Zeitschrift: Archives des sciences [2004-ff.]

Herausgeber: Société de Physique et d'histoire Naturelle de Genève

Band: 61 (2008)

Heft: 1

Artikel: Mettre les sciences humaines en équations : une méthode simple et robuste

Autor: Dieterlen, Florent

DOI: https://doi.org/10.5169/seals-738303

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 01.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Mettre les sciences humaines en équations: une méthode simple et robuste

Florent DIETERLEN[1]

Ms. reçu le 8 février 2006, accepté le 20 mai 2008

Abstract

Bringing equations into humans sciences: a simple and robust method. - *A system of* n *nonlinear coupled differential equations is constructed out of a set of* n *time series. The method is based on a nonlinear fit of* n *time derivatives, each function of* n *variables, which can be enhanced by the Newton-Gauss method. Verification of the validity of the method is done with data constructed using the Lotka-Volterra and Lorenz systems. The method is also applied to psychology (addictions), giving clues for therapy, an explanation of the phenomenon and classification of various types of addictions. A set of two differential equations is constructed out of data for competition between two bacteria types, and obtains a good similarity with data. Additionally, this method is applied to macroeconomics (growth prediction) and finance (exchange rates). The method also applies to* n *times series and* n *series over* m *regions, for at least two different moments: two differential equations are constructed, using data for birth rate and alphabetisation, for four different years and 137 countries. Two solutions are obtained, the first giving a low birth rate and high alphabetisation, corresponding to post-industrialised countries, the second (unstable) giving a high birth rate and low alphabetisation, corresponding to peoples in a pre-industrialised status.*

Keywords: *time series, differential equations, non linear dynamical system, demography, chaos, psychology.*

Résumé

On construit un système de n *équations différentielles non linéaires couplées, à partir de données évoluant dans le temps. La méthode est basée sur un fit non linéaire de* n *dérivées par rapport au temps, en fonction de* n *variables, qui peut être amélioré par la méthode de Newton-Gauss. Nous vérifions la validité de la méthode avec des données issues de systèmes d'équations différentielles de Lotka-Volterra (oscillations) et de Lorenz (chaos et état stationnaire stable). Nous avons appliqué la méthode en psychologie, en testant des personnes dépendantes à l'alcool, ce qui permet de donner des indications thérapeutiques ciblées, d'expliquer mathématiquement le phénomène et de classifier les types de dépendances. Nous avons construit deux équations différentielles à partir des données de bactéries en compétition, et obtenu une bonne similarité avec les données. Nous avons appliqué la méthode en économie à la prédiction de la croissance, et en finance à la recherche d'instabilités. La méthode s'applique aussi à des ensembles de* n *séries chronologiques ou à* n *données sur* m *régions, à au moins deux dates différentes: nous avons construit deux équations différentielles à partir de données sur la natalité et l'alphabétisation, pour 4 dates différentes et 137 pays. Nous avons obtenu deux solutions principales, la première donnant une natalité basse et une grande alphabétisation, correspondant au cas des pays post-industrialisés, et une deuxième (instable) donnant une natalité élevée et une alphabétisation basse, pouvant correspondre à un état pré-industriel.*

Mots-clefs: *séries chronologiques, équations différentielles, systèmes dynamiques non-linéaires, démographie, chaos, psychologie.*

I. Introduction

Les mathématiques sont le langage des sciences physiques, elles le deviennent dans les sciences naturelles, mais sont encore peu utilisées en sciences humaines. La physique est la base des sciences naturelles en général, de par son domaine d'analyse (de «l'infiniment» petit à «l'infiniment» grand), mais aussi de par ses méthodes, expérimentales et mathématiques. Si la physique théorique utilise presque tous les domaines des mathématiques, beaucoup de ses grandes théories sont exprimées avec des équa-

[1] *4, quai du Cheval Blanc; CH-1227 Carouge; Université de Lausanne*; Florent.dieterlen@bluewin.ch

tions différentielles. On peut ainsi dire que les équations différentielles forment la base principale de la représentation scientifique actuelle.

L'outil mathématique principal des sciences humaines sont les statistiques, qui sont en marge du corps des mathématiques. On a assisté depuis les années 70 à des essais d'introduire les équations différentielles en psychologie et en économie (systèmes dynamiques). En sociologie et sciences politiques cela semble moins fréquent. On peut lire depuis, de nombreux articles de psychologie décrivant les systèmes dynamiques, d'une façon ou d'une autre, mais on en reste, en général, à des voeux pieux, avec de temps en temps une équation type, à laquelle on adapte le système à décrire, plutôt que de trouver les équations différentielles qui lui correspondent.

Les économistes et les financiers ont vraiment commencé à explorer ce domaine lors de la mode du «chaos», à la fin des années 80. La théorie des catastrophes a fourni quelques modèles simples (Zeeman 1974; Ho et Saunders 1980; Rand 1977). On a réalisé aussi des modèles non linéaires simples et souvent seulement qualitatifs (Zhang 1999; Greiner 2007; Stutzer 1980; Zellner et Israelevich 2005). Plus nombreux sont les articles qui portent sur la nécessité des systèmes dynamiques et au fait que l'économie et la finance sont non linéaires par essence (Nichols 1993; Grandmont et Malgrange 1986; LeBaron 1994; Rosser 1991) et sur la recherche de non linéarités, de chaos ou de structures fractales, en économie (Frank et al. 1988; Frank et Stengos 1988) ou en finance (Hsieh 1991; DeCoster et al. 1992; Varson et Jalilvand 1994; Peters 1989; Ambrose et al. 1992).

Je vais montrer ici qu'il est facile de construire des systèmes dynamiques hors équilibre originaux, spécifiques au problème étudié, à partir de données ou de tests. Les différentes méthodes permettant de construire des équations différentielles à partir de données temporelles sont nombreuses et en général assez compliquées: intégration par prédicteur-correcteur d'Adams (Brown et al. 1994), analyse fonctionnelle de données (Ramsay et Silverman 2002), méthodes non paramétriques (Jost et Ellner 2000; Peifer et al. 2002), méthode des «splines» (Bellman et Roth 1986), addition de constantes et résolution de systèmes d'équations algébriques non linéaires (Palaniyandi et Laksgmanan 2004), programmation génétique (Ahalpara et Parikh 2006a), itérations discrètes (Gouesbet et Maquet 1992), ondelettes (Ahalpara et Parikh 2006), et presque toujours sur des données incomplètes (Eisenhammer et al. 1991; Cremers et Hübler 1987; Crutchfield et Mcnamara 1987; Varah 1996; Boggs et al. 1987). Ces méthodes sont complémentaires de celle présentée ici, pour des données incomplètes, en sciences humaines. Elle est tellement simple que personne ne semble avoir pris la peine de la publier. Nous partirons de l'hypothèse qu'elle est connue, mais qu'elle n'a pas été exploitée en sciences humaines, ni validée.

Cette méthode permet de construire des équations différentielles ordinaires de façon simple, mais quand même robuste, quand on a des données complètes et contenant du bruit de façon raisonnable, ce qui est souvent le cas en sciences humaines. Elle permet non seulement de dire si un système dynamique est chaotique, oscillant, ou stable, mais aussi de fournir les équations différentielles du système, et de ce fait, de prévoir son évolution dans les cas non chaotiques, d'analyser ses états stationnaires et leur stabilité, et de les simuler pour prédire son futur. Un autre avantage est que la connaissance des équations différentielles donne le graphe des influences mutuelles entre les variables, c'est à dire la connaissance de la structure du système étudié.

Cette méthode est basée sur un fit non linéaire des dérivées par rapport au temps. La plupart des recherches en sciences humaines ont des données qui évoluent plutôt lentement dans le temps, et sont assez lisses, ce qui rend cette méthode facile à utiliser. Une exception notable est celle des changements brusques de dynamique (il y en a beaucoup en sciences politiques, par exemple) la chute d'un empire, les révolutions économiques ou politiques, un changement de majorité, le déclenchement et la fin d'une guerre, le passage à l'indépendance, la transition vers la démocratie, les changements d'opinion, etc.

II. Description de la méthode et validation

II.1. Choix du système et des variables

Les premiers chercheurs appliquant les systèmes dynamiques utilisaient les influences mutuelles entre les composants du système (un «graphe») pour construire des équations différentielles, ou bien ils prenaient des équations différentielles connues et essayaient d'adapter leur modèle à ces équations (Haken et Stadler 1990; Tschacher et Dauwalder 2003). Nous verrons que le choix des approches est plus large, et qu'on peut utiliser le graphe des influences, les données temporelles (la dynamique observée, deux époques ou une estimation de l'évolution du système dans différentes configurations peuvent suffire) ou une combinaison de ces approches (voir graphe 1).

Quelle que soit l'approche, l'utilisateur choisit d'abord un système qui évolue dans le temps: un organe, un être vivant, un groupe d'humains, un marché financier, une banque, une économie, un pays, la Terre, etc., et un phénomène à décrire: dysfonctionnement de l'organe, dépendance à une drogue chez une personne, chute d'un empire, démographie de la Terre, risque de crash financier, évolution du taux de change ou de matières premières, relation entre emploi et facteurs économiques d'un pays, lien entre économie et taux moyen de CO_2 (et donc température moyenne de la terre), etc.

Il choisit ensuite une ou plusieurs variables (il vaut mieux commencer par peu de variables) dont on pense qu'elles s'influencent peut-être mutuellement, et qui sont caractéristiques du système et du phénomène à étudier.

Pour le nombre de variables, idéalement, on doit faire un compromis entre un nombre suffisant (pour que le système décrit soit assez complet), et pas trop (pour qu'elles participent à la description du système, c'est à dire que leur apport au pourcentage de la variance soit suffisamment élevé).

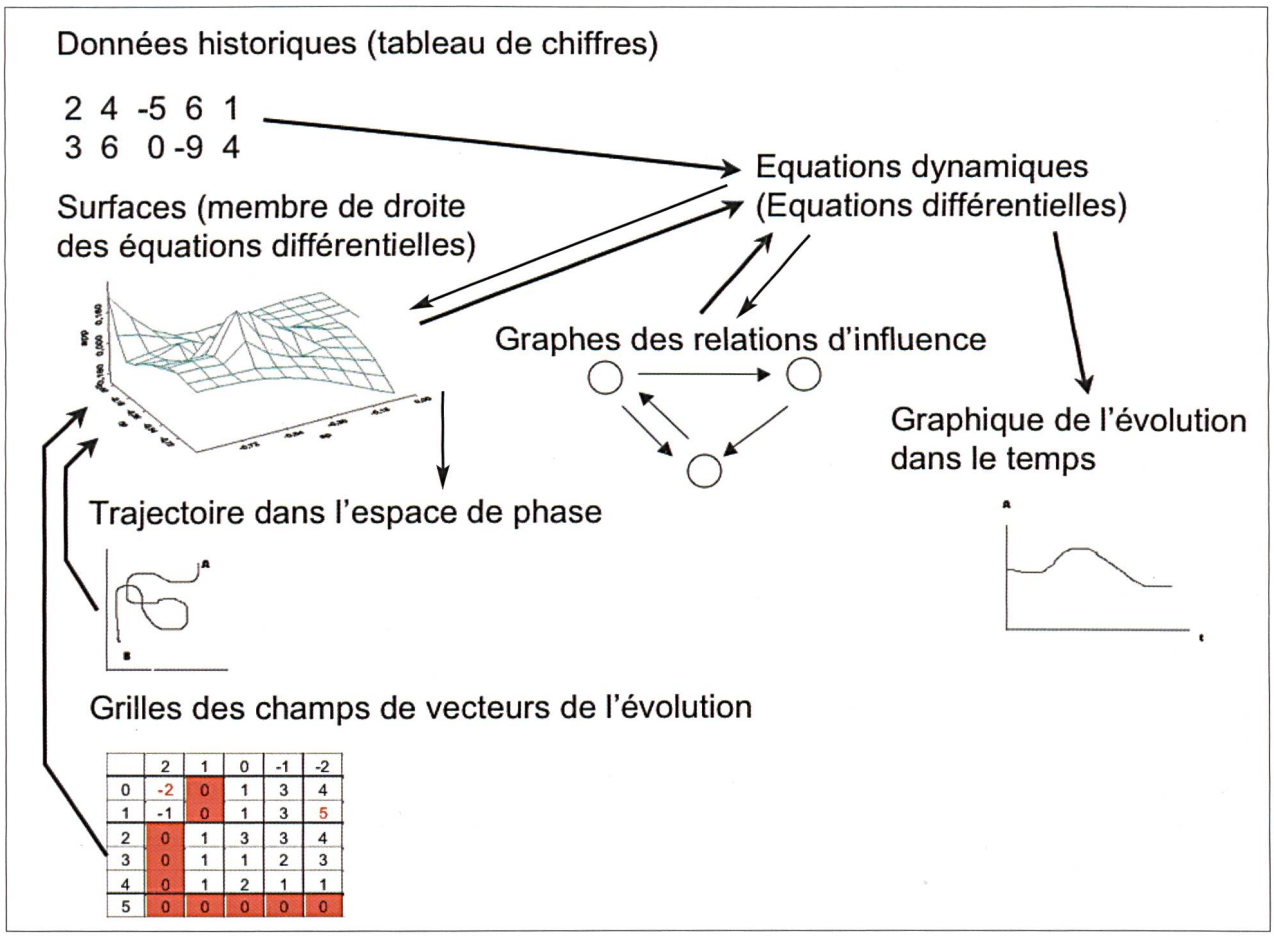
II. 2. Présentation des données

Les valeurs en question, c'est à dire les données, doivent être soit:

- a) Représentées dans plusieurs endroits à deux mêmes dates suffisamment rapprochées (deux est un chiffre minimum, mais plusieurs dates valent mieux, voir la démographie plus loin)
- b) Soit en un même lieu mais à de nombreuses dates simultanées (voir compétition entre bactéries, économie et finance plus loin).
- c) Une troisième possibilité est de connaître l'évolution (qui correspond à la dérivée par rapport au temps) des variables pour une partie importante des configurations du système (voir psychologie plus loin).

On calcule pour chaque valeur des variables la dérivée par rapport au temps. On utilisera la méthode des différences finies à trois ou cinq valeurs ou, dans

Graphe 1: Graphique des relations entre les différents aspects de la modélisation dynamique en sciences humaines.Le système dynamique, c'est-à-dire les équations différentielles, est construit à partir de données historiques directement (ici l'exemple en économie), ou à partir du champ de vecteurs (ici l'exemple en psychologie). Ces deux méthodes peuvent être aidées par la connaissance partielle du graphe des relations et par celle de la trajectoire du système. Les équations différentielles permettent de retrouver le graphe des influences.



le pire des cas, à deux valeurs. Même des variables variant rapidement (cours des changes en finance) peuvent être modélisées ainsi.

On obtient ainsi, pour n variables, $2n$ colonnes: n colonnes de variables, et n colonnes de dérivées par rapport au temps de ces variables.

II. 3. Traitement des données, construction des équations différentielles

L'étape suivante consiste à faire une régression non linéaire avec ces $2n$ colonnes: on doit trouver pour chaque colonne de dérivées, une fonction non linéaire, en général polynomiale, dépendant des n variables. Pour cela, on peut rentrer suffisamment de non linéarités dans un logiciel faisant une régression linéaire (SPSS, Splus, etc.). Suivant les cas, on gardera les termes donnant le meilleur fit, on peut aussi utiliser les méthodes «forward» et «backward», qui donnent directement ces termes. Le pourcentage de variance expliquée R^2 est important, il peut permettre de savoir si le nombre de variables est suffisant ou non.

On trouve alors les termes de droite des équations différentielles, il suffit alors de rajouter les dx_i/dt à gauche et l'on a les n équations différentielles:

$$dx_i/dt = f_i(x_J), \text{ avec } i \text{ et } j \text{ allant de 1 à } n.$$

II. 4. Analyse des équations différentielles

La connaissance des n équations différentielles nous permet de connaître beaucoup de caractéristiques sur le système considéré:

- Son évolution (solution du système), simulée par un logiciel adéquat (Mapple, Matlab, etc.), sur la durée des données et sur l'avenir. L'obtention d'équations différentielles permet naturellement de la simuler avec un logiciel (Mapple, Matlab, etc.). Cela permet d'une part de prédire l'évolution, d'autre part de comparer cette évolution prédite avec la réalité. L'application de la méthode Newton-Gauss permet d'affiner les coefficients et exposants des équations différentielles en fittant au plus près la courbe prédite et la courbe d'évolution réelle.
- Ses états stationnaires (équilibres dynamiques vers lesquels il tend). Cet état «final» ne l'est pas vraiment, car le système peut dépendre d'une ou plusieurs variables supplémentaires à partir d'un certain moment, et donc changer d'évolution (Nottale et al. 2000).
- La stabilité des états stationnaires (Glansdorff et Prigogine 1971).
- L'existence éventuelle d'oscillations (obtenue grâce au calcul de stabilité) ou de chaos.
- Les bifurcations vers d'autres états stationnaires, liées à une perte de stabilité.

II. 5. Prédictions

Les prédictions temporelles se manifestent de plusieurs façons:

- La forme explicite de la solution donne l'évolution du système dans le temps (toujours en fonction de la précision des données).
- Le calcul de stabilité, dans le cas stable, donne une indication sur le temps que le système mettra pour atteindre l'état stationnaire.
- De même, dans le cas oscillant, le calcul de stabilité donne la période du cycle.
- Description géométrique (Abraham et Shaw 1983): Les termes de droite des équations différentielles, trouvés grâce à cette méthode, sont des surfaces (ou hypersurfaces) dans l'espace des dérivées temporelles en fonction des variables. Chaque point des données se trouve idéalement sur cette surface. Dans la réalité, ces points se trouvent près de ces surfaces, et ce sont eux qui permettent de trouver ces surfaces par régression non linéaire. Quand le système part d'un point éloigné de l'état stationnaire, les points seront au début éloignés des surfaces en question, et ces points particuliers sont de peu d'utilité pour appliquer la présente méthode.

III. Validation et robustesse du modèle

III.1. Modèle de Lorenz

III.1.a. Cas chaotique

Nous avons pris le cas classique du modèle de Lorenz (Lorenz 1963):

$$dx/dt = s\,(y - x)$$
$$dy/dt = rx - xz - y$$
$$dz/dt = xy - bz$$

Les paramètres valent, pour le cas chaotique: $s = 10$, $r = 28$, $b = 8/3$

On retrouve les équations de Lorenz à partir de peu de données si les dérivées sont bonnes et les données réparties sur un grand intervalle, avec plus de données si les dérivées sont moins précises. J'ai utilisé la méthode des différences finies à 5 points pour calculer les dérivées.

Les points de la courbe de l'erreur en fonction du bruit ne sont pas sur une droite mais sur une parabole. Pour trouver son équation, on a pris les logarithmes des valeurs (en omettant l'origine, Fig. 1), ce qui nous donne l'équation suivante, qui est une loi en puissance:

$$Erreur = 0{,}3116\ (bruit)^{1.295}$$

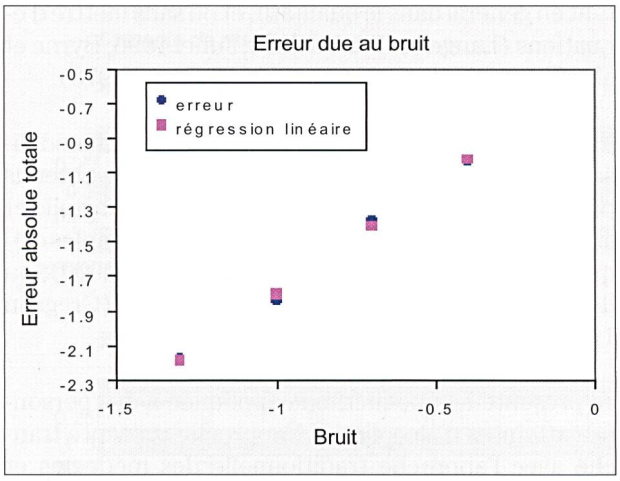
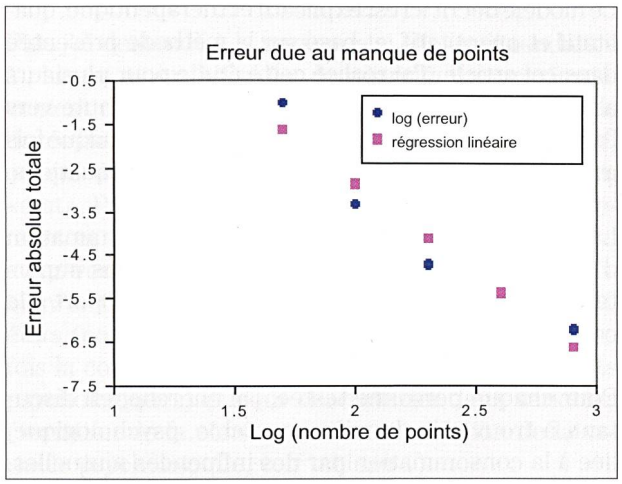


Fig. 1: Logarithme de la somme des erreurs absolues en fonction du logarithme du bruit (losanges bleus), comparées à la régression linéaire (carrés roses), aussi sur les logarithmes. Cette dernière définit une loi en puissance, d'exposant 1.295. Unités: bruit: logarithme du pourcentage de bruit ajouté (0.1=10% de bruit); erreur: logarithme des unités arbitraires des paramètres des équations.

On a fait directement le graphe des logarithmes des erreurs en fonction du logarithme du nombre de points utilisés, toujours tous les 0.01 unités de temps (arbitraire, dépend des unités des paramètres de l'équation, eux aussi arbitraires ici), à partir du temps 1. On a aussi une loi en puissance (Fig. 2):

$$Erreur = 238097\ (nombre\ de\ points)^{-4.137}$$

Fig. 2: Logarithme de la somme des erreurs absolues (losanges bleus) due au manque de point (l'erreur diminue avec le nombre de points) en fonction du logarithme du nombre de points (carrés roses), comparées à la régression linéaire, aussi sur les logarithmes. Cette dernière définit une loi en puissance, d'exposant –4.137. Unités: Nombre de points: logarithme du nombre de points: erreur: logarithme des unités des paramètres des équations



Remarquons que nous avons pris des points équidistants, ne couvrant pas tout l'espace de phase, mais seulement une petite partie, en particulier lorsque nous avons pris peu de points.

III.1.b. Cas stable

La dernière remarque est encore plus vraie lorsque l'on veut retrouver les équations dans le cas non chaotique, avec des états stationnaires (par exemple en prenant les 3 paramètres $s = b = r = 1$. Le cas stable est, de façon étonnante de prime abord, plus difficile à résoudre pour un modèle comme celui de Lorenz que pour une application quelconque. En effet, on doit simuler les équations à divers moments, à partir de valeurs initiales. Or des valeurs initiales quelconques donneront des points qui seront, pour des petits t, loin des hypersurfaces des équations, et pour des t grands, seront trop près de l'état stationnaire, et les x, y et z obtenus seront trop proches les uns des autres. Cela n'arrive pas dans la grande majorité des applications, sauf lorsque le système a été perturbé et est dans la phase de rapprochement des surfaces des équations, comme pour un point initial pris au hasard.

Ainsi, pour 10 points pris aux temps 1 à 10 avec la condition initiale (5, 5, 5), avec les dérivées calculées à partir des équations, on obtient une équation pour dz/dt très proche de Lorenz (à un millième près), mais les 2 autres équations sont mauvaises. Comme on calcule les dérivées par la méthode des différences finies, le fait que les valeurs varient beaucoup moins que dans le cas chaotique, entraîne que le calcul des dérivées par la méthode des différences finies est bien meilleur, cela donne des valeurs très proches des valeurs «réelles», et ne nécessite pas beaucoup plus de données pour une même précision dans les résultats.

III.2. Cas oscillant: Lotka-Volterra

Considérons le modèle de Lotka-Volterra pour le cas oscillant:

$$dx/dt = x - xy$$
$$dy/dt = -y + xy$$

Le point initial n'est pas très important, comme dans le cas chaotique, pour autant que l'on n'ait pas des valeurs pour des très petits t, car le système varie assez vite.

a) 12 points entre $t = 6$ et $t = 13$, avec 3 chiffres significatifs donnent l'équation exacte, à moins d'un millième près.

b) Les résultats avec plus de points sont proches de ceux obtenus avec Lorenz: le premier critère pour avoir de bons résultats est la répartition des points, puis le nombre de points et la précision des données.

On ne s'approche pas d'un état stationnaire qui enlève de l'information pour la méthode, en rapprochant les points. Pour les applications, le cas oscillant est comme les autres cas, il faut seulement que les valeurs des variables soient suffisamment différentes entre elles pour que l'on puisse construire les hypersurfaces (dans un espace à n dimensions, où n est le nombre de variables) correspondant aux équations différentielles.

IV. Applications

IV.A. Données uniquement temporelles ou connaissance de l'évolution de chaque état:

Ce type de données intervient dans les domaines suivants:

- Finance: Dow Jones, change, etc.
- Climat: température moyenne de la terre, CO2 moyen, etc.
- Médecine, biologie, chimie
- Psychologie, sociologie

IV.A.1. Test sur l'évolution du système (connaissance des dérivées approximatives pour chaque état): exemple en psychologie - psychiatrie

Les psychologues savent que les statistiques ne peuvent pas résoudre tous leurs problèmes, surtout ceux où apparaît le temps. Ils parlent depuis longtemps d'utiliser les systèmes dynamiques (Guntern 1982; Haken et Stadler 1990; Robertson et Combs 1995; Van Gelder 1998; Smith et Thelen 1993; Bütz 1992; Pezard et al. 1996; Barton 1994; Priel et Schreiber 1994; Paulus 2003). En 1976, lors du 80e anniversaire de Piaget, Prigogine avait été invité. Mais le dialogue n'est pas passé, et l'école de Prigogine, prolifique dans d'autres applications des systèmes dynamiques, n'a pas percé.

Ils pourraient pourtant utiliser un système d'équations différent pour chaque application, mais n'ont pas les instruments pour construire ces systèmes, contrairement aux biologistes et certains chercheurs en médecine qui ont trouvé des équations originales, correspondant à leur problèmes respectifs, avec souvent des calculs de bifurcations (Goldberger et al. 1988; Kahn et Hobson 1993; Stam et al. 1999; Erdi et al. 1993; Vaidya et al. 1991; Fussmann et al. 2000; Eyal et Akselrod 2000; Cavalcanti 2000; Sturis et Brons 1997; Porcher et Gatto 2000; Goushcha et al. 2000; Abbiw-Jackson et Langford 1998).

D'autres chercheurs en psychologie ont mis leurs théories dans un autre cadre, celui de la théorie des catastrophes. Là aussi, on a affaire en général à une sorte d'équation-type (celle de la fronce), à laquelle les psychologues ont essayé de restreindre leurs modèles, en restant en général dans le qualitatif, et ou sans mettre d'équations (Lange et Houran 2000; Juhel 1996; Byrne et al. 2001; Clair 1998; Craigie 1980; Zeeman 1973).

Par contre, des résultats intéressants pour la modélisation, la dimension des attracteurs, ont été obtenus par certains auteurs (Hornero et al. 1999; Guastello et al. 1999). D'autres ont utilisé des systèmes simples d'équations différentielles (par ex. Partridge 2000), ou des systèmes dynamiques à temps discret (Gregson 1999; Warren et al. 2003).

Je présente ici une méthode appliquée à des personnes atteintes d'alcoolisme. L'approche présente tranche avec l'approche traditionnelle des médecins et psychologues: au lieu de faire des statistiques (par ex. Jost et Ellner 2000; Daeppen et al. 1996; Rosenberg et al. 1992) sur une population, et de comparer les patients à la moyenne, nous étudions les comportements individuels des patients.

La plupart des modèles explicatifs ou thérapeutiques de l'alcoolisme sont ou bien seulement quantitatifs (statistiques, par ex. Ellis et McClure 1992), ou bien qualitatifs (par ex. Galanter 1993). Il en va de même pour les modèles classifiant les types d'alcoolisme (Reynolds et al. 1992; Babor et al. 1992). Mohan et al. (1992) discutent des choix de variables liées à l'alcoolisme.

Deux modèles ont abordé l'alcoolisme par la dynamique: un modèle discret non linéaire (Warren 2002, 2003) qui n'utilise qu'une seule variable, la consommation, et est donc trop simple pour expliquer, guérir ou classifier. L'autre modèle est plus proche du notre (An der Heiden et al. 1998): il construit deux équations différentielles, proches des modèles de la théorie des catastrophes, mais il n'est que qualitatif, avec deux variables, la consommation et la variable psychologique «frustration». Aucune donnée expérimentale n'est utilisée. Il donne cependant des comportements intéressants, et présente un début de classification des comportements alcooliques.

Le modèle décrit ici est explicatif et thérapeutique, qualitatif et quantitatif, et basé sur la méthode présentée dans cet article. J'ai réalisé cette étude pour plusieurs alcooliques, et pour différentes situations: chute vers l'alcoolisme, arrêt de l'alcoolisme, et rechute, chaque fois avec seulement deux variables et donc deux équations.

La première variable est évidemment la consommation d'alcool, en unités standard. L'unité n'est pas importante, car on compte les proportions par rapport à la consommation maximale.

Pour chaque personne testée, j'ai cherché, en discutant, à trouver la deuxième variable, psychologique, liée à la consommation par des influences mutuelles,

Exemple de résultat de test:

variation de la consommation d'alcool:

consommation \ conscience du problème	0	0.25	0.5	0.75	1
0	3	2	2	1	1
0.25	2	1	1	1	0
0.5	1	0	0	0	-1
0.75	0	-1	-1	-2	-2
1	-2	-3	-3	-4	-4

variation de la conscience du problème:

consommation \ conscience du problème	0	0.25	0.5	0.75	1
0	0	-1	-1	-2	-3
0.25	1	0	-1	-1	-2
0.5	2	1	0	-1	-1
0.75	2	2	1	0	-1
1	3	3	2	1	0

4= augmentation maximum, -4=diminution maximum

Tableau 1: Tableau des résultats des questions posées à la personne testée. Le tableau du haut correspond aux chiffres d'augmentation ou diminution de la consommation d'alcool. Le tableau du bas donne les variations de conscience du problème. En abscisse, la variable psychologique (conscience du problème dans ce cas), en ordonnées, du haut vers le bas, la consommation d'alcool. Les unités de ces deux variables sont les proportions par rapport à la valeur maximum (consommation maximum par jour, et conscience aiguë du problème). Les unités des valeurs du tableau, c'est à dire des variations de consommation et de conscience, sont notées de –4 (diminution maximum) à +4 (augmentation maximum). Elles sont approximatives puisque subjectives, mais la méthode est suffisamment robuste pour accepter ces incertitudes.

J'ai montré ici un exemple simple de test sur un ancien alcoolique, décrivant sa dynamique avant l'arrêt. Sa variable psychologique est la conscience du problème, qui a été à l'origine de son arrêt de l'alcool. Le tableau 1 montre le résultat du test. Chacun des chiffres représente la dérivée par rapport au temps, c'est-à-dire le degré d'augmentation ou de diminution, de la consommation sur le tableau supérieur, et de la conscience du problème sur le tableau inférieur.

La Fig. 3 montre la superposition de ces deux tableaux, qui donne le champ de vecteurs sur le système. Les flèches rouges ont une longueur proportionnelle aux chiffres du tableau supérieur (consommation d'alcool) et les flèches vertes ont une longueur proportionnelle aux chiffres du tableau inférieur, la conscience du problème. Les flèches noires sont la somme des vecteurs rouges et verts pour chaque point de l'espace de phase. Chaque état est représenté par 2 coordonnées, la consommation et la conscience du

c'est à dire par feedbacks positifs et négatifs. Les variables trouvées pour ces quelques sujets ont été:

- avant la chute: l'anxiété, la dépression, trop de confiance en soi;
- avant l'arrêt: la conscience du problème, le bien-être du à la gratification, la pression pour arrêter;
- avant la rechute: l'anxiété, le stress, les problèmes socio-économiques.

Chacune de ces variables influençait et était influencée par la consommation.

Puis j'ai demandé à la personne si elle avait tendance à plus ou moins boire (et de quantifier approximativement cette augmentation ou diminution), sur une durée de quelques jours (plutôt que dans la journée, ce qui donnerait un tout autre modèle) selon son «état». Par exemple: «lorsque vous buviez moyennement et que vous étiez très conscient du problème, aviez-vous tendance à boire plus ou moins, et de combien?». Ce genre de question était répété pour 25 états (consommation égale à 0, 0.25, 0.5, 0.75 et une fois la consommation maximum, idem pour les états psychologiques) et pour chacune des deux variables (variation de la consommation et variation de la variable psychologique, voir tableau 1).

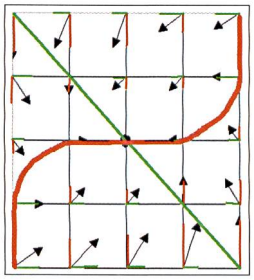


Fig. 3: Graphique du champ de vecteurs construit grâce au tableau 1. Axe horizontal: «Conscience du problème» (unités. Proportions par rapport au maximum); Axe vertical: Consommation d'alcool (du haut vers le bas, comme pour le tableau, unités identiques au tableau 1). Les vecteurs rouges (proportionnels aux chiffres du tableau supérieur, variations de la consommation d'alcool) et verts (proportionnels aux chiffres du tableau inférieur, conscience du problème) ont une longueur proportionnelle aux chiffres du tableau 1 (unités du tableau 1). Les vecteurs noirs représentent le champ de vecteurs, et sont égaux à la somme des vecteurs rouges et verts. Les lignes rouges et vertes en gras sont les états du système pour lesquels la consommation d'alcool (rouge) et la conscience du problème (vert) sont constantes. Le point noir au milieu, leur intersection, est l'état stationnaire du système, c'est à dire l'état vers lequel tend la personne testée. Dans les équations qui suivent, la ligne sinueuse (non linéaire) rouge a été approximée en une ligne droite (linéaire) par soucis de simplicité. Cette approximation ne change pas le problème dans ce cas précis. Les non linéarités doivent être prises en compte en général.

problème. La figure montre donc tous les états possibles de cette personne pour ce qui est de sa situation par rapport à ces 2 variables ou coordonnées. On a ainsi le champ de vecteur du système. Un modèle plus sophistiqué aurait plus de variables, et le champ de vecteurs aurait plus que 2 dimensions, mais serait difficile à interpréter visuellement.

Les deux lignes (rouge pour le tableau des variations de consommation et verte pour le tableau des variations de conscience du problème) représentent les états stationnaires pour chacune des variables: elles montrent les états (psychique et consommation) pour lesquels les dérivées par rapport au temps sont nulles (consommation et état psychique). Ici les formes de ces 2 courbes donnent le terme de droite des équations différentielles du système. L'état stationnaire du système (en l'occurrence la personne testée) est l'intersection entre les deux courbes (voir Fig. 3). Il peut y avoir en général plusieurs intersections et donc plusieurs états stationnaires.

Dans ce cas précis, on a un seul état stationnaire, vers le milieu du graphe: avant l'arrêt de l'alcool, cette personne tendait à boire en moyenne la moitié de sa dose maximale, et à avoir une conscience du problème réduite de moitié. Pour cette variable psychologique, il s'agit évidemment d'une quantification subjective. Le fait que la réalisation de modèles sur plusieurs personnes ait donné des résultats chaque fois correspondant à la réalité montre que l'on peut utiliser de telles variables, avec une marge d'erreur assez grande, sans que le résultat soit faussé. La méthode montre donc une certaine robustesse.

Un examen visuel du champ de vecteur montre que le système va vers cet état stationnaire en un mouvement tournant, effectuant une spirale. Ce résultat, comme toutes les informations réunies sur ce graphique, se retrouvent dans l'analyse des équations obtenues: le calcul de stabilité donnera des valeurs propres complexes, ce qui signifie que le système oscille.

Les équations différentielles sont données par la forme des courbes rouges et vertes du graphique. Nous allons les approximer ici par des droites, par souci de simplification. Dans ces modèles appliqués à la psychologie, on a en général à faire à des courbes non linéaires, ce qui donne des équations différentielles non linéaires. Leur étude n'est pas beaucoup plus compliquée que pour des équations linéaires, car les non linéarités sont en général des carrés ou des cubes. Le cas linéaire permet des oscillations mais exclut une bifurcation, qui nécessite des non linéarités.

Pour être précis, on aurait du prendre ici une courbe de degré 3 pour l'alcool. Cela ne change pas beaucoup l'étude du cas, car il aurait fallu que les courbes soient assez différentes (comportement différent pour la personne testée) pour qu'on ait 3 intersections au lieu d'une. Certains des alcooliques testés étaient dans ce dernier cas, avec 3 états stationnaires, les 2 extrêmes stables et celui du milieu instable. On peut alors avoir un saut d'un état alcoolique extrême à un état sobre (les 2 états extrêmes stables), ou réciproquement, lorsqu'on change un des paramètres (un des aspects du comportement).

Ici, les droites ont approximativement pour équations:

$$y = -3x/2 + 12$$
$$x = y$$

On en déduit facilement les équations différentielles:

$$dx/dt = -3x - 2y + 24$$
$$dy/dt = x - y$$

Notons que ces équations ressemblent aux célèbres équations de Lotka-Volterra, et ont un comportement similaire. La différence essentielle est que les oscillations sont ici atténuées (comme dans les exemples du paragraphe suivant sur la compétition entre bactéries), d'où une trajectoire en spirale, au lieu d'une trajectoire constituée de cercles concentriques dans le cas des oscillations pures.

L'analyse des équations donne un état stationnaire à $x = y = 4.8$, comme trouvé approximativement sur le graphique. Le calcul de stabilité donne des valeurs propres complexes à partie réelle négative, caractéristiques d'un système stable oscillant atténué et se dirigeant en spirale vers un attracteur. La valeur des parties imaginaires est liée à la période des oscillations (que l'on peut retrouver dans le comportement de la personne en question) et les parties réelles sont liées au temps moyen de ces oscillations, jusqu'à l'état stable. Toute fluctuation du système (petites perturbations qui affectent la personne) va perturber le système (la personne est «déstabilisée»), qui reviendra ensuite vers son état stable en oscillant.

A tout système d'équations différentielles comme celui décrit ici, correspond un graphe des influences mutuelles entre les variables. Ici, ce graphe est représenté dans le graphe 2. Chacun des paramètres (chiffres) correspond à une flèche.

On peut aussi voir avec ces équations, le graphe des influences et le graphe du champ de vecteurs, comment on peut aider le patient (si on avait réalisé ce modèle avant sa guérison, et donc trouvé la bonne variable psychologique). Voyons-en un exemple. Remplaçons chacun des coefficients des équations par des lettres:

$$dx/dt = -ax - by + c$$
$$dy/dt = dx - ey + f$$

L'état stationnaire pour la consommation d'alcool est:

$$x = (c - bf)/(db + a)$$

On voit que la consommation, dans ce modèle simplifié, sera diminuée si

- *c*, l'influence extérieure sur la boisson (situations favorisant la prise de boisson par ex.), diminue
- *f*, l'influence extérieure sur la prise de conscience (le conseil de médecins par ex., augmente
- *b*, la diminution de boisson due à la prise de conscience, augmente
- *d*, la prise de conscience due aux excès de boisson, augmente.

Ces conclusions peuvent paraître évidentes, mais pour des modèles plus complexes, on arrive à des conseils plus sophistiqués, avec par exemple une bifurcation d'un état buveur excessif à un état non buveur ou buveur modéré.

La comparaison des résultats pour plusieurs patients permet de classer les types d'alcoolisme, et par là même les types de traitement.

IV.A.2 Compétition entre populations

Nous donnerons ici un aperçu de la méthode appliquée aux oscillations dans la compétition entre 2 sortes de bactéries. Nous avons pris les données de Gause (Jost et Ellner 2000). Ces données sont assez courtes (18 valeurs, voir annexe), et donnent des oscillations qui sont loin d'être parfaites. En conséquence, on ne peut espérer obtenir des équations donnant des oscillations parfaites. Nous avons obtenu des oscillations dont l'amplitude décroissait au cours du temps, à partir d'amplitudes correctes. Pour avoir un bon résultat, la qualité des données importe plus que leur longueur.

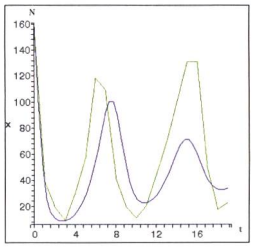


Fig. 4: Modèle de compétition entre bactéries en xy, variable x, comparaison des données (en vert) avec la simulation (en bleu). Unités: ordonnées = nombre de bactéries, abscisses = temps (jours)

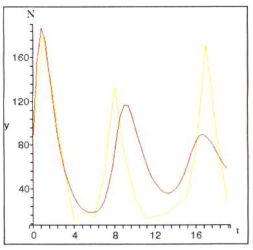


Fig. 5: Modèle de compétition entre bactéries en xy, variable y, comparaison des données (en jaune) avec la simulation (en rouge). Unités: ordonnées = nombre de bactéries, abscisses = temps (jours)

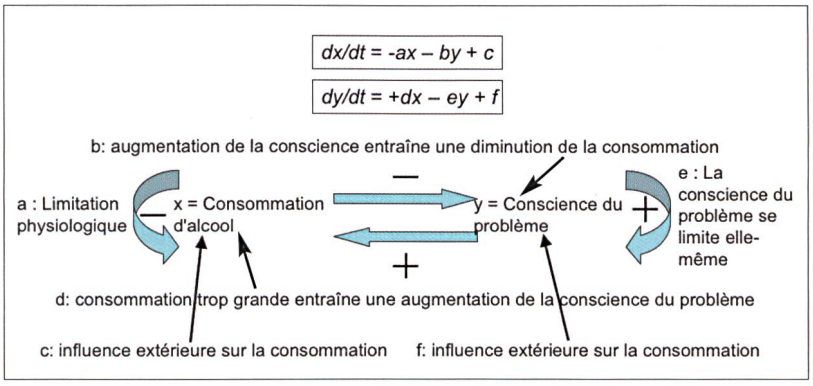
J'ai pris les termes en x, x^2, y, y^2, xy pour commencer, puis ai rajouté les termes de degré supérieur. Pour les valeurs des dérivées, nous avons pris les dérivées sur 2 points pour la deuxième et l'avant-dernière mesure, et les dérivées sur 5 points pour les mesures comprises entre ces 2 mesures.

Pour les termes jusqu'au deuxième degré, on obtient les équations suivantes:

$$dx/dt = 1.108x - 0.00377x^2 - 0.0145xy$$
$$(R^2 = 86\%)$$
$$dy/dt = -0.812y + 0.0173xy \ (R^2 = 83\%)$$

(R^2 est la variance expliquée)
On constate sur les Figs. 4 et 5 que la première oscillation est bien reproduite, que la période est assez bonne, mais que les oscillations diminuent, ce qui sera toujours le cas. La variance expliquée par le modèle est élevée, au dessus de 80%. Nous verrons plus loin qu'une variance élevée n'est pas toujours synonyme de succès.

Graphe 2: Graphique des relations d'influence mutuelles entre les deux variables du modèle simplifié d'un alcoolique. A chaque coefficient des équations différentielles (a, b, c, d, e, f) *correspond une flèche du graphique, représentant un effet psychologique.*



i) Termes jusqu'au troisième degré:
On obtient les équations suivantes:

$$dx/dt = 1.299x - 0.00599x^2 - 0.01705xy + 3.526\ 10^{-5}x^2y$$
$$(R^2 = 86\%)$$
$$dy/dt = -1.34y + 0.0387xy - 1.508\ 10^{-4}x^2y$$
$$(R^2 = 93\%)$$

On voit sur les Figs. 6 et 7 que l'ajustement est encore meilleur que pour les équations en xy: le deuxième pic coïncide presque totalement, et les amplitudes baissent moins, et cela pour les 2 variables.

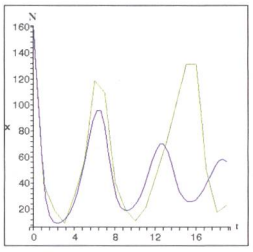


Fig. 6: Modèle de compétition entre bactéries en x^2y, variable x, comparaison des données (en vert) avec la simulation (en bleu). Unités: ordonnées = nombre de bactéries, abscisses = temps (jours)

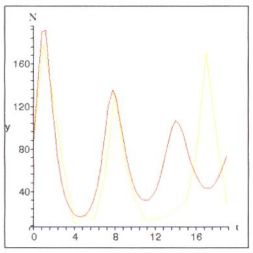


Fig. 7: Modèle de compétition entre bactéries en x^2y, variable y, comparaison des données (en jaune) avec la simulation (en rouge). Unités: ordonnées = nombre de bactéries, abscisses = temps (jours)

ii) Termes jusqu'au quatrième degré:
On obtient les équations suivantes:

$$dx/dt = 1.3066x - 5.973\ 10^{-3}x^2 - 0.01126xy + 1.478\ 10^{-4}x^2y + 1.0699\ 10^{-6}x^3y \quad (R^2 = 87\%)$$
$$dy/dt = -5.166y + 0.04151xy - 2.089\ 10^{-4}x^2y + 2.99\ 10^{-7}x^3y \quad (R^2 = 92\%)$$

Le fit est mauvais dans ce cas. On pouvait le deviner en voyant que pour x l'alternance des signes n'est pas respectée. On constate aussi que la variance expliquée est toujours haute, et que ce fait n'est pas une condition suffisante pour un bon ajustement. Le fait que celle-ci soit trop basse dans certaines applications entraîne par contre un mauvais fit, et indique surtout qu'on a besoin d'au moins une variable supplémentaire et/ou d'appliquer la méthode de Newton-Gauss.
Le fit est aussi mauvais lorsqu'on ajoute un terme de degré cinq.

En conclusion, bien que ces données étaient très courtes et avec beaucoup de bruit, la méthode a donné quand même des résultats satisfaisants.

IV.A.3. Macro-économie: prévision de la croissance

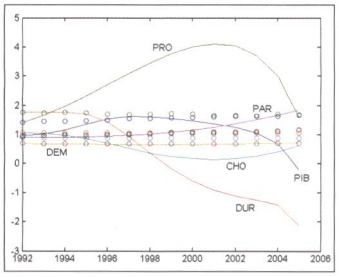
Les modèles macro-économiques ne sont presque jamais continus et endogènes (la méthode présente peut cependant introduire des facteurs exogènes, en faisant un système ouvert). Je montre ici qu'un modèle continu obtenu grâce à la présente méthode peut donner des prévisions comparables aux modèles classiques. Je présente ici un modèle de prévision de la croissance.

Plusieurs choix de variables sont envisageables pour la précision de la croissance: variables constitutives du PIB (consommation, investissements, importations et exportations, variations de stock), variables non constitutives, ou mélange des deux. Je présente ici un exemple de prévision à bonne précision basé sur des variables non constitutives, suggérées par Sollberger de l'Office Fédéral de la Statistique (Berne) et basé sur un rapport de cet office (Rais et Sollberger 2007). Ce choix est doublement original, étant ni des variables constitutives ni les variables habituellement les plus représentatives de l'influence sur le PIB (salaires, taux d'intérêt, etc.), mais il représente une action désirée, c'est à dire une influence sur le PIB. Nous ne prétendons pas qu'il s'agisse des meilleures variables possible, mais dans ce cas précis, elles ont donné une bonne prévision.

Les 6 variables annuelles sont le PIB (sa mesure du 4e trimestre ici), la productivité (PRO) horaire du travail, la durée (DUR) moyenne du travail par actif occupé, le chômage (CHO), le taux de participation (PAR) sur le marché du travail et la démographie (DEM). On a changé d'unités pour que toutes les variables soient autour de l'unité. La méthode a été appliquée à ces variables sur les séries chronologiques fournies par l'OFS, en incluant seulement les non linéarités d'ordre 3, avec le méthode backward (jusqu'au quatrième niveau chaque fois) sur SPSS. Les équations différentielles ainsi obtenues ont été simulées sur matlab de 1992 à 2005, et donnent un mauvais fit (voir Fig. 8). On a ensuite appliqué la méthode de Newton-Gauss sur ces équations pour les fitter aux valeurs données par l'OFS (avec Matlab) et le fit est devenu bon (voir Fig. 9). Les équations sont (on a enlevé des décimales pour une question de place):

$$d(PIB)/dt = -1.47 + 0.43PRO^{1.08} + 0.43DUR^{0.95} - 0.13PIB^{3.07} + 0.86DEM^{3.11}$$
$$d(PRO)/dt = -0.18 + 0.091PRO^{0.82} + 0.11DUR^{1.02} - 0.36PAR^{2.97} + 0.51DEM^{3.48}$$
$$d(DUR)/dt = -1.51 + 1.52PIB^{1.1} + 0.043DUR^{0.96} - 0.4PIB^{3} + 1.15DEM^{2.99}$$

Fig. 8: Graphique de la simulation en économie, avant d'avoir appliqué la méthode de Newton-Gauss. Unités: PIB (4e trimestre): 10^{11}CHF; Productivité horaire (PRO): 10CHF/heure; Durée moyenne du travail (DUR): 1000 heures/personne; Chômage (CHO): 10^5 personnes; Participation (PAR): rapport de nombres de personnes; effet démographique: rapport de nombres de personnes.



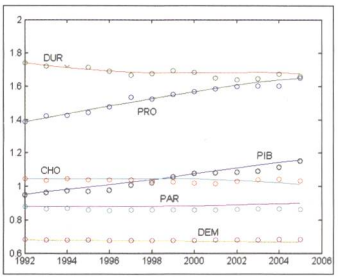


Fig. 9: Graphique de la simulation en économie, après avoir appliqué la méthode de Newton-Gauss (mêmes symboles et unités que pour la fig. 13).

$$d(CHO)/dt = 0.26 + 0.46PIB^{1.01} - 0.2PRO^{0.9} - 0.19DUR^{1.02} - 0.1PIB^{3}$$
$$d(PAR)/dt = -0.015 - 0.022PIB^{1.01} + 0.025PRO + 0.033PAR^{2.7} - 0.077DEM^{3.08}$$
$$d(DEM)/dt = 0.049 + 0.017PIB^{0.99} - 0.016PRO^{1.01} - 0.017DUR^{1.02} - 0.043DEM^{2.63}$$

Remarquons que dans cet exemple, l'équation différentielle pour le chômage n'est pas nécessaire pour la prévision du PIB, mais elle peut permettre de prévoir la valeur du chômage. Ces équations correspondent comme toujours à un graphe des influences mutuelles entre ces variables.

Ces équations ont été simulées pour prédire les valeurs futures. Comme le calcul des dérivées prenait en compte les valeurs de 2006 (modèle fitté de 1992 à 2005), on doit simuler le système sur deux ans, pour prédire ainsi la valeur pour 2007. La valeur du PIB trimestriel (quatrième semestre) pour 2007 est 1.215, proche de la valeur réelle 1.233 (erreur relative: 1.5 %)

Il est possible d'ajouter à ce genre de modèle d'autres variables, non seulement économiques et financières, mais aussi environnementales, énergétiques, climatiques et sociales, pour en faire un modèle global à temps continu (Caetano et al. 2008; Meadows et al. 1972), complémentaire des modèles à temps discret, réalisables aussi avec cette méthode (voir IV.A.4, taux de change).

IV.A.4. Finance

La finance utilise des modèles complexes, mais beaucoup s'accordent pour dire que des modèles plus simples peuvent être aussi compétitifs. Je propose ci-dessous, sans entrer dans trop de détails, une application où les systèmes dynamiques continus peuvent modéliser des instabilités en finance:

Cercle «vicieux» de la crise des subprimes

Le prix Nobel Stiglitz a cité en 2008, lors d'une conférence, le cercle vicieux à la base, selon lui, de la crise de l'immobilier d'août 2007. Le feedback positif consistait selon lui en une hausse des prix de l'immobilier, qui a entraîné une hausse des taux d'intérêt, qui a entraîné une hausse des dettes (des acheteurs puis des banques), qui elle-même entraînait la hausse des prix, ainsi de suite. Ce cercle vicieux a fini par rompre, entraînant une baisse des prix et des taux d'intérêts et une crise due au non remboursement des dettes.

J'ai modélisé ce cercle vicieux en prenant les chiffres de 2000 à 2007 pour l'immobilier américain, les taux à 5 ans, et la dette des particuliers. Ce qui est intéressant ici, c'est de comparer le calcul de stabilité effectué sur les valeurs d'août 2007 et d'avant (par exemple au début de la montée des cours de l'immobilier).

Plusieurs modèles sont possibles avec ces données, selon les feedbacks négatifs que l'on garde. En effet, dans tout modèle avec un feedback positif, il faut avoir des feedback négatifs qui compensent et empêchent le système de «diverger» (s'emballer, exploser).

Voici un exemple simplifié de calcul de stabilité, où on a enlevé la troisième variable (dette) et on a gardé le prix (p) de l'immobilier et le taux d'intérêt (r). Nous avons réalisé huit systèmes, chacun avec des non linéarités différentes. Les équations différentielles sont, après approximation:

$$dp/dt = f_{pp}(p) + f_{pr}(r)$$
$$dr/dt = f_{rp}(p) + f_{rr}(r)$$

Où $f_{pp}(p)$ est la dynamique interne (positive dans les huit cas) de la poussée des prix de l'immobilier (la demande augmente, ce qui fait augmenter les prix, ce qui fait augmenter la demande, probablement par contagion).

$f_{pr}(r)$ est le changement de prix dû aux taux d'intérêts. La méthode donne dans six cas sur huit une valeur négative. On n'a donc pas de feedback positif dans cette moitié du cycle du graphe d'influences mutuelles, ce qui empêche de trouver une instabilité, et montre la limite de ce modèle trop simplifié. Les taux ne font pas monter les prix directement, il faudrait inclure la dette.

$f_{rp}(p)$ est le changement de taux d'intérêt dû au prix de l'immobilier. On a ici des valeurs positives dans les huit cas, cette deuxième moitié du cycle du feedback mutuel est donc positive. Les prix font augmenter les taux, par décisions humaines.

$f_{rr}(r)$ est la dynamique interne des taux d'intérêts. Elle est positive dans les huit cas: cela veut dire que les décideurs (banque fédérale par exemple) ne laissent pas trop monter les taux, ils les limitent, les stabilisent.

La stabilité est obtenue en résolvant l'équation pour les exposants L (Glansdorff 1971):

$$(df_{pp}/dp - L)(df_{rr}/dr - L) - (df_{pr}/dr)(df_{rp}/dp) = 0$$

Les solutions sont:

$$L = ((df_{pp}/dp + df_{rr}/dr) +/- \sqrt{\Delta})/2$$

où:

$$\Delta = (df_{pp}/dp + df_{rr}/dr)2 - 4((df_{pp}/dp)(df_{rr}/dr) - (df_{pr}/dr)(df_{rp}/dp)),$$

et où les derivées sont calculées sur les valeurs de p et r qui nous intéressent: nettement avant août 2007, et en août 2007. On s'attendrait, si le modèle simplifié

était suffisant, à ce que les L calculés pour «avant la crise», disons en 2000, soient négatifs (solution stable) et inférieurs au L calculés en août 2007, qui pourraient être positif (instabilité). Nous avons appliqué la méthode pour trouver les quatre dérivées citées ci-dessus, puis calculé les deux L, pour les valeurs du taux d'intérêt à cinq ans et le prix de l'immobilier en 2000 et en août 2007 (sources: taux d'intérêts à cinq ans: Federal Reserve statistical release; prix de l'immobilier: economagic.com, economic time series page, «median price of homes sold including land price»).

Nous avons pris plusieurs possibilités de non linéarités, et toutes, ou presque toutes, donnaient des L légèrement négatifs (système stable mais proche de l'instabilité), mais donnaient un système plus stable pour août 2007 que pour 2000, contrairement à la réalité. L'explication provient du fait que ces deux variables ne sont pas suffisantes: c'est l'ajout de la troisième variable, une mesure de la dette des particuliers et/ou des banques, qui amène l'instabilité en août 2007. Si l'idée de Stiglitz est juste, on doit retrouver l'instabilité réelle par le calcul.

On peut changer la valeur des dérivées, et donc de la stabilité, en tenant compte d'influences extérieures (régulation étatique ou toute autre influence), ce qui permet de voir comment mieux prévenir une instabilité. Ainsi, si la banque fédérale ne stabilisait pas les taux, on aurait $f_{rr}(r)$ qui serait nul. Un calcul facile montre que cela rend le système instable.

On peut aussi utiliser cette méthode pour quantifier des modèles basés sur la théorie des catastrophes : crash du marché (Zeeman 1974) et faillite d'une banque (Ho et Saunders 1980) par exemple.

IV.B. Données à la fois spatiales et temporelles: démographie

Dans ce cas, on connaît les valeurs dans suffisamment de lieux, pour au moins deux dates pas trop éloignées: démographie (natalité, taux d'alphabétisation, etc.) sociologie, économie sur plusieurs régions ou pays, etc.

a) Méthode

Je présente ici un modèle simple de démographie, pour l'interaction entre alphabétisation et natalité (voir Courbage et Todd 2007, par ex.). J'ai pris les données des Nations Unies pour l'alphabétisation («a») et la natalité («n»), de 1970 à 2005. J'ai pu calculer les dérivées par rapport au temps à partir de la formule pour 2 ou 3 valeurs, pour 1980, 1990, 1995 et 2000, pour 137 pays. J'ai ensuite calculé les termes non linéaires pour a et n: a^2, n^2, an pour le degré 2, auxquels on rajoute a^3, n^3, a^2n et an^2 pour les termes de degré 3. J'ai calculé l'état stationnaire pour chaque couple d'équations différentielles obtenues. Comme il y a plusieurs termes non linéaires, on a plusieurs solutions, certaines ne sont pas pertinentes (valeurs complexes ou négatives par exemple), mais on obtient presque toujours au moins une solution correspondant à la réalité. Ces solutions se classent en 2 catégories principales: une solution post-industrielle et une pré-industrielle.

b) Exemple concret

On donne ici un exemple de combinaison de termes non linéaires (donc un couple d'équations différentielles) correspondant à tous les termes de degré inférieur ou égal à deux:

$$dn/dt = -0.0469 + 6.57\ 10^{-4}a - 0.01n + 2.43\ 10^{-3}n^2 - 2.15\ 10^{-4}a.n$$
$$da/dt = -0.348 + 0.051a - 0.056n + 0.0136\ n^2 - 1.4\ 10^{-4}a.n$$

c) Etats stationnaires

On obtient deux types d'états stationnaires:

- Solution «moderne»: a = 101.048%, n = 0.993 enfants par femme
  C'est une des 2 solutions type trouvées pour différents couples d'équations différentielles. Notons qu'ici, on a une valeur supérieure à 100%, alors qu'on a en général des valeurs comprises entre 96% et 100%. On peut mettre cette petite erreur sur le compte des fluctuations dans les valeurs de chaque pays. La valeur pour la natalité peut sembler très basse, elle est un peu plus basse que la moyenne des valeurs obtenues. Est-ce là notre avenir, une humanité ne se reproduisant plus assez pour le renouvellement de la population, il est peut-être trop tôt pour le dire, mais la tendance de certains pays occidentaux montre que cela n'est pas impossible.
- Solution «ancienne»: a = 1.92%, n = 7.04 enfants par femme
  C'est l'autre solution type. Cette solution peut paraître bizarre, puisque le monde entier semble promis à l'alphabétisation massive et à la baisse de la natalité. Cependant, il y a peut-être une explication: certains peuples autochtones (qui comptent 300 millions d'habitants sur terre) pourraient peut-être garder leurs traditions, et pour cela garder une alphabétisation minimum et une grande natalité. Ce phénomène serait peu stable, comme nous le verrons.

d) Stabilité

J'ai calculé la stabilité des 2 couples de solutions trouvés pour le système précédent, de degré inférieur ou égal à 2:
la solution «moderne» est stable pour des valeurs très proches de l'état stationnaire, pour une natalité de 1.02. La simulation (voir plus loin) montre qu'elle est stable pendant au moins 200 ans, ce qui est suffisant pour que le système calculé ne soit plus valable, et qu'il faille changer de paradigme.

La solution «ancienne» est instable.
Pour avoir une meilleure précision sur la stabilité et sur les états stationnaires, il faudrait au moins une équation supplémentaire.

V. Conclusion

Cette méthode d'analyse de données permet de recréer un système (ses équations différentielles et son graphe de relations d'influences mutuelles) et son évolution future à partir de données chronologiques, dans les conditions réunies dans de nombreux cas d'applications en sciences humaines. Nous n'avons pas cité la possibilité de trouver les solutions analytiques, car les systèmes étudiés en sciences humaines sont en général trop complexes pour que ce soit faisable.
C'est pourquoi elle devrait pouvoir résoudre nombre d'applications dans des domaines divers, y compris en sciences politiques où les exemples de systèmes dynamiques possibles ne manquent pas: chute d'un empire, révolutions économiques ou politiques, changement de majorité, déclenchement et fin d'une guerre, passage à l'indépendance, transition vers la démocratie, changements d'opinion, etc.

De très nombreux problèmes peuvent être modélisés avec les équations différentielles en psychologie (le nombre important d'articles dédiés à cette possibilité le laissait supposer), mais aussi en sociologie, sans citer les sciences économiques.

Notons que la plupart de ces applications potentielles comportent une bifurcation, une transition brusque, phénomène où les systèmes dynamiques sont particulièrement bien adaptés, et que ces modèles sont en général (et en particulier en sciences humaines) dissipatifs. On pourrait peut être par contre calculer la production d'entropie générée par le système en redéfinissant une nouvelle production d'entropie. Cela à partir de la définition de la production d'entropie pour une réaction chimique, donnée par exemple dans la formule 14.39 de Glansdorff et Prigogine (Glansdorff 1971) pour le système de Lotka-Volterra.

Les sciences humaines ont adhéré à deux paradigmes au moins, le structuralisme et les statistiques, depuis le milieu du 20e siècle. Nous pensons, comme beaucoup d'autres, que les systèmes dynamiques seront le prochain paradigme, suite en partie aux travaux de Prigogine. Ils tranchent avec les statistiques (tout en leur étant complémentaires) en incluant l'évolution temporelle et les graphes d'influences mutuelles. Mais ils sont le prolongement du structuralisme: celui-ci avait insisté sur les relations entre les variables (graphe), mais ne disposait pas encore des outils pour modéliser ces graphes, leur évolution, et les relations entre graphes, la dynamique (équations différentielles), l'analyse et la simulation, cette dernière nécessitant l'informatique.

«Gause3»		
T	X	Y
0	157.7753	89.2168
1	39.56363	178.7237
2	19.19536	121.352
3	9.496784	59.44003
4	31.12625	10.4679
5	57.75003	18.47648
6	118.653	14.22626
7	109.6355	54.69794
8	40.9131	133.0811
9	19.63678	67.99092
10	11.5273	32.18565
11	20.89797	13.17947
13	70.74032	16.93782
15	131.2524	32.04692
16	131.5424	76.15084
17	50.33422	171.1061
18	17.70715	91.48695
19	23.21856	28.66692

Annexe: Données pour la compétition entre bactéries, selon Gause (1935).

Remerciements

L'auteur remercie Jean Jaquet (CMU Genève) pour ses suggestions, Gerhard Wanner et Martin Gander (département de mathématiques de l'université de Genève) pour leur aide informatique, ainsi que Robert Degli Agosti (SPHN et Université de Genève) pour ses précieux conseils pour la publication de cet article.

References

- **Abbiw-Jackson RM, Langford WF.** 1998 Gain-induced oscillations in blood pressure. J.Math.Biol. 37: 203-34.
- **Abraham RA, Shaw C.** 1983. Dynamics, the geometry of behaviour. Aerial press, Santa Cruz.
- **Ahalpara DP, Parikh JC.** 2006. Modeling time series data of real systems. Int. J. Mod. Physics, Arxiv nlin.CD/0607027.
- **Ahalpara DP, Parikh JC.** 2006a. Modeling time series of real systems using genetic programming. Arxiv nlin.CD/0607029.
- **Ambrose BW, Ancel E, Griffiths MD.** 1992. The fractal structure of real estate investment trust returns: the search for evidence of market segmentation and nonlinear dependency. J. of the American real estate and urban economics association, 20: 25-54.
- **An der Heiden U, Schwegler H, Tretter F.** 1998. Patterns of alcoholism: a mathematical model. Mathematical models and methods in applied science, 8: 521-41.
- **Babor TF, Dolinsky ZS, Meyer RE, Hesselbrock M, Hofman M, Tennen H.** 1992.Types of alcoholics: concurrent and predictive validity of some common classification schemes» B. J. Addict. 87: 1415-1431.
- **Barton S.** 1994. Chaos, self-organization, and psychology. Amer. Psychologist, 49: 5-14.
- **Bellman RE, Roth RS.** 1986. Methods in approximation. Reidel publishing, Dordrecht.
- **Boggs PT, Byrd RH, Schnabel RB.** 1987. A stable and efficient algorithm for nonlinear orthogonal distance regression. SIAM J. Sci. Stat. Comp. 8: 1052-1076.
- **Brown R, Rulkov NF, Tracy ER.** 1994. Modeling and synchronizing chaotic systems from time-series data. Physical Rev. 49: 3784-3800.
- **Bütz MR.** 1992. The fractal nature of the development of the self. Psychol. Reports, 71: 1043-63.
- **Byrne DG, Mazanov J, Gregson RA.** 2001. A cusp catastrophe analysis of changes to adolescent smoking behavior in response to smoking prevention programs. Nonlinear dynamics, psychology and life sciences, 5: 115-137.
- **Caetano MAL, Gherardib DFM, Yoneyama T.** 2008. Optimal resource management control for CO_2 emission and reduction of the greenhouse effect. Ecological modelling, 213: 119-126.
- **Cavalcanti S.** 2000. Arterial baroreflex influence on heart rate variability: a mathematical model-based analysis. Med. Biol. Eng. Comput. 38:189-97
- **Clair S.** 1998. A cusp catastrophe model for adolescent alcohol use: an empirical model. Nonlinear dynamics, psychology and life sciences, 2: 217-241.
- **Courbage Y, Todd E.** 2007. Le rendez-vous des civilisations, Seuil, Paris, p. 32.
- **Craigie NS.** 1980. Catastrophe theory and disorders of the family system. Internal report, IC/80/46 ICTP Trieste
- **Cremers J, Hübler A.** 1987. Construction of differential equations from experimental data. Z. Naturforsch. 42: 797-802.
- **Crutchfield JP, McNamara BS.** 1987. Equations of motion from a data series. Complex Syst. 1: 417-452.
- **Daeppen J-B, Burnand B, Schnyder C, Bonjour M, Pécoud A, Yersin B.** 1996. Validation of the addiction severity index in french-speaking alcoholics patients. J. Stud. on Alcohol, 57: 585-9.
- **DeCoster GP, Labys WC, Mitchell DW.** 1992. Evidence of chaos in commodity futures prices. The journal of futures markets, 12: 291-305.
- **Eisenhammer T, Hübler A, Packard N, Kelso JAS.** 1991. Modeling experimental time series with ordinary differential equations. Biol.Cyber. 65: 107-112.
- **Ellis D, McClure J.** 1992. In-patient treatment of alcohol problems-predicting and preventing relapse. Alcohol-alcohol 27: 449-456.
- **Erdi P, Grobler T, Barna G, Kaski K.** 1993. Dynamics of the olfactory bulb: bifurcations, learning, and memory. Biol. Cybern. 69: 57-66.
- **Eyal S, Akselrod S.** 2000. Bifurcation and a simple model of the cardiovascular system. Methods Inf. Med. 39: 118-21.
- **Frank M, Gencay R, Stengos T.** 1988. International chaos? European Economic Review, 32:1569-1584.
- **Frank M, Stengos T.** 1988. Chaotic dynamics in economic time-series. Journal of economic surveys, 2: 103-133.
- **Fussmann GF, Ellner SP, Shertzer KW, Hairston NG Jr.** 2000. Crossing the Hopf bifurcation in a live predator-prey system. Science, 290: 1358-1360.
- **Galanter M.** 1993. Network therapy for addiction: a model for office practice. Am. J. Psychiatry, 150: 28-36.
- **Gause GF.** 1935. Experimental demonstration of Volterra's periodic oscillations in the numbers of animals, Br. J. Exp. Biol. 12: 44-48.
- **Glansdorff P, Prigogine I.** 1971. Thermodynamic theory of structure, stability and fluctuation. John Wiley and Sons, London.
- **Gouesbet G, Maquet J.** 1992. Construction of phenomenological models from numerical scalar time series. Physica D, 58: 202-215.
- **Goldberger L, Rigney DR, Mietus J, Antman EM, Greenwald S.** 1988. Nonlinear dynamics in sudden cardiac death syndrome: heart rate oscillations and bifurcations. Experientia, 44: 983-987.
- **Goushcha AO, Kharkayanen VN, Scott GW, Holtzwarth AR.** 2000. Self-regulation phenomena in bacterial reaction centres. I. General theory. Biophys J. 79: 1237-1252.
- **Grandmont JM, Malgrange P.** 1986. Nonlinear economics dynamics: introduction. Journal of economic theory, 40: 3-12.
- **Gregson RAM.** 1999. Confidence judgements for discrimination in nonlinear psychophysics. Nonlinear dynamics, psychology and life sciences, 3: 31-47.
- **Greiner A.** 2007. The dynamic behaviour of an endogenous growth model with public capital and pollution. Studies in Nonlinear Dynamics & Econometrics, Vol. 11: No. 2, Article 4.
- **Guastello SJ, Johnson E, Rieke ML.** 1999. Nonlinear dynamics of motivation flows. Nonlinear dynamics, psychology and life sciences, 3: 259-273.

- **Guntern G.** 1982. Auto-organization in human systems. Behavioral Sc. 27: 323-337.
- **Haken H, Stadler M** (eds.) 1990. Synergetics of cognition. Springer, Heidelberg.
- **Ho T, Saunders A.** 1980. A catastrophe model of bank failure. Journal of finance, 35:1189-1207.
- **Hornero R, Alonso A, Jimeno N, Jimeno A, Lopez M.** 1999. Estimation of correlation dimension to evaluate cognitive performance in schizophrenic patients using a new computer technique. Nonlinear dynamics, psychology and life sciences, 3: 49-63.
- **Hsieh DA.** 1991. Chaos and nonlinear dynamics: application to financial markets. The journal of finance, 46: 1839-1877.
- **Jost C, Ellner SP.** 2000. Testing for predator dependence in predator-prey dynamics: a nonparametric approach. The Royal Society, electronic supplement, doi 10.1098/rspb.2000.1186.
- **Juhel J.** 1996. Les systèmes dynamiques non linéaires dans l'étude de la variabilité en psychologie: une introduction. Présentation au séminaire du LEACM, Université de Lyon Lumière, 19/2/1996.
- **Kahn D, Hobson JA.** 1993. Self-organization of dreaming. Dreaming, 3:151-178.
- **Lange R, Houran J.** 2000. Modeling Maher's attribution theory of delusion as a cusp catastrophe. Nonlinear dynamics, psychology and life sciences, 4: 235-254.
- **LeBaron B.** 1994. Chaos and nonlinear forecastability in economics and finance. Phil. Trans. Roy. Soc. Lond. A, 348: 397-404.
- **Lorenz EN.** 1963. Deterministic nonperiodic flow. Journal of Atmos Sci. 20: 130-141.
- **Meadows DH, Meadows DL, Randers J, Behrens WW.** 1972. Rapport su les limites de la croissance. In Halte à la croissance, Fayard, Paris.
- **Mohan D, Ray R, Sethi H.** 1992. Unidimensionality of alcohol dependence syndrome ? Drug Alcohol-depend. 31: 51-55.
- **Nichols NA.** 1993. Efficient ? Chaotic ? What's the new finance ? Harvard business review 3-4, p. 50.
- **Nottale L, Chaline J, Grou P.** 2000. Les arbres de l'évolution. Hachette, Paris.
- **Palaniyandi P, Laksgmanan M.** 2004. Estimation of system parameters and predicting the flow function from time series of continuous dynamical systems. Arxiv: nlin. CD/0406027.
- **Partridge T.** 2000. Temperament development modeled as a nonlinear complex adaptative system. Nonlinear dynamics, psychology and life sciences, 4: 339-357.
- **Paulus MP.** 2003. Chaos and schizophrenia: does the method fit the madness ? Biol. Psychiatry, 53 3-11.
- **Peifer M, Timmer J, Voss HU.** 2002. Nonparametric nonlinear system modelling. International symposium on nonlinear theory and its applications, Xi'an, PRC, october 7-11/2002.
- **Peters EE.** 1989. Fractal structure in the capital markets. Financial Analysts Journal, 45: 32-37.
- **Pezard L, Nandrino JL, Renault B, El Massioui F, Allilaire J-F, Müller J, Varela F Martinerie J.** 1996. Depression as a dynamical disease. Biol. Psychiatry, 39: 991-999.
- **Porcher E, Gatto M.** 2000. Quantifying the dynamics of prion infection: a bifurcation analysis of Laurent's model. J. Theor. Biol. 205: 283-96.
- **Priel B, Schreiber G.** 1994. On psychoanalysis and non-linear dynamics: the paradigm of bifurcation. Br. J. Med. Psychol. 67: 209-218; et discussion p. 219-225.
- **Rais G, Sollberger P.** 2007. Analyse structurelle de l'économie suisse. Office fédéral de la statistique, Neuchâtel.
- **Ramsay JO, Silverman BW.** 2002. Applied functional data analysis. Springer, New York.
- **Rand D.** 1977. Catastrophes and economic models. Colloques internationaux du CNRS, 259 p.77.
- **Reynolds DL, Chambers LW, DeVillaer MR.** 1992. Measuring alcohol abuse in the community: consumption, binge-drinking, and alcohol-related consequences («alcoholism»). Can. J. Public Health, 83: 441-447.
- **Robertson R, Combs A.** 1995. Chaos theory in psychology and the life sciences. Lawrence Erlbaum, Mahwah.
- **Rosenberg H, Melville J, Levell D, Hodge JE.** 1992. A ten years follow-up survey of acceptability of controlled drinking in Britain. J. Stud. Alcohol, 53: 441-446.
- **Rosser JB.** 1991. From catastrophe to chaos: a general theory of economic discontinuities. Kluwer, Boston.
- **Smith LB, Thelen E.** 1993. A dynamic systems approach to development, application. MIT press, Cambridge.
- **Stam CJ, Pijn JP, Suffczynski P, Lopes da Silva FH.** 1999. Dynamics of the human alpha rhythm: evidence for non linearity? Clin. Neurophysiol. 110: 801-813.
- **Sturis J, Brons M.** 1997. Local and global bifurcations at infinity in models of glycolytic oscillations. J. Math. Biol. 36:119-132.
- **Stutzer MJ.** 1980. Chaotic dynamics and bifurcation in a macro model" Journal of Economic Dynamics and Control, 2: 353-376.
- **Tschacher W, Dauwalder J-P.** 2003. The dynamical systems approach to cognition, concepts and empirical paradigms based on self-organization, embodiment, and coordination dynamics. Studies in nonlinear phenomena in life sciences, Vol.10, World Scientific, New Jersey.
- **Vaidya PG, Vaidya VG, Martin DG.** 1991. An application of the non-linear bifurcation theory to tumor growth modelling. Int. J. Biomed. Comput. 27: 27-46.
- **Van Gelder T.** 1998. The dynamical hypothesis in cognitive science. Behav. Brain Sci. 21: 615-665.
- **Varah JM.** 1996. Least square data fitting with implicit functions. BIT Numerical Mathematics, 36: 842-854.
- **Varson P L, Jalilvand A.** 1994. Evidence on deterministic chaos in TSE-300 monthly data. Canadian Journal of Administrative Sciences, 11:43-53.
- **Warren K, Hawkins RC, Sprott JC.** 2003. Substance abuse as a dynamical disease: evidence and clinical implications of nonlinearity in a time series of daily alcohol consumption. Addictive Behaviors, 28: 369-374.

- **Warren K, Sprott JC, Hawkins RC.** 2002. The spirit is willing: nonlinarity, bifurcations and mental control. Nonlinear dynamics, psychology and life sciences, 6: 55-70.
- **Zeeman EC.** 1973. Applications of catastrophe theory. Manifolds-Tokyo 1973, University of Tokyo Press, Tokyo.
- **Zeeman EC.** 1974. On the unstable behavior of stock exchanges. Journal of mathematical economics, 1: 39-49.
- **Zellner A, Israelevich G.** 2005. The mashallian macroeconomic model: a progress report. International journal of forecasting, 21: 627-645.
- **Zhang W-B.** 1999. Endogenous opening policy and nonlinear economic development-physical capital, human capital and economic structure. Nonlinear dynamics, psychology and life sciences, 3: 275-299.