

Sujet de thèse IFI-MSI / UPMC-LIP6

ED UPMC 130 EDITE

Titre de la thèse:

# Modélisation, Simulation multi-niveau pour l'optimisation de politiques de vaccination

**Direction de thèse:** Pr. Yann Chevalerey (PR, Université Paris 13)  
Dr. Jean-Daniel Zucker (DR1, UMI UMMISCO, IFI-MSI, IRD)  
**Co-encadrement:** Dr. Marc Choisy (CR1, MIVEGEC, NIHE, IRD)  
Dr. Vu Dinh Thiem (NIHE)  
**Etudiante :** Mme TRAN Thi Cam Giang  
**Financement :** Programme Doctoral International Modélisation des Systèmes Complexes  
**Début :** 1<sup>er</sup> Décembre 2012 ou 1<sup>er</sup> Janvier 2013

## 1. Cadre général du projet de recherche

Les événements récents en matière d'épidémiologie (SRAS en 2003, grippe aviaire en 2004 ou grippe porcine en 2009), de catastrophe naturelle (tremblement de terre –Haïti, Chili, Chine–, tsunamis –Océan Indien 2004, France 2010, Japon 2011– inondations, etc), tempête d'insectes (nuée de criquet pèlerin), révèlent d'importants risques environnementaux. La propagation rapide du virus A-H1N1 (grippe porcine) illustre combien il est important de pouvoir anticiper les phénomènes épidémiologiques pour mieux les gérer lorsqu'ils surviennent. Dans ce contexte, de nombreux travaux menés notamment par l'OMS<sup>2</sup>, l'Institut Pasteur<sup>3</sup> et l'Inserm<sup>4</sup> dans le domaine de la « **sécurité environnementale** » tentent de comprendre l'émergence de ces phénomènes et leur propagation sur un territoire, afin de mieux gérer les situations de crise. Ces recherches consistent en des études mathématiques ou statistiques appliquées à des mesures de terrain obtenues via des réseaux de surveillance [Chauvin 2005]. Elles sont effectuées soit a priori pour la prévision, soit en temps réel pour le suivi des épidémies, soit a posteriori pour la validation des modèles. Ces études constituent l'un des axes des thèmes de recherche du laboratoire UMMISCO (UMI IRD 209).

## 2. Vers la modélisation-simulation hybride de la complexité

L'avènement des TIC connu ces 20 dernières années favorise aujourd'hui la conception, le développement et l'exécution de modèles, de véritables outils scientifiques et/ou d'aide à la décision dont l'objectif est de reproduire, par simulation, un phénomène environnemental en vue de le comprendre et/ou d'en faire des études prospectives. Ainsi, le développement croissant des techniques de simulation a contribué à la mise en place d'une pléthore de simulateurs reposant sur des approches mathématiques [Atzrouni *et al.* 2001 ; Bacaer *et al.* 2008 ; Chattopadhyay *et al.* 2009 ], statistiques [Abbas *et al.* 2004 ; Durand *et al.* 1999] et/ou informatiques [Amouroux *et al.* 2008 ; Badariotti *et al.* 2005].

A l'origine simple caricature d'une réalité, les modèles ont évolué avec les besoins croissants des communautés scientifiques en terme d'expressivité et de réalisme vis à vis du système

complexe étudié. Ainsi dans un modèle épidémiologique de transmission de maladie on peut décrire la dimension spatiale des rencontres inter-individuels et représenter explicitement les déplacements et contacts des agents. Ainsi en complément des modèles agrégés parfois caricaturaux ou « KISS » (Keep it simple and stupid) (Axelrod 1997) sont apparus des modèles souvent individu-centré très détaillés « KIDS » (Keep It Descriptive Stupid) (Edmonds 2004) apportant une plus grande fidélité fonctionnelle et structurelle vis à vis du système complexe modélisé. Cependant ces modèles KIDS sont aussi plus difficiles à calibrer, à analyser analytiquement compte tenu de l'explosion du nombre de paramètres et de la méconnaissance de leur rôle dans le modèle. Un calibrage particulièrement intéressant qui nous intéressera dans ce travail est celui de l'action exogène sur le système. Concrètement dans le cas d'un modèle épidémiologique il s'agira de la vaccination qui permet de contrôler la dynamique du système.

Un compromis semble donc nécessaire et c'est l'un des enjeux de cette thèse, à savoir de combiner des modélisations à différents niveaux de granularité. Ceci devrait permettre d'une part d'obtenir une modélisation suffisante pour prendre en compte la dimension spatiale dans la modélisation épidémiologique. Mais d'autre part offrir une résolution suffisamment efficace qui permette la mise en oeuvre d'algorithmes d'apprentissage permettant d'optimiser le contrôle possible sur un tel système dynamique.

### 3. Apprentissage et Analyse statistique pour l'exploration de modèles

L'automatisation de l'exploration de l'espace des modèles engendré par un espace de paramètres peut être utilisée à plusieurs fins : analyser la sensibilité d'un modèle proprement dit, calibrer un modèle, découvrir des propriétés émergentes d'un modèle, étudier la viabilité d'un modèle, etc. L'élaboration d'un système autonome d'exploration des modèles qui déterminerait, par lui même, l'espace des paramètres à explorer en fonction d'expériences antérieurs et de nouvelles données recueillies (par un réseau de capteur par exemple) pose plusieurs problèmes de recherche qui mêlent apprentissage statistique multiagent et parallélisation. L'aspect parallélisation de cette question fait l'objet d'un travail de doctorat sur lequel le présent travail s'appuiera. La dimension de l'apprentissage apparaît dès qu'il s'agit d'optimiser les simulations pour atteindre un but (optimiser ou contrôler un paramètre ou une variable du système). Dans le cadre de cette thèse, il s'agira en particulier de mettre en oeuvre des techniques permettant d'optimiser les politiques de vaccination des modèles dont l'objectif est de modéliser les épidémies en Asie du Sud-Est.

### 4. Objectifs

Le premier travail de cette thèse sera de faire un état de l'art couvrant les différentes approches de modélisation hybrides mathématique et informatique notamment dans le domaine de la modélisation épidémiologique. Dans un second temps il s'agira de concevoir ou d'étendre un langage de représentation permettant à un modélisateur d'exprimer d'une part les objectifs des simulations du modèle mais aussi des contraintes liées à celles-ci. D'un point de vue concret, l'ambition de ce travail est de permettre l'utilisation de modèles complexes hybrides (mathématiques et individu-centré) développés au sein d'UMMISCO pour être capable d'optimiser le contrôle de leur dynamique et permettre une analyse prédictive. Les simulations s'appuieront sur des algorithmes distribués profitant ainsi de compétences dans ce domaine au sein de MSI.

## 5. Application à la vaccination dans un contexte spatial

La politique la plus ancienne (début dans les années cinquante pour les pays riches) et encore aujourd'hui la plus utilisée est la politique dite "*de masse*" qui consiste à vacciner le maximum d'enfants avant un certain âge (c'est par exemple la politique actuellement utilisée par le programme EPI de l'OMS contre la rougeole au Vietnam). C'est une politique qui porte ses fruits dans la mesure où elle a conduit à une nette diminution de l'incidence dans de nombreux pays. Toutefois, c'est une politique tout à fait inefficace, extrêmement coûteuse et strictement impossible à mettre en œuvre dans nombre de pays pauvres, notamment en Afrique, pour des raisons à la fois financières et logistiques. Ce constat a stimulé dans les années 1990-2000 des recherches théoriques sur des politiques de vaccination alternatives efficaces, donc basées sur des critères d'optimisation. Une des politiques proposées est la politique dite "*par pulsation*" qui consiste à vacciner à intervalles de temps réguliers une certaine proportion de la population (Agur et al. 1993, Nokes & Swinton 1997). La durée entre deux événements de vaccination ainsi que la couverture vaccinale au cours de chaque événement sont choisies de telle sorte à ce qu'elles maintiennent la proportion de susceptibles dans la population en-dessous du seuil critique nécessaire au démarrage d'une épidémie. C'est donc une politique qui, sur le papier, s'avère extrêmement efficace (du moins beaucoup plus que la politique de vaccination de masse) et également particulièrement aisée à mettre en œuvre sur le terrain au niveau logistique. Son inconvénient tient plus à des raisons éthiques dans la mesure où il est très difficile de justifier que l'on ne vaccine pas tout le monde et pas tout le temps. C'est pour cette raison que cette politique de vaccination reste très anecdotique. Dans ce contexte, l'idée serait de quantifier l'efficacité de la politique de vaccination par pulsations par rapport à la vaccination de masse, afin de renforcer les arguments en faveur de cette première. Un des aspects qui n'a en particulier jamais été étudié est les effets de telles politiques de vaccination dans un contexte spatial. En effet, les populations humaines ne sont pas réparties de façon homogène sur le territoire mais plutôt agrégées dans des villes connectées par des réseaux de transport. Nous avons donc une structure à deux niveaux : les épidémies apparaissent au niveau d'une ville (niveau local) mais peuvent se propager de ville en ville au niveau du pays (niveau global). Le lien entre les dynamiques locales et la dynamique globale est dépendant en particulier du degré de synchronicité entre les dynamiques locales (Grenfell et al. 2001). En effet, la probabilité que des éradications locales de la maladie conduisent à une éradication globale augmente avec le degré de synchronicité entre les dynamiques locales. Or les politiques de vaccination de masse et par pulsation ont des effets opposés sur ce degré de synchronicité (Earn et al. 1998). Alors que la première tend à le diminuer, la seconde peut-être faite en sorte qu'elle l'augmente. C'est cet effet de la vaccination dans un contexte spatial qu'il serait intéressant de quantifier dans un premier temps. Dans un deuxième temps, il serait intéressant d'aller plus loin et d'utiliser un algorithme d'apprentissage par renforcement qui nous dirait, pour une structure de population donnée, où et quand il faut vacciner afin de diminuer au maximum l'incidence globale ou augmenter au maximum la probabilité d'éradication globale. (A noter que ces deux notions ne sont pas nécessairement corrélées et ce sera également quelque chose d'intéressant à montrer.) Ce travail théorique sera couplé à des analyses sur données épidémiologiques réelles issues des systèmes de surveillance nationaux du Vietnam (en partenariat avec le NIHE).

## Protocole de recherche et plan de travail

Cette thèse se déroulera dans le cadre d'une codirection devant déboucher sur un diplôme de thèse français, elle est prévue dans le cadre d'une **bourse du Programme Doctoral International**

**Modélisation des Systèmes Complexes.** Le déroulement de la thèse est prévu pour **11 mois par année au Vietnam** au sein de l'équipe de recherche MSI (Modélisation et Simulation Informatiques) de l'IFI, et **1 mois en France** à l'IRD.

### A contacter:

Pr. Yann Chevaleyre  
Email : [yann.che@gmail.com](mailto:yann.che@gmail.com)

Dr. Jean-Daniel Zucker  
email : [jdzucker@gmail.com](mailto:jdzucker@gmail.com)

Dr. Marc Choisy  
email : [mchoisy@gmail.com](mailto:mchoisy@gmail.com)

Md. Vu Dinh Thiem  
email : [thiemvd@gmail.com](mailto:thiemvd@gmail.com)

## 6. Références

- (Abbas et al. 2004) K. Abbas, A. Mikler, A. Ramezani, S. Meneze, Computational epidemiology: Bayesian disease surveillance, Proceedings of the International Conference on Bioinformatics and its Applications (ICBA'04), 2004
- (Agur et al., 1993) Agur, Z.; Cojocar, L.; Mazor, G.; Anderson, R. M. & Danon, Y. L. Pulse mass measles vaccination across age cohorts. *Proceedings of the National Academy of Sciences of the USA*, **1993**, 90, 11698-11702
- (Amouroux et al. 2008) E. Amouroux, S. Desvaux et A. Drogoul Towards Virtual Epidemiology: An Agent-Based Approach to the Modeling of H5N1 Propagation and Persistence in North-Vietnam, Intelligent Agents and Multi-Agent Systems, page 26-33, 2008
- (Atzrouni et al. 2001) M. Artzrouni, et J.P. Gouteux, Population dynamics of sleeping sickness : A micro simulation, *Simulation & Gaming*, Vol. 32, No. 2, 215-227, 2001
- (Badariotti et al. 2005) D Badariotti, A Banos, V Laperrière, Vers une approche individuelle centrée pour modéliser et simuler l'expression spatiale d'une maladie transmissible : la peste à Madagascar, SAGEO 2005
- (Chattopadhyay et al. 2009) J. Chattopadhyay, P. Auger, R. Mchich, T. Chowdhury, G. Sallet and M. Tchuente. Effects of a disease affecting a predator on the dynamics of a prey-predator system. In *Journal of Theoretical Biology*, 2009
- (Chauvin 2005) Pierre Chauvin, Constitution and monitoring of an epidemiological surveillance network with sentinel general practitioners, *European Journal of Epidemiology*, 2005.
- (Deffuant et al. 2007) Deffuant, G., Chapel, L. et Martin S. 2007. Approximating viability kernel with Support Vector Machines, *IEEE transactions on automatic control*. 52 (5), p.933-937.
- (Durand et al. 1999) B. Durand, O. Mahul , An extended state-transition model for foot-and-mouth disease epidemics in France, *Prev. Vet. Med.* 47, 1999
- Earn, D. J.; Rohani, P. & Grenfell, B. T. Persistence, chaos and synchrony in ecology and epidemiology. *Proceedings of the Royal Society of London B*, **1998**, 265, 7-10
- Grenfell, B. T.; Bjørnstad, O. N. & Kappey, J. Travelling waves and spatial hierarchies in measles

epidemics. *Nature*, **2001**, 414, 716-723

Nokes, D. J. & Swinton, J. Vaccination in pulses: a strategy for global eradication of measles and polio?  
*Trends Microbiol*, **1997**, 5, 14-19