

ANNE CORCOS

CYRILLE PIATECKI

## **CAC 40 et chaos : les diagrammes de proximité**

*Journal de la société statistique de Paris*, tome 135, n° 3 (1994),  
p. 73-98

[http://www.numdam.org/item?id=JSFS\\_1994\\_\\_135\\_3\\_73\\_0](http://www.numdam.org/item?id=JSFS_1994__135_3_73_0)

© Société de statistique de Paris, 1994, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

# CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ

par Anne CORCOS  
Université Paris-II & CEFIB  
par Cyrille PIATECKI  
Université d'Orléans et CRESEP

## INTRODUCTION<sup>1</sup>

Les découvertes récentes dans le domaine du Chaos – voir Bergé & al. (1984), Eckmann & Ruelle (1985), Lorenz (1989), Ruelle (1989) –, ont mis en avant des *systèmes déterministes à petit nombre de variables endogènes* générant des comportements complexes. Ceux-ci s'expliquent, en partie, par le *caractère non-linéaire du processus* et surtout par la propriété de *sensibilité aux conditions initiales – S.C.I. –* qui donne un aspect irrégulier, et donc imprédictible, au processus : ces systèmes sont tels qu'une petite perturbation dans les conditions initiales génère une trajectoire très nettement distincte de celle du processus non perturbé.

Ainsi, depuis une dizaine d'années, physiciens et statisticiens ont développé divers protocoles pour déceler la présence de chaos dans une série : distinction entre série chaotique et aléatoire – voir Brock & Dechert (1988) –, évaluation de la dimension de corrélation, des exposants de Lyapunov, de l'entropie... – pour des précisions, voir Bergé (1989), Eckmann & al. (1986), Eckmann & Ruelle (1985) –.

On peut ainsi s'interroger, à juste titre, sur l'apport de la théorie du Chaos dans l'étude des séries temporelles économiques et financières – voir Anderson & al. (1988), Baumol & Benhabib (1989), Brock, Dechert, Grandmont & Malgrange (1986), Hsieh, Kelsey (1989), Mirowski (1990), Scheinkman... –. Peut-on mettre en évidence dans la chronique, une composante *déterministe* structurée par un petit nombre de variables ? La présence de phénomènes déterministes en finance pourrait

---

1. Cet article doit énormément à J.P. Eckmann et à D. Ruelle qui en ont assuré la direction. Je leur exprime à cette occasion toute ma reconnaissance. Bien entendu, toute erreur ou omission qui pourrait subsister serait mon propre fait. Il a été présenté lors du colloque de l'ISEG en 1991.

expliquer l'aptitude des professionnels de la finance à faire des gains plus substantiels qu'un non-initié<sup>1</sup>.

Néanmoins, il est peu vraisemblable que les évolutions des marchés financiers puissent être imputables exclusivement à une structure déterministe, car même s'ils intègrent la totalité de l'information passée et à venir, ils semblent sujets, dans un degré qui doit encore être précisé, à des phénomènes qui ne peuvent être représentés que par des chocs aléatoires. Il ne semble donc pas abusif d'envisager une représentation mixte des cours boursiers dont la complexité des évolutions temporelles émaneraient ainsi tout à la fois d'une composante déterministe et d'une composante stochastique.

De nombreux modèles empruntant à la théorie économique ont été construits – Benhabib & Day (1981), Boldrin (1989, 1991), Day (1982, 1983), Day & Walter (1989), Deneckere (1986), Weddepohl (1990), Goodwin (1990) – ; des applications à l'économie ont été réalisées – voir notamment Barnett & Chen (1988) pour leur étude sur les agrégats monétaires, Grassberger (1986) pour ses travaux sur les attracteurs climatiques, Scheinkman & Lebaron (1989a et 1989b) pour leurs investigations sur le niveau général des prix et les rendements des obligations... –. Mais l'utilisation de ces outils en économie où les séries temporelles sont trop courtes et trop imprécises reste encore fragile – voir Eckmann & Ruelle (1990) –. On a longuement fait état des outils d'analyse développés dans la théorie du chaos. Malheureusement, les séries économiques et financières semblent impropres à révéler du chaos à partir de ce type d'outils. Les séries sont trop courtes et limitent sérieusement la validité des tests.

Dans cette optique, la *méthode des proximités*, qui est employée ici, est plus appropriée à ce type de données, elle constitue un palliatif. Elle a été mise au point par Eckmann, Kamphorst & Ruelle (1987) et se fonde sur les concepts élaborés dans la théorie du chaos. La philosophie de cette approche – *diagramme de proximité* ou *Récurrence Plot* – sera d'extraire le sens temporel d'un signal en dépassant la connaissance détaillée de la dynamique sous-jacente. *On recherche, en particulier, les propriétés qui caractérisent l'évolution temporelle du signal en terme de critères géométriques.* L'enjeu est de déceler des similitudes de comportement des points de la série à travers le temps. Pratiquement, le principe, à cheval entre les mesures statistiques de corrélation ou d'analyse spectrale et celles d'analyse des données, consiste à établir des indices de proximité entre différents *agents* ou équilibres temporels et d'en inférer des principes de comportement, des règles de conduite ou des évolutions habituelles. Cette analyse procède davantage d'une analyse dynamique que d'une analyse statique : on recherche dans les processus d'ajustements temporels des comportements récurrents. La présence de tels schémas nous permettrait de croire que l'évolution temporelle étudiée provient d'une dynamique comportant une structure, une loi.

---

1. Certains pourront répliquer à cette affirmation que de manière récurrente depuis Treynor (1965), il a été démontré que les fonds communs de placement – qui sont gérés par des spécialistes – n'ont pas une rentabilité supérieure à celle du marché. Mais, Roll (1977) a montré que ces résultats sont conditionnels à l'efficacité moyenne-variance de l'indice qui représente le marché.

Ce travail se présente donc sous la forme suivante : la méthodologie de construction du diagramme de proximité est développée dans une première section où il en est fait une application à des séries simulées. La seconde section met en perspective cette méthodologie avec des outils d'analyse plus traditionnels – autocorrélation, analyse spectrale, analyse des données –. Le troisième point est consacré à l'étude plus spécifique de l'indice CAC40.

## 1. Description du test

La *méthode des proximités*, développée dans ce qui suit, est essentiellement *géométrique*. Elle permet une appréciation des propriétés de la série étudiée au travers de graphiques.

La nature de la récurrence Plot évoque celle de l'analyse des données. En effet, ces techniques de traitement des données n'impliquent apparemment aucune hypothèse sur les phénomènes étudiés. Ainsi, la validité de la méthode des proximités ne dépend pas d'hypothèses invérifiables sur la nature de la série que l'on va utiliser : la condition de stationnarité de la série n'est pas imposée parce qu'aucune structure probabiliste ne lui est associée. En évitant ainsi de faire entrer la série dans une structure mathématique formelle, on élargit le champs d'investigation des liaisons entre les points de la série. Néanmoins, il faut souligner que dans ce cadre d'analyse, les tests statistiques traditionnels ou les intervalles de confiance qui permettent de tester la robustesse et la fiabilité des hypothèses ne peuvent être utilisés.

### 1.1. Reconstruction de la chronique dans un pseudo-espace des phases

Dans sa conception, cette méthode s'inspire de pratiques employées pour étudier les aspects chaotiques des séries, notamment la reconstruction de la série dans un pseudo-espace des phases, le choix de certains paramètres, l'emploi de la norme du maximum...

Supposons qu'un système dynamique soit généré par un ensemble d'équations telles que :

$$x_{t+1}^i = g_i(x_t), \quad x \in \mathfrak{R}^n, \quad i = 1, \dots, n. \quad (1)$$

En économie, on ne connaît ni la forme structurelle de  $g_i(x_t)$ , ni la dimension  $n$  de l'espace dans lequel évolue la trajectoire globale de  $x_t$ . En l'absence d'information sur le nombre et la nature des variables intervenant dans le système, le principe d'analyse sera de *construire des variables indépendantes à partir de l'évolution temporelle d'un signal unidimensionnel*  $u_t$ , qui est une certaine fonction observée de  $x_t$  soit  $u_t = h(x_t)$ .

Une des façons de générer ces séries est la méthode des retards qui consiste à partir de l'observation d'une série unidimensionnelle,  $\{u_t\}_{t=1}^N$ , pour construire un système dynamique artificiel  $\{X_t^m\}_1^N$  qui évoluera dans un *pseudo-espace des phases*

de dimension  $m$ . Ce dernier permet de bien reproduire linéairement la dynamique asymptotique du système (1).

Le passage de la chronique  $\{u_t\}_{t=1}^N$  à  $\{X_t^m\}_1^{N'}$  va correspondre à un changement d'espace, de repère : *il va s'agir de projeter une chronique de dimension très grande vers un espace de dimension finie  $m$* . Les paramètres de la transformation sont  $m$  et  $t_0$  et définissent respectivement une dimension et un écart temporel. En partant de la chronique observée  $\{u_t\}_{t=1}^N$ , on compose des vecteurs  $X_t^m$ , de dimension  $m$ , appelés *m-historiques* des observations. Chaque vecteur  $X_t^m$ , de la nouvelle chronique  $\{X_t^m\}_1^{N'}$  dépend des valeurs observées de la variable  $\{u_t\}$ , à l'instant  $t$  et aux  $(m-1)$  instants décalés de  $\tau_1, \tau_2, \tau_3, \dots, \tau_{m-1}$ . En pratique, on choisit les temps de retard tels que  $t_0$  soit fixé et que  $\tau_k = k \cdot t_0$ .

Les valeurs de la nouvelle chronique s'écrivent :

$$\begin{cases} X_1^m = (u_1, u_{1+t_0}, u_{1+2t_0}, \dots, u_{1+(m-1)t_0}) \\ X_2^m = (u_2, u_{2+t_0}, u_{2+2t_0}, \dots, u_{2+(m-1)t_0}) \\ \vdots \\ X_N^m = (u_{N-(m-1)t_0}, u_{N-(m-2)t_0}, u_{N-(m-3)t_0}, \dots, u_N) \end{cases}$$

où

$$N' = N - (m-1)t_0.$$

*Cette procédure revient ainsi à projeter l'évolution du signal d'un espace de très grande dimension voire de dimension infinie, sur un espace de dimension finie  $m$ .*

## 1.2. Calcul des indices de proximité

Le diagramme de proximité est construit sur la base de la nouvelle série : il s'attache à nous fournir une représentation visuelle des liaisons temporelles. En effet, pour tout  $X_t^m$  de  $\{X_t^m\}_1^{N'}$ , on s'intéresse aux points de cette même série qui lui sont le plus *proches*. Plusieurs remarques s'imposent :

a) Ici, la *notion de proximité* entre les points est entendue au sens de la *distance dite du maximum*. Ainsi, la distance entre deux points  $X_i^m$  et  $X_j^m$  est définie par :

$$d(X_i^m, X_j^m) = \text{Max}_{k \in \{1 \dots m\}} |X_i^m(k) - X_j^m(k)|.$$

où  $X_i^m(k)$  est la valeur de la  $k^{\text{ième}}$  coordonnée du point  $X_i^m \in \mathfrak{R}^m$ .

En d'autres termes, la distance du maximum correspond au plus grand écart absolu entre les coordonnées de deux points.

b) Dans notre étude, le nombre de points voisins<sup>1</sup> employés pour la construction du diagramme de proximité, est choisi égal à six dans un souci de clarté des graphiques<sup>2</sup>.

De façon plus formelle, on note  $\Omega$  l'ensemble des indices  $\{s_m(i, k) \mid k = 1, 2, \dots, 6\}$  des six points les plus proches du point  $X_i^m$ ,  $i \in \{1, \dots, N - t_0(m - 1)\}$ . Cet ensemble vérifie les conditions :

$$\left\{ \begin{array}{l} s_m(i, 1), \text{ un index pour lequel } d(X_i^m, X_{s_m(i, 1)}^m) \text{ est minima} \\ s_m(i, 2), \text{ un index } \neq s_m(i, 1), \text{ pour lequel } d(X_i^m, X_{s_m(i, 2)}^m) \text{ est minima} \\ \cdot \\ \cdot \\ s_m(i, 6), \text{ un index } \neq \{s_m(i, 1), \dots, s_m(i, 5)\}, \text{ pour lequel } d(X_i^m, X_{s_m(i, 6)}^m) \text{ est minimal.} \end{array} \right.$$

Les distances associées s'écrivent :  $d_{ik} = d(X_i^m, X_{s_m(i, k)}^m)$  pour  $k \in \{1, \dots, 6\}$  .

Puis, à partir de  $\{X_i^m\}_{i=1}^{N'}$ , on examine la répartition géographique des six voisins les plus proches de chaque  $X_i^m$ .

Il est à noter que cette répartition géographique est étonnamment robuste au changement de la distance à partir de laquelle le diagramme de récurrence est construit, ce qui a pu être mis en évidence à partir de simulations parallèles utilisant tour à tour la distance du maximum et la distance euclidienne<sup>3</sup>.

### 1.3. Diagrammes de proximité

Nous allons voir comment ils se construisent avant d'analyser l'application de la méthode à la série du CAC40.

1. Le nombre de points voisins auxquels on devrait s'intéresser devrait être calculé sur la base de critères de détermination des exposants de Lyapunov – voir Eckmann & al. (1986) – : si  $d_M$  désigne la dimension de la matrice à partir de laquelle on calcule les exposants de Lyapunov –  $d_M$  supérieure au nombre d'exposants de Lyapunov positifs –, il faudrait choisir un nombre de points voisins supérieur à  $\text{Min}(2d_M, d_M + 4)$ . Cependant, les exposants de Lyapunov pour des séries courtes ne sont pas fiables – voir à ce propos Eckmann & Ruelle (1990) –, de ce fait, nous nous sommes limités à l'étude des six plus proches voisins.

2. On a également redéfini l'ensemble des proximités en retenant tous les points contenus dans une boule de rayon fixé, à l'instar de ce qui se fait dans le calcul des dimensions de corrélation. Cette procédure présente le double avantage de laisser inchangé le diagramme de proximité suite à un allongement de la série et d'être informatif du point de vue de sa densité. Mais il est nécessaire de préciser qu'elle n'a pas modifié sensiblement les résultats trouvés pour le CAC40 quand on ne conservait que les six points les plus proches. C'est est pourquoi, pour des raisons de commodité informatique, la dernière procédure a été retenue.

3. La distance euclidienne est obtenue par sommation des carrés des écarts entre les points :

$$d(X_i^m, X_j^m) = \sum_{k=1}^m (X_i^m(k) - X_j^m(k))^2$$

**A. Construction**

Dans un cadre carré ayant  $N'$  points de côté,  $N'$  correspondant à la longueur de la série reconstruite, on repère l'ensemble  $\omega$  : on marque d'un point, chacun des six indices  $s_m(i, k)$ ,  $k \in \{1, \dots, 6\}$ , des points les plus proches de  $X_i^m$ . En d'autres termes, pour chaque point d'abscisse, on noircit en ordonnée les six points qui lui sont le plus proches et ceux qui sont éloignés sont laissés en blanc. Ce schéma est ce que l'on appelle un *diagramme de proximité*.

**B. Analyse**

Nous allons nous familiariser avec cette méthodologie à partir de chroniques simulées. Nous étudierons plusieurs types de chroniques :

- des séries déterministes non-linéaires ;
- des processus stochastiques ;
- des chroniques combinant déterminisme et hasard ;
- des séries générées par deux processus successifs différents.

Le tableau 1 présente l'ensemble des séries simulées. On note que toutes les séries ne sont pas stationnaires.

Tableau 1.

I - Séries déterministes	II - Séries aléatoires pures	III - Combinaisons additives	IV - Combinaisons multiplicatives
Lien de dépendance quadratique	$A1 : Y_t \sim \mathcal{N}(0,1)$	Bases quadratiques	Base quadratique
$DQ1 : Y_t = \alpha t^2 + \beta t + \gamma$	$A2 : Y_t \sim \mathcal{U}([0,1])$	$AQ1 : DQ1 + A1$	$MQ1 : Y_t \sim \text{Modèle ARCH}(1)$
$DQ2 : Y_t = \alpha Y_{t-1}(1 - Y_{t-1})$			
$AT1 : Y_t = \begin{cases} \mathcal{N}(0,1) & \text{si } t \leq t_0 \\ \mathcal{N}(0, \sigma) & \text{sinon } -\sigma \neq 1 - \end{cases}$		Modèle à changement de régime	

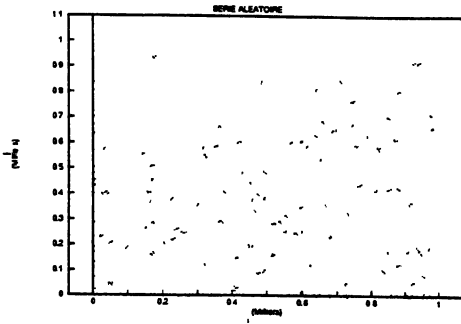
On trace les diagrammes de proximité de chacune de ces séries (voir ci-dessous).

À partir de ces schémas se dégagent des enseignements de plusieurs natures sur le système qui génère le phénomène.

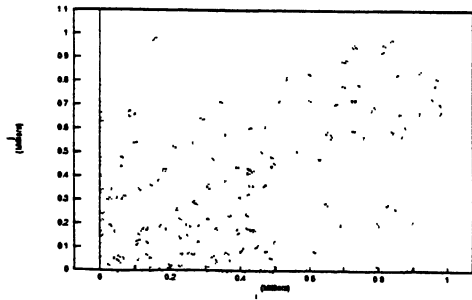
**C. Conclusions**

a) La présence d'un *trend monotone* dans la série *ne permet pas d'exploiter* utilement et avantageusement *la méthode* dans la mesure où proximités temporelle et géographique sont dans ce cas synonymes : les six voisins géographiques d'un point sont situés temporellement de part et d'autre de ce dernier. Il est préférable d'étudier une série sinon stationnaire du moins non divergente, à l'image des séries chaotiques.

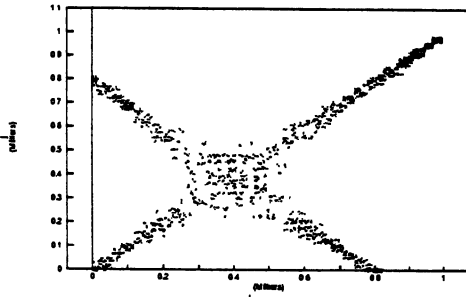
## CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ



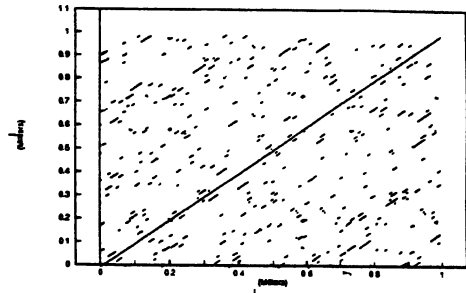
A1, A2 et MQ1



AT1



AQ1



DQ2

b) L'aspect localement diffus d'un diagramme de proximité peut être la conséquence d'un affaiblissement de la tendance d'une série ou bien celle d'un changement de structure.

c) Néanmoins, ce ne saurait constituer un outil pour distinguer le caractère déterministe ou stochastique d'un processus.

d) La distribution des proximités peut attester sinon de l'existence d'une structure sous-jacente unique – ou de rupture de structure –, du moins de liens temporels systématiques ou transitoires – on pourra consulter Scheinkman et Lebaron (1989a) qui mettent en évidence une telle rupture sur le produit national brut des USA –.

e) Cas non linéaire.

Le diagramme de proximité de la courbe logistique se distingue de celui du bruit blanc. Bien que dans les deux cas, l'évolution de la série soit très erratique, le diagramme de proximité est beaucoup moins confus et diffus pour la série *chaotique*. En cela, il évoque un certain ordre, inhérent à cette série.



## CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ

- Des lignes apparaissent sur ce dernier graphique : elles constituent un premier pas vers l'existence d'une structure sous-jacente de détermination de cours boursiers. La présence de lignes atteste en effet, d'un phénomène de répétition dans les évolutions de cours. Elles témoignent de la présence de séquences de cours qui se ressemblent par instants : une ligne – de longueur 10 – démarrant au point  $(i, j)$  traduit le fait que la séquence des cours des jours  $(i, i + 1, \dots, i + 9)$  est semblable à celle des cours des jours  $(j, j + 1, \dots, j + 9)$ .

1) La longueur de la ligne est un indicateur de la durabilité du phénomène de répétition : plus elle est longue, plus l'analogie entre les séquences temporelles est stable et la comparaison pertinente.

2) La fréquence des lignes à la verticale de l'abscisse  $i$  est également évocatrice : si le diagramme de proximité révèle la présence de plusieurs lignes initialisées en  $i$ , par exemple aux points  $(i, j)$ ,  $(i, k)$ ,  $(i, l)$  et  $(i, m)$  alors cela signifiera que la même séquence est répétée plusieurs fois – 5 fois dans notre cas – à partir des jours  $i, j, k, l$  et  $m$ .

Si les lignes sont toutes de même longueur alors l'intégralité de la séquence est reproduite. Si elles sont de longueurs diverses il est alors probable que la séquence répétée a été initialisée et qu'un événement l'a interrompue.

3) Les positions relatives des diverses lignes, si elles suivent une règle précise, révèlent les périodes, les délais-clé, initiateurs de cycles<sup>1</sup>. Si les 4 lignes s'obtiennent les unes des autres par des translations de 50 jours  $(i, j)$ ,  $(i, k = j + 50)$ ,  $(i, l = k + 50)$  et  $(i, m = l + 50)$ , cela laisse supposer que les séquences homogènes sont distantes de 50 jours. Il s'écoule 50 jours avant que deux schémas d'évolution semblable ne se produisent.

- Un graphique présentant des comportements de cours boursiers homogènes permet d'attester de l'unicité de la structure sous-jacente. En effet, un changement de structure ferait apparaître des distorsions, des ruptures dans les comportements de cours.

Ainsi, l'allure du diagramme de proximité renseigne correctement sur l'existence, la persistance, la stabilité et l'unicité de séquences répétitives de cours boursiers. L'emploi de ces remarques sera étudié plus loin. Néanmoins, on mesure dès à présent la façon dont ce type d'enseignements peut conditionner les anticipations des agents, et par là leurs comportements d'achats et de ventes de titres.

f) Le diagramme de proximité peut permettre dans une certaine mesure d'identifier un processus d'aspect totalement désordonné mais dont la suite de variables qui le constituent sont interdépendantes, d'un mouvement où toutes les variables sont indépendantes. En effet, bien que les deux séries présentent un comportement totalement erratique et désordonné, dans le second cas, le graphique est uniformément gris et diffus ; dans le premier tous les points du cadran n'ont pas la même probabilité

---

1. Le terme de cycle s'entend ici en terme de répétition de séquences de cours erratiques ; il n'est pas question ici de cycle économique au sens où la théorie traditionnelle l'entend.

d'être visités. Il peut alors apparaître des lignes, de longueur plus ou moins importante, permettant de tirer des conclusions sur d'éventuelles répétitions de séquences de cours de bourse.

g) Pour déterminer les paramètres  $m$  et  $t_0$ , on itère cette procédure de construction de diagrammes en faisant varier les paramètres  $m$  et  $t_0$ . Comme il a été précisé plus haut, le théorème de Takens (1981) indique que la dimension optimale  $m$  est approximativement de l'ordre de deux fois la dimension de l'objet ou de la trajectoire que l'on étudie. Or, comme on ne connaît pas *a priori* la dimension de l'objet, les paramètres  $m$  et  $t_0$  sont choisis de façon empirique. Le couple optimal de paramètres est sélectionné selon le résultat informatif du diagramme de proximité obtenu.

Il existe néanmoins des bornes liées à la méthode qui contraignent les paramètres. Des valeurs trop basses ne permettent pas de faire des conclusions et de tirer des enseignements, les diagrammes présentant un aspect trop diffus.

Mais, pour apprécier la portée de cet outil, il serait très instructif de le mettre en perspective avec des outils traditionnels tels les coefficients d'autocorrélation, l'analyse spectrale ou l'analyse des données.

## 2. Apports de la méthode

La *Réurrence Plot* cherche à déceler la présence d'une structure, même ténue. Elle se caractériserait par des similitudes intermittentes de comportements des prix. Or d'autres outils étaient employés jusqu'à présent pour détecter des périodicités et des corrélations, notamment les coefficients d'autocorrélation, l'analyse spectrale et l'analyse des données.

### 2.1. L'autocorrélation en question

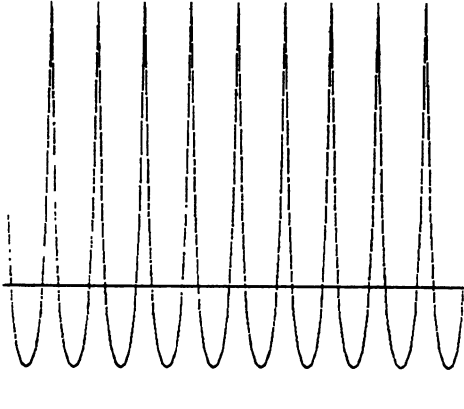
Une corrélation entre les termes d'une série témoigne du fait que les observations successives sont dépendantes les unes des autres. Or, les méthodes linéaires traditionnelles d'estimation s'accommodent mal des corrélations. Plusieurs indicateurs statistiques ont été conçus dont le calcul des coefficients d'autocorrélation.

Soit une série temporelle  $X_t$ . Les coefficients d'autocorrélation permettent entre autres choses de détecter des composantes saisonnières. La méthode de *réurrence Plot* est à même de déceler des séquences de points similaires sans que cette similitude soit systématique. Afin d'apprécier la portée du calcul des coefficient d'autocorrélation et celle de la *réurrence Plot*, raisonnons à partir d'un cas d'école.

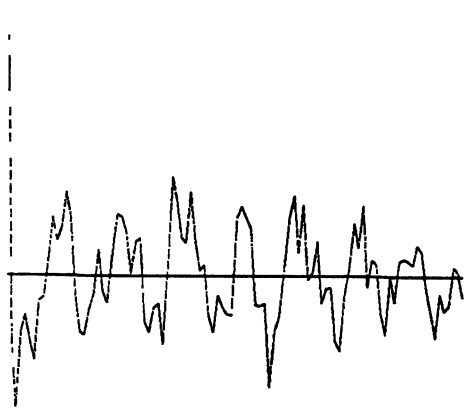
On a construit deux séries. La première (A) est cyclique de période 10. La seconde (B) répète une même séquence de longueur variable : à des intervalles de temps irréguliers on reprend la séquence à son début. En d'autres termes, on convient initialement d'une séquence – récurrente par exemple –. À un instant choisi de manière aléatoire, le déroulement de cette séquence est interrompu et repris à son début

## CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ

et ainsi de suite. La série présente donc des répétitions par morceaux, sans régularité. Sur chacune des séries (A) et (B) on calcule les coefficients de corrélation et le diagramme de proximité.

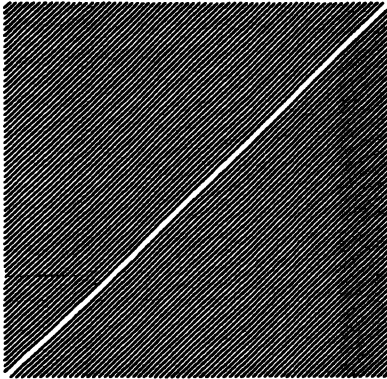


Coefficients autocorrélation A

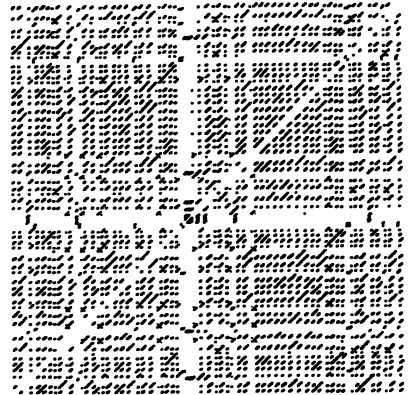


Coefficients autocorrélation B

Les coefficients font apparaître très nettement la périodicité de la série (A) par les forts coefficients obtenus toutes les 10 pas. Dans le cas de la série (B), le résultat est plus mitigé. Il ne ressort pas de période en particulier. Les coefficients se situent pour la plupart à l'intérieur des bornes signifiant par là leur *nullité* au seuil significatif de 5 %. Les graphiques de récurrence Plot nous apportent des résultats plus fins.



Graphique Recurrence Plot A



Graphique Recurrence Plot B

Dans le premier graphique, la périodicité de 10 de la série (A) apparaît très précisément. Le graphique présente de longues lignes parallèles distantes de 10 unités.

L'interprétation en est simple : les points les plus semblables à un point  $i$  quelconque sont distants de  $i$  de 10 en 10. *L'écart entre les lignes indique la période ; la longueur des lignes précise la stabilité, la durabilité, la pérennité de cette périodicité.* Il n'est donc pas surprenant de trouver des lignes aussi longues que le graphique lui-même attestant d'une périodicité constante, conformément à la construction de cette série.

S'agissant de la série (B), le diagramme des proximités révèle des lignes de longueurs variables : la série reproduit des séquences similaires mais ces séquences sont de longueurs diverses. En outre, elles surgissent à des moments irréguliers : les écarts entre les lignes sont variables alors que dans le cas de la série (A) ils étaient de 10 jours. Ainsi, l'aspect irrégulier d'apparition de séquences répétées de longueur variable est mis en évidence dans le diagramme. Les coefficients d'autocorrélation n'y étaient pas arrivés.

À travers ces cas simples on a pu apprécier la souplesse d'analyse des diagrammes de proximité qui leur confère un domaine d'application plus étendu. Par ailleurs, on s'intéressera à sa comparaison avec l'analyse spectrale.

### 2.2. L'analyse spectrale

Originellement, l'analyse spectrale s'est développée dans les domaines confrontés aux phénomènes vibratoires, au sens physique du terme. Les études ont généré une méthode de décomposition des séries chronologiques. L'idée est en effet, de *mettre en évidence l'ensemble des fréquences présentes dans la chronique et d'en estimer leur amplitude.*

La puissance spectrale peut être définie comme la *contribution de chaque fréquence à la totalité du mouvement sur la période.* L'existence de grands pics dans le spectre est signe de périodicité ou de quasi-périodicité. En présence d'un mouvement *apériodique ou quasi-périodique à nombre de fréquence très élevé,* le spectre ne décèlera aucune périodicité. Ainsi, bien que l'analyse spectrale nous indique, une quelconque cyclicité ou au contraire un mouvement chaotique, selon que le spectre présente ou non des pics, elle ne laisse, en revanche, rien deviner du caractère déterministe ou purement aléatoire de l'évolution des points. Cette distinction nécessite une analyse complémentaire sur le nombre de degrés de liberté du système.

En d'autres termes, l'analyse spectrale ne permettra pas de détecter la structure sous-jacente d'un modèle chaotique non plus que son caractère déterministe. Le spectre de la série sera constitué d'une large bande continue qui caractérise les séries aléatoires – ou certaines séries quasi-périodique à nombre de fréquences très élevées –.

Dans la section qui suit on fera le lien avec l'analyse des données qui présente dans ses fondements certains éléments de comparaison avec les diagrammes de proximité. L'analogie s'arrêtera là : une étude plus approfondie montrera leurs multiples divergences.

### 2.3. L'analyse des données

L'analyse des données gère un grand nombre de données et de variables sans formuler d'hypothèses sur les phénomènes étudiés. Elle a pour objet de décrire les données contenues dans un tableau de  $n$  individus- $p$  caractères. Une représentation graphique de  $n$  points dans un espace à  $p$  dimensions est impossible dès que  $p$  excède 2 ou 3.

Dès lors, on peut difficilement se faire une idée des ressemblances entre les caractères ou les individus à moins de ne les considérer que deux à deux.

Le principe de l'analyse des données va consister à chercher dans  $\mathbb{R}^p$  où sont représentés les  $n$  individus, un sous-espace de dimension 2 ou 3 dans lequel projeter orthogonalement – selon la métrique euclidienne par exemple – de manière à représenter le plus fidèlement possible le nuage de  $n$  points. L'analyse en composantes principales réalise ce résultat en fournissant en outre la suite des sous-espaces de dimension 1, 2, 3, ...,  $p - 1$  dans lesquels le nuage projeté est le plus proche du nuage initial. *On montre qu'il s'agit d'une famille de sous-espaces engendrés par les vecteurs propres de la matrice  $V$  des covariances.* De manière plus précise, le sous-espace de dimension  $k$  qui donne le nuage projection le plus proche du nuage initial est engendré par les vecteurs propres correspondant aux  $k$  plus grandes valeurs propres de la matrice  $V$ .

La proximité entre les deux nuages sera repérée par le rapport entre l'inertie de l'un et l'autre nuage. On appelle inertie totale du nuage de points, la moyenne des carrés des distances des  $n$  points au centre de gravité. Cette quantité caractéristique du nuage mesure d'une certaine manière, l'éloignement des points par rapport à leur centre de gravité, c'est-à-dire la dispersion globale du nuage. Une inertie presque nulle traduira la proximité des points avec leur centre de gravité. Ainsi, le rapport des inerties sera d'autant plus proche de l'unité que la projection aura conservé les distances. Le sous-espace reproduira alors de manière presque exhaustive le nuage initial.

Le nuage projeté permet ensuite de faire des rapprochements entre les individus, ou entre les caractères.

- a) Ainsi, les *points communs* entre les deux principes d'analyse sont bien balisés :
- la problématique est de décrire un grand nombre de données ;
  - aucune des deux analyses ne formule d'hypothèses *a priori* sur les phénomènes ;
  - dans l'un et l'autre cas il est question de projections optimales de données d'un espace de grande dimension vers un espace de dimension moindre.

b) Si les deux méthodes semblent se réduire à la recherche d'un espace dans lequel projeter un nuage de points de manière à le déformer le moins possible, les moyens d'y parvenir seront très divergents car *les données initiales sont fondamentalement différentes et renversent la problématique.*

- En effet, au-delà d'une différence de méthodologie, il se trouve une *opposition de fond dans la formulation du problème* : alors qu'en analyse des données, les dimensions des vecteurs initiaux –  $n$ , nombre d'individus et  $p$  leurs caractéristiques – sont connues et qu'on cherche à les réduire en minimisant la perte d'information, le vecteur dont on dispose dans la méthode des proximités est unidimensionnel – série temporelle de cours boursiers –. La problématique consiste à le transformer en plusieurs vecteurs de dimension  $m$  à définir de manière à *inférer la dimension du phénomène dont la série temporelle est issue*, puis à effectuer des recherches de similarités et des effets de cycles.
- *La méthode des proximités étudie des évolutions dynamiques alors que l'analyse des données s'attache en premier lieu aux phénomènes statiques*. Les  $n$  vecteurs d'individus aux  $p$  caractéristiques n'ont pas d'éléments de comparaison avec les  $N'$  vecteurs reconstruits par les proximités. La suite des vecteurs reconstruits par la méthode des proximités a un sens que n'ont pas les vecteurs de l'analyse des données : ils suivent une logique chronologique qui a une certaine valeur puisqu'elle décrit les mouvements de la dynamique sur l'attracteur. Par ailleurs, c'est grâce à elle que l'on identifie les cycles et que l'on en repère les début et fin. Les liaisons et les règles que recherche la méthode des proximités sont avant tout temporelles. L'ordonnement des vecteurs a donc son importance. Ce n'est pas le cas des individus de l'analyse des données qui peuvent être réordonnés dès lors qu'ils le sont avec leurs propres caractères.
- De même que les  $n$  individus n'avaient pas le même rôle que les  $N'$  vecteurs reconstruits, de même, les caractères des individus se distinguent des composantes des vecteurs reconstruits. Les premiers décrivent plusieurs éléments du caractère de l'individu, les seconds représentent plusieurs facettes d'une même notion : la valeur du titre à des instants différents. Ce sont toutes des réalisations de cours boursiers et décrivent le même phénomène. La méthode des proximités ne cherche pas les corrélations éventuelles entre ces différentes coordonnées des vecteurs.

Ainsi, les critères diffèrent nécessairement du fait que les buts recherchés sont fondamentalement dissemblables. Appliquer les méthodes d'analyse des données plutôt que celle des proximités aux séries temporelles supposerait la connaissance préalable du fait générateur de la série.

### 3. Applications et résultats

La transformation des données, dans le cadre d'études non-linéaires, fait l'objet d'un traitement spécifique.

#### 3.1. Transformation des données

Dans les analyses de cours boursiers, on étudie généralement les rendements plutôt que les prix eux-mêmes. En effet, ces derniers présentent des propriétés d'autocorrélation : chaque prix dépend de celui de la veille, violant par là l'hypothèse de loi

normale nécessaire à l'analyse jusqu'à présent. Les rendements eux transformant les données et les rendent plus appropriées aux analyses linéaires et aux hypothèses d'indépendance qui en découlent.

En revanche, dans le cadre d'études non-linéaires, une transformation en taux de rendement n'est plus adaptée : elle peut éliminer les éléments de structure présents dans les données. Or c'est précisément ce que l'on cherche à mettre en évidence. L'intérêt des transformations est de rendre *stationnaire* l'évolution tout en conservant sa structure. Les scientifiques étudiant sur ces domaines travaillent sur les données brutes et non sur des taux. En finance ou en économie, les évolutions avec tendances ont conduit à employer des rendements. Les dynamiques non-linéaires doivent nous mener vers de nouveaux concepts, de nouvelles transformations.

Les premières études des marchés financiers sous l'angle de dynamiques non-linéaires se fondaient sur les rendements – Scheinkman and Lebaron (1986 publié en 1989a) –. Outre les problèmes relatifs à un nombre insuffisant de données, ces auteurs ont rencontré des problèmes avec l'emploi de rendements, difficultés qui se simplifieraient en raisonnant sur les prix.

La croissance des prix<sup>1</sup> peut être sans limite. Il importe donc de la neutraliser : on s'intéresse aux mouvements des prix, pas à la croissance inflationniste. Chen (1988) a éliminé le trend des agrégats monétaires en leur soustrayant un taux interne de croissance selon l'expression suivante :

$$\varepsilon_t = \text{Ln}(P_t) - at + b.$$

$\varepsilon_t$  désigne la série *détrendée* ;

$P_t$  désigne la série originale des prix et  $t$  le temps.

D'autres modes de transformations peuvent être envisagés. Ce qui importe n'est pas tant rendre les successions de point indépendantes les unes des autres que de disposer d'une série de prix apurée de facteurs exogènes de croissance, c'est-à-dire telle qu'elle ne diverge pas.

### 3.2. L'indice CAC40

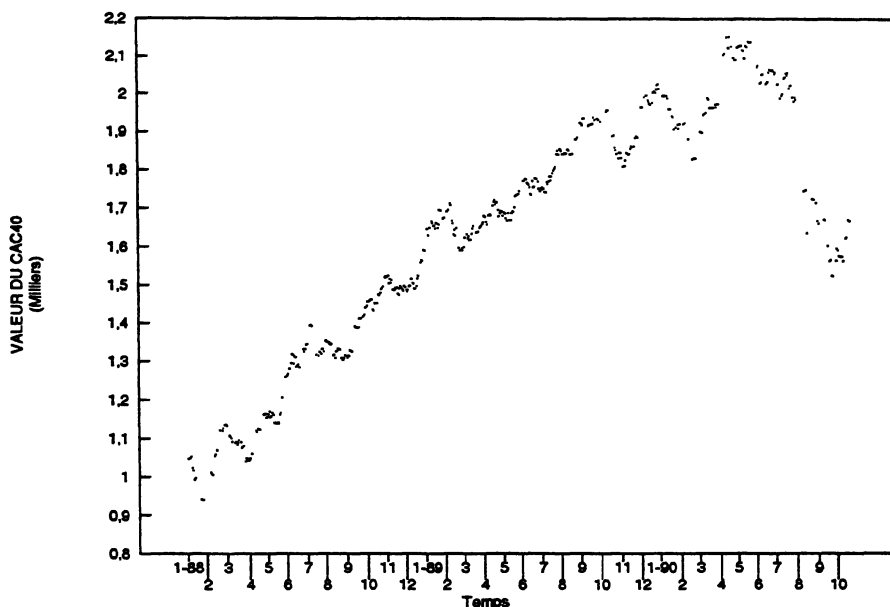
La série analysée correspond au cours de clôture journalier du CAC40, indice du marché boursier français. L'indice CAC40 est calculé sur un échantillon de 40 actions françaises. Il exprime à tout moment le rapport de la capitalisation boursière – produit du cours de l'action par le nombre de titres – instantanée des 40 valeurs avec leur capitalisation boursière de référence – base 1000 au 31 décembre 1987 –. Le cas échéant, cette capitalisation est ajustée lors des opérations sur titres. Cette série, initialisée début janvier 1988, compte, jusqu'à fin octobre 1990, 700 points (voir figure 2.1.a). L'analyse d'un indice plutôt que d'une évolution individuelle de titres a été motivée par l'aspect en principe moins erratique du premier qui permet ainsi, quand elles existent, de mieux appréhender certaines relations structurelles.

---

1. Qu'elle résulte de l'inflation ou de la croissance de l'économie.

## CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ

Figure 2.1.a  
Évolution du CAC40 du 01-88 au 10-90



La croissance des deux dernières années semble caractérisée par une concavité marquée laissant apparaître un épuisement de son potentiel.

La crise du Golfe a un statut ambigu : en effet, a-t-elle provoqué la récession brutale de l'indice ou a-t-elle uniquement servi d'amplificateur de la récession qui semblait poindre – rappelons que tous les signes d'une grave récession imminente semblent rassemblés dès le début 1991 : dérégulation du système financier et bancaire américain passant par un accroissement des chiffres du chômage et du nombre de faillites bancaires, graves difficultés économiques et sociales sous-jacentes dans l'Allemagne réunifiée, le passage plus long et difficile des pays de l'Est à une économie de marché... – ?

Exprimée en d'autres termes, la question posée est de savoir dans un premier temps si l'évolution du cours possède une structure déterminée ou si les variables sont indépendantes. Dans un second temps, il faudra s'interroger sur le point de savoir si la crise du Golfe a été un phénomène destructurant ou, à l'opposé, une perturbation momentanée que le marché français peut absorber plus ou moins rapidement : dans le premier cas, la série ne peut être analysée sans incorporer une rupture de structure qui s'expliquerait par un dérèglement des modes anticipatifs utilisés par les intervenants des marchés financiers. Dans le second cas, une seule et unique structure permet d'expliquer l'évolution temporelle de la série.

On a réalisé sur cette chronique, une régression par les moindres carrés ordinaires. On opte pour un modèle de régression simple qui cherche à expliquer la tendance en



fonction du temps uniquement. Par ailleurs, le passage au logarithme du CAC40 réduit l'influence des valeurs aberrantes et *extra-ordinaires*. Posons  $y_t = \text{Ln}(\text{CAC}(t))$ .

La régression d'une série admettant un trend quadratique en  $t$  s'écrit :

$$y_t = at^2 + bt + c + \varepsilon_t \quad (2)$$

Dans le modèle, la non-linéarité permet d'éliminer le trend de façon appréciable même si l'analyse des résidus  $\varepsilon_t$  révèle une autocorrélation d'ordre 1.

### A. Résultats et analyse

Un balayage des différentes valeurs de  $m$  et  $t_0$ , paramètres de reconstruction de la trajectoire dans le pseudo-espace des phases a été réalisé. Plusieurs types d'analyses sont effectués sur la série reconstruite. Ils permettent de dégager les propriétés essentielles de la série  $\varepsilon_t$ .

a) De l'examen des différents diagrammes de récurrence il en ressort deux, assez significatifs (voir figures 2.2.a et 2.2.b) obtenus pour différentes valeurs des paramètres  $m$  et  $t_0$ .

Le premier laisse présager d'un certain ordre, d'une structure, de par les figures qui se dessinent ; les points semblent se regrouper et s'articuler principalement autour d'une région centrale qui correspond, dans le temps, à l'infléchissement constaté en 1989. En fin d'évolution, il apparaît un point d'accumulation dont l'origine serait le choc d'août 1990.

Ces caractéristiques, zone centrale d'articulation en 1989 et guerre du Golfe, se retrouvent dans le second graphique ( $m = 4$  et  $t_0 = 5$ ). Mais il est jugé *optimal* pour plusieurs raisons :

- la configuration apparaît pour une petite dimension du pseudo-espace des phases ;
- il présente des lignes obliques relativement longues et moins de points épars que pour d'autres valeurs des paramètres. Or, la présence, dans le diagramme de proximité, de lignes diagonales parallèles à la première bissectrice, correspond à des *mouvements d'événements qui se ressemblent à des moments différents*<sup>1</sup>. Cela permet de déceler le caractère dynamique de la série par la répétition de certains phénomènes dans le temps. De ce fait, la présence de lignes obliques dans la *distribution des diagrammes de proximité* laisse *présumer l'existence d'une structure dans la série des résidus*.

Dans ce qui suit,  $j = s_m(i, k)$  désignera l'indice de l'un *quelconque* des six points les plus proches du point d'indice  $i$  :  $k$  peut signifier indifféremment une valeur dans  $\{1, \dots, 6\}$ .

---

1. Notons que la corrélation qui apparaît en dimension 4 ne se résume pas à une autocorrélation d'ordre 1 puisque le pas temporel est de 5 jours : les résidus mis en correspondance sont distants de 5 jours. Ce premier résultat vient confirmer les corrélations à 5 jours constatées par les financiers (voir par exemple le fameux effet du lundi). De plus, les résidus ne sont pas mis en relation un à un, mais par blocs de  $m$  points.

# CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ

Figure 2.2.a  
 $m = 2 \quad t_0 = 1$

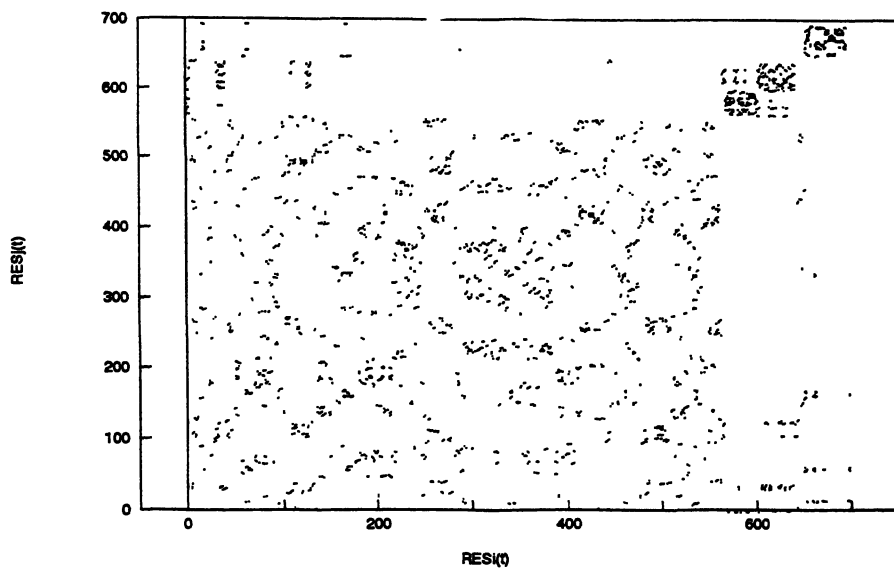
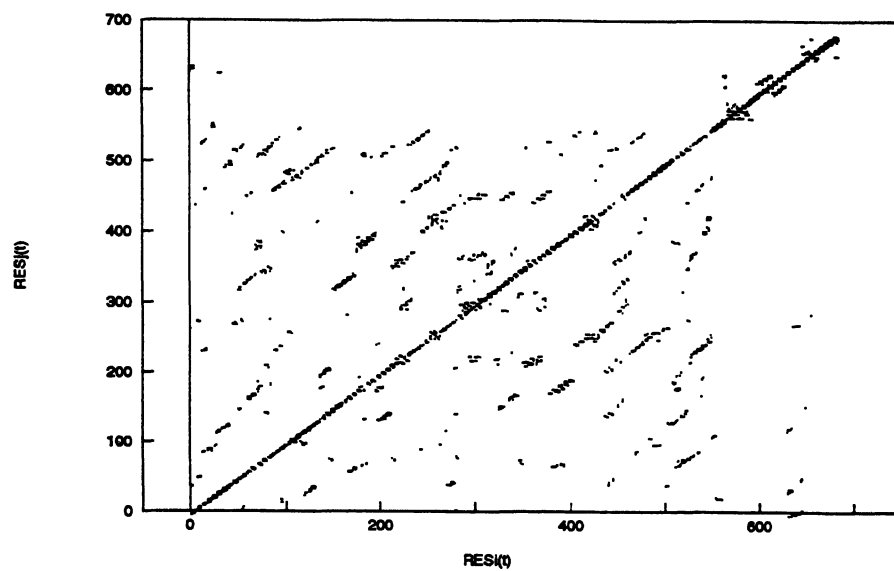


Figure 2.2.b  
 $m = 4 \quad t_0 = 5$



## CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ

• *Le long d'une diagonale A* : Si  $X_i^m$  et  $X_j^m$  sont proches et sont tels que  $j = s_m(i, k)$  appartient à une ligne diagonale, alors la probabilité pour que  $X_{i+1}^m$  et  $X_{j+1}^m$  soient proches, est beaucoup plus élevée que pour un couple quelconque de points  $X_i^m$  et  $X_j^m$  proches, mais dont l'indice  $l = s_m(i, k)$  ne se situe pas sur une diagonale.

De manière formelle :

$$\text{Prob} [j + 1 = s_m(i + 1, k) \in A | j = s_m(i, k) \in A] >$$

$$\text{Prob} [l + 1 = s_m(i + 1, k) \in A]$$

En d'autres termes, s'il existe un indice  $j = s_m(i, k)$  situé sur une diagonale, alors la présomption de proximité entre les couples consécutifs  $X_{i+1}^m$  et  $X_{j+1}^m$  s'accroît fortement. La validité de cette règle s'étend sur un horizon  $h$ , dont la longueur dépend de celle de la ligne  $A$ .

• *Le long d'une coupe verticale* – coupe instantanée de l'ensemble  $\Omega$  : cette optique est en réalité le corollaire de l'interprétation précédente –. On relève les six plus proches voisins de  $X_i^m$ . On peut toujours noter, pour deux des points situés le long de la coupe,  $j = s_m(i, k) = (i + a)$  et  $l = s_m(i, k') = (i - b)$ , par exemple. En d'autres termes, si le vecteur  $X_i^m$  prend certaines valeurs, les points situés  $a$  jours après ou  $b$  jours avant auront des valeurs *similaires*, la similarité étant entendue au sens de la distance définie plus haut. Cette correspondance peut se poursuivre dans le temps si chacun des deux indices appartient à une diagonale et si, en outre, les diagonales subsistent un certain temps. Le point  $X_{s_m(i, k)}^m = X_{i-b}^m$  sera alors un indicateur d'avance pour  $X_i^m$ . De même,  $X_i^m$  et  $X_{i-b}^m$  pour  $X_{i+a}^m$ .

On visualise ainsi des liaisons temporelles sur des périodes temporaires et non nécessairement régulières. Compte tenu des liens qu'elle permet d'appréhender, cette approche est assez large.

• *En termes d'anticipations et d'extrapolations* : Si la longueur de la ligne parallèle à la première bissectrice est supérieure au pas temporel  $t_0$ , de reconstruction de la série, la structure en dimension  $m$  est susceptible de se conserver en dimension  $m + 1$ , puisque le passage d'un espace à l'autre se réalise précisément par l'allongement des vecteurs  $X_i^m$  du point  $x(i + mt_0)$  situé  $t_0$  instants plus tard. Ainsi, d'une certaine façon on est à même d'*extrapoler* la structure des corrélations en dimension  $m + 1$ , compte tenu de celle en dimension  $m$  et par là, d'anticiper l'évolution du système.

En outre, la longueur des lignes – supérieure à 5 points – laisse à penser que les liens unissant les points les plus proches de la chronique en dimension 4, ont tendance à persister pendant la période qui suit : *au regard du diagramme en dimension 5, qui n'est pas reproduit, la structure en dimension 4 présente la caractéristique d'être stable*. Les deux diagrammes montrent en effet une forte ressemblance.

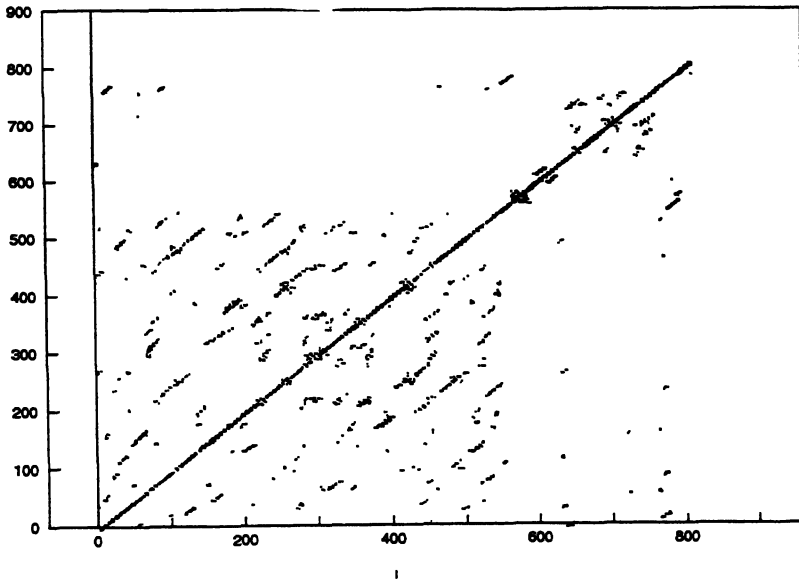
• *La position géographique* des lignes nous informe des dates et des périodicités des répétitions. Sur la figure 2.2.b les écarts oscillent autour de 60 jours selon les

## CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ

périodes. Il semblerait donc qu'une *périodicité* de 60 jours apparaisse dans l'évolution du CAC40. Si l'on considère un mois boursier de 20 jours, le phénomène de *cycle* s'étend sur trois mois soit un trimestre.

b) Les événements du Golfe laissent apparaître un point d'accumulation sur le graphique des proximités. Peut-on l'attribuer à une rupture de structure ou à une perturbation passagère du système ? Cela ne paraît pas correspondre à une rupture de structure similaire au graphe de récurrence des deux lois uniformes. La brutalité de la transition évoque davantage un choc exogène perturbant un temps le système dynamique. Après l'avoir intégré, ce dernier retrouverait ensuite son comportement habituel. L'évolution actualisée du CAC40 dont on dispose aujourd'hui et le diagramme de récurrence associé (voir figure 2.2.c) semblent confirmer ce verdict.

Figure 2.2.c  
 $t_0 = 5$   $m = 4$



Il apparaît en effet que la configuration qui semble se profiler est similaire à celle qui était apparue en 1989 : dans un premier temps, il se développe une structure autour d'un point d'accumulation central. Les lignes apparaissent ensuite au fur et à mesure de l'allongement de la série. Les conséquences du choc ne sont pas analysées avec suffisamment de recul pour que les résultats soient réellement probants, mais l'homogénéité des réactions face à ces deux événements peut constituer dans une certaine mesure un argument en faveur de l'hypothèse d'unicité de structure.

Trois autres aspects du phénomène permettent d'entériner ces conclusions comme nous le verrons par la suite (cf. infra).

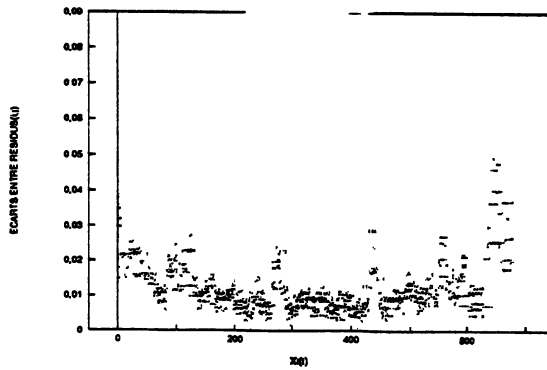
### B. Test de validité des analyses

Les précédentes analyses des diagrammes de proximité peuvent être complétées par la mise en évidence de plusieurs phénomènes qui permettent de valider ou d'invalider les conclusions élaborées.

#### a) Distribution des distances et écarts d'indices

• La distribution des six *distances minimales* associée au diagramme de proximité en dimension 4 (figure 2.3.a) indique les distances séparant les points voisins. Elle nous fournit une *évaluation du rayon de la boule* dans laquelle sont situés les points voisins. La valeur des rayons fournit une approximation de la densité des points de la chronique dans l'espace.

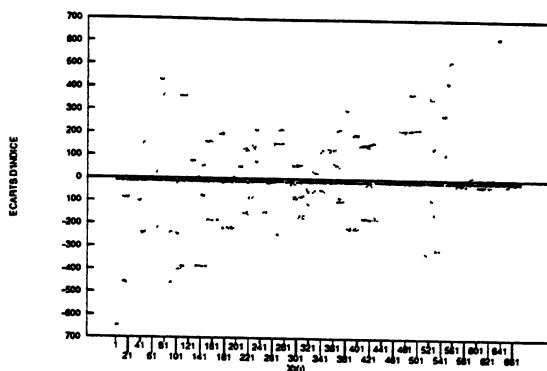
Figure 2.3.a



La grande similitude constatée dans la distribution des distances en dimension 4 et en dimension 5 plaide en faveur d'une stabilité de la structure dans le temps.

• Variations d'indices de proximité : de même, la *comparaison entre la dimension  $m$  et  $m + 1$ , des six indices et des points les plus proches d'un point de même indice  $X_i^m$  et  $X_i^{m+1}$*  (voir figure 2.3.b) peut être instructive. L'*analogie entre les voisins* – que l'on étudie par l'écart entre les indices des voisins de même rang  $e_{m+1,m}(i, k) = s_{m+1}(i, k) - s_m(i, k)$  – donne une *évaluation de la sensibilité des points* du diagramme des proximités à la dimension de reconstruction. La mesure  $e_{m+1,m}(i, k)$  de l'écart peut prendre des valeurs comprises entre  $[-N^m ; N^m]$ . Ainsi, plus l'écart entre les indices est important, plus les indices des points les plus proches diffèrent d'une dimension à l'autre. Or l'accroissement de la dimension de reconstruction revient à considérer l'évolution de la chronique  $t_0$  instants plus tard. De ce fait, si l'écart entre les structures de proximité sont faibles, la structure de la série présentera une relative stabilité dans le temps.

Figure 2.3.b



On note, en dépit d'un certain mouvement et d'une légère réorganisation des structures de proximité, que les écarts d'indice

$$e_{m+1,m}(i, k) = s_{m+1}(i, k) - s_m(i, k),$$

sont fortement regroupés autour de zéro. Ce phénomène corrobore les résultats de stabilité établis précédemment.

#### b) Brouillage par perturbation chronologique

Il peut être judicieux de tester en outre l'existence de la structure et sa robustesse en ayant recours à une pratique employée en théorie du chaos. Elle consiste à *réorganiser aléatoirement la chronologie des résidus et à lui faire subir les mêmes traitements qu'à la série ordonnée*. Cette méthodologie s'inspire de celle utilisée pour valider le calcul de la dimension de corrélation d'une série chaotique. Quand les indices temporels sont brouillés de façon arbitraire, si le nouveau calcul de la dimension conduit à une valeur qui croît de façon significative, alors il y a une forte présomption pour que le système générateur soit plutôt déterministe. En revanche, si la dimension reste sensiblement la même que la série soit brouillée ou non, le système sera plutôt stochastique – LeBaron et Scheinkman (1986) –.

Dans le cadre de la série du CAC40, en transposant le raisonnement aux diagrammes de récurrence, *si la structure qui semblait transparaître dans la série ordonnée disparaît avec la série brouillée, on peut présumer que le CAC40 est régi par une dynamique temporelle déterministe. En effet, cela sera un témoignage de l'importance de l'ordre temporel des points. En revanche, si, en dépit de la désorganisation temporelle aléatoire des points de la chronique des résidus, la structure persiste, on pourra présumer du caractère stochastique du processus générateur.*

## CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ

Le diagramme des proximités des  $\varepsilon_t$  brouillés offre la même apparence que celui fourni par un processus stochastique du type loi normale ou loi uniforme. Pour des dimensions raisonnables le diagramme est gris et homogène.

Le signal émis par le diagramme de proximité de la série brouillée diffère donc sensiblement de celui émanant de la série ordonnée. L'ordre chronologique importe donc. Il en résulte que la structure établie précédemment a un sens, une signification.

L'ensemble de ces considérations plaide en faveur de l'existence et de la stabilité d'une structure sous-jacente à l'évolution des résidus issus de la régression quadratique. L'étude des extrapolations ne fera que corroborer ces assertions.

### C. Extensions et développements : le modèle autorégressif d'ordre 5

De l'analyse précédente, il ressort des phénomènes remarquables : la série  $\varepsilon_t$  présente des séquences d'événements qui se répètent à cinq jours. L'idée serait donc d'exploiter les liens de la série  $\varepsilon_t$  au profit d'une transformation de la série du CAC40 de type autoregressif.

Il s'agirait en d'autres termes de s'intéresser aux résidus  $\mu_t$  issus d'un modèle autorégressif d'ordre cinq :

$$y_{t+5} = ay_t + b + \mu_t \quad (3)$$

L'intérêt d'une telle manipulation serait de retrouver avec la série  $\mu_t$  des différences d'ordre cinq, les résultats de la série  $\varepsilon_t$ . Si les variations quotidiennes ne sont que bruit, les variations hebdomadaires – en effet, cinq jours correspondent à une semaine boursière – contiennent probablement des éléments de structure, compte tenu des répétitions à cinq jours qui sont apparues dans la série des  $\varepsilon_t$ .

Les diagrammes de proximité de la série  $\mu_t$  ont donc été construits pour différentes valeurs des paramètres  $t_0$  et  $m$ .

Le couple optimal de paramètres est  $m = 4$  et  $t_0 = 5$ , conformément à ceux retenus pour la série  $\varepsilon_t$ . Pour ces valeurs, le diagramme de proximité offre des lignes transversales très proches de celles mises en évidence dans la série  $\varepsilon_t$ . Ainsi, bien que les variations à cinq jours soient très erratiques, les diagrammes de proximité offrent des caractéristiques de même nature que ceux de la série  $\varepsilon_t$  : présence de lignes, plus nombreuses dans la série des  $\mu_t$ .

- La dimension optimale est très évocatrice : les vecteurs sont composés de 4 coordonnées espacées de 5 jours – par exemple lundi 1<sup>er</sup>, lundi 8, lundi 15 et lundi 22 –. Il règne donc une homogénéité dans les données qui est de l'ordre du mois et qui pourrait tout à fait correspondre au rythme mensuel qui scande le marché à règlement mensuel. Les données semblent donc se répandre, de mois en mois.

- Par ailleurs, sont présentes un nombre important de lignes transversales qui attestent d'une similarité, pendant plusieurs jours consécutifs, de comportement des cours sur un mois : par exemple, la séquence des vecteurs de cours (lundi 1<sup>er</sup>, lundi 8, lundi 15 et lundi 22), (mardi 2, mardi 9, mardi 16 et mardi 23), (mercredi 3, 10, 17 et 24), (jeudi 4, 11, 18 et 25) du même mois peut être très semblable à la séquence

## CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ

des vecteurs (lundi 12, 19, 26 et 3), (mardi 13, 20, 27 et 4), (mercredi 14, 21, 28 et 5), (jeudi 15, 22, 29 et 6) d'un autre mois. La longueur des séquences dépend de celle des lignes. En l'occurrence, celles-ci semblent vouloir s'étendre sur plusieurs mois – entre 1 mois et 5 mois –. Pour interpréter ceci, il faudrait pouvoir le mettre en relation avec des événements boursiers ou des effets d'annonces présentant une régularité de l'ordre de ces durées.

- L'écart entre ces symétries dépend des positions relatives des lignes entre elles. Une mesure fournit un écart d'ordre mensuel : les vecteurs de cours des débuts de mois par exemple, correspondent aux vecteurs de cours obtenus au même moment du mois suivant.

*Au regard du diagramme de proximité des résidus  $\varepsilon_t$  et  $\mu_t$ , tout porte à croire en l'existence d'une structure sous-jacente au CAC40.*

Ainsi, cette dernière étude concrétise la synthèse et la coordination de l'ensemble des considérations établies avec les différentes séries.

## CONCLUSION

En définitive, les diagrammes de proximité établissent des parallèles entre les faits. Ils nous permettent d'accéder à des informations qui confirment les résultats empiriques et que l'on peut approfondir. En effet, l'analyse est *didactique* – certains résultats indiquent les voies à explorer et les démarches à réaliser –. De ce fait, cette procédure permet d'aboutir à des diagnostics présentant une certaine robustesse : les résultats obtenus d'une certaine manière peuvent être précisés et confirmés d'un autre point de vue.

Concernant la série du CAC40, l'aspect *évolutif et interactif* de la procédure est bien mis en évidence.

- Les résidus  $\varepsilon_t$  du modèle (2) témoignent de l'existence d'une structure relativement stable : des séquences d'événements distants de cinq jours se répètent dans le temps. Les résultats viennent confirmer ceux établis empiriquement. Concernant les événements du Golfe, bien que le recul ne soit pas suffisant, il est probable qu'ils n'aient pas bouleversé les liens temporels existants : la crise se présenterait tout au plus comme un choc aléatoire exogène absorbé ensuite plus ou moins rapidement par l'économie. La nature de la structure autoriserait, en outre, dans une certaine mesure, des extrapolations temporelles. Mais il convient de les mettre en perspective avec celles obtenues par d'autres voies. Les conclusions établies pour ce premier modèle sont ainsi très intéressantes.

En mettant à profit les conclusions tirées de l'étude des résidus  $\varepsilon_t$ , – indiquant la présence de liens temporels à cinq jours –, on s'est intéressé à la série des résidus issus d'un autorégressif d'ordre 5. Là, en dépit d'une évolution très erratique des résidus, une structure à cinq jours est apparue à nouveau. L'accroissement de l'ordre



## CAC 40 ET CHAOS : LES DIAGRAMMES DE PROXIMITÉ

des différences a permis que des effets de structure soient présents dans les écarts  $\mu$ , et notamment des effets *mois*.

Le paramètre *mois* apparaît en effet à trois reprises :

- dans la dimension de l'espace des phases (qui vaut 4 avec un pas temporel de 5 ce qui couvre le mois) ;
- dans la longueur des lignes qui est souvent un multiple du mois (1,2, voire 5 mois) ;
- dans les écarts qui séparent les lignes (de l'ordre du mois).

Les cours boursiers se correspondent ainsi au rythme des deux catégories d'effets : des effets semaine et des effets mois. Ces derniers représentent une part très importante de la structure qui semble agir et dynamiser le CAC40 puisqu'aussi bien des effets d'un mois apparaissent que des effets bimensuels. Ces effets, outre leur fréquence élevée, sont d'autant plus importants qu'ils durent à leur tour pendant plusieurs mois – effets trimestriels et quelquefois semestriels –. Ainsi, il existerait des *cycles* mensuels qui s'étendraient pendant plusieurs mois qu'il conviendrait de mettre en corrélations avec les événements politiques et économiques présentant la même périodicité – annonce du déficit américain, d'une baisse des taux ou encore effets de reports et de déports... –.

Finalement cette méthodologie semble très prometteuse à plusieurs titres. Elle autorise l'étude d'un grand nombre de séries. La démarche par étapes qui la caractérise en permet l'étude et le développement. De nombreux compléments peuvent être envisagés et notamment une extension aux extrapolations et aux prévisions.

Cette étude consacre le rapprochement entre l'analyse technique et l'analyse numérique. Elle participe de l'analyse technique en s'intéressant aux mouvements de marché et non aux raisons qui les conditionnent. De plus, elle ne tend pas à faire entrer la chronique dans un cadre mathématique et analytique rigide : aucune hypothèse *a priori* sur l'existence d'une liaison entre les points, sur la nature de cette corrélation et sur son ordre n'est établie. La place accordée aux graphiques est également une caractéristique de l'analyse technique.

Dans un même temps, la nécessité des calculs rapproche la méthode de Récurrence Plot de l'analyse numérique. Finalement, tous ces aspects confèrent à la méthode la *propriété d'universalité* car elle s'applique indifféremment à toutes les séries.

Il faut néanmoins préciser que cet outil relève d'appréciations graphiques qui, comme le chartisme laissent une place non négligeable à la subjectivité. Il n'existe pas encore d'instrument ni de mesure statistique qui permettraient d'évaluer et d'analyser l'information contenue dans les diagrammes. Plus encore, il serait impératif de tester la robustesse et la pertinence des paramètres retenus et des conclusions apportées.

## BIBLIOGRAPHIE

- ANDERSON P.W. & KENNETH J.A. & PINES D. (Eds) (1988) *The Economy as an Evolving Complex System*, Santa Fe Institute.
- BARNETT W.A. & CHEN P. (1988) "The Aggregation-Theoretic Monetary Aggregates are Chaotic and Have Strange Attractors: an Econometric Application of Mathematical Chaos", in BARNETT W.A., BERNDT E.R. & WHITE H. (Eds), *Dynamic Econometric Modeling*, pp. 199-245.
- BAUMOL W. & BENHABIB J. (1989) "Chaos: Significance, Mechanism and Economics Applications", *J. Economics Perspectives*, Winter, pp. 77-106.
- BENHABIB J. & DAY R.H. (1981) "Rational Choice and Erratic Behaviour", *R. of Economics Studies*, July.
- BERGÉ P. & POMBAU Y. & VIDAL Ch. (1984) *L'ordre dans le chaos*, Hermann.
- BERGÉ P. éd. (1989) *Le chaos*, Eyrolles.
- BOLDRIN M. (1989) "Path of Optimal Accumulation in Two-sector Models", in BARNETT W.A., GEWEKE J. & SHELL K. (Eds), *Economic Complexity : Chaos, Sunspot, Bubbles and Nonlinearity*, Cambridge University Press.
- BOLDRIN M. (1991) "Chaotic Dynamics and General Equilibrium Theory", *Actes du Symposium Sanjo*, Conference Hall, University of Tokyo, April.
- BROCK W.A. (1988) "Nonlinearity and Complex Dynamics in Economics and Finance" in ANDERSON P.W., ARROW K.J., PINES D. (Eds), *The Economy as an Evolving Complex System*, Addison-Wesley.
- CHEN P. (1988) "Empirical and Theoretical Evidence of Economic Chaos", *System Dynamics Review*.
- DAY R.H. (1982) "Irregular Growth Cycles" *American Economic Review*, June, pp. 406-414.
- DAY R.H. (1983) "The Emergence of Chaos from Classical Economic Growth", *Quarterly Journal of Economics*, May, pp. 201-213.
- DAY R.H. & WALTER J.L. (1989) "Economic Growth in the Very Long Run: on the Multiple-phase Interaction of Population, Technology, and Social Infrastructure", in BARNETT W.A., GEWEKE J. & SHELL K. (Eds), *op. cit.*
- DAY R.H. & SHAFER W. (1985) "Keynesian Chaos", *J. of Macroeconomics*.
- DECHERT W. (1984) "Does Optimal Growth Preclude Chaos? A Theorem on Monotonicity", *Zeitschrift für Nationalökonomie*.
- DECHERT W. (1988) *A Characterization of Independance for Gaussian Process in terms of the Correlation Dimension*, SSRI Working Paper 8812, University of Wisconsin at Madison.
- DENECKER R. (1986) "Competitive Chaos", *J. of Economic Theory*, Oct., pp. 13-25.
- ECKMANN J.P., OLIFFSON KAMPHORST, RUELLE D. & SCHEINKMAN J. (1988) "Lyapunov Exponents for Stock Returns", in ANDERSON P.W., ARROW K.J., PINES D. (Eds), *op. cit.*
- ECKMANN J.P., OLIFFSON KAMPHORST D., RUELLE D. & CILIBERTO S. (1986) "Lyapunov Exponents from Times Series", *Physical Review*, pp. 4971-4979.
- ECKMANN J.P., OLIFFSON KAMPHORST S., & RUELLE D. (1987) "Recurrence Plot of Dynamical Systems", *Europhys. Lett.*, pp. 973-977.
- ECKMANN J.P. & RUELLE D. (1985) "Ergodic Theory of Chaos and Strange Attractors", *R. of Modern Physics*, pp. 617-656.

- ECKMANN J.P. & RUELLE D. (1990) *Fundamental Limitations for Estimating Dimensions and Lyapunov Exponents in Dynamical Systems*, preprint.
- GOODWIN R.M. (1990) *Chaotic Economic Dynamics*, Clarendon Press Oxford.
- GRANDMONT J.P. & MALGRANGE P. (1986) "Nonlinear Economic Dynamics: Introduction", *J. of Economic Theory*, oct., pp. 3-11.
- GRASSBERGER P. (1986) "Are There Really Climatic Attractor ?", *Nature* n° 323, pp. 609-612.
- HSIEH D. & LEBARON B. (1988) *Finite Sample Properties of the BDS Statistic*, Working Paper, University of Chicago and University of Wisconsin.
- HSIEH D. (1988) "Statistical Properties of Daily Exchange Rates", *J. of International Economics*.
- HSIEH D. (1989) "Testing for Nonlinearity in Daily Foreign Exchange Rate Changes", *J. of Business*.
- HSIEH D. (1991) "Chaos and Nonlinear Dynamics: Application to Financial Markets", *J. of Finance*, Déc.
- HSIEH D. (1992) *Implications of Nonlinear Dynamics for Financial Risk Management*, Workshop on Nonlinear Dynamics in Economics, European University Institute, July.
- KELSEY D. (1989) "An Introduction to Nonlinear Dynamics and its Application to Economics", in HAHN F., *The Economics of Missing Markets, Information and Games*, pp. 410-434.
- LORENZ H.W. (1989) *Nonlinear Dynamical Economics and Chaotic Motion*, Springer-Verlag.
- MIROWSKI P. (1990) "From Mandelbrot to Chaos in Economic Theory", *Southern Economic Journal*, n° 2, October.
- RUELLE D. (1989) *Chaotic Evolution and Strange Attractor*, Cambridge University Press.
- SCHEINKMAN J.A. & LEBARON B. (1989a) "Nonlinear Dynamics and GNP Data", in BARNETT W.A., GEWEKE J. & SHELL K. (Eds), *Economic Complexity: Chaos, Sunspot, Bubbles and Nonlinearity*, Cambridge University Press.
- SCHEINKMAN J. & LEBARON B. (1989b) "Nonlinear Dynamics and Stock Returns", *J. of Business*, (3).
- TAKENS F. (1981) in *Dynamical Systems and Turbulence*, éd. D. Rand and L.S. Young, *Lectures notes in Math.*, 898, Springer.
- WEDDEPOHL C. (1990) "Overlapping Generations Models, an Introduction", in Van der Ploeg (Eds), *Advances Lectures in Quantitative Economics*, pp. 249-312.