

MODELISATION EN EPIDEMIOLOGIE : OBJECTIFS ET METHODES*

Marc A. Dubois¹

En souvenir de Frédérique Messud-Petit

RESUME : Plutôt qu'une introduction à la modélisation, j'ai cherché à donner un bref mode d'emploi de la modélisation (et des modélisateurs) pour l'épidémiologie. En particulier, j'insiste sur l'importance de décider vraiment ce que l'on cherche : compréhension ou prédiction ? Car les outils ne sont en général pas les mêmes.

Mots-clés : Modélisation, modèles scientifiques, épidémiologie, systèmes complexes.

SUMMARY : Rather than an introduction to modelling for epidemiology, we give here a short user's notice to modelling and modellers in epidemiology. We emphasize the necessity to know in advance what one really seeks : understanding or prediction ? Because the tools are usually not the same.

Keywords : Modelling, scientific models, epidemiology, complex systems.



I - INTRODUCTION

Les mots « modèle » et « modélisation » méritent une brève discussion. Partons de la définition du dictionnaire (Petit Larousse Grand Format 2002) :

Modèle n. m. (ital. *modello*)

1. Ce qui est donné pour servir de référence type. *Modèle d'écriture.*
2. Ce qui est donné, ou choisi, pour reproduit. *Copier un modèle.*
3. *Spécial.* Personne qui pose pour un être photographe, un peintre, un sculpteur, etc.
4. Personne ou objet qui représente idéalement une catégorie, un ordre, une qualité etc. *Un modèle de classicisme. Un modèle de loyauté, d'hypocrisie.*
5. Prototype d'un objet. - *Spécial.* Modelage en terre, en cire, en plâtre, etc., constituant le prototype d'une sculpture.
6. *Métall.* Pièce en bois, en métal, en matière plastique, en cire, ayant, au retrait près, la même forme que les pièces à mouler et destinée à la réalisation de moules de fonderie.
7. *Modèle réduit* : reproduction à petite échelle d'une machine, d'un véhicule, d'un navire, etc.
8. *Didact.* Structure formalisée utilisée pour rendre compte d'un ensemble de phénomènes qui possèdent entre eux certaines relations. *Modèle mathématique* : représentation mathématique d'un phénomène physique, économique, humain, etc., réalisée afin de pouvoir mieux étudier celui-ci.
9. Représentation schématique d'un processus, d'une démarche raisonnée. *Modèle linguistique.*

* Texte de la conférence présentée à la Journée AEEMA, 19 mai 2005

¹ Service de physique de l'état condensé, CEA Saclay –Orme des Merisiers, 91191 Gif sur Yvette cedex France (mail : mad@cea.fr)

Il est évident que dans le cadre d'un travail scientifique c'est la huitième acception de la définition qui nous intéresse. On remarquera pourtant que l'acception 7 recouvre les sphères armillaires (en anglais : « orrery ») qui sont en un sens un modèle scientifique descriptif. On s'étonnera aussi de la distinction de l'acception 9 qui nous paraît être confondue avec l'acception 8.

Revenons à l'acception 8 et plus spécifiquement au concept de modèle scientifique (terme que nous préférons de loin à modèle mathématique, trop restrictif). On divisera ce concept en trois sous-concepts suivant l'objectif poursuivi :

i) **les modèles descriptifs** – par exemple la représentation statique des coquillages

avec trois angles par D'Arcy Thompson (« On growth and form », 1917) ;

ii) **les modèles prédictifs**, dont le but est de décrire le devenir d'un système physique à partir de données sur son comportement passé ;

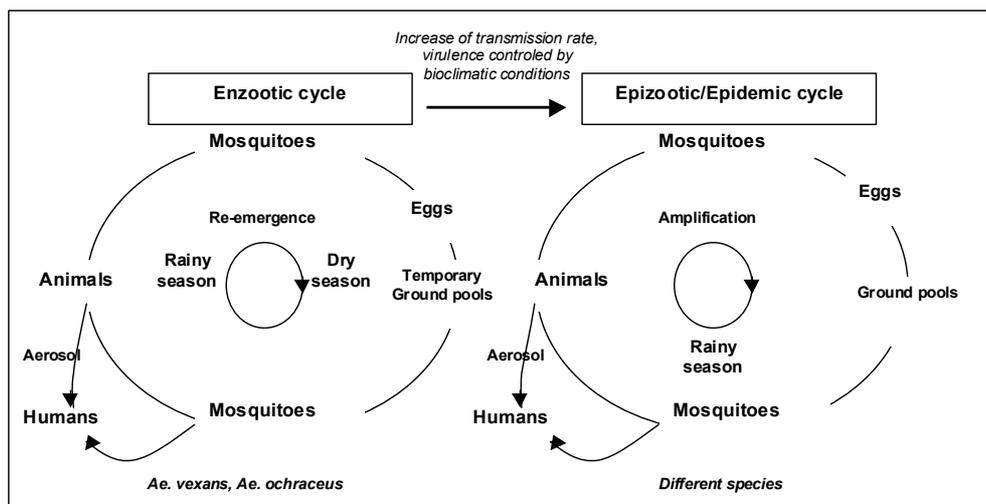
iii) **les modèles explicatifs** qui visent à simuler un système réel en rassemblant des hypothèses sur ses mécanismes élémentaires.

Les trois types de modèles évoqués ci-dessus sont utilisés en épidémiologie.

➤ Les schémas classiques de cycle d'infection virus - vecteur - hôte sont des modèles descriptifs (type i) (figure 1).

Figure 1
Schéma du cycle de la fièvre de la vallée du Rift en région sahélienne

[Fontenille *et al.*, 1998]



Ces schémas sont statiques, en ce qu'ils ne décrivent pas quantitativement l'évolution temporelle des phénomènes, mais ils soulignent déjà les processus qui permettront de construire un modèle mécaniste.

De même, une analyse statistique en composantes principales (ACP) est, prise isolément, un modèle descriptif qui peut suggérer des mécanismes sous-jacents.

➤ Si l'on combine une analyse en composantes principales (ACP) avec des méthodes de régression (linéaire ou non, simple ou multiple), on peut construire des modèles prédictifs opérationnels de type ii du

moins tant que le champ des paramètres est bien couvert par les données acquises. Les larges fourchettes des prédictions initiales des victimes de la vache folle montrent bien les limites de ces méthodes...

Les modèles à base de réseaux de neurones artificiels ont bien entendu une visée prédictive (type ii). Leur capacité de prédiction tient à l'intégration d'un nombre élevé d'observations. Ils sont en ce sens proches des méthodes statistiques traditionnelles, et si leurs performances sont en général supérieures à celles de ces dernières, on leur reproche souvent leur côté "boîte noire", en d'autres termes, on leur reproche de ne pas aider à comprendre. Ils sont donc essentiellement

prédictifs (type ii), même si de plus en plus on développe des méthodes pour « ouvrir la boîte » et en tirer des suggestions de mécanismes. De la même façon, et avec un certain avantage qualitatif sur les méthodes statistiques traditionnelles en raison de meilleures capacités à l'extrapolation, les réseaux de neurones artificiels (ANN) du type perceptron multicouches à rétro-propagation (en anglais : back propagating neuronal networks – BPN) ont une utilité prédictive avérée. Le cas des réseaux de neurones de type « carte de Kohonen », ou carte auto-organisée (en anglais : self-organizing map – SOM) qui cartographient et classifient des objets par affinités est un peu spécial : ils relèvent d'abord du type i (descriptifs), mais là aussi, la mise en évidence des similitudes peut suggérer des mécanismes, aidant ainsi à l'élaboration de modèles explicatifs (type iii).

On classera enfin dans les modèles prédictifs (type ii) les modèles matriciels élaborés à partir de bases de données en suivi temporel. Là encore, si la base de données est suffisamment complète, et si les paramètres extérieurs ne sortent pas trop du champ dans lequel elle a été acquise, leur fonction prédictive est utilisable. Mais on n'a toujours pas d'explication des phénomènes étudiés.

➤ La compréhension d'un système réel, en physique, en écologie ou en épidémiologie, ne peut passer que par un modèle explicatif (type iii) qui rassemble les mécanismes

élémentaires supposés par les spécialistes concernés. Le mot "comprendre" (*cum-prehendere* : prendre avec, prendre ensemble) décrit bien le fait que la modélisation mécaniste (ou explicative) est un carrefour de disciplines.

Je parlerai essentiellement ici des modèles explicatifs (type iii), et chercherai à montrer comment ne pas se perdre dans le vaste arsenal des méthodes et outils, et comment la nature des questions posées doit orienter le modélisateur et les spécialistes qui l'entourent dans la construction d'un modèle utile à la compréhension d'un problème épidémiologique, et à terme, à l'aide à la décision des politiques sanitaires à mettre en oeuvre.

L'activité de modélisation est un travail d'équipe, où le modélisateur, à l'écoute des différents spécialistes (virologistes, médecins, vétérinaires, entomologistes, éleveurs, etc.) doit aider à isoler les acteurs et mécanismes principaux, puis choisir dans sa « boîte à outils » les techniques mathématiques et numériques à mettre en oeuvre pour bâtir un modèle utile et pertinent. D'ores et déjà, je vous mets en garde contre les modélisateurs qui prétendraient bâtir un modèle sans « mettre les mains dans le cambouis », c'est-à-dire sans aller sur le terrain voir de leurs propres yeux le problème, et surtout sans passer beaucoup de temps à écouter les différents spécialistes impliqués.

II - MODELES MECANISTES : VOCABULAIRE ET BOITE A OUTILS

Nous parlerons d'abord de quelques termes souvent utilisés : le premier sera « système complexe » : se caractérise-t-il par le fait qu'il a des composantes multiples (comme une horloge), ou bien par le fait qu'il s'auto-organise (comme un tas de sable), ou bien encore parce qu'il a un comportement imprévisible (je n'ai pas d'exemple politiquement correct) ? Il n'y a pas de définition claire. Par contre, un concept bien défini est celui de « système dynamique ».

Un **système dynamique** comprend :

- un espace de phase **S** qui représente les états possibles du système ;
- un temps **t**, qui peut être continu ou discret ;

c. une loi d'évolution, c'est-à-dire une règle qui permet d'établir l'état du système au temps **t** à partir de la connaissance de tous les (ou partie des) états antérieurs (causalité).

S peut avoir une ou plusieurs variables, spatialisées ou non, et **t** peut ne pas être un temps au sens strict (par exemple, c'est une coordonnée longitudinale) [Hennequin *et al.*, 1992]. Par contre, la loi d'évolution DOIT être causale.

Voici une liste (non exhaustive) des concepts et méthodes employés par les modélisateurs de systèmes dynamiques.

hamiltonien, jacobien, intégrabilité, chaos, ergodicité, tores de KAM,

exposants de Lyapunov, séparatrices [Hennequin *et al.*, 1992] ;

bifurcations, transition de phase [Favier *et al.*, 2004] ;

systèmes dynamiques, systèmes complexes, systèmes conservatifs, systèmes dissipatifs, attracteurs (étranges) [Dubois et Yiou, 1999 ; Dubois *et al.*, 1992] ;

fractals, percolation [Vocka et Dubois, 2000] ;

modèle déterministe, stochastique, Monte Carlo, probabiliste, champ moyen, robustesse [Dubois *et al.*, 2002 ; Sabatier *et al.*, 2004 ; Favier *et al.*, 2004] ;

modèle 0-D, modèle spatialisé [Favier *et al.*, 2005a] ;

changement d'échelle, agrégation [Favier *et al.*, 2005b ; Durand *et al.*, 2004] ;

ODEs, PDEs, théorie des perturbations, FDEs, CA [Favier *et al.*, 2004], **réseaux couplés** [Dubois *et al.*, 1992], **small world** [Favier *et al.*, 2005a].

Ce petit catalogue à la Prévert ne doit pas faire peur : ce n'est qu'une **boîte à outils**, et c'est le modélisateur qui est censé savoir s'en servir. Si je n'ai choisi des illustrations que parmi mes propres publications, ce n'est pas par vanité,

mais pour insister sur la nécessité de maîtriser de nombreux outils afin de bien adapter le modèle à la question posée.

Des systèmes très simplement définis peuvent avoir des comportements complexes. Prenons par exemple la conjecture de Collatz :

« Pour tout entier positif N_0 , la suite définie par :

$$N_{n+1} = N_n / 2 \text{ si } N_n \text{ est pair, et } N_{n+1} = (3 \cdot N_n + 1) / 2 \text{ si } N_n \text{ est impair}$$

converge vers 1 ».

Cette conjecture n'est pas démontrée, mais elle n'a pas été infirmée numériquement, et est très probablement vraie. Le nombre d'itérations avant convergence est montré sur la figure 2 pour les 10 000 premiers entiers.

A contrario, des systèmes apparemment « compliqués » peuvent avoir un comportement très simple : dans un projet pédagogique du MIT, des « termites » marchent au hasard sur un pavage où sont semés au hasard des « copeaux » ; quand une « termite » marche sur un carré où il y a un copeau isolé, elle le « ramasse », et ne le re-dépose que si elle passe par un carré où il y a déjà au moins un copeau. La figure 3 montre que le système s'auto-organise en formant un tas.

Figure 2

Nombre d'itérations (ordonnées) pour les entiers de 1 à 10000 (abscisses)

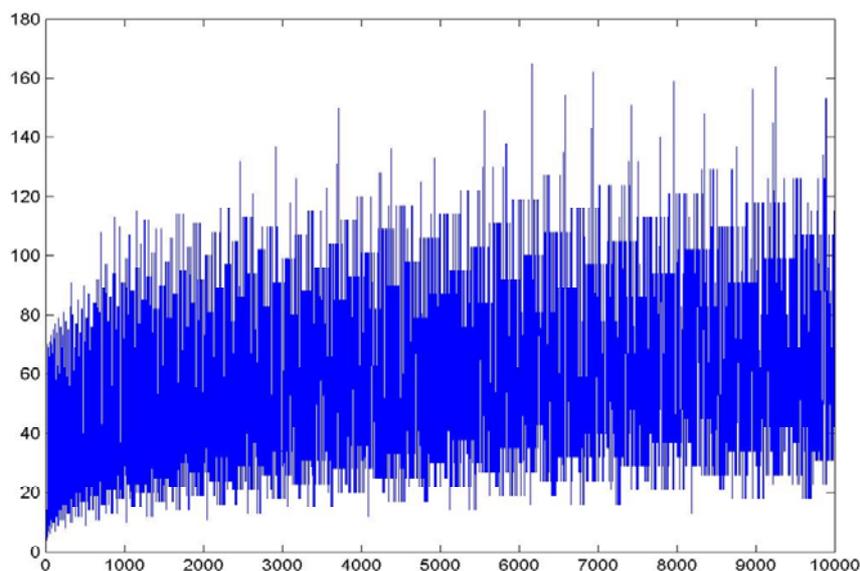
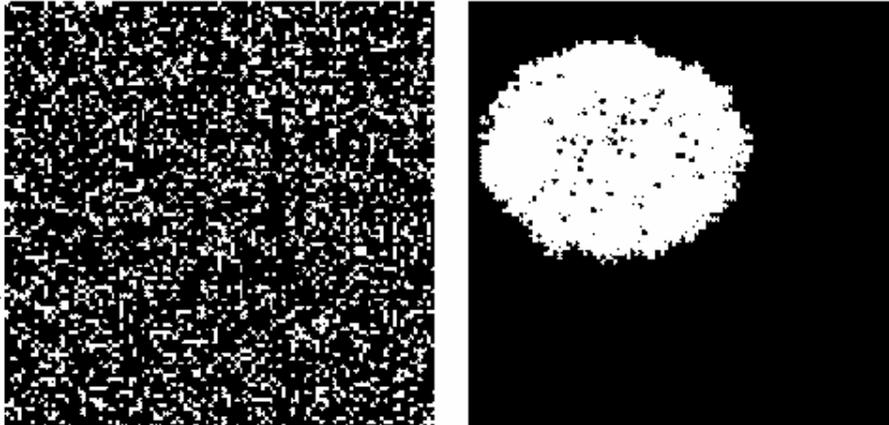


Figure 3

Auto-organisation d'un tas de « copeaux » par des « termites » (MIT)



La mécanique classique est décrite par des équations différentielles dans un espace continu et avec des vitesses et un temps continus : systèmes dynamiques totalement continus. A l'opposé, on peut définir un système avec des variables d'état entières et un temps discret : les deux exemples précédents sont des systèmes dynamiques totalement discrets (FDDS). Et chacun connaît les « automates cellulaires » ou « C.A. », dans lesquels chaque maille M d'un réseau à une ou plusieurs dimensions prend parmi un nombre fini d'états (par exemple 0 ou 1) à tout instant t (le temps est lui aussi discret, c'est-à-dire qu'il progresse par sauts d'égale durée) un état déterminé par l'état des mailles au voisinage de M au temps $t-1$. On pourrait alors se demander : « **qu'est-ce qui est le mieux ?** ». C'est l'archétype d'une **question stupide** : il n'y a pas de réponse unique, la méthode doit être adaptée au problème posé. Pour choisir une méthode de modélisation mécaniste d'un phénomène réel que l'on veut comprendre, il

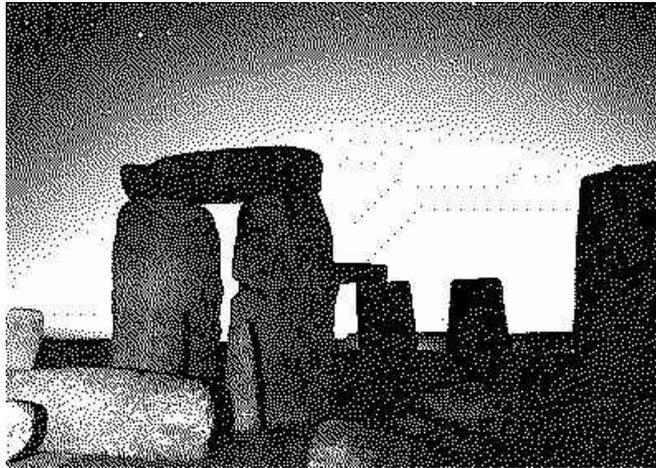
faut d'abord bien définir la(les) question(s) que l'on se pose ; il faut ensuite faire la liste des données accessibles à la connaissance et à la mesure (si un modélisateur vous demande pour construire son modèle de lui donner les valeurs de milliers de paramètres inaccessibles, il est en général urgent de s'en séparer). Puis, il faut rassembler les observations suggérant des mécanismes du système étudié. Alors, et alors seulement, il est temps d'opter pour un cadre de formalisation mathématique. La **démarche de modélisation** est une démarche **d'équipe**, qui se fait au **carrefour de disciplines** différentes. Une règle essentielle à garder à l'esprit est qu'un modèle mécaniste doit être le plus simple possible : « small is beautiful » (que l'on pourrait traduire par « c'est dans les petits pots qu'on met les bons onguents »). Il n'est utile de compliquer un modèle que quand la version minimaliste est impuissante à expliquer les phénomènes observés. Voyons quelques exemples hors du champ de l'épidémiologie.

III - QUELQUES EXEMPLES DANS D'AUTRES DISCIPLINES

Avant de donner des exemples de modélisation en épidémiologie (science du vivant), je voudrais détailler une histoire de démarche scientifique aboutissant à des théories (modèles explicatifs) : l'histoire abrégée d'une science « dure » ou « exacte », à savoir l'astronomie. Je parlerai ensuite un peu de deux sciences de la nature, à savoir la météorologie et la climatologie, qui, si on ne peut les qualifier de sciences « molles », sont intermédiaires entre l'astronomie et les sciences du vivant.

L'**astronomie** fut d'abord astrologie avant de devenir mécanique céleste : les premiers écrits babyloniens relatent les mouvements des planètes et les éclipses en même temps que les relevés des moissons ou les observations météorologiques. Les bâtisseurs des mégalithes alignaient ceux-ci avec les positions du soleil au solstice, et on y peut trouver d'autres usages astronomiques (figure 4).

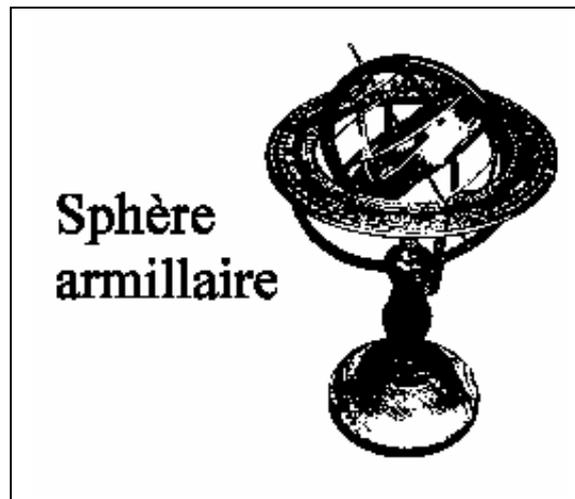
Figure 4

Les mégalithes de Stonehenge : un « observatoire » préhistorique

Déjà du temps des grecs se posait le problème de savoir quel « modèle » était le bon, du géocentrisme (défendu par exemple par Ptolémée) ou de l'héliocentrisme (proposé d'abord par Philolaos, puis défendu par Aristarque et d'autres). Et des planétaires,

sphères armillaires (figure 5) et autres instruments furent construits pendant des siècles, « modèles » (de type i, ou descriptifs) dont la finalité était de visualiser les modèles abstraits du mouvement des sphères.

Figure 5

Sphère armillaire

Ce sont les observations fines du mouvement des planètes (mouvements rétrogrades, variations de luminosité) qui mirent en difficulté le géocentrisme, obligeant Ptolémée et ses successeurs à le compliquer jusqu'à l'absurde (théorie des épicycles). La Renaissance bouleversa les certitudes aristotéliennes, et Copernic, bientôt relayé par Giordano Bruno

(qui paya fort chèrement son prosélytisme), remit à l'honneur l'héliocentrisme, au grand dam de l'église. L'observation par Galilée des phases de Vénus et la découverte et l'étude du mouvement des satellites de Jupiter ne suffirent pas à bouleverser l'orthodoxie. Puis les observations de Tycho Brahé furent mises en équations par son élève Képler, portant un

coup fatal à l'idée de corps célestes en orbites circulaires parfaites. Il faut bien voir que l'enjeu était en fait la distinction entre un monde céleste et l'univers sublunaire où se trouvent la Terre et ses habitants. Tant que cette distinction existait, les docteurs en théologie pouvaient recourir à des anges pour guider les planètes : l'astronomie relevait plus de la métaphysique que de la physique. C'est d'ailleurs la même terreur de voir disparaître l'inexplicable qui a justifié les réticences de l'Eglise catholique envers la théorie de l'évolution (et si Teilhard de Chardin ne fut pas brûlé, c'est parce que ce n'était déjà plus politiquement correct). Notons au passage que la démarche inverse, qui cherche à établir la vacuité de la métaphysique à partir des avancées de la physique (et de la science en

général) est aussi vaine et puérite : la démarche de Monod dans « Le hasard et la nécessité » rappelle souvent la phrase de Titov, le second cosmonaute, qui - en descendant de sa capsule - déclara que Dieu n'existait pas car il ne l'avait pas vu dans le ciel...

La révolution Newtonienne mérite ce nom car elle permet d'apporter un sens aux équations de Képler : un mécanisme simple (du moins à première vue), l'attraction universelle des corps massifs, suffisait à construire une théorie des mouvements du système solaire.

Mais si ce système d'équations a le mérite de pouvoir être formulé de façon compacte :

The general n -body problem can be stated in the following way.

For each body i , with mass m_i , let $c_i(t)$ be its [trajectory](#) in 3-dimensional space, where the parameter t is interpreted as time. Then the acceleration $c_i''(t)$ of each mass m_i satisfies by the law of gravity:

$$c_i''(t) = \gamma \sum_{1 \leq j \leq n, i \neq j} m_j \frac{c_j(t) - c_i(t)}{\|c_j(t) - c_i(t)\|^3}$$

The solutions of this system of differential equations give the positions as a function of time.

The force on each mass m_i is

$$F_i = c_i''(t)m_i$$

il a un grand défaut, il n'est pas intégrable en général, au sens qu'on ne peut donner une expression analytique de la position et de la vitesse d'un corps en un temps t quelconque. Les astronomes ont alors bâti des modèles approximatifs (théorie des perturbations) qui permettaient, même avant l'ère de l'ordinateur, de calculer avec précision les trajectoires des corps célestes. Les écarts entre prédictions et observations ont conduit à la découverte de la planète Neptune. D'autres écarts, plus fins (avance du périhélie de Mercure), ont conduit à revoir la théorie Newtonienne de la gravitation, et ont été expliqués par la relativité générale. Maintenant, les ordinateurs permettent de calculer les trajectoires avec une grande précision, au point de pouvoir heurter volontairement une comète avec un engin fabriqué par l'homme.

Mais le problème à n -corps n'est toujours pas intégrable. Les travaux de Hénon et Heiles ont montré qu'il pouvait y avoir des comportements chaotiques dans les trajectoires d'un amas, même en régime « hamiltonien », c'est-à-dire « conservatif », ou encore sans dissipation (en première approximation, la mécanique céleste est hamiltonienne ; à un niveau plus fin, il faut

tenir compte des collisions et des phénomènes de marée : c'est un bon exemple de la démarche de modélisation qui doit d'abord essayer les approches les plus simples et négliger ce qui n'est pas essentiel pour répondre à la question posée). Notons que Poincaré, au début du siècle dernier, avait déjà entrevu le concept de « chaos déterministe ».

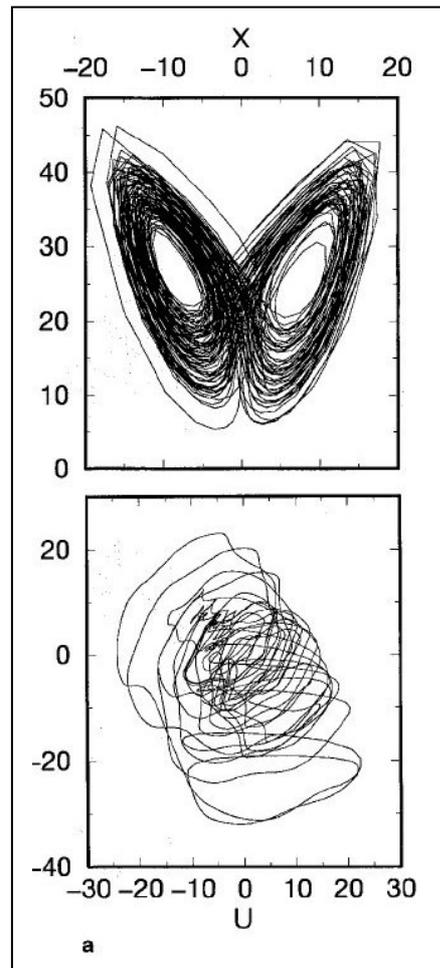
Dans un domaine différent de la connaissance scientifique (encore que les astres - et singulièrement le soleil et la lune - y jouent un rôle essentiel), des progrès immenses ont été réalisés au cours des dernières décennies : je veux parler ici de la **météorologie** et de la **climatologie**. On entre là dans le domaine des « systèmes dissipatifs » (en jargon, le « jacobien » est inférieur à l'unité), car la viscosité des fluides (air et eau) transforme les mouvements en chaleur. Une caractéristique des systèmes dissipatifs est que, en partant de conditions initiales réparties dans une large fraction de l'espace des variables du système, celui-ci évolue vers un sous-ensemble de cet espace des phases, que l'on appelle un « attracteur » : l'attracteur peut être un point (on a alors évolution vers un équilibre) ou une courbe fermée (on a alors un système périodique) ou un tore (système multi-périodique). Il peut aussi s'agir d'un objet plus complexe, de dimension non-entière (fractale) ; on parle alors d'« attracteur étrange » : nous en voyons deux

exemples sur la figure 6 ; celui du haut est le célèbre attracteur de Lorentz, qui de par sa forme a donné naissance à la légende de l' « effet papillon » (un papillon battant des ailes en Amazonie pourrait déclencher une tornade à Tokyo...), alors que celui du bas est l'attracteur de Rössler. Les deux systèmes dynamiques correspondant à ces attracteurs sont des « jouets » utiles en sciences du climat (en fait, sur ces images, ils sont un peu déformés, car on les a « couplés », afin de faire un modèle jouet d'interaction atmosphère-océan).

La **forme** de l'attracteur caractérise les propriétés du système, mais pour connaître à un moment donné la valeur des variables, il faut connaître la **position** du système sur l'attracteur. Et si l'attracteur est un attracteur étrange (c'est-à-dire si le système est chaotique), des conditions initiales très voisines vont donner des trajectoires qui s'écartent exponentiellement. C'est la raison pour laquelle la météorologie, même avec un réseau dense de capteurs très performants et des calculateurs surpuissants, ne pourra jamais faire des prédictions précises et exactes au delà de quelques semaines. La climatologie quant à elle a l'ambition de décrire l'évolution du climat non pas à quelques semaines, mais à quelques centaines d'années, voire des millénaires : l'apparente contradiction entre les limites des prévisions météorologiques et climatologiques se comprend aisément si l'on voit que pour les besoins de la climatologie il suffit de calculer la forme de l'attracteur et son évolution en fonction des paramètres de forçage, mais que, contrairement à la météorologie, on n'a pas besoin de connaître la position du système sur l'attracteur au temps t . En effet, on ne cherche pas à savoir si le 11 avril 2962 il fera ou non beau à Toulouse, mais plus simplement si les températures et la pluviométrie seront plus ou moins élevées, et si les saisons seront plus ou moins marquées.

On voit là combien il est important de bien cerner la ou les questions que l'on se pose : en épidémiologie, il sera souvent possible de prédire les zones à risque pour telle ou telle maladie, mais il sera la plupart du temps impossible de prédire à l'avance la date et le lieu de nouveaux foyers.

Figure 6
Attracteurs de Lorentz et de Rössler



IV - TROIS EXEMPLES DE MODELES EN EPIDEMIOLOGIE

1. LA TREMBLANTE DU MOUTON : CHANGEMENT D'ECHELLE

La tremblante du mouton (« scrapie » en anglais) est connue depuis plus de deux siècles. Elle pose des problèmes économiques aux éleveurs, mais n'était pas considérée comme pouvant être dangereuse pour la santé humaine. Depuis l'apparition des encéphalopathies humaines dues à l'absorption de viande bovine contaminée par le prion de l'ESB (encéphalopathie spongiforme bovine), la tremblante fait l'objet d'un regain d'attention, tant comme modèle bien documenté de maladie à prions, qu'en raison d'une suspicion d'existence d'ovins porteurs du prion ESB, avec les risques de passage à l'homme que cela suggère. Le modèle que nous allons évoquer part du risque au niveau de l'animal individuel, décrit la dynamique de la maladie au sein d'un élevage, et décrit enfin son évolution à l'échelle régionale. Il fait donc appel à une technique de **changement d'échelle**.

Au niveau individuel, et malgré l'évidence expérimentale d'une origine infectieuse de la tremblante du mouton, l'existence d'une variabilité importante de la forme naturelle de la maladie au sein des élevages suggère une influence des facteurs génétiques. Il a été établi que le polymorphisme du gène de la protéine prion (PrP) pouvait être corrélé à l'incidence de la tremblante ainsi qu'à la durée de survie des individus infectés. Cinq formes alléliques déterminantes ont ainsi été décrites à partir des variations des séquences d'acides aminés aux codons 136, 154 et 171. L'étude de l'appariement des cinq variants connus sur la sensibilité à la tremblante montre que la sensibilité d'un individu donné à la tremblante est le produit des risques élémentaires associés à chacun des deux allèles hérités [Dubois *et al.*, 2002].

On décrit ensuite la dynamique d'un troupeau et de la maladie au sein de ce troupeau, à l'aide d'un modèle individu-centré (chaque brebis est suivie au cours du temps, avec une probabilité de

changement d'état à chaque pas de temps : le modèle est donc mécaniste, mais non strictement déterministe) [Sabatier *et al.*, 2004]. Le modèle reproduit bien les différents scénarios possibles : cas sporadique, épizootie, enzootie ; il permet aussi d'étudier les effets d'une conduite spécifique du troupeau, en particulier dans le choix des reproducteurs (lutte naturelle ou insémination artificielle par des béliers de souche résistante à la tremblante).

Il serait impossible de suivre individuellement toutes les brebis de tous les troupeaux d'une région entière comme les Pyrénées-Atlantiques, pour lesquels la tremblante a été modélisée [Durand *et al.*, 2004]. On a donc utilisé une **agrégation de variables** pour rendre possible ce changement d'échelle. En pratique, on a établi un système d'équations qui décrivent l'évolution de variables moyennées sur un troupeau. La comparaison avec les résultats du modèle individu centré étant satisfaisante, on peut alors établir un modèle régional, tenant compte de la topographie (vallées parallèles) des échanges (commerce d'agnelles, choix des reproducteurs) et des pratiques de transhumance. La figure 7 montre l'évolution naturelle de la tremblante au cours des derniers siècles dans les Pyrénées. Il est intéressant de remarquer qu'on observe une auto-sélection des allèles de résistance qui aboutit à long terme à une raréfaction des cas.

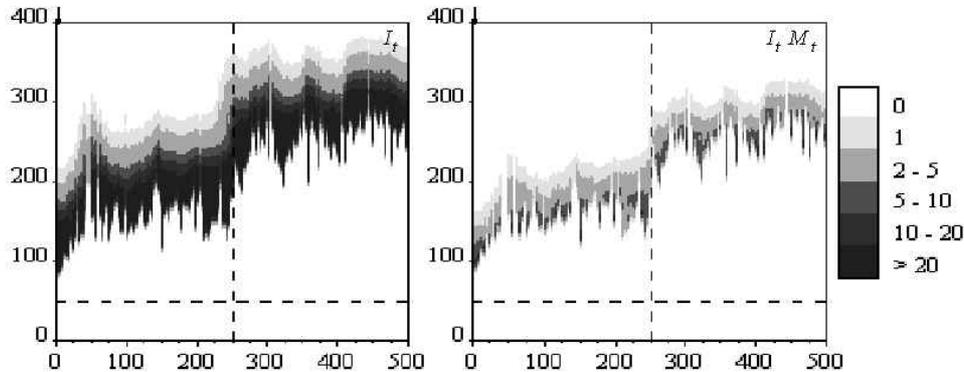
2. LA FIEVRE DE LA VALLEE DU RIFT : SPATIALISATION

La fièvre de la vallée du Rift est une arbovirose émergente en Afrique de l'Ouest qui est décrite depuis le début du XX^e siècle en Afrique de l'Est et du Sud ; elle se traduit par des avortements et de la mortalité chez les ruminants, et atteint occasionnellement l'homme (impasse épidémiologique). Après quelques épisodes épidémiques intenses en Mauritanie et au Sénégal, elle semble s'être établie sous forme enzootique au Sahel, alors qu'il ne semble pas y avoir de réservoir sauvage du virus. L'hypothèse d'une transmission verticale chez le vecteur (survie d'oeufs contaminés d'*Aedes vexans* sur une période d'une à plusieurs années) a alors été envisagée.

Figure 7

Evolution spatio-temporelle de la tremblante dans les Pyrénées

A gauche : taux de brebis contaminées, à droite, taux de mortalité.
En abscisses : répartition des élevages en hiver (500 élevages dans deux vallées contiguës),
en ordonnées : temps (années). L'échelle de grisé donne les pourcentages.



Mais si l'on construit un modèle de la maladie dans un village (troupeau plus point d'eau type mare ou étang), on s'aperçoit qu'il faudrait que le climat soit quasi périodique (une année humide suivie de deux années sèches, etc.) pour qu'une enzootie soit possible (il faut en effet que le troupeau ait à nouveau des jeunes non immunisés). On a alors étudié l'effet de la spatialisation sur le maintien de la maladie à l'échelle régionale avec un modèle généralisé d'« automate cellulaire » : dans un premier temps, on a obtenu des résultats identiques au modèle non spatialisé, avec une diffusion de la maladie vers les troupeaux adjacents au premier cas, puis extinction plus ou moins rapide suivant les séquences climatiques. On a ensuite incorporé dans le modèle le fait qu'en saison sèche une partie des troupeaux transhume vers le sud, plus humide. Ce type de modèle avec des connexions distantes se rapproche des modèles « small worlds » en physique statistique. Le résultat est concluant : alors que le modèle non spatialisé n'exhibe une enzooticité que pour des séquences climatiques très peu réalistes, cette enzooticité est assurée à l'échelle régionale pour une large gamme de séries climatiques dans le modèle spatialisé (figure 8). Localement (au niveau d'un village), on observe des épidémies

espacées de quelques années à quelques dizaines d'années. La prise en compte de l'espace est essentielle pour la compréhension du phénomène. En général on retiendra que les modèles non-spatialisés (ou modèles 0-D) sont peu réalistes et donnent souvent des résultats erronés.

3. LA DENGUE : CHANGEMENT D'ECHELLE, SPATIALISATION, ROBUSTESSE

Nous évoquons enfin une autre démarche de modélisation en épidémiologie (humaine cette fois et non vétérinaire), où la spatialisation et les changements d'échelles sont aussi importants. Ce sera l'occasion de parler d'une autre caractéristique des modèles : leur **robustesse**.

La dengue est elle aussi une arbovirose et, à part quelques souches qui affectent des singes en Afrique et dans le sud ouest asiatique, c'est une maladie humaine. Connue depuis des siècles, elle est en pleine expansion, en particulier en raison de la banalisation du transport aérien. Endémique en Thaïlande, elle resurgit au Pérou (cas récents à Lima) et dans l'ensemble de l'Amérique du Sud.

Le cas de l'île de Pâques est intéressant : la dengue y était inconnue jusqu'en 2002, année où un touriste étranger en phase infectieuse la visita. Il y eut alors une épidémie qui fut étudiée attentivement.

Figure 8

Evolution temporelle de la fièvre de la vallée du Rift pour un modèle spatialisé avec transhumance des troupeaux

A gauche, comportement moyenné à l'échelle régionale, à droite, évolution pour un village. De haut en bas, niveau de la (des) mare(s), stock d'oeufs infectés d'*Aedes vexans*, séroprévalence pour le bétail, taux d'avortements, pourcentage du bétail infecté et enfin occurrence des épisodes épidémiologiques. Années en abscisses de chaque figure. On voit sur la partie gauche que l'enzooticité au niveau régional est atteinte en quelques années.

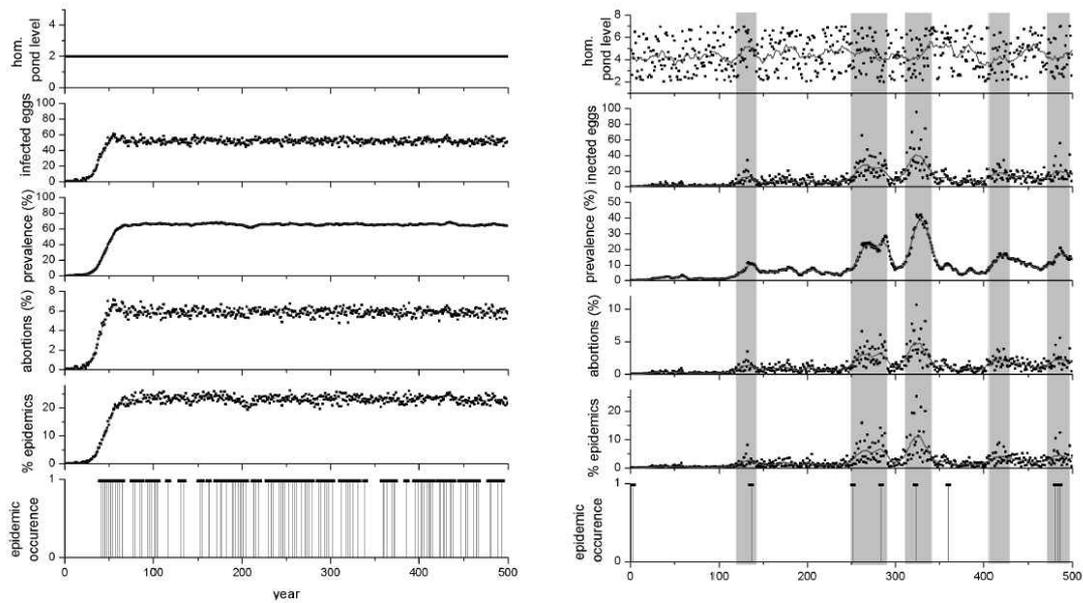
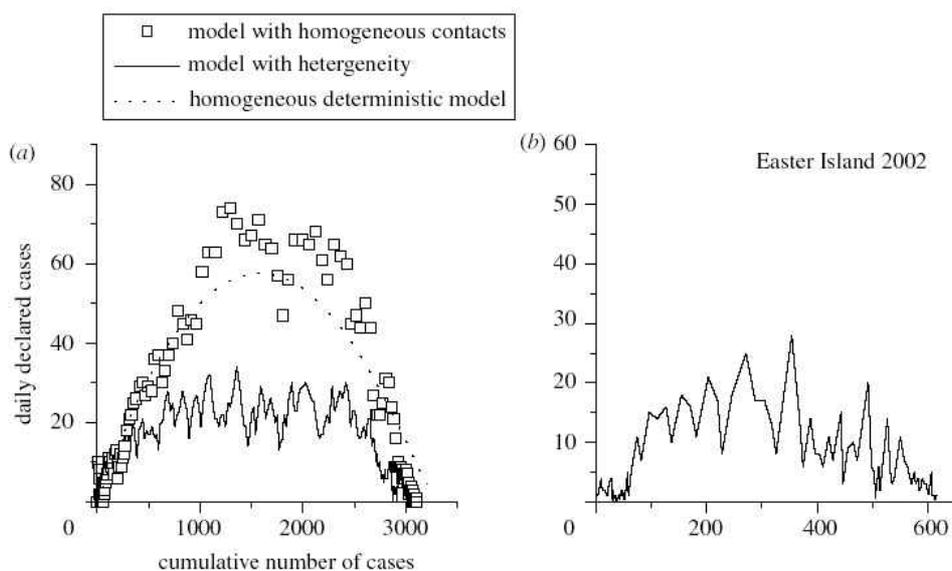


Figure 9

Dengue : modèles et épidémie de l'île de Pâques (2002)

A gauche : en pointillé, modèle 0-D analytique ; carrés : modèle individu centré non spatialisé ; en trait continu, modèle individu centré spatialisé. A droite : dynamique de l'épidémie de l'île de Pâques en 2002.



Les modèles analytiques 0-D classiques donnent les mêmes résultats que les modèles individu-centrés non spatialisés, avec une évolution du nombre de cas qui suit une quasi parabole (figure 9). Dans le cas réel, on observe une montée rapide, suivie d'une saturation puis d'une décroissance. Seul un modèle individu centré spatialisé, qui prend en compte l'hétérogénéité de l'habitat (alors même que l'île de Pâques semble un point dans l'océan !) est capable de reproduire la dynamique observée.

S'il faut considérer en détail la géographie de chaque village, on conçoit que la modélisation de la dengue à l'échelle d'un pays, voire d'un continent, fera nécessairement appel à des techniques de changement d'échelle. On doit donc élaborer un modèle analytique qui porte sur des variables agrégées, et qui inclut les effets de l'hétérogénéité spatiale pour chaque agglomération. Alors seulement on pourra passer à l'échelle régionale.

A propos de la tremblante, nous avons dit qu'un modèle mécaniste pouvait ne pas être strictement déterministe, par exemple si le passage d'un état à un autre pour une cellule ou un élément du modèle est géré par un tirage au sort. Il est évident qu'aucun modèle ne pouvait prédire qu'en 2002 un touriste infecté irait visiter l'île de Pâques. On peut par contre estimer une probabilité pour qu'un voyageur infecté arrive à un moment donné en un endroit donné. Il sera impossible de prédire que tel village d'Amérique du Sud subira une épidémie de

dengue à tel moment, mais il sera possible d'en estimer le risque. Le modèle peut être **robuste** à une échelle suffisamment étendue, alors qu'il ne l'est pas au niveau local. Et à une échelle nationale ou continentale, il sera possible de définir les zones à risques, et d'utiliser ces prédictions pour conduire les politiques sanitaires. On voit ici la similitude avec le cas de la météorologie par rapport avec la climatologie.

L'importance des systèmes d'information géographiques (SIG) : passage à une validité opérationnelle

Le rôle des modèles mécanistes pour la compréhension est indéniable, mais il ne faut pas perdre de vue qu'on doit attendre d'eux qu'ils servent aussi à la prise de décision et à la gestion des risques épidémiques ; contrairement aux modèles prédictifs non mécanistes, ils ont la possibilité d'explorer des domaines de paramètres non encore observés. Pour remplir ces objectifs, il faut que leur valeur prédictive soit bonne, et elle ne peut l'être (comme dans le cas de la météorologie) que si les données que l'on entre dans le modèle sont précises et bien spatialisées. La tendance actuelle est d'essayer de coupler des SIG aux modèles, en dégradant la grille de spatialisation des SIG pour l'adapter à celle des modèles, ou en adaptant le modèle au SIG. Cette étape est assez laborieuse, et ne peut se faire qu'avec un effort et des modélisateurs (un théoricien a tendance à considérer cette étape comme indigne de ses compétences) et des gens de terrain (qui doivent faire accepter d'« entrer » dans le modèle pour être efficaces). Mais cela en vaut la peine...

V - CONCLUSION : DU BON USAGE DE LA MODELISATION ET DES MODELISATEURS

Ce petit exposé n'a pas vocation à être un texte de référence : il est incomplet, provocateur par moments, certainement partial et contestable sur de nombreux points, mais il est pragmatique et son but est d'ôter au lecteur une éventuelle timidité ou méfiance devant les modèles et les modélisateurs. Encore une fois, un bon modèle est un travail d'équipe, et ce sont les spécialistes de terrain qui élaborent les

questions et suggèrent les mécanismes : le modélisateur doit en permanence être à l'écoute, au carrefour de la discussion, et s'il s'isole hautainement dans sa tour d'ivoire, c'est qu'il est inutile. Enfin, un modèle n'est jamais fini : même s'il a fait progresser la compréhension d'un phénomène, il est rare qu'il ne fasse pas aussi apparaître des failles ou de nouvelles questions. Et s'il n'est sans doute pas possible de tout comprendre, il est souhaitable d'essayer...

BIBLIOGRAPHIE

Quelques ouvrages à lire pour en savoir plus

- Anderson R.M. and May R.M. - Infectious Diseases of Humans- Dynamics and Control. Oxford Science Publications, 1991.
- Anderson R.M. and May R.M. - Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, 1991.
- Boccara N. - Modelling Complex Systems. Springer-Verlag, 2004.
- D'Arcy Thompson - On growth and form. Réédité par Cambridge University Press, 1917.
- Franc A., Gourlet-Fleury S., Picard N. - Une introduction à la modélisation des forêts hétérogènes. ENGREF, 2000.
- Lek S., Guégan J.F., editors - Artificial neural networks : Application to Ecology and evolution. Springer-Verlag, 2000.
- Murray J.D. - Mathematical Biology. Springer-Verlag, 1989.

Les références des articles cités en illustrations du catalogue à la Prévert, ou en exemple dans le texte

- Dubois M.A., Beaufumé P., Fromont B. - Model of anomalous transport with backreacting localized perturbations. *Nature*, 1992, **358**, 133-136.
- Dubois M.A., Sabatier P., Durand B. *et al.* - Multiplicative genetic effects in scrapie disease susceptibility. *Comptes rendus biologie*, 2002, **325** (5), 565-570.
- Dubois M.A., Yiou P. - Testing asynchronous coupling on simple « ocean-atmosphere » dynamic systems. *Climate dynamics*, 1999, **15** (1), 1-7.

- Durand B., Dubois M.A., Sabatier P. *et al.* - Multiscale modelling of scrapie epidemiology - II. Geographical level: hierarchical transfer of the herd model to the regional disease spread. *Ecological modelling*, 2004, **179** (4), 515-531.
- Favier C., Chalvet-Monfray K., Sabatier P., Lancelot R., Fontenille D., Dubois M.A. - Rift Valley Fever in West Africa: the role of space in enzooticity, submitted to *Proc. R. Soc. B*, 2005a.
- Favier C., Chave J., Fabing A., Schwartz D., Dubois M.A. - Modelling forest-savanna mosaic dynamics in man-influenced environments: effects of fire, climate and soil heterogeneity. *Ecological modelling*, 2004, **171** (1-2), 85-102.
- Favier C., Schmit D., Muller-Graf C., Cazelles B., Degallier N., Mondet B., Dubois M.A. - Influence of spatial heterogeneity on an emerging infectious disease: the case of dengue epidemics, in press, *Proc. R. Soc. B*, 2005b.
- Fontenille D., Traore-Lamizana M., Diallo M., Thonnon J., Digoutte J.P. and Zeller H.G. - New vectors of Rift Valley fever in West Africa. *Emerging Infectious Diseases*, 1998, **4**(2), 289-93.
- Hennequin P., Dubois M.A., Nakache R. - Anharmonicity in resonant perturbations and its effect on stochasticity extension. *Phys. Lett.*, 1992, A **164**, 259-262.
- Sabatier P., Durand B., Dubois M.A. *et al.* - Multiscale modelling of scrapie epidemiology - I. Herd level: a discrete model of disease transmission in a sheep flock. *Ecological modelling*, 2004, **180** (2-3), 233-252.
- Vocka R., Dubois M.A. - Pore network as a model of porous media: Comparison between nonhierarchical and hierarchical organizations of pores. *Physical review*, 2000, E **62** (4): 5216-5224.



Remerciements

Je tiens à remercier les organisateurs des « Journées Scientifiques de l'AEEMA : Intérêt et limites de la modélisation en épidémiologie pour les décisions de santé », ainsi que les participants à ces journées : le nombre des questions et la qualité des discussions ont montré que cette réunion était utile et attendue. Quant aux autres personnes que je voudrais remercier, trop nombreux pour être tous nommés, on en trouvera un échantillon dans les coauteurs des papiers que je cite en illustration de mes propos : apportant leur enthousiasme depuis divers champs de la science, ils me permettent de vivre cette aventure passionnante qu'est le travail de modélisation - et de compréhension - du monde où nous vivons. Et peut-on aimer sans comprendre - au moins un peu ?