

**Outils mathématiques, outils technologiques :
lorsque le choléra rencontre le satellite**

Guillaume Constantin de Magny & Marc Choisy
Génétique et Évolution des Maladies infectieuses
UMR CNRS-IRD 2 724, Montpellier

« Epidemiology, concerned as it is with variation of disease from time to time or from place to place, must be considered mathematically (...), if it is to be considered scientifically at all. »

Sir Ronald Ross, M.D.

The Prevention of Malaria. Churchill, London. 1911

{10} De façon générale, un modèle est un objet d'imitation. Il en est de même en science. Le réel est complexe et l'utilisation de modèles comme images simplifiées de la réalité aide le scientifique à comprendre les phénomènes qu'il observe. Parmi la variété des modèles utilisés, les modèles mathématiques tiennent une place de choix du fait de leur relative simplicité d'élaboration (une feuille de papier et un crayon suffisent) et de leur puissance inhérente aux théories mathématiques (résolution d'équations).

Contrairement à certains domaines de théorie pure comme l'astrophysique, les modèles épidémiologiques sont caractérisés par leur étroite intimité avec les données réelles. Les données épidémiologiques sont généralement des nombres de malades en une localité et à une date donnée. L'utilisation conjointe des modèles mathématiques et des données épidémiologiques présente plusieurs avantages :

Les données permettent d'estimer les paramètres du modèle. Un modèle mathématique décrit la réalité sous forme d'équations c'est-à-dire de phrases dont les noms représentent des concepts (souvent notés par des lettres plus ou moins exotiques : λ , ξ , β , χ , etc.) et les verbes sont des sommes, des variations, des multiplications (là encore notés par des signes cabalistiques : +, =, >, -, \times , \neq , δ , Σ , f , etc.). On appelle paramètres du modèle les grandeurs λ , ξ , β , χ , et leur estimation consiste à leur affecter une valeur numérique. Le bon choix de ces valeurs numériques se fait grâce aux données réelles : la valeur du paramètre retenue est celle qui permet au modèle de rendre compte le mieux des données observées ^{Encart (1)}. On pourra ainsi estimer des paramètres tels que le taux de contact ou la force d'infection d'un virus ou d'une bactérie, par exemple.

Les données permettent de sélectionner les paramètres importants.
« Tout ce qui est simple est faux mais ce qui est compliqué est inutilisable. » Cette phrase de Paul Valéry résume assez bien le compromis dans le choix d'un modèle. Tous les modèles sont faux ! Plus les modèles sont compliqués (en terme de nombre de paramètres), moins ils sont faux mais moins ils sont

universaux également ^{Fig. 1}. Les données aident à choisir, parmi plusieurs modèles, le plus parcimonieux et mettre ainsi en avant les paramètres les plus importants dans la compréhension du phénomène étudié. Pour paraphraser Einstein, « *everything should be made as simple as possible, but no simpler* ».

Les données permettent de tester des scénarios. Un modèle est basé sur des hypothèses. L'idée ici est de confronter plusieurs modèles en utilisant les données comme arbitres pour sélectionner le meilleur modèle. Les hypothèses correspondant au modèle sélectionné représentent alors le scénario le plus probable. Par exemple pour tester l'hypothèse d'un forçage climatique sur les épidémies, on pourra confronter deux modèles avec et sans forçage.

Le meilleur modèle permet de prédire. Jusqu'à présent nous avons utilisé des données historiques pour élaborer des modèles mathématiques. Après avoir estimé les paramètres et sélectionné le meilleur modèle, nous gagnons en compréhension du phénomène étudié. Par exemple nous pouvons avoir une idée de la force d'infection ou savoir si le climat joue un rôle dans la dynamique de certaines maladies. Plus que ça, en faisant l'hypothèse que le futur ressemblera au passé, notre modèle acquiert un pouvoir prédictif, très important d'un point de vue pratique. Nous pourrions ainsi avoir des idées sur les conséquences de différentes politiques vaccinales ou sur les effets des changements climatiques et/ou environnementaux.

L'exemple du choléra. Les bactéries du type *Vibrio cholerae*, responsables d'épidémies de choléra dans les populations humaines, sont des procaryotes ubiquistes présents naturellement à l'état commensal dans les écosystèmes marins, lagunaires et estuariens. Elles vivent fixées à des micro-algues constituant du phytoplancton ou à des arthropodes du zooplancton (Islam *et al.*, 1994), et sont au cœur d'un réseau d'interactions et de chaînes trophiques – les *Vibrio* assurent notamment une fonction de dégradation des protéines de leurs hôtes arthropodes. Il existe ainsi une dépendance étroite entre les dynamiques de phyto- et zoo-plancton et de *Vibrio*.

La saisonnalité marquée du choléra et le constat de l'apparition simultanée des cas à différents endroits du globe ont été les principales raisons de la recherche de facteurs à l'échelle globale et en particulier de forçages climatiques et environnementaux (Colwell, 1996). En effet, une épidémie de choléra est provoquée par le passage de la bactérie *Vibrio cholerae* du milieu aquatique aux populations humaines. Les facteurs climatiques et/ou environnementaux influent sur la dynamique du plancton, cette dernière détermine la dynamique des bactéries *V. cholerae* dans le milieu marin. Enfin la densité de *Vibrio* à proximité des populations humaines conditionne directement le risque de passage aux populations humaines et donc l'émergence d'épidémies de choléra ^{Fig. 2}. Une forte puissance de prédiction sur les paramètres environnementaux et/ou climatiques peut donc se répercuter sur la puissance de prédiction des épidémies de choléra. Une application envisageable de ces

recherches serait un système d'aide à la décision pour une meilleure organisation des moyens sanitaires à l'échelle locale.

La télédétection spatiale nous permet aujourd'hui de suivre l'évolution de paramètres environnementaux et climatiques avec une excellente résolution spatiale et temporelle et apporte un continuum d'échelles d'observations dans l'espace et dans le temps des structures de l'environnement océanique comme terrestre. Cette nouvelle source de données, couplée aux données épidémiologiques nous permet aujourd'hui de tester l'importance des facteurs environnementaux dans l'émergence des épidémies de choléra en mesurant la capacité qu'ont nos modèles à prédire le passé (études rétrospectives). Dans un deuxième temps, après identification des variables environnementales et/ou climatiques à fort pouvoir prédictif, nous serons en mesure de proposer des cartes de risque potentiel à choléra.

Figure 1. Un exemple du compromis entre réalisme et universalité. Un éléphant (à gauche) est modélisé par une fonction polynomiale (à droite) de nombre de paramètres croissant : 5, 10, 20, 30 de haut en bas. Plus le modèle est compliqué, plus la modélisation est réaliste. En revanche, plus le nombre de paramètres augmente, plus le modèle tient compte des particularités spécifiques à un jeu de données, et devient donc compliqué. Le choix du modèle optimal se fait par un compromis entre simplicité et réalisme.

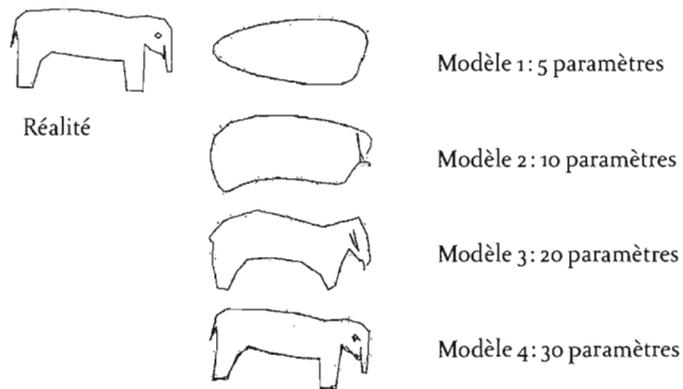
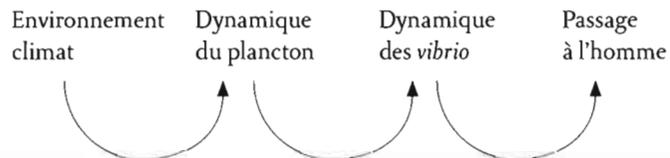


Figure 2. Cascade d'événements régissant l'émergence d'épidémie de choléra dans les populations humaines (modèle conceptuel à la base des modèles mathématiques utilisés en épidémiologie écologique).



Constantin de Magny Guillaume, Choisy Marc (2005)

Outils mathématiques, outils technologiques : lorsque le choléra rencontre le satellite

In : Barbault R. (ed.), Chevassus-au-Louis B. (ed.),
Teyssède A. (coord.). *Biodiversité et changements globaux : enjeux de société et défis pour la recherche*

Paris : ADPF, p. 126-128

Biodiversité Sciences et Gouvernance : Conférence
Mondiale, Paris (FRA), 2005/01/24

ISBN 2-914935-27-7