle des résidus iid, la dynamique de la série est plus complexe pour être captée par un modèle hétéroscédastique.

4. Test de déterminisme :

La dernière propriété qu'on doit tester à présent est le déterminisme de la série, autrement dit, la fonction qui génère la dynamique de la série suit une règle bien définie qui n'obéit pas aux lois stochastiques. La contamination par un bruit stochastique est possible. Mais l'ossature doit être de nature déterministe et le degré de contamination doit être faible. C'est l'essence de la théorie du chaos qui tend à expliquer les phénomènes déterministes qui apparaissent évoluer d'une manière aléatoire.

Kaplan a mis en place en 1994 un outil afin de mettre en évidence cette propriété. La logique du test repose sur la question suivante : si deux points x_i et x_k sont proches, leurs images x_{i+1} , et x_{k+1} sont-elles proches ?

Kaplan propose l'examen de toutes les paires de points x_i, x_k et de calculer la distance $\delta_{i,k} = |x_i - x_k|$ et la distance entre leurs images $\epsilon_{i,j} = |x_{i+1} - x_{j+1}|$. Il calcul ensuite la moyenne de cette distance : $e(r) = \bar{\epsilon}_{i,j}$ pour $r < \delta_{i,j} < r + \Delta r$.

Pour un processus déterministe, Kaplan a démontré que $\lim_{(r+\Delta r) \rightarrow 0} e(r) = 0$, alors que pour un processus stochastique, cette limite atteint une borne inférieure $\lim_{(r+\Delta r) \rightarrow 0} e(r) = \varepsilon$.

L'inconvénient de ce test est qu'il présente une sensibilité pour le bruit qui peut donner une borne inférieure à la statistique même si la série est déterministe.

Les étapes pour effectuer le test sont au nombre de trois

1) La génération d'un ensemble de séries qui ont la propriété d'être des mélanges aléatoires de la série brute tout en préservant les propriétés de la série originale (autocorrélation, densité spectrale...)

2) Le calcul de la Statistique E de Kaplan pour la série originale et les séries mélanges aléatoires, $\delta_{i,j} = |x_i - x_j|$ étant la distance euclidienne entre les points $x(i)$ et $x(j)$, et $\epsilon_{i,j} = |x_{i+1} - x_{j+1}|$ la distance euclidienne entre leurs images. On calcule une moyenne $e(r) = \bar{\epsilon}_{i,j}$ pour $r < \delta_{i,j} < r + \Delta r$. La statistique est donc :

$$E = \bar{\epsilon}_{i,j} \text{ pour } \delta_{i,j} < r$$

3) La comparaison des résultats de la statistique pour la série originale et pour l'ensemble des mélanges aléatoires, si $\lim_{r \rightarrow 0} E = 0$, cela veut dire que la série est déterministe. Si la série est affectée d'un bruit faible, alors $\lim_{r \rightarrow 0} E < \varepsilon$.

On reporte sur le graphique suivant les courbes mettant en relation la E-statistique par rapport à l'échelle r pour la série originale (représenté en points-tirés) et les mélanges aléatoires (sous forme de boîtes à moustaches) pour différentes valeurs de la dimension de plongement m (de 1 jusqu'à 9).

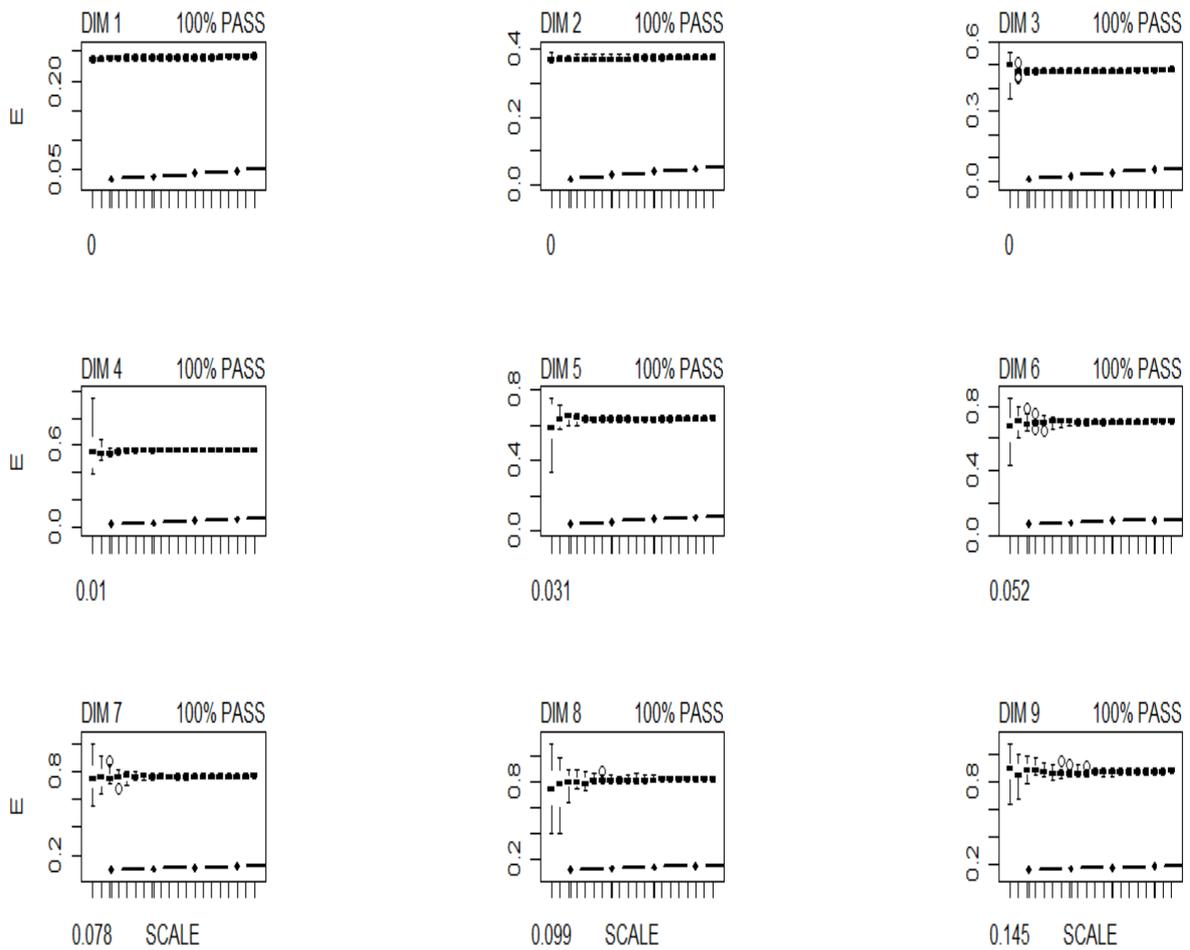


Figure 6. Test de Kaplan

Il y a une séparation nette entre la E-statistique de la série originale et celle des mélanges aléatoires quand r tend vers 0, mais il faut souligner que cette statistique atteint une borne inférieure très faible indiquant toujours la présence d'une composante stochastique dans la série même après le filtrage. Un problème additionnel réside dans le fait que cette statistique ne croit pas exponentiellement avec r , ce qui est le cas pour un processus chaotique.

III. Reconstruction de la dynamique dans l'espace des phases :

La logique de la reconstruction découle du théorème de Takens (1981). Le problème posé par cet auteur est la reconstruction de la dynamique d'un système inconnu à partir de la seule information contenue dans son signal. La solution consiste à plonger l'attracteur de ce signal dans un espace avec une dimension deux fois supérieure à sa dimension.

On reconstruit pour cette fin $(T-m+1)$ vecteurs à m dimensions, ou les m -historiques, dont les composantes sont les valeurs consécutives de la série observée décalée d'un retard τ fixe :

$$X_t^m = (X_{t-(m-1)\tau}, X_{t-(m-2)\tau}, \dots, X_t) \quad (3)$$

m est la dimension de plongement : c'est la dimension de l'espace où l'attracteur est reconstruit. τ est le retard optimal. Il est nécessaire de déterminer d'une manière très fiable les paramètres de la reconstruction.

1. Le retard optimal :

Ce paramètre correspond au premier retard tel qu'il n'existe plus de dépendance temporelle dans la série, le choix de ce paramètre est très important comme le note Kantz et Schreiber (2004). S'il est très petit, il n'y aura plus de différence entre les éléments des m historiques, en sorte que tous les points seront groupés autour de la 1^{er} bissectrice de l'espace de plongement, ceci est appelé redondance par Casdagli et al (1991). Et s'il prend une valeur trop large, il n'y aura plus de corrélation géométrique entre l'ensemble des vecteurs du m historique.

On choisit pour cette fin la technique de l'information mutuelle contenue dans la série :

$$I_\epsilon(\tau) = \sum_{i,j} P_{ij}(\tau) \ln P_{ij}(\tau) - 2 \sum_i P_i \ln P_i \quad (4)$$

P_i est la probabilité que le signal a une valeur i , et $P_{ij}(\tau)$ est la probabilité jointe que le signal $X(t)$ a une valeur i et le signal $X(t+\tau)$ a une valeur j .

Cette statistique décrit la quantité d'information dont on dispose sur $X(t+\tau)$ sachant $X(t)$.

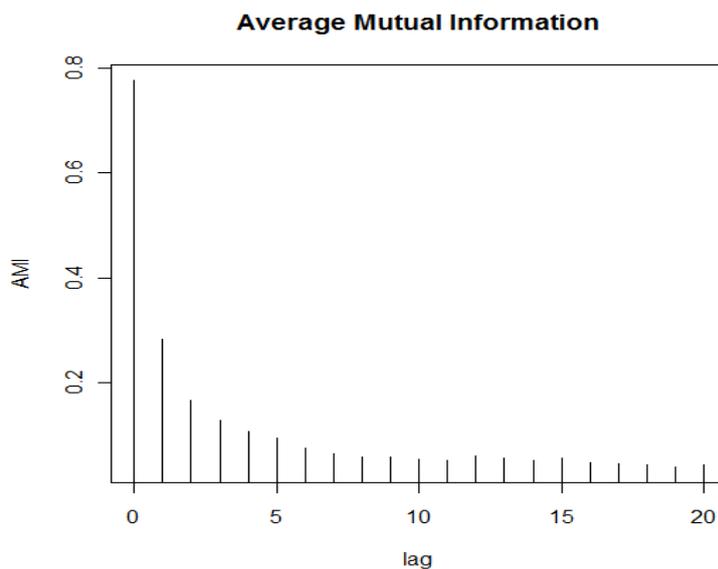


Figure 7. Graphe de l'information mutuelle

Le retard optimal correspond au premier minimum de l'information mutuelle, il s'agit de $\tau = 11$.

2. La dimension de plongement :

La dimension de plongement représente le nombre d'axes nécessaires pour représenter l'attracteur. Le choix de cette dimension découle du théorème de Takens

(1981) : si $m \geq 2D + 1$ ou D est la dimension inconnue du système sous-jacent, alors le comportement des m -historiques imitera celui du vrai système.

Ceci implique un bon choix de la dimension m , or cette dimension ne doit pas être très grande. Si c'est le cas, cela induit une redondance et dégrade la qualité des algorithmes de calcul des caractéristique de l'attracteur.

La technique du plus proche voisin introduite par Kennel et al (1992) constitue une bonne méthode pour son estimation. Le principe de cette technique repose sur le fait que, si m vérifie le théorème de Takens, alors le voisinage dans l'attracteur reconstitué est semblable à celle de l'évolution temporelle. Le principe de l'algorithme consiste à calculer le nombre de faux voisins pour chaque dimension m , jusqu'à ce qu'on atteigne une valeur où le pourcentage de faux voisin est nul.

Pour deux éléments k et n tel que $k \neq n$, et pour un seuil r :

$$X_{fmm}(r) = \frac{\sum_{n=1}^{N-m-1} \Theta \left(\frac{|x_n^{m+1} - x_{k(n)}^{m+1}|}{|x_n^m - x_{k(n)}^m|} - r \right) \Theta \left(\frac{\sigma}{r} - |x_n^m - x_{k(n)}^m| \right)}{\sum_{n=1}^{N-m-1} \Theta \left(\frac{\sigma}{r} - |x_n^m - x_{k(n)}^m| \right)} \quad (5)$$

σ étant l'écart-type, et Θ est la fonction de Heaviside. La formule ci-dessus calcul le nombre de faux voisins, c'est-à-dire, le nombre de points qui ne se trouvent pas dans le voisinage r , on utilise la fenêtre de Theiler pour éliminer l'effet de la corrélation temporelle.

<i>M</i>	1	2	3	4	5	6	7	8	9
<i>Fraction des faux voisins</i>	0.199	0.077	0.041	0.024	0.017	0.012	0.010	0.008	0.007

Tableau 3. Fraction des faux voisins en fonction de la dimension de plongement

La fraction des faux plus proches voisins devient quasiment nulle à partir de la dimension de 4, l'attracteur a besoin de 4 dimensions au moins pour être représenté.

3. Plongement de l'attracteur :

Une fois qu'on a déterminé les paramètres de la reconstruction de l'attracteur, il est très intéressant de le représenter afin d'avoir une idée sur l'aspect géométrique de l'attracteur.

Le problème qui se présente dans notre cas, c'est que l'attracteur doit être plongé dans une dimension > 3 , ce qui ne peut pas être représenté dans les diagrammes usuelles. On propose alors de faire la représentation de la série en 2 et 3 dimensions :

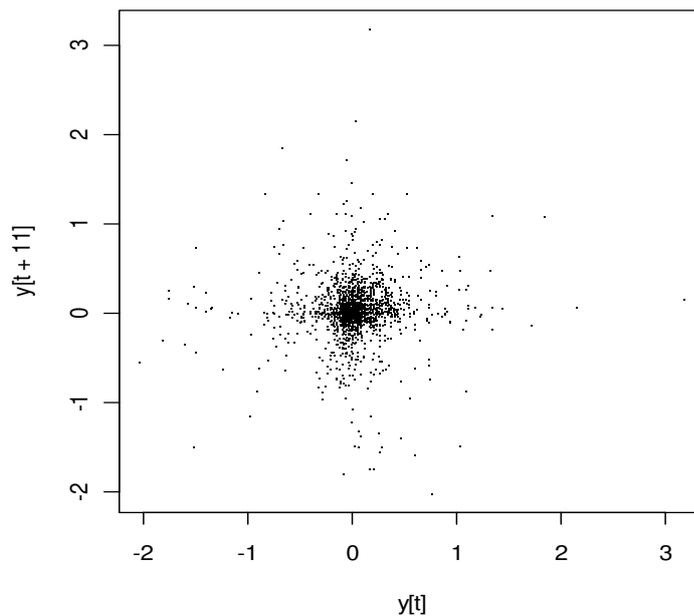


Figure 8. Plongement de la dynamique en dimension 2

Le nuage de point représente le lien dynamique entre $X(t)$ et $X(t-\tau)$, un premier coup d'œil sur le graphique nous donne l'impression que les points tendent à se regrouper autour de la moyenne, un résultat qu'on a déjà rencontré avec la statistique Kurtosis. Mais cette représentation ne nous amène pas très loin. L'insuffisance de nombre d'observation et la

représentation en deux dimensions seulement peuvent nous conduire à de mauvaises conclusions.

La même chose peut être dite de la représentation 3D, la structure paraît très complexe, et on peut comprendre la notion de faux voisin, en passant de la représentation 2D à 3D.

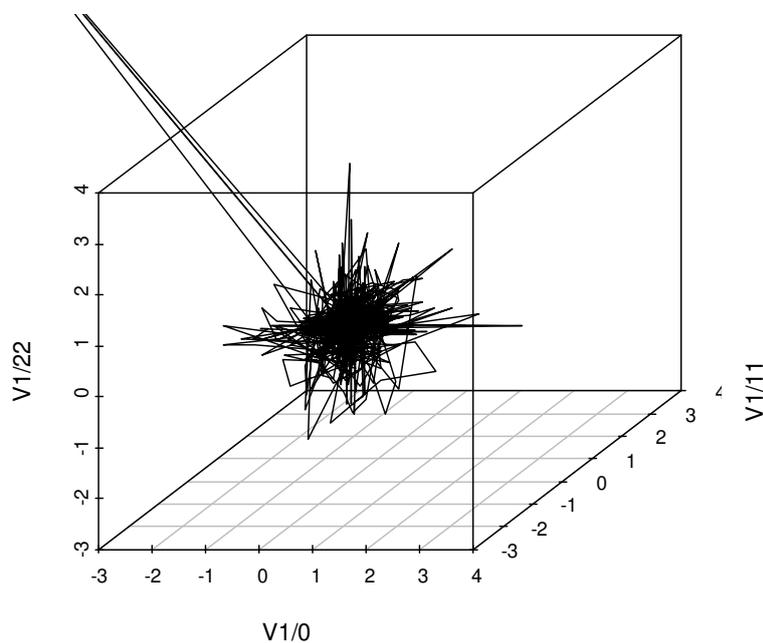


Figure 9. Plongement de la dynamique en dimension 3

Afin de mieux caractériser la nature de ce signal, on doit recourir aux mesures invariantes.

IV. Calcul des mesures invariantes et validation

Il est temps maintenant de mettre en évidence la nature de la dynamique en question, la représentation graphique de l'attracteur en 2D et 3D ne nous a pas été d'une très grande

utilité. On propose ci-dessous un ensemble de statistiques qui nous permettront de juger la nature de la dynamique de l'attracteur de la bourse de MASI.

1. La dimension de corrélation

La dimension de corrélation est très importante dans l'étude des dynamiques non linéaires, Elle permet d'avoir une information topologique concernant le système sous-jacent : elle quantifie la complexité de l'attracteur. D'autre part, elle peut servir à distinguer entre un processus déterministe et un processus stochastique.

Grassberger et Procaccia (1983) ont introduit cette mesure. L'objectif est de mesurer la dimension d'un ensemble de points. La technique proposée se base sur le calcul de l'intégrale de corrélation. La somme de corrélation suit une loi d'échelle de type $C_m(\epsilon) \propto \epsilon^D$, où D est la dimension de corrélation.

La dimension de corrélation augmente d'une manière monotone avec la dimension de plongement pour un processus stochastique, elle ne se stabilise pas. Pour un processus déterministe, la dimension se stabilise à un certain niveau et devient indépendante de m .

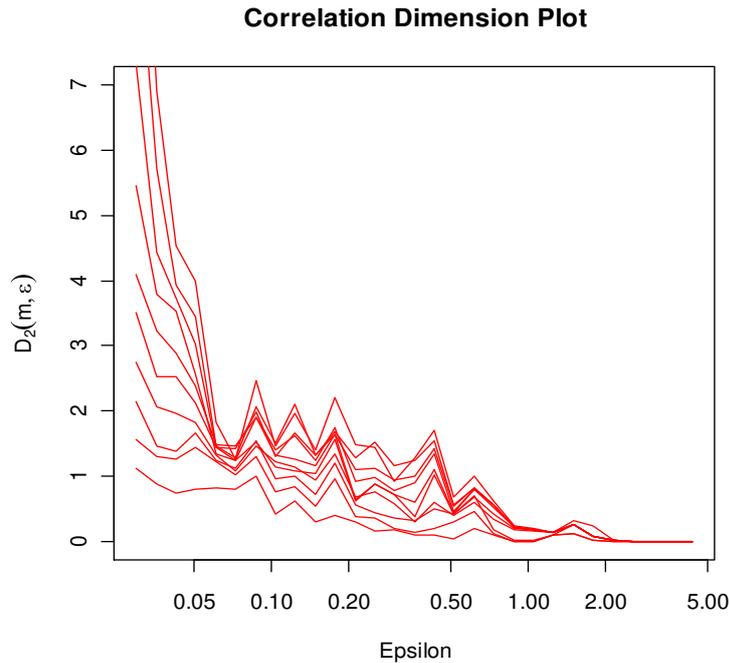


Figure 10. Graphe de la dimension de corrélation en fonction de epsilon, pour différentes valeurs de m

Le graphique ci-dessus représente les courbes de dimension de corrélation associées aux différentes dimensions de plongement allant de 1 à 10. L'objectif est de trouver une « scaling region » qui nous permettra de déterminer la dimension de corrélation. On remarque que les courbes ne sont pas stables, et il existe trois régions :

- la région d'epsilon inférieur à 0.06 qui est caractérisée par la divergence des courbes : il s'agit essentiellement de l'effet du manque de voisins.
- La région avec epsilon supérieur à 0.10, la dimension de corrélation est croissante.
- C'est dans l'intervalle $0.06 \leq \text{epsilon} \leq 0.9$ que les courbes convergent vers une valeur $D = 1.4 \pm 0.15$. Le calcul étant imprécis du fait de l'existence des fluctuations dues à une composante stochastique, il s'agit effectivement d'un ensemble avec une dimension fractale.

2. Les exposants de Lyapunov :

Les exposants de Lyapunov sont un moyen direct de détection du « chaos » dans la série. Ils donnent une information sur l'instabilité locale de l'attracteur et quantifie la sensibilité aux conditions initiales.

L'exposant de Lyapunov s'exprime ainsi :

$$\lambda_i = \lim_{T \rightarrow \infty} \frac{1}{T} \ln |j_i^T| \quad (6)$$

j_i^T sont les valeurs propres de la matrice jacobienne de la fonction en $X(0)$.

Il s'agit d'une quantité qui mesure le taux de séparation de deux trajectoires avec une différence infinitésimale, autrement dit

$$\delta_{\Delta T} \approx \delta_0 e^{\lambda_i \Delta T} \quad (7)$$

$\delta_{\Delta T}$ étant la distance après ΔT , δ_0 est la distance initiale.

Le taux de séparation peut être différent pour des orientations différentes du vecteur de séparation initiale. Cela implique l'existence d'un spectre qui regroupe tous les exposants de Lyapunov.

Plusieurs algorithmes sont conçus pour estimer le plus grand exposant, on propose dans ce cadre l'algorithme de Hegger, Schreiber et Kantz, cette technique a l'avantage d'être robuste même avec un petit nombre d'observations. Un intérêt particulier est accordé au plus grand exposant. Comme le note Kantz et Schreiber (2004), Si

$\lambda_{max} < 0$: Il s'agit d'un point fixe stable.

$\lambda_{max} = 0$: Il s'agit d'un cycle-limite

$\lambda_{max} > 0$: Le processus est chaotique.

$\lambda_{max} = \infty$: Il s'agit d'un bruit.

Dans le tableau suivant, on reporte l'estimation du plus grand exposant de Lyapunov en fonction de la dimension de plongement :

M	λ_{max}
1	0.5306
2	0.4027
3	0.2731
4	0.2159
5	0.1481
6	0.1679
7	0.1489
8	0.1537
9	0.1857

Tableau 4. Le grand exposant de Lyapunov en fonction de la dimension de plongement

Le plus grand exposant a un signe positif. Cela veut dire que pour deux points $X(i)$ et $X(j)$, la distance entre ces deux points croît d'une manière exponentielle selon la formule $\delta_{\Delta T} \approx \delta_0 e^{\lambda_{max} \Delta T}$, δ_0 étant la distance initiale entre les deux points. Et Δt est le temps écoulé. La valeur de cet exposant se stabilise à une valeur de $\lambda_{max} \approx 0.155 \pm 0.015$. Il s'agit bien d'un attracteur « chaotique ».

3. Test sur la base de la prédiction : Le test de Casdagli

L'idée du test est simple : un processus chaotique a la propriété de la divergence exponentielle des trajectoires avec des conditions initiales différentes. Sur la base de cette propriété, Casdagli construit en 1992 son test sur la base d'un modèle local de prévision. L'algorithme proposé par l'auteur est le suivant

- (1) Diviser la série en deux ensembles : un ensemble d'estimation et un ensemble de prévision.

- (2) Choisir la dimension de plongement et le retard optimal.
- (3) Construire les m-historiques et choisir les étapes T de prévision.
- (4) Calculer les distances entre le vecteur de test et le vecteur d'estimation.
- (5) Classer les distances et choisir les k plus proches voisins $x_{j(1)}, \dots, x_{j(k)}$ de x_i et estimer par les moindres carrés le modèle : $x_{j(l)+T} = \alpha_0 + \sum_{n=1}^m \alpha_n x_{j(l)-(n-1)\tau}$, avec $l = 1, \dots, k$.
- (6) Utiliser le modèle pour tester les vecteurs de test x_i et calculer l'erreur de prévision.
- (7) Calculer l'erreur quadratique moyenne : Si un minimum d'erreur est atteint pour une taille ϵ de voisins minimale ceci est un signe que le processus est chaotique, si le minimum n'est atteint que pour un ϵ élevé, alors le modèle le plus approprié à la série est un modèle linéaire stochastique, et pour une taille ϵ intermédiaire, le processus en question est un cycle-limite ou un processus non-linéaire stochastique.

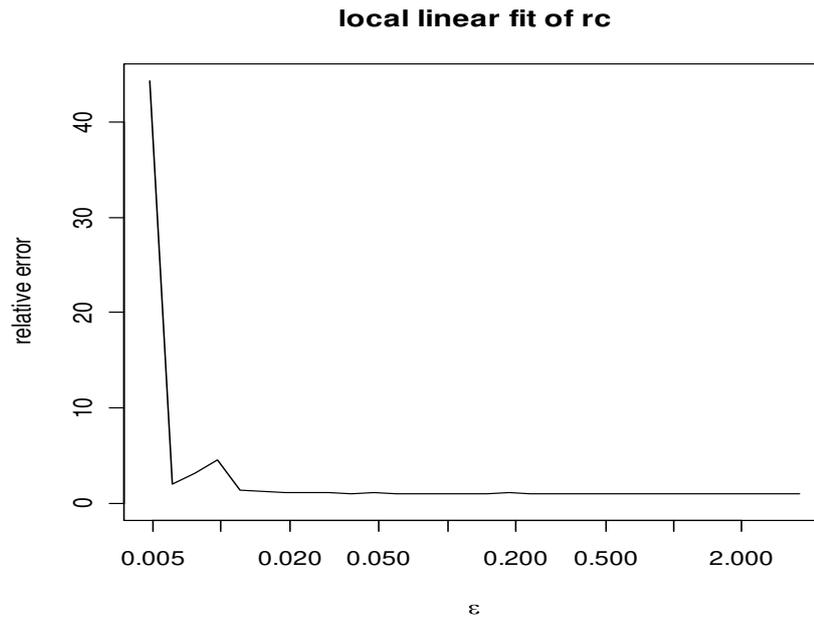


Figure 11. Erreur relative de prévision en fonction d'epsilon

On remarque que l'erreur relative de prévision atteint son minimum pour une taille ϵ faible, se stabilise et ne suit pas un rythme exponentielle, il s'agit d'un cas que Casdagli qualifie de non-linéaire stochastique.

Conclusion:

Nous avons défini dans cet article une méthodologie qui fixe les modalités de l'application des tests empiriques, il s'agit de répondre au préalable à trois questions, à savoir, la série est-elle stationnaire ? La série est-elle linéaire ? La série est-elle stochastique ? Et nous avons montré que la série est stationnaire, non-linéaire, en se basant sur la statistique de la dimension de corrélation. Mais nous avons aussi conclu que la série contient une composante stochastique qui influençait presque toutes les statistiques. Cette réponse est essentielle pour notre étude et elle a confirmé notre hypothèse de départ, mais le caractère stochastique de la série ne nous a pas permis d'exploiter tous les outils de reconstruction de la dynamique dans l'espace des phases, des outils qui sont adaptés essentiellement à des processus déterministes. Mais nous avons tout de même abouti à un résultat fondamental, à savoir le caractère « chaotique » et la sensibilité aux conditions initiales de la série.

Ces résultats ont des implications importantes sur notre compréhension du marché boursier, et du marché financier en général, et plusieurs postulats doivent être remis en cause. Le premier postulat est celui de l'efficacité du marché et la rationalité des anticipations, ces deux principes impliquent une évolution des cours en termes de marche aléatoire. Or, nous avons démontré que ce n'est pas le cas.

Le deuxième postulat est la convergence vers un point d'équilibre stable, alors que nous avons montré que le marché exhibe un comportement de « chaos stochastique » très complexe et qui dépasse les propriétés d'un marché convergent, à travers l'exposant de

Lyapunov, et qu'une simple perturbation peut avoir des conséquences dramatiques sur l'évolution du marché.

Un troisième postulat n'est toujours pas admis définitivement, celui de la nature stochastique du marché. La limite imposée par le nombre d'observations nous amène à accepter l'hypothèse de la nature stochastique avec prudence. Comme le montre Kantz et Schreiber (2004), dans une dynamique Hyper-chaotique ou un chaos de haute dimension, le comportement de la série ainsi que les mesures invariantes est semblable à un processus stochastique, l'exemple du générateur de nombre aléatoire « processus déterministe » en est un exemple. On tend aussi à associer les bruits à des chaos de haute dimension, la condition posée par les deux auteurs sur le nombre d'observation nécessaire pour le traitement d'un processus chaotique est $N = 10^D$ où D : est la dimension du processus. Mais la réponse à cette question dépendra de l'accumulation des observations.

Les interrogations de cette thèse peuvent être appliquées aussi à d'autres types de marché, et le l'hypothèse de non-linéarité ne concerne pas seulement le marché financier.

Les implications de la non-linéarité sont révolutionnaires, et elles nous emmènent à réfléchir profondément sur la méthodologie utilisée en économétrie pour tester les résultats d'une théorie économique. La véracité des conclusions des études économétriques sont donc remise en cause, mais la pertinence de cette accusation est conditionnée par le développement de nouvelles techniques de tests empiriques moins sensibles au nombre d'observations.

References:

Arnold L. (23 avril 1974), *Stochastic Differential Equations: Theory and Applications*, Wiley Interscience, 1ère édition.

Bakker R., Schouten C., Coppens M-C., Takens F., Lee Giles C. et Van Den Bleek M. (2000), « Robust Learning of Chaotic Attractors », MIT Press, pp. 879—885.

Benhabib J. et Day H. (1981), « Rational Choice and Erratic Behaviour », *The Review of Economic Studies*, Vol. 48, No. 3, pp. 459-471.

Bonache A. (2008), « Les ventes de produits innovants à la mode sont-elles chaotiques ? Le cas des ventes de Game Boy au Japon », Munich Personal RePEc Archive.

Brock W., Dechert W.D., Scheinkman J. et Lebaron B. (1991), « A test for independence based on the correlation dimension », *Econometric Reviews*, vol.15, n°3.

Brusse M. et Weihs C. (2004), « Lyapunov Exponent for Stochastic Time Series ».

Campbell J. Y., Lo A. W. et Mackinlay A. C. (9 décembre 1996), *The Econometrics of Financial Markets*, Princeton University Press.

Casdagli M. (1992), « Chaos and Deterministic versus Stochastic Non-Linear Modeling », *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 54, No. 2, pp. 303-328

Chan K. S. et Tong H. (1994), « A Note on Noisy Chaos », *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 56, No. 2, pp. 301-311.

Dennis B., Desharnais A., Cushing J. M., Henson M. et Constantin R. F. (2003), « Can noise induce chaos? », *Oikos* 102, pp. 329—339.

Eckmann J.-P., Kamphorst O., Ruelle D. et Ciliberto S. (1986), « Lyapunov Exponent from time series », *Physical Review A*, Vol. 34, No. 6.

Eckmann J.-P. et Ruelle D. (1985), « Ergodic Theory of Chaos and Strang Attractor », *Reviews of Modern Physics*, Vol. 57, No. 3, Part 1.

Enescu B., Ito K., et Struzik Z. R. (2004), « Wavelet-Based Multifractal Analysis of Real and Simulated Time Series of Earthquakes », *Annals of Disas. Prev. Res. Inst., Kyoto Univ.*, No. 47 B.

Franses P. H. et Van Dijk D. (4 septembre 2000), *Nonlinear Time Series Models in Empirical Finance*, Cambridge University Press, 1ère édition.

Girerd-Potin I. et Taramasco O. (1994), « Les rentabilités à la bourse de Paris sont-elles chaotiques ? », *Revue économique*, Volume 45, Numéro 2, pp. 215 – 238.

Grassberger P. et Procaccia I. (1983), « Estimation of the Kolmogorov Entropy from Chaotic Signal », *Physical Review A*, Vol. 28, No. 4.

Guégan D. (2003), *Les Chaos en Finance : Approche Statistique*, Economica.

Gundlach M. (1991), « Chaos in Random Dynamical Systems ».

Halsey C., Jensen H., Kadanoff P., Procaccia I. et Shraiman I. (1985), « Fractal Measures and their Singularities, The Characterization of Strange Sets », *Physical Review A*, Vol. 33, No. 2.

Heiner A. (2001), « Origin of predictable behavior », *The American Economic Review*, Vol. 73, No. 4.

Kantz H., Schreiber T. (26 janvier 2004), *Nonlinear Time Series Analysis, Cambridge University Press*, 2ème édition.

Kaplan T. (1994), « Exceptional Events as Evidence for Determinism », *Physica D* 73: 38-48.

Kennel B., Brown R. et Abarbanel D. I. (1991), « Determining Embedding Dimension for Phase-space Reconstruction Using a Geometrical Construction », *Physical Review A*, Vol. 45 No. 6.

Marwan N., Romano M. C., Thiel M. et Kurths J. (2006), « Recurrence Plots for the Analysis of Complex Systems », *Physics Reports* 438, pp. 237 - 329.

NELSON C. ZIVOT E. ET PIGER J. M. (2001) « Markov Regime Switching and Unit Root Tests », Federal Reserve Bank of ST. LOUIS.

Peters E. E. (1991), « A Chaotic Attractor for the S&P 500 », *Financial Analysts Journal*, Vol. 47, No. 2, pp. 55-62+81.

Peters E. E (12 janvier 1994), *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*, Wiley, 1ère édition.

Peters E. E. (août 1996), *Order and Chaos in the Capital Market: a New View of Cycles, Prices, and Market Volatility*, Wiley, 2ème édition.

Parlitz U. (1992), « Identification of True and Spurious Lyapunov Exponents from Time Series », *International journal of bifurcation and chaos*, Vol. 2, No. 1, pp. 155-165.

Priestley M. B. et Subba Rao T. (1969), « A Test for Non-Stationarity of Time-Series », *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 31, No. 1, pp. 140-149

Schreiber T. et Schmitz A. (1997), « Discrimination power of measures for nonlinearity in a time series », Physical Review E, Vol. 55, No. 5.

So P., Ott E., Schiff J., Kaplan T., Sauer T. et Grebogi C. (1995), « Detecting Unstable Periodic Orbits in Chaotic Experimental Data », Physical Review letters, Vol. 76, No. 25.

Strozzi F. (2002), « Application of Non-Linear Time Series Analysis Techniques to High Frequency Currency Exchange Data », Liuc Papers n. 99, Serie Metodi quantitative 14.

Theiller J. (1987), « Efficient Algorithm for Estimating the Correlation Dimension from a Set of Discrete Points », Physical Review A, Vol. 36, No. 9.