

La modélisation de l'expression des maladies transmissibles, vers une intégration de l'individualité et de l'hétérogénéité spatiale le tournant de la complexité

L'un des initiateurs du *tournant de la complexité* en épidémiologie, J. S. Koopman (2005a), rappelait l'aspect fondamental, pour une science dont la finalité est l'explication et si possible la prédiction en vue d'une prise de décision destinée à contrôler l'infection, de disposer à la fois de bases de données empiriques, de théories donnant une représentation abstraite et simplifiée des mécanismes causaux de la maladie et enfin de méthodes de modélisation permettant de faire le lien entre ces théories et les observations : d'un côté, les théories sont construites et testées sur des observations du monde réel à travers des modèles, qui sont des représentations simplifiées de la réalité. Ces modèles peuvent être implicites, correspondant alors à des modèles mentaux, ou explicites, formalisés par des équations mathématiques ou des programmes informatiques. D'un autre côté, nos observations sont quant à elles bien souvent guidées par nos théories (Gilbert, 1995 ; Batty et Torrens, 2005 ; Edmonds, 2005 ; Epstein, 2008).

Aussi, J. Koopman rappelait la nécessité, pour la science des maladies transmissibles, de ne pas se limiter à élaborer des modèles statistiques prédictifs et non causaux du risque épidémiologique. Il invitait à se démarquer d'une tradition de l'inférence

causale observationnelle (Koopman et Longini, 1994) et à développer une épidémiologie théorique qui puisse être reliée, à travers des modèles dynamiques explicatifs, aux données empiriques et serve à comprendre l'expression des maladies transmissibles : « *nous devons aller plus loin dans le sens d'une compréhension des déterminants des flux d'infections dans des populations, de la même manière que la science a permis de comprendre les déterminants du temps qu'il fait ou ceux des courants des océans* » (Koopman, 2004, p. 305, trad. pers.)

L'objet de ce chapitre est d'évoquer l'apport des différentes théories et méthodes de modélisation des systèmes complexes permettant une meilleure compréhension des mécanismes endémo-épidémiques des maladies transmissibles comme la peste, dans le sens invoqué par Koopman. Nous commencerons par préciser les grands principes du *tournant de la complexité* que sont notamment la reconnaissance des non-linéarités et l'attention portée à la notion d'émergence (1.) Nous verrons ensuite comment l'évolution des démarches de modélisation a permis d'aborder, en intégrant progressivement des détails de la réalité et en remettant en cause certaines hypothèses simplificatrices, la complexité des systèmes pathogènes comme la peste (2.) Enfin, nous présenterons les différentes expériences de modélisation consacrées à la peste bubonique, réalisées par nos prédécesseurs (3.)

1. La peste : un système pathogène complexe appréhendé par la modélisation

1.1 La complexité : une nouvelle orientation donnée au paradigme systémique

Le paradigme systémique a jeté les bases des théories de la complexité en ce qu'il a permis, en réaction au paradigme réductionniste de Descartes, de « *restituer aux phénomènes toutes leurs solidarités* » (Bachelard, 1934, p. 134 ; Le Moigne, 1999). Son apport majeur, en ce qui concerne les maladies transmissibles, tient dans la considération de la *toile des causalités* (Krieger, 1994) ou du *complexe de causes* (Rothman, 1976) : nous avons exposé précédemment (chapitre 2), l'évolution des recherches ayant conduit à considérer le déclenchement d'une maladie au niveau individuel, ainsi que le déclenchement d'une épidémie au niveau d'une population, comme résultant de l'intervention de multiples facteurs relevant de différents niveaux, liés les uns aux autres, aux effets multiples, apportant un caractère indéterminé à la maladie. N'est-ce pas là la reconnaissance de la complexité apparente d'une maladie comme la peste, si l'on considère que le terme complexité désigne, au premier abord, un tissu (*complexus* = ce qui est tissé ensemble) de constituants hétérogènes inséparablement associés ?

Aussi, la théorie des systèmes est, depuis longtemps, concernée par l'étude de systèmes que l'on qualifie aujourd'hui de complexes. Elle contenait les germes des théories de la complexité par sa considération des interrelations entre les facteurs et la possibilité de boucles de rétroaction : « *il est indispensable d'étudier non seulement les parties et les processus d'un système de manière isolée, mais également de résoudre le problème crucial de l'organisation, de l'ordre qui les réunit sous l'influence de leur interaction dynamique* » (von Bertalanffy, 1968, p. 31, trad. Phipps, 2000). Dès lors, on peut se demander quels sont les apports des théories de la complexité et en quoi celles-ci, associées aux méthodes d'analyse des systèmes complexes, font que la complexité se révèle être davantage qu'une propriété attribuée

à un système, et orientent la science vers un nouveau paradigme unificateur des disciplines (Manson, 2001 ; Urry, 2003 ; Gatrell, 2005).

D'une manière générale, la recherche sous l'angle des théories de la complexité s'intéresse à la dynamique des systèmes et à la manière dont leur structure globale évolue au cours du temps, dû aux interactions entre leurs parties constituantes (Manson, 2001). Différents champs de la science ont leur propre définition et établissent leur propre mesure de la complexité, si tant est que l'on puisse la quantifier. Ces définitions et mesures, orientées vers leurs propres applications, ne font pas forcément sens lorsqu'elles sont transposées vers d'autres champs (Couclelis, 1988 ; Gell-Mann, M., 1995). Cependant, les théoriciens de la complexité concourent à une vue commune : un système complexe est un tout cohérent dont les interactions entre ses nombreux constituants conduisent à une dynamique de fonctionnement global difficilement prédictible et à l'émergence de propriétés nouvelles, non déductibles de la seule connaissance des éléments isolés. Ainsi, le tout est plus que la somme de ses parties (Morin, 1990 ; Batty et Torrens 2001 ; Manson, 2001 ; Reitsma, 2003 ; Gatrell, 2005 ; Zwirn, 2006). La complexité d'un système repose donc en grande partie sur les interactions entre ses éléments. Une condition essentielle pour qu'un système devienne complexe est que ces interactions soient non linéaires. Pour cette raison, il est impossible, malgré une connaissance parfaite de ses composants, de prévoir son comportement autrement que par l'expérience ou la simulation, à l'aide d'un modèle dynamique représentant explicitement les mécanismes causaux à l'œuvre (Wolfram, 1984 ; Holland, 1995 ; Epstein, 1996 ; Le Moigne, 1999 ; Batty et Torrens, 2001 ; Edmonds, 2005 ; Epstein, 2008).

1.1.1 La complexité dans une perspective holiste : la reconnaissance des effets non-linéaires sur la dynamique globale

L'appréhension d'un ingrédient essentiel à la complexité que sont les non-linéarités a débuté par l'utilisation d'équations différentielles décrivant la dynamique du système d'une manière holiste, c'est-à-dire dans sa globalité. Des travaux d'H. Poincaré (1912), à ceux d'E. N. Lorenz (1963), la contribution majeure des mathématiques à l'étude des systèmes complexes est l'élaboration de la théorie du chaos déterministe (Gleick, 1988) : dans une équation différentielle simple et déterministe, l'existence

d'un terme non-linéaire fait que les effets ne sont pas proportionnels aux causes et que de petites variations sur un paramètre ou sur les conditions initiales peuvent avoir pour conséquence un changement drastique de la trajectoire du système et l'instauration d'un régime chaotique totalement imprédictible.

Une illustration classique de la théorie du chaos en écologie est la dynamique d'une population animale décrite par l'équation logistique. Cet exemple, relevant de l'écologie théorique, n'est pas sans rapport avec notre questionnement sur l'expression d'une zoonose vectorielle comme la peste dans la mesure où celle-ci résulte de l'association parasitaire entre plusieurs populations dynamiques, hôtes, vecteur et germes provoquant la maladie. Le modèle logistique est très simple car il ne comporte qu'une seule variable d'état dynamique, celle de la population d'une seule espèce. Il se veut toutefois plus satisfaisant que le modèle linéaire malthusien d'une croissance exponentielle, en faisant l'hypothèse que les ressources disponibles, sur un espace donné, sont limitées et qu'il existe une population maximale au-delà de laquelle la population décroît. Cela induit l'existence d'un terme rétroactif limitant la hausse de population quand la densité devient trop élevée, qui rend le modèle non-linéaire. Dans les années 1970, on s'interroge sur le fait qu'indépendamment de facteurs externes tels la présence de maladie ou de prédateurs, certaines populations se stabilisent, tandis que d'autres suivent des cycles réguliers et d'autres encore fluctuent d'une manière totalement aléatoire. Bénéficiant des progrès du calcul numérique, plusieurs chercheurs établissent alors la forte sensibilité du modèle à la valeur d'un paramètre, le taux de croissance effectif de la population. Selon les valeurs prises par ce paramètre-clé, le comportement du système change drastiquement d'une dynamique régulière vers un régime chaotique (Li et Yorke, 1975 ; May, 1976 ; Feigenbaum, 1978).

L'introduction de la théorie du chaos et ses concepts dérivés de sensibilité aux conditions initiales, de bifurcations et d'attracteurs étranges marque donc un premier pas vers le paradigme de la complexité, en rupture avec la conception linéaire des causalités et des dynamiques de la science analytique traditionnelle. J. Gleick écrivait en 1988 (p. 3) : « *où commence le chaos cesse la science classique* ». Des équations mathématiques reposant sur des hypothèses simples et déterministes, mais intégrant des effets non-linéaires, permettent d'appréhender une réalité complexe et

imprévisible. Cela pose les limites de la prédiction et la dissocie de l'explication : ce qui peut être expliqué ne peut pas forcément être prédit (Philippe et Mansi, 1998 ; Manson, 2001 ; Batty, Torrens, 2005 ; Epstein, 2008).

1.1.2 Le traitement de la complexité liée à l'articulation des niveaux par la modélisation bottom-up

La reconnaissance de l'importance des non-linéarités entre les parties du système sur sa dynamique de fonctionnement global marque donc un premier pas franchi vers un paradigme de la complexité. Cependant, une question essentielle adressée par les théories de la complexité est celle de l'articulation des niveaux, des mécanismes reliant l'organisation d'un système observé à un niveau donné au comportement des entités et à la manière dont elles interagissent à un niveau inférieur. Or, l'adoption d'un point de vue holiste sur le système simplifie considérablement les interactions multiples non-linéaires se produisant entre des entités nombreuses, hétérogènes et évolutives. De ce fait, avec une telle perspective, on laisserait de côté une part de la complexité du système liée à la manière dont les entités individuelles interagissent et génèrent des propriétés nouvelles au niveau global, ces dernières exerçant un effet rétroactif sur les actions individuelles.

La théorie du chaos déterministe nous a révélé une complexité qui réside dans l'apparition d'une dynamique en apparence totalement irrégulière et désordonnée, issue d'un mécanisme en réalité totalement réglé. En contraste, la complexité d'agrégation, pour reprendre la terminologie de S. M. Manson (2001), traite de comment de l'ordre reconnaissable au niveau de la globalité du système peut émerger de l'interaction plus ou moins aléatoire et désordonnée de ses éléments au niveau inférieur (figure 14).

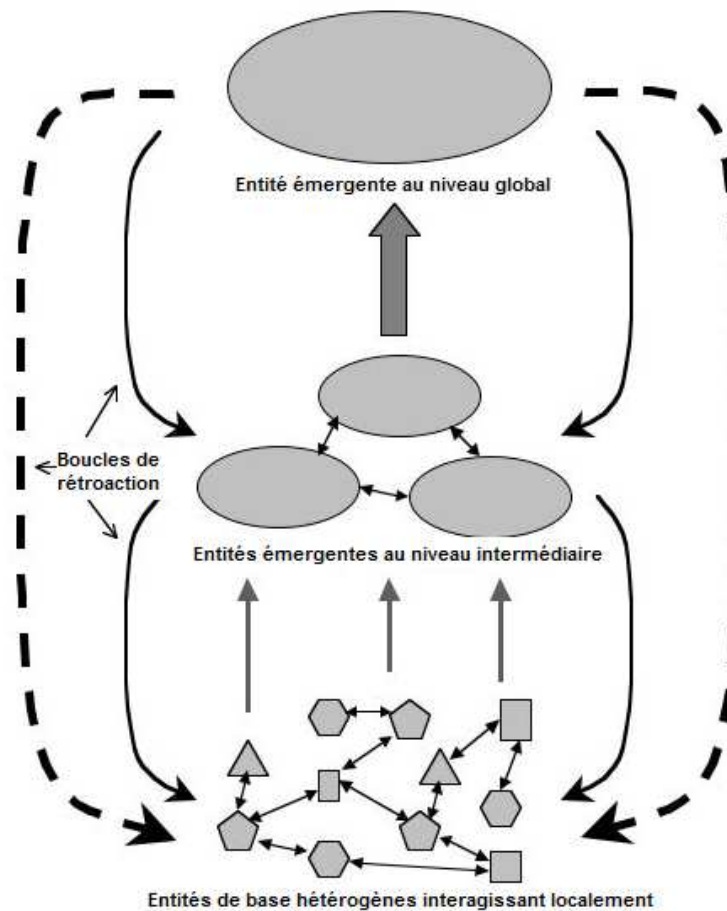


Figure 14. Articulation des niveaux dans un système complexe (source : Parrott, 2002)

Ainsi, la vision exprimée par R Lewin paraît-elle s'inscrire en rupture avec celle de J. Gleick (1988) cité plus haut : « *la science de la complexité a à voir avec la structure et l'ordre* » (Lewin, 1992, p. 10, trad. pers.) Pourtant, du fait de la non-linéarité des interactions, beaucoup d'éléments présents dans la théorie du chaos se retrouvent dans la complexité d'agrégation, tel le caractère imprévisible du système et sa forte sensibilité à certains paramètres ou conditions initiales (Bak et Chen, 1991).

Si les atomistes grecs en avaient une première intuition, la mise en évidence et la popularisation d'une telle complexité d'agrégation, créatrice d'*ordre par le bruit* (von Foerster, 1960 ; Atlan, 1979), doit beaucoup à la cybernétique, elle-même inspirée de la thermodynamique et des travaux sur les structures dissipatives (Nicolis et Prigogine, 1977). Celles-ci ont permis d'illustrer la notion d'auto-organisation, traduisant l'émergence spontanée d'une structure se développant dans l'espace et le

temps, sous l'effet conjoint d'un apport extérieur d'énergie et des interactions à l'œuvre entre les éléments du système considéré. La notion a par la suite été transposée à l'étude de processus variés du monde physique et du vivant, se réalisant à tous les niveaux d'organisation, de la cellule, de l'être organisé (Camazine et al., 2003), aux écosystèmes (Schneider et Kay, 1994), aux systèmes économiques (Witt, 1997) et sociaux (Allen et Sanglier, 1981 ; Pumain et al., 1989 ; Sanders, 1992). Proche de l'auto-organisation mais se voulant plus général et universel, le concept d'émergence est progressivement apparu au centre des interrogations sur la complexité des systèmes, défini comme le processus par lequel un ensemble d'entités interagissantes à un niveau font apparaître au niveau supérieur une propriété nouvelle (une structure particulière, une fonctionnalité, une loi) reconnaissable par un observateur et irréductible à la superposition des contributions individuelles (Holland, 1998 ; Goldstein, 1999 ; Laughlin, 2005 ; Dessalles et al., 2007).

1.1.2.1 Des automates cellulaires...

Bien qu'aucune théorie quantitative de l'émergence n'ait été formulée, son exploration a été rendue possible par le développement récent et rapide de l'informatique, qui a donné naissance à une nouvelle classe de modèles dynamiques permettant de considérer des systèmes que l'on qualifiera d'une manière générale, comme Phipps (2000), de distribués dans le sens où ils sont basés sur les entités constitutives du système, qui sont explicitement formalisées et distribuées dans un réseau d'interactions. Nous les qualifierons d'individus-centrés lorsque les entités du système correspondent effectivement aux individus d'une population, suivant le sens étroit défini, en écologie, par J. Uchmański et V. Grimm (1996 ; Grimm et Railsback, 2005).

La démarche sous-tendue par l'approche de modélisation distribuée est de type *bottom-up*, volontiers qualifiée de constructiviste (Goldstein, 1999) ou *générative* (Epstein, 1999). Elle consiste à formaliser en entrée du modèle, sur la base d'hypothèses simplificatrices, la description des entités discrètes et la manière dont elles se comportent et interagissent pour générer, par la simulation, la structure et la dynamique globale du système étudié, observée en sortie du modèle. Historiquement, les premiers modèles de systèmes distribués qui ont commencé à explorer la complexité d'agrégation sont les automates cellulaires (Wolfram, 1984). Un automate

cellulaire est défini comme un réseau de cellules où chacune prend un état possible parmi un ensemble d'états possibles. A chaque itération, des *règles de transition* sont appliquées à toutes les cellules, qui spécifient comment les différents états des cellules réagiront aux configurations d'états se trouvant dans le *voisinage* immédiat de chaque cellule. Le principal apport de ces modèles, théorique, a été de préciser comment des propriétés globales émergent d'interactions strictement individuelles et locales, tandis que l'utilisation de règles simples, déterministes et invariantes a permis de réitérer le message de la complexité déterministe.

Ainsi, pour reprendre l'exemple de l'écologie théorique développé plus haut, H. Couclelis (1988) a utilisé un automate cellulaire déterministe pour simuler la dynamique d'une population animale sur un espace isotrope, en appliquant une fonction de densité de population à chaque cellule de l'automate. Cette fonction exprime l'hypothèse que la croissance locale de la population au niveau d'une cellule au pas de temps $t + 1$ dépend de sa densité et de celle des cellules environnantes à l'instant t . A partir de cette règle initiale, qui reprend l'hypothèse simplificatrice du modèle logistique holiste, mais en la formalisant au niveau des entités discrètes du système, les simulations génèrent des structures spatio-temporelles complexes : outre le fait qu'elles reproduisent la gamme variée des dynamiques temporelles obtenues à l'aide du modèle holiste (May, 1976), l'originalité de l'approche distribuée est qu'elle permet d'observer au niveau global l'émergence de formes particulières d'organisation de la population, plus ou moins stables, très sensibles à la distribution spatiale initiale de la population. La reconnaissance de l'importance de ce facteur, jusqu'alors peu étudié parmi l'ensemble des facteurs susceptibles d'influencer la dynamique d'une population, constitue en un sens un plaidoyer pour l'élaboration de modèles intégrant davantage de détails de la réalité, à commencer par la considération d'entités discrètes, distribuées dans un environnement spatial explicite (Uchmański et Grimm, 1996 ; Grimm et Railsback, 2005).

1.1.2.2 ... aux modèles à base d'agents adaptatifs : diversité des agents et structure des interactions peuvent faire la différence

Après les automates cellulaires, modèles distribués les plus simples, l'exploration de la complexité d'agrégation dans les systèmes distribués se poursuit avec l'idée que des comportements complexes émergent dans des systèmes possédant un haut degré

d'hétérogénéité parmi leurs composants et un arrangement particulier de ces composants (Levin, 1999 ; Parrott, 2002). Ainsi a-t-on vu ces dernières années se développer un intérêt porté aux systèmes complexes qualifiés d'adaptatifs. Ce terme, au demeurant très proche des concepts d'auto-organisation et d'autopoïèse, fait référence à la tendance qu'ont de nombreux systèmes à rester globalement stables, malgré les changements permanents qui s'observent parmi ses composants individuels (Holland, 1995). Dans cette optique, pour modéliser de tels systèmes, besoin est de rompre avec l'homogénéité et l'invariance des règles de comportement formulées au niveau des entités du système.

Les modèles à base d'agents adaptatifs vont dans ce sens. Issus de l'Intelligence Artificielle, relevant du domaine des Systèmes Multi-Agents (Drogoul, 2005), ils forment une nouvelle classe de modèles distribués dans lesquelles les agents adaptatifs sont des entités diversifiées et autonomes, programmées de sorte qu'elles soient capables d'apprendre de leurs interactions locales avec les autres entités du système et puissent changer leur comportement en vue d'atteindre un but préalablement défini (Epstein, 1999 ; Bonabeau, 2002 ; Macal et North, 2007 ; Parrott, 2008). Dans la littérature, une distinction s'est ainsi constituée entre ces modèles où les agents sont dits cognitifs, capables d'anticiper et planifier leur comportement et des modèles plus simplement constitués d'agents réactifs (Ferber, 2006). A propos de ces derniers, certains auteurs se réservent d'utiliser le terme *agent* (Macal et North, 2005 ; Parrott, 2008), d'autant qu'il est perçu différemment selon que l'on se place du point de vue du thématique ou de l'informatique (Drogoul et al., 2003). Pour ces raisons, lorsque les agents formalisés dans un Système Multi-Agents sont purement réactifs et désignent des individus, nous préférons qualifier le modèle d'individu-centré. Notons toutefois que dans l'étude des dynamiques de populations et des organisations socio-spatiales, les entités élémentaires du système peuvent, en fonction de nos hypothèses sur les mécanismes moteurs du phénomène étudié, désigner des entités infra-individuelles ou supra-individuelles, comme des groupes d'individus, ou bien encore des objets *agentifiés* (Sanders, 2007).

1.1.3 Le statut épistémologique du modèle et de la simulation en question

Finalement, le *tournant de la complexité* invoqué par E. Morin (1990), en rupture avec le réductionnisme de Descartes qui réduit le tout étudié à ses éléments les plus simples, se démarque aussi progressivement de la démarche systémique classique qui aborde le système de manière holiste sans tenir compte de l'hétérogénéité et des différences de comportements des entités qui le composent. Au cœur de ce nouveau paradigme qui se met en place, se trouve l'idée que la modélisation des systèmes dynamiques est un moyen essentiel de la compréhension de la réalité et de la construction théorique. Elle permet d'étudier les mécanismes internes à la *boîte noire* (Susser et Susser, 1996b), reliant facteurs causaux et configurations observées.

Il apparaît donc que la science de la complexité repose entièrement sur des modèles de simulation de la réalité, qui impliquent des décisions de la part du modélisateur sur *ce qui est*, sur *ce qui est utile à représenter ou à laisser de côté* à travers les hypothèses simplificatrices, mais également *sur comment le formaliser*. (O'Sullivan, 2004 ; Becker et al., 2005 ; Livet et al., 2008). L'exploration de la complexité liée à l'articulation des processus entre les niveaux nous pousse à élaborer des modèles selon une approche *générationnelle* ou constructiviste, dont les propriétés observées à un niveau sont dérivées des propriétés de ses entités constitutives formalisées à un niveau inférieur. Or aujourd'hui, grâce au calcul distribué sur une grande quantité d'ordinateurs, l'informatique permet de simuler des systèmes complexes qui mettent en jeu des millions de constituants hétérogènes en interaction (Bonabeau et Theraulaz, 1994 ; Epstein et Axtell, 1996).

Cela ne va pas sans poser de question sur l'évolution du statut épistémologique du modèle, dans son rôle intermédiaire entre théorie et empirie (Le Moigne, 1999 ; Becker et al., 2005 ; Varenne, 2007). Les premiers modèles dynamiques, par leur fort degré d'abstraction, ont longtemps été considérés comme relevant du domaine de la théorie, permettant l'*exploration conceptuelle* des propriétés internes au modèle dans une démarche hypothético-déductive, à l'opposé des méthodes empiriques observationnelles et expérimentales (Phan et Varenne, 2009). En contraste, les dernières avancées des sciences computationnelles offriraient la possibilité nouvelle d'appliquer une démarche expérimentale classique à une classe de phénomènes

macroscopiques jusqu'alors non manipulables, au travers de simulations réalisées sur des modèles volontiers perçus comme représentant une *réalité virtuelle* ou une *empirie simulée* (Axelrod, 1997 ; Epstein, 1999 ; Deffuant et al., 2003 ; Janssen et Ostrom, 2006). En effet, le développement d'un modèle procède d'une même démarche d'isolement du ou des mécanismes causaux présumés que celle réalisée dans le cadre du montage d'une expérience (Mäki, 2005). La différence, de taille, est qu'un modèle simule la réalité *in silico*, au contraire d'une expérience classique où les entités utilisées pour l'expérience sont bien les mêmes que dans la réalité.

Nous nuancerons les deux positions contrastées que nous venons d'exposer en rappelant le rôle assigné au modèle par M. L. Minsky (1965, p. 45, trad. pers.) et qui demeure inhérent à toute entreprise de modélisation : « *aux yeux d'un observateur B, un objet A* est un modèle d'un objet A dans la mesure où B peut utiliser A* pour répondre à des questions qu'il se pose sur A* ». La modélisation impliquera donc toujours une simplification de la réalité, réalisée d'après nos hypothèses sur les entités, leurs propriétés et leurs relations, suffisantes à intégrer pour répondre à une question posée. Par ailleurs, quel que soit le degré d'abstraction du modèle, l'utilisation d'une démarche expérimentale est justifiée dès lors que les règles de fonctionnement du modèle, bien que totalement spécifiées, ne permettent pas de déduire son comportement global, ce qui est inhérent à tout système complexe. Mais, pour que le résultat d'une expérience réalisée sur un modèle puisse être inféré à la réalité, il reste un pas à franchir, celui de la validité du modèle à l'égard de ce qu'il est censé représenter. Cette question n'est pas évidente à traiter dans la mesure où un modèle n'a pas vocation, justement, à représenter toute la réalité, nous y reviendrons (chapitre 5, 3.3.1).

Les différents éléments évoqués ci-dessus vont dans le sens de la nécessité d'adopter une démarche de modélisation privilégiant la parcimonie plutôt que l'exhaustivité, intégrant pas à pas les détails de la réalité en vue de mesurer leurs effets sur le comportement global. C'est de cette manière qu'a progressé la science de la complexité, c'est de cette manière que peut progresser la connaissance sur l'expression d'une maladie transmissible comme la peste à travers la modélisation, comme nous nous attacherons à le montrer en suivant.

1.2. Reconnaissance et modélisation de la complexité de l'expression des maladies transmissibles

1.2.1 La reconnaissance ancienne des non-linéarités dans un système pathogène

Le *tournant de la complexité*, dont nous venons d'esquisser les contours, passe par une rupture avec le paradigme linéaire, qui réduisait la connaissance d'un tout à la connaissance des éléments qui le composent. Pour comprendre la complexité de l'expression globale d'une maladie transmissible, un premier pas essentiel à franchir consiste donc à préciser d'où proviennent les non-linéarités du système de transmission, qui mettent à bas l'hypothèse réductionniste selon laquelle une population se résume à la somme de ses individus et qu'il suffit de comprendre les déterminants des cas de manière isolée pour comprendre la distribution globale de la maladie (Gatrell, 2005 ; Koopman, 2005a ; Pearce et Merletti, 2006).

La question de la non-linéarité peut-être traitée sous deux angles particuliers qu'il convient de ne pas confondre (Koopman et Longini, 1994) : un premier angle, largement discuté dans notre deuxième chapitre, concerne la recherche de causalités établies, dans la tradition observationnelle, d'après l'association entre un ou plusieurs facteurs de risque et un niveau de présence de la maladie. De ce point de vue, les relations non-linéaires sont celles qui ne satisfont pas à deux principes. Celui de proportionnalité, d'une part, entre un facteur de risque X et un niveau d'infection Y ; celui de superposition, d'autre part : si Y est linéairement associé à plusieurs facteurs $X_1...X_n$, la probabilité de Y est la somme de ses composantes affectées de leur coefficient respectif (Philippe et Mansi, 1998). La nature multiple, multi-niveaux et non-linéaire de la causalité est aujourd'hui largement prise en compte dans les modèles statistiques de l'épidémiologie observationnelle.

Un deuxième angle d'analyse des non-linéarités dans un système pathogène concerne la dynamique de l'infection. Or, dans la plupart des études épidémiologiques observationnelles classiques, l'on peut traiter de causalités non-linéaires entre facteurs de risque et niveaux de maladie tout en supposant implicitement que la maladie est

générée par des processus qui sont linéairement dynamiques, c'est-à-dire qu'on considère que les états épidémiologiques des individus qui composent la population étudiée sont des faits indépendants les uns des autres (Koopman, 1996). Pourtant, la reconnaissance que la probabilité pour un individu de développer la maladie dépend de celle des autres individus de la population, et donc rétroactivement de la prévalence globale, est très ancienne et d'une évidence particulière pour les maladies transmissibles (Ross, 1916) : ce sont précisément les contacts contagieux qui se produisent entre les individus infectieux et susceptibles au cours du processus de transmission qui rendent l'expression d'une maladie transmissible non-linéaire et complexe.

Cela fait qu'au niveau individuel, un individu peut être affecté par les facteurs de risque d'un autre individu, avec lequel il entre directement ou indirectement en contact. Il existe en effet de nombreuses façons, pour un facteur de risque, d'exercer une influence indirecte sur un individu sain, outre une influence directe sur sa susceptibilité biologique, qui va modifier sa probabilité de développer l'infection. Un facteur de risque peut :

- altérer la contagiosité d'un individu infectieux avec lequel l'individu sain entre en contact (la quantité d'agents pathogènes excrétés par cet individu infectieux) ;
- modifier la quantité ou la viabilité des pathogènes entre le moment où ils sont excrétés par l'individu contagieux et le moment où ils atteignent l'individu susceptible ;
- modifier le nombre de contacts réalisés par l'individu susceptible ;
- modifier les individus contactés par cet individu sans en changer le nombre (Koopman et Longini, 1994).

Au niveau de la population, la non-linéarité du processus de transmission, qui repose sur les contacts individuels, fait que la prévalence de l'infection dans la population ne dépend pas seulement de l'exposition indépendante des individus à des facteurs de risque identifiés mais aussi de la structure des contacts qui se réalisent entre eux (Koopman et Lynch, 1999). Certains facteurs de risque concernent donc des propriétés structurelles des populations, irréductibles au niveau individuel.

On perçoit là toute la limite des traditionnelles études observationnelles à saisir la complexité de l'expression d'une maladie transmissible, qui, en plus des difficultés liées à la disponibilité de données suffisamment détaillées, peinent à reconnaître les non-linéarités générées par le processus de transmission (Koopman, 1996). Réalisées au niveau individuel ou au niveau écologique, elles se révèlent d'abord impuissantes à examiner simultanément les effets de facteurs de risque individuels et structurels sur les états épidémiologiques observés. Le développement des modèles statistiques multi-niveaux permet de contrer en partie cette limitation, par le fait qu'ils permettent de prendre en compte des variables définies au niveau du groupe au côté de variables individuelles. Mais leur limitation majeure est qu'ils ne considèrent pas facilement la dimension temporelle. Ils ne prennent donc pas facilement en compte les relations dynamiques et rétroactives entre facteurs de risque et états épidémiologiques (Diez-Roux et Aiello, 2005). Au niveau individuel pourtant, l'infection à l'instant t d'un individu constitue, à l'instant $t+1$, un facteur de risque supplémentaire pour un autre individu de son entourage, tandis qu'au niveau de la population, l'incidence de la maladie entre t et $t+1$ dépend de la prévalence de la maladie au temps t .

La reconnaissance de la complexité de l'expression endémo-épidémique d'une maladie transmissible comme la peste, qui réside dans les interactions, définies par les contacts contagieux, entre ses entités constitutives, rend donc son approche par la modélisation dynamique pertinente. Elle passe par l'élaboration de modèles qui intègrent au minimum la dimension temporelle d'une part, une hypothèse simplificatrice relative à la distribution spatiale des entités individuelles et à leurs contacts d'autre part, pour simuler les processus dynamiques de la maladie dans la population. De cette manière, comme le résument J. S. Koopman et J. W. Lynch (1999, p. 1173, trad. pers.), « *la caractéristique qui distingue le mieux les modèles dynamiques du système de transmission du modèle statique de l'ensemble de causes suffisant [Rothman, 1976 ; Rothman, 2005] est qu'ils intègrent les effets non-linéaires au niveau de la population.* »

1.2.2 Modélisation de la dynamique non-linéaire d'une maladie transmissible dans une perspective holiste : l'hypothèse simplificatrice de l'action de masse

La reconnaissance ancienne des non-linéarités liées à la dynamique des maladies transmissibles s'est faite à travers le développement d'une première classe de modèles dynamiques, dont la formalisation remonte à environ un siècle, avec les travaux fondateurs de R. Ross (1916) et W. O. Kermack et A. G. McKendrick (1927) notamment, succédant de peu aux découvertes pasteuriennes (Heesterbeek, 2005). Ces premiers modèles, mathématiques, ont joué un rôle essentiel dans la construction de la théorie en épidémiologie (Anderson et May, 1991) et plus largement en biologie des populations (Murray, 1993). Pourtant, l'épidémiologie théorique a représenté, jusque dans les années 1970-80, une part faible de la recherche épidémiologique, évoluant dans l'ombre et entretenant peu de rapports avec l'épidémiologie observationnelle (Koopman, 2005a ; Valleron et Schwartz, 2006). La raison principale de cette relative déconnexion entre formalisation théorique et données de terrain est le fort degré d'abstraction de modèles abordant la maladie de manière holiste, directement au niveau de la population. S'ils ont été conçus par des épidémiologistes, ils ont par la suite été repris et analysés par des bio-mathématiciens pas toujours soucieux de les confronter aux données de terrain, tandis que les épidémiologistes focalisaient de plus en plus leurs recherches sur l'identification de causes aux niveaux individuel et inférieur (Koopman, 2005b).

Sans entrer de suite dans sa description détaillée, le modèle épidémique le plus simple (Kermack et McKendrick, 1927) décrit la dynamique d'une maladie transmissible dans une population d'hôtes par un système d'équations à variables d'état. Chaque variable décrit le nombre d'individus se trouvant aux différents états épidémiologiques en fonction du temps. Tout comme les premiers modèles de l'écologie théorique, du modèle logistique, décrivant la variation de la taille d'une seule population, aux modèles décrivant la dynamique de population de multiples espèces interagissantes, les premiers modèles épidémiques sont fondés sur une hypothèse simplificatrice principale, celle de densité-dépendance. L'incidence est décrite par la loi de l'action de masse qui simplifie considérablement le processus de rencontre entre les individus en les assimilant à des molécules se déplaçant dans un gaz ou une solution. Selon cette hypothèse empruntée à la chimie réactionnaire (Ross,

1916 ; Heesterbeek, 2005), les individus se rencontrent au hasard et le nombre de nouveaux cas d'infection à l'instant $t+1$ est proportionnel au produit des densités des sous-populations susceptible et infectieuse à l'instant t . C'est la manière la plus simple de représenter directement au niveau de la population, dans un système d'équations différentielles, les effets non-linéaires des interactions se produisant au niveau individuel.

L'épidémiologie théorique des maladies transmissibles, entièrement fondée sur des modèles du système de transmission, procède donc d'une démarche qui consiste à ignorer volontairement certains aspects de la réalité de sorte que l'essentiel du phénomène étudié puisse être saisi et mis en lumière (Mollison, 1995). A partir de là, différentes voies de complexification du modèle mathématique de base sont possibles, par l'intégration progressive de différentes sources d'hétérogénéité inhérentes aux entités individuelles, à leur distribution et à leurs contacts, pour explorer, à travers la simulation, des questions relatives à l'expression des différentes maladies. Parmi celles-ci figure la question de l'instauration de régimes périodiques réguliers ou chaotiques (Grenfell et al., 1995 ; Earn et al., 2000 ; May, 2000).

1.2.3 Le potentiel unificateur de la modélisation individu-centrée dans l'étude de l'expression des maladies transmissibles

Fondamentalement, les processus de contact et de transmission d'une maladie sont des conséquences des comportements d'individus distribués spatialement. Or, l'adoption d'un point de vue holiste sur la dynamique de la maladie amène à simplifier considérablement les interactions multiples se produisant localement entre des entités nombreuses, hétérogènes et évolutives. L'influence grandissante du paradigme de la complexité, posant la question de l'articulation des processus selon les niveaux, conduit l'épidémiologie théorique à évoluer dans le sens d'une meilleure prise en compte de la diversité biologique des individus et de la structure de leurs interactions socio-spatiales (Koopman, 2004 ; Keeling, 1999 ; Keeling, 2005a ; Bansal et al., 2007). L'idée sous-jacente est que cette diversité peut profondément affecter la dynamique de l'infection. Dans ce sens, les méthodes récentes de modélisation informatique distribuée, des automates cellulaires simples aux modèles à base d'agents adaptatifs, conduisent à l'élaboration de modèles épidémiologiques

avec une approche *bottom-up*, dont les propriétés émergentes observées au niveau de la population sont dérivées des propriétés formalisées au niveau individuel. Ces modèles individus-centrés participent donc à une classe de modèles dits *structurés*, qui contrastent sur deux aspects essentiels avec les modèles holistes *non structurés* fondés sur l'hypothèse de l'action de masse :

- les individus sont uniques ; ils diffèrent les uns des autres selon différents aspects considérés dans le modèle ;
- leurs interactions écologiques ne se résument pas à des rencontres aléatoires ; le principe de la localité prime, ce qui signifie qu'un individu particulier est affecté avant tout par les organismes qui se trouvent dans son voisinage spatio-temporel (DeAngelis et Gross, 1992 ; de Roos et Persson, 2005 ; Odenbaugh, 2005).

Le développement de modèles devenant progressivement plus réalistes, considérant des individus distribués dans un réseau de contacts, pousse à un rétablissement du lien, dans l'étude des maladies transmissibles, entre formalisation théorique et observations de terrain. Il pousse à une convergence de deux traditions qui ont longtemps évolué séparément au sein de la discipline épidémiologique : d'un côté, celle d'une épidémiologie observationnelle et expérimentale qui identifie les facteurs de risque et mécanismes de développement de la maladie au niveau individuel sans considérer les interdépendances des états individuels, en lien avec la dynamique globale ; d'un autre côté, celle d'une épidémiologie théorique qui, à ses débuts, formalisait des modèles du système de transmission au niveau de la population sans considération des variations de susceptibilité et d'exposition individuelle, avec un taux de contact et une probabilité de transmission constants (Mollison et al., 1994 ; Koopman et Lynch, 1999 ; Koopman, 2004 ; Koopman, 2005a). Par ailleurs, le développement de modèles de plus en plus détaillés, qui reposent sur une description naturelle et flexible de *qui est susceptible d'entrer en contact avec qui*, les rend plus crédible pour aborder les problèmes liés au contrôle de l'infection. Ils apportent la possibilité d'évaluer, par la simulation, les effets de différentes mesures d'action potentielles qui tiennent compte de la structure des contacts.

Au-delà de la discipline épidémiologique, la modélisation de type *bottom-up* permet l'intégration des connaissances issues des différents champs disciplinaires concernés

par l'étude des maladies transmissibles à différents niveaux d'organisation. En reconnaissant le niveau individuel comme le niveau fondamental auquel le processus de transmission se réalise, elle favorise par exemple l'intégration des connaissances issues des recherches biomoléculaires en évaluant les effets, sur la dynamique globale, des composants génétiques et des mécanismes internes à l'individu susceptibles de modifier sa susceptibilité ou sa contagiosité. Tout comme elle permet d'évaluer les effets, sur la dynamique de la maladie, de facteurs environnementaux, par exemple climatiques, en supposant que leurs effets écologiques sur les mécanismes individuels, dans le sens *top-down*, soient connus. Enfin, la modélisation distribuée et les Systèmes Multi-Agents particulièrement, dans la mesure où ils comportent une dimension socio-spatiale explicite et offrent la possibilité d'implémenter des comportements individuels diversifiés et adaptatifs, amènent à intégrer les connaissances issues des sciences sociales, susceptibles d'influencer les réseaux de contacts et donc les processus endémo-épidémiques (Daudé, 2004 ; Daudé, 2005 ; Daudé et Eliot, 2005 ; Eliot et Daudé, 2006 ; Sanders, 2007 ; O'Sullivan, 2008).

Comme nous l'évoquions en introduction de notre deuxième partie, l'épidémiologie spatiale, utilisant les méthodes traditionnelles de l'analyse spatiale, était habituée à caractériser les configurations endémo-épidémiques observées au niveau global pour en rechercher les causes et processus sous-jacents (Mayer, 1983 ; Gesler, 1986). La modélisation distribuée lui apporte une nouvelle orientation constructiviste et expérimentale, où l'on explore les configurations simulées de la maladie au regard de nos hypothèses sur la distribution de la population, les comportements individuels et les structures d'interaction (Gatrell, 2005). Une telle démarche paraît à même de rendre compte des effets non-linéaires, au niveau global, dus au processus de transmission au niveau individuel et qui font :

- qu'un même résultat peut être produit à partir d'un grand nombre de configurations initiales possibles ;
- qu'à partir d'une même situation initiale, les lois d'interactions ont des effets globaux non triviaux et parfois inattendus.

1.3 Synthèse du complexe pathogène et du système pathogène : le système pathogène complexe

En guise de bilan, le *tournant de la complexité* apporte de nouveaux fondements théoriques et méthodologiques à l'étude de l'expression des maladies transmissibles. Dans une évolution du questionnement poussant à étudier les mécanismes endémo-épidémiques de la maladie à une résolution fine, les méthodes observationnelles classiques échouent à prendre en compte de manière adéquate l'articulation du risque individuel et du risque au niveau de la population. En réponse à cette limite, à laquelle il faut ajouter la difficulté liée au suivi, sur le terrain, des populations concernées, le tournant de la complexité se fonde sur une longue tradition de modélisation des systèmes dynamiques pour un dépassement du réductionnisme et une meilleure prise en compte des non-linéarités du processus de transmission.

Cela nous amène à opérer un glissement conceptuel vers la considération d'un système pathogène qualifié de complexe (Gatrell, 2005 ; Koopman, 2005a ; Pearce et Merletti, 2006). Le système pathogène complexe propose en quelque sorte une synthèse du complexe pathogène (Sorre, 1933) et du système pathogène (Picheral, 1982), concepts que nous avons successivement présentés dans notre première partie (chapitre 1, 3.1 ; chapitre 2, 1.2) : en effet, il réhabilite d'une part la notion d'interaction parasitaire entre les entités individuelles des différentes populations d'hôtes et de vecteurs qui forment le complexe pathogène. D'autre part, il associe à ces entités individuelles, aux populations qu'ils forment et à leur environnement les multiples facteurs de risque, identifiés à différents niveaux, qui composent le système pathogène.

Ce glissement conceptuel étant fait, la simulation permet d'expérimenter *in silico* sur les mécanismes endémo-épidémiques de la maladie, ce qui apparaît impossible dans la réalité. Elle repose sur des modèles qui formalisent, selon des hypothèses simplificatrices, les comportements des individus, leurs contacts, la transmission et l'évolution de la maladie chez ces individus. En référence à la célèbre recommandation d'Einstein de garder les choses aussi simples que possibles mais non simplistes (Calaprice, 2000, p. 314), une bonne raison pour qu'un modèle

épidémiologique reste aussi simple que possible est que la complexité réaliste peut être contreproductive à la compréhension et la formalisation théorique. Déterminer les effets de chaque détail consomme en effet de la ressource et détourne l'attention sur les éléments essentiels. A l'opposé, une autre bonne raison pour qu'un modèle ne soit pas simpliste est qu'il peut amener à ignorer des détails essentiels ce qui, au final, peut avoir pour conséquence de prendre de mauvaises décisions qui peuvent coûter des vies. Pour éviter l'un et l'autre des travers, il faut développer les modèles pas à pas, comparer les modèles entre eux (Rahmandad et Sterman, 2008), de sorte qu'il soit possible d'évaluer la robustesse des inférences réalisées, en regard des hypothèses simplificatrices introduites (Durett et Levin, 1994 ; Manson, 2002 ; Koopman, 2005a).

2. Les apports de l'épidémiologie théorique : intégration progressive de l'hétérogénéité spatiale et de l'individualité dans les modèles endémo-épidémiques

L'exposé des grands principes du *tournant de la complexité* appliqué à l'étude des processus endémo-épidémiques des maladies transmissibles nous amène à examiner plus en détail les principaux apports d'une épidémiologie théorique fondée sur le développement et l'exploration des propriétés de modèles dynamiques du système de transmission. Sa progression s'est faite dans un sens *top-down*, qui a consisté à introduire progressivement dans les modèles différentes sources d'hétérogénéité inhérentes aux entités individuelles et à leur distribution spatiale notamment, pour en explorer les effets sur l'expression de la maladie au niveau de la population.

Les travaux de cette épidémiologie théorique ont été principalement guidés par un questionnement relatif aux conditions pour qu'une maladie introduite dans une population puisse d'une part évoluer en une épidémie, et d'autre part s'y pérenniser à

un niveau endémique, là où précisément les méthodes observationnelles ont montré leurs limites (chapitre 3, 3.3).

2.1 Un apport des modèles compartimentaux : le théorème du seuil endémo-épidémique

Comme nous l'avons déjà évoqué plus haut, les premiers modèles dynamiques de maladies transmissibles sont des modèles compartimentaux, formalisés par des équations différentielles qui modélisent des populations continues plutôt que des individus discrets et simplifient le processus de contact et de transmission au niveau individuel par la loi de l'action de masse. Chaque équation décrit la variation de la taille des différentes classes de population se trouvant aux différents états épidémiologiques en fonction du temps, en considérant l'hypothèse d'homogénéité à l'intérieur de chaque classe, c'est-à-dire le fait que tous les individus sont strictement identiques.

2.1.1 Le modèle basique de Kermack-McKendrick à une seule population d'hôtes et ses extensions

2.1.1.1 Le modèle épidémique SIR

2.1.1.1.1 Les hypothèses simplificatrices du modèle

Le modèle épidémiologique de base, initialement formulé par W. O. Kermack et A. G. McKendrick (1927), est un modèle compartimental déterministe de type SIR, qui considère la dynamique d'une maladie directement transmissible dans une seule population d'hôtes divisée en trois classes selon les différents états épidémiologiques individuels considérés (figure 15) :

- la classe des individus susceptibles S
- celle des infectés I , où tous les individus infectés sont supposés infectieux, c'est-à-dire capables de transmettre la maladie sans distinction d'un état où l'infection est

latente entre le moment de l'infection et le début de la contagiosité

- celle des retirés R , qui comprend indistinctement les individus morts de la maladie et ceux rétablis et immunisés, de sorte qu'ils ne puissent plus être à nouveau infectés.

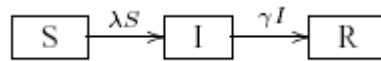


Figure 15. Classes de population et flux du modèle épidémique classique (d'après Kermack et McKendrick, 1927)

Ce modèle de base est un modèle épidémique, conçu pour mettre l'accent sur la dynamique épidémique d'une maladie à court terme, dans une population de taille N constante qui ne se renouvelle pas, sans naissance ni décès liés au cycle de vie naturel. La dynamique épidémique est représentée par les trois équations différentielles suivantes, qui décrivent la variation d'abondance des trois classes de population et les flux d'individus de l'une à l'autre :

$$\begin{aligned}\frac{dS}{dt} &= -\lambda S \\ \frac{dI}{dt} &= \lambda S - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

Concernant le flux de la classe susceptible à la classe infectieuse, l'incidence, ou nombre de nouveaux cas d'infection par unité de temps, est déterminée par le taux d'infection λ auquel les individus susceptibles deviennent infectieux. Dans le modèle traditionnel, ce taux est formulé dans l'hypothèse d'une transmission réalisée suivant une loi de pure action de masse, c'est-à-dire qu'il est proportionnel à la prévalence ou nombre des individus infectieux :

$$\lambda = \alpha I$$

α est un coefficient d'action de masse sans signification épidémiologique particulière, mais qui traduit implicitement le fait que le taux de contact β , ou nombre moyen de contacts par personne et par unité de temps, augmente linéairement avec la taille de la population N .

La plupart des auteurs qui fondent leurs travaux sur ce modèle original utilisent une version de la fonction de transmission différente en considérant l'hypothèse de pseudo action de masse, fréquence-dépendante (Hethcote, 2005 ; Keeling, 2005a ; Choisy et al., 2007). Selon cette hypothèse alternative à la pure action de masse, dite densité-dépendante, l'incidence est liée à la proportion de la population infectieuse (I/N), de sorte que chaque individu susceptible rencontre à chaque pas de temps un nombre fixe d'individus β indépendant de N . C'est la proportion d'entre eux qui sont infectieux qui détermine le taux d'infection :

$$\lambda = \beta \frac{I}{N}$$

Pour les maladies humaines et animales, cette hypothèse apparaît plus plausible que la première évoquée, dans la mesure où la transmission de la maladie se réalise avant tout localement, si bien que le nombre et la configuration des rencontres journalières sont peu liés à la taille de la population totale (Hethcote, 2005). Mais lorsque la population reste constante, les résultats de la version fréquence-dépendante varient peu de ceux de la version densité-dépendante (Keeling, 2005a). Quelle que soit la version de ce modèle de base, une conséquence fondamentale du formalisme à équations différentielles et de la loi de l'action de masse est que les rencontres, qui se font au taux β , sont supposées aléatoires. Cela signifie que la population est constamment mélangée de façon homogène et que tout individu a la même probabilité d'entrer en contact avec n'importe quel autre individu. De plus, le contact est un processus instantané sans considération de sa possible inscription dans la durée. Enfin, tout contact $S-I$ cause une transmission, après quoi la population est remélangée.

Le flux de la classe des susceptibles à celle des retirés, qui agrège, au niveau global, les évolutions et issues individuelles de la maladie, est déterminé par un taux de retrait γ constant. Cela revient à supposer une distribution exponentielle des périodes

infectieuses individuelles, dont la moyenne est $1/\gamma$: à chaque unité de temps, un nombre constant d'individus quitte la classe des infectieux sans considération de la période qui s'est écoulée depuis l'infection.

2.1.1.1.2 Un seuil épidémique défini sur la population susceptible initiale

Dans sa forme la plus simple, le modèle de Kermack-McKendrick permet d'aborder une série de questions qui concernent la dynamique temporelle d'une épidémie (Anderson et May, 1991, Diekmann et Heesterbeek, 2000) : étant spécifiés le taux d'infection et le taux de retrait, l'introduction de l'agent pathogène dans une population d'hôtes conduira-t-elle au développement d'une épidémie ? Quelle sera son cours et notamment quand s'amorcera son déclin jusqu'à son extinction ?

Une variable-clé peut alors être définie, le taux de reproduction épidémique R (à ne pas confondre avec la densité R des retirés). Défini à un instant t , cette variable mesure le nombre d'infections secondaires causées par un individu infectieux typique introduit à t , durant sa période entière d'infectiosité. Elle s'écrit, pour le modèle fréquence-dépendant :

$$R(t) = s_t \beta \frac{1}{\gamma}$$

où $s_t = \frac{S(t)}{N}$ est la proportion d'individus susceptibles à l'instant t , β est le taux de contact et $\frac{1}{\gamma}$ est la période infectieuse moyenne. Le taux de reproduction épidémique initial $R(0)$ désigne le nombre moyen d'infections secondaires causées par l'introduction d'un individu infectieux à l'instant initial $t = 0$. Le taux de reproduction de base noté R_0 , lui, désigne le taux de reproduction initial dans le cas particulier où un individu est introduit dans une population entièrement susceptible ($s_0 = 1$) (Hethcote, 2005). Il s'écrit alors :

$$R_0 = \beta \frac{1}{\gamma}$$

A partir de là, il est possible de dégager un effet de seuil sur le taux de reproduction

épidémique, qui conditionne le développement ou non d'une épidémie ainsi que son déclin. Ce seuil constitue, autrement dit, une bifurcation générée par les non-linéarités du système d'équations déterministes (Nåsell, 1995). A l'instant initial, où l'on introduit un petit nombre d'individus infectieux dans la population :

- si $R(0) > 1$, un infectieux typique se remplace initialement par plus d'un nouvel individu infectieux. Il se produit donc une épidémie (figure 16) : la proportion i_t des infectieux augmente jusqu'à une valeur maximale calculable tandis que la proportion des susceptibles s_t diminue. Lorsque s_t passe en dessous d'une valeur égale à $1/R_0$ de sorte que $R(t)$ passe en dessous de 1, s'amorce le déclin de l'épidémie. L'extinction est inévitable parce que la population des susceptibles n'est pas renouvelée et s_t décroît toujours.
- Si $R(0) \leq 1$, alors aucune épidémie ne se développera et la population infectieuse ne fera que décroître jusqu'à extinction finale.

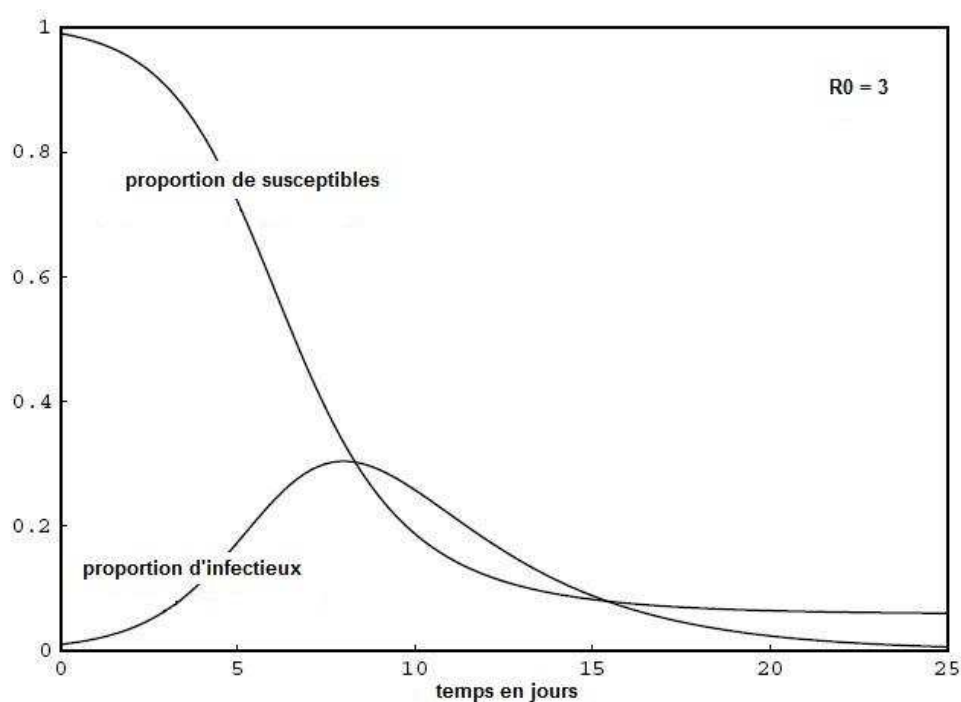


Figure 16. Solutions du modèle épidémique classique SIR avec $R_0 = 3$ et une période infectieuse moyenne $1/\gamma = 3$ jours (source : Hethcote, 2005)

Le taux de reproduction épidémique initial sert donc, dans ce modèle d'action de masse déterministe, à mettre en évidence un seuil épidémique, ou seuil d'invasion, sur la population susceptible initiale. En effet, dans sa version fréquence-dépendante, puisque $R(0)$ dépend de la proportion des susceptibles s_0 , il existe une proportion critique γ/β de cette population au-delà de laquelle, pour un taux de contact et une période moyenne infectieuse fixés, une épidémie se développe. Dans le modèle densité-dépendant, où R à l'initialisation est fonction de la densité ou taille de la population de susceptibles $S(0)$, c'est une densité seuil égale à γ/α qui est nécessaire pour initier une épidémie (Anderson et May, 1991).

2.1.1.2 Le modèle endémique SIR

Pour aborder la question de l'endémicité et étudier les conditions limites de persistance à long terme d'une maladie transmissible dans une population d'hôtes, il faut considérer un système ouvert et introduire la possibilité pour cette population de se régénérer. Pour cela, une évolution est apportée au modèle épidémique de base (figure 17). Elle consiste à introduire un flux de naissances ou d'émigration dans la classe des susceptibles à un taux μ , et un flux de sortie de décès ou d'immigration, concernant toutes les classes S , I , R , au même taux μ , de sorte que la population N reste constante (Anderson et May, 1991, Diekmann et Heesterbeek, 2000 ; Hethcote, 2005 ; Choisy et al., 2007).

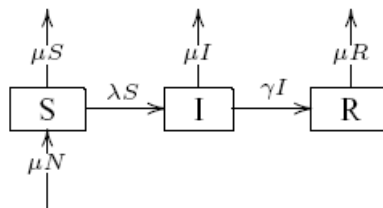


Figure 17. Classes de population et flux du modèle endémique classique

Dans ce cas, le taux de reproduction épidémique s'écrit $R(t) = s_t \beta \frac{1}{\gamma + \mu}$ et le système

possède deux points d'équilibre stable. A l'initialisation, où l'on introduit un petit nombre d'individus infectieux dans la population, l'atteinte de l'un ou l'autre dépend alors précisément du taux de reproduction de base R_0 (Hethcote, 2005) :

- si $R_0 \leq 1$, la maladie décroît et atteint le point d'équilibre donné par $s_t = 1$ et $i_t = 0$, c'est-à-dire celui qui correspond à l'extinction de l'infection.
- Si $R_0 > 1$, avec i_0 suffisamment petit et s_0 suffisamment grand pour que $R(0)$ dépasse l'unité, alors la proportion i_t des infectieux augmente jusqu'à un premier pic épidémique, puis décroît jusqu'à un minimum, le temps que la population susceptible se reconstitue. Lorsque s_t atteint un niveau suffisamment élevé pour que $R(t)$ atteigne 1, une deuxième épidémie se produit, de moindre intensité que la première. Ce processus d'oscillations épidémiques d'amplitude allant diminuant se poursuit jusqu'à ce que la maladie atteigne un point d'équilibre endémique stable. Le taux de reproduction épidémique $R(e)$ est alors égal à un et les proportions d'individus susceptibles s_e et infectieux i_e sont déterminées en fonction des trois paramètres : le taux de contact β , le taux de retrait γ et le taux de renouvellement de la population μ du modèle.

Dans ce modèle endémique SIR déterministe, le devenir d'une maladie introduite dans une population d'hôtes et sa capacité à se maintenir sont ainsi complètement contrôlés par ces trois paramètres qui déterminent la valeur du taux de reproduction de base de l'infection R_0 . Dans la version fréquence-dépendante du modèle, où chaque individu établit des contacts avec un nombre fixe d'individus β indépendant de la taille de la population N , R_0 est indépendant de N . Au contraire, dans la version densité-dépendante où R_0 dépend de N , le modèle permet de définir un seuil critique sur la population initiale entièrement susceptible au-delà duquel, si on y introduit un individu infectieux, la maladie pourra se développer jusqu'à se stabiliser à un niveau endémique.

2.1.1.3 Une déclinaison possible du modèle SIR : le modèle SEIR

Une voie possible de complication du modèle SIR de base est la remise en cause de l'hypothèse irréaliste selon laquelle tout individu infecté peut immédiatement transmettre l'infection, tout en gardant celle intacte d'un processus de transmission se

réalisant selon la loi de l'action de masse. Cela se fait par la considération d'une classe supplémentaire au modèle mathématique, celle des individus exposés, c'est-à-dire infectés mais non encore infectieux, chez qui l'infection est latente (figure 18). La considération d'un taux de transfert ϵ constant de la classe des exposés à celle des infectieux implique alors l'hypothèse d'une distribution exponentielle des périodes de latence individuelles, tout comme l'est celle des périodes infectieuses. Cette évolution du modèle SIR affecte la dynamique globale de la maladie par l'introduction d'un décalage temporel. Elle ne modifie cependant pas le théorème du seuil défini sur la population initiale pour qu'une épidémie se développe et, dans le cas où la population se renouvelle, se maintienne à l'état endémique.

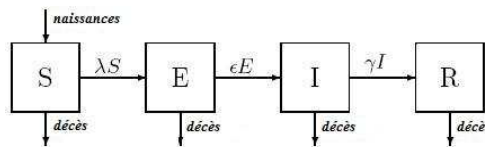


Figure 18. Classes de population et flux du modèle endémique SEIR

Ajoutons que le modèle SEIR n'est qu'une des déclinaisons possibles du modèle SIR de base. D'autres sont possibles, selon les caractéristiques des maladies modélisées. La considération d'une immunité temporaire amène à ajouter un flux dirigé de la classe des retirés vers celle des susceptibles (modèles endémiques de type SIRS, SEIRS). Guérison totale et absence d'immunité donnent par ailleurs des modèles de type SIS ou SEIS (Hethcote, 2005).

2.1.2 Le modèle de Ross-MacDonald d'une maladie transmissible vectorielle

Le modèle de Kermack-McKendrick évoqué précédemment, dans la mesure où il ne considère qu'une seule population d'hôtes et considère les contacts contagieux au hasard, convient le mieux aux maladies directement transmissibles par voie aérienne. Il leur convient d'autant mieux qu'elles nécessitent un contact moins proche et moins prolongé entre les individus. A l'inverse, il apparaît fortement simplificateur du

processus de transmission des maladies indirectement transmissibles, faisant intervenir au minimum une population d'hôtes et une autre de vecteurs, comme la peste. Un autre modèle a donc été formulé, initialement proposé par R. Ross (1911) à propos du paludisme, repris par G. McDonald (1957) puis d'autres (Anderson et May, 1991 ; Rodriguez et Torres-Sorando, 2001 ; Auger et al., 2008 ; Ruan et al., 2008). Ce modèle représente explicitement la dynamique de la maladie dans deux populations, hôtes et vecteurs, ainsi que le processus de transmission de l'une à l'autre, par piqûre d'un vecteur sur un hôte. L'incidence de l'infection dans la population d'hôtes entre t et $t+1$ dépend alors de sa prévalence dans la population de vecteurs au temps t et vice versa pour l'incidence chez le vecteur vis à vis de la prévalence chez l'hôte (figure 19). La transmission se réalise selon un processus qui conserve l'hypothèse de l'action de masse, c'est-à-dire d'un mélange homogène des individus.

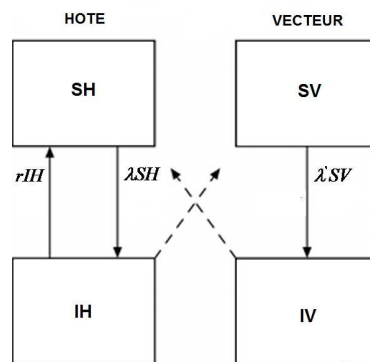


Figure 19. Classes de population et flux du modèle de transmission vectorielle de Ross (SI pour le vecteur, SIS pour l'hôte)

Plus précisément, dans le modèle de R. Ross fréquence-dépendant, le taux d'infection de la population d'hôtes s'écrit : $\lambda = \beta \frac{IV}{NV}$.

β désigne le taux de contact contagieux du vecteur à l'hôte, synthèse de trois paramètres : une fréquence des piqûres a , c'est-à-dire le nombre de piqûres réalisées par un vecteur à un hôte par unité de temps, une efficacité de la transmission b , qui traduit le fait que toutes les piqûres réalisées par un vecteur infectieux ne résulteront

pas forcément en une infection de l'hôte susceptible et m , le ratio du nombre de vecteurs sur le nombre d'hôtes. Ainsi, un hôte aura d'autant plus de chance de s'infecter qu'il subit de multiples piqûres réalisées par de nombreux vecteurs infectieux et $\beta = a * b * m$.

Le taux d'infection de l'hôte au vecteur s'exprime différemment : $\lambda' = \beta' \frac{IH}{NH}$.

En effet, s'il dépend également de la fréquence des piqûres a , un vecteur s'infectera sur un seul hôte et l'efficacité de la transmission c dans ce sens hôte-vecteur n'est pas forcément la même que dans le sens inverse. Ainsi, $\beta' = a * c$.

A partir des hypothèses réalisées sur le processus de transmission, la condition d'invasion peut toujours être déterminée dans le modèle de Ross de manière analytique, par l'expression mathématique du taux de reproduction de base : nombre moyen de cas secondaires chez l'hôte (ou le vecteur) indirectement produits *via* le vecteur (ou l'hôte) suite à l'introduction d'un premier cas infectieux dans une population entièrement susceptible (Tran et al., 2005 ; Ruan et al., 2008). En fonction des facteurs de la transmission définis précédemment, pour une population de vecteurs qui ne se remet pas de l'infection et se renouvelle à un taux μ constant et pour une guérison totale de la population d'hôtes au taux r , R_0 s'exprime alors ainsi :

$$R_0 = \frac{a^2 b c m}{r \mu}$$

Cette expression vaut dans le cas d'un modèle qui n'inclut pas les états latents, ni chez l'hôte, ni chez le vecteur. La latence de la maladie chez le vecteur sera ajoutée au modèle par MacDonald (1957). En effet, dans la mesure où le vecteur est généralement un insecte, dont la période de latence est du même ordre que l'espérance de vie, la considération d'un état latent peut largement modifier l'estimation de R_0 , à la baisse.

Dans la mesure où les deux populations bénéficient d'un renouvellement de leur classe susceptible, le système possède, comme pour le modèle endémique à une seule population, deux points d'équilibre stable : un premier si $R_0 \leq 1$, où l'infection

disparaît, un deuxième, si $R_0 > 1$, où l'infection atteint un point d'équilibre endémique. L'un des principaux résultats du modèle fondateur de R. Ross (1911), à propos du paludisme où la population d'hôtes est une population humaine et les vecteurs sont des moustiques, tient dans l'établissement du fait que la densité des vecteurs, reliée au taux de reproduction de base, doit dépasser un seuil théorique critique pour que la maladie puisse persister à un état endémique dans la population d'hôtes.

2.1.3 Intégration de l'hétérogénéité temporelle, statistique et spatiale dans les modèles compartimentaux

Les modèles les plus simples présentés précédemment, qui concernent l'expression d'une maladie transmissible au sein d'une seule population ou bien entre une population de vecteurs et une autre d'hôtes, sont des modèles déterministes non structurés, qui considèrent un processus de transmission se réalisant selon la loi de l'action de masse. En ce sens, ils reposent sur l'hypothèse d'un mélange constamment homogène des individus et sont en quelque sorte *aspatiaux*. Selon qu'ils incluent ou non une dynamique reproductrice vitale et plus largement un renouvellement de la population des susceptibles, ils modélisent l'occurrence ou non, sur le court terme, d'une épidémie isolée ou bien l'établissement ou non, sur le long terme, d'un équilibre endémique après atténuation d'oscillations épidémiques. Cependant, par leurs hypothèses simplificatrices, ils se révèlent insuffisants pour reproduire les mécanismes générant une configuration endémo-épidémique de la maladie, caractérisée par l'apparition plus ou moins régulière d'épidémies sur fond endémique.

2.1.3.1 L'introduction de l'hétérogénéité temporelle et statistique

Plusieurs voies de complexification sont possibles pour développer des modèles véritablement endémo-épidémiques. Une première voie possible est d'introduire de l'hétérogénéité temporelle, dans le sens où l'on fait varier temporellement, tout en restant dans un cadre déterministe, le taux de contact ou d'autres paramètres reflétant les caractéristiques évolutives de la maladie dans l'organisme. Cette voie a par exemple été suivie pour introduire de la saisonnalité dans les modèles de maladies infantiles, par la variation saisonnière du taux de contact, reflétant les périodes

d'ouverture et de fermeture des écoles, ou encore par la variation du taux de recrutement des nouveaux susceptibles, reflétant la variation des naissances et des vaccinations (Grenfell et al., 1995 ; Earn et al., 2000 ; Conlan et Grenfell, 2007). Cette voie convient également aux maladies vectorielles (Lord et al., 1996), particulièrement sensibles aux variations saisonnières du climat, comme nous avons pu le préciser à propos de la peste (chapitre 2, 2.2.1.1). A partir de là et selon le degré de forçage saisonnier, il est possible de représenter une large gamme de comportements dynamiques de la maladie, allant de la coexistence de cycles périodiques réguliers au chaos, les transitions de l'un à l'autre étant expliquées par des bifurcations du système dynamique non linéaire.

Une autre voie possible est d'introduire de la stochasticité dans le modèle de sorte que les événements liés à la dynamique démographique ou à la transmission de la maladie dans la population se produisent à des probabilités dérivées du modèle déterministe (Chen et al., 2005). Cette hétérogénéité statistique peut jouer un rôle important au moment du démarrage de l'épidémie où le nombre initial d'individus infectieux est faible et faire en sorte que $R_0 > 1$ ne garantira pas le développement d'une épidémie (Choisy et al., 2007). Dans un modèle endémique à saisonnalité, les fluctuations stochastiques introduites seront spécialement déterminantes, pour la persistance de la maladie, au moment des creux épidémiques (Grenfell et al., 1995).

2.1.3.2 L'introduction de la dimension spatiale et de l'hétérogénéité spatiale

Une autre forme de complexification des modèles compartimentaux, qui se superpose aux formes que nous venons d'évoquer, concerne l'intégration explicite de la dimension spatiale dans le modèle. Cela amène à remettre en cause l'hypothèse simplificatrice d'un mélange constamment homogène de l'ensemble de la population, qui fait qu'un individu a la même chance d'entrer en contact avec n'importe quel autre individu.

L'introduction des effets liés à la distribution spatio-temporelle de la population dans le modèle de base de Kermack-McKendrick peut se faire selon différentes méthodes (Murray, 1993 ; Lloyd et May, 1996 ; Hagenaars et al., 2004 ; Ruan, 2007). Chacune repose sur des hypothèses différentes à propos de cette distribution, qui détermine la structure des contacts et la dynamique globale de l'infection. Nous évoquerons deux

évolutions particulières du modèle de base : l'une conduisant à étudier les conditions de diffusion épidémique d'une maladie introduite dans une population susceptible, par la propagation d'un front continu dans l'espace ; l'autre amenant à considérer les effets possibles de relocalisation de la maladie et de diffusion hiérarchique, plus à même de reproduire les mécanismes de persistance endémo-épidémique de l'infection.

2.1.3.2.1 Les modèles réaction-diffusion

Une première évolution a consisté à transformer le modèle de Kermack-McKendrick en un modèle spatio-temporel décrivant la variation des densités locales des populations susceptibles $S(x,t)$, infectieuses $I(x,t)$ et retirées $R(x,t)$ au temps t et à la localité x .

A partir de là, une possibilité est de représenter l'interaction entre les trois populations en supposant des individus fixes et une infection se diffusant d'une localité à l'autre (Kendall, 1965 ; Thieme, 1977 ; Ruan, 2007). Un terme de l'équation traduit les contributions des individus infectés dans la localité y à l'infection des susceptibles dans la localité x . Une autre approche consiste à représenter explicitement les déplacements des susceptibles, infectieux et retirés d'une localité à l'autre par les équations de réaction-diffusion suivantes, pour un modèle SI :

$$\begin{aligned}\frac{dS(x,t)}{dt} &= -\lambda S + D\nabla^2 S \\ \frac{dI(x,t)}{dt} &= \lambda S - \gamma I + D\nabla^2 I\end{aligned}$$

Ces équations sont les mêmes que celles du modèle épidémique de base (chapitre 5, 2.1.1.1) où λ est le taux d'infection, γ le taux de retrait, auxquelles on rajoute un terme de diffusion, supposant que les individus susceptibles et infectieux sont caractérisés par un même coefficient de diffusion D . La prise en compte explicite des mobilités permet ainsi de considérer le fait que des individus infectieux localisés en x au temps t ont été infectés dans une autre localité y au temps $t - s$ (Noble, 1974 ; Murray, 1993 ; Kuperman et Wio, 1999 ; Ruan, 2007).

Les apports théoriques des premiers modèles réalisés selon ces deux approches, considérant l'introduction en un point source d'un nombre d'infectieux dans une population susceptible uniformément répartie dans un espace unidimensionnel, ont été de préciser les conditions de développement d'une onde épidémique se déplaçant à une vitesse constante, dont le front sépare une région susceptible d'une région infectée. Plus tard, la modélisation théorique par réaction-diffusion sera étendue pour intégrer les effets dus à l'hétérogénéité spatiale de l'habitat et des mouvements individuels sur le potentiel endémique de la maladie (Gudelj et White, 2004 ; Allen et al., 2008). Ce sont toutefois les approches statistiques, fondées sur les données épidémiologiques, qui ont été les plus utilisées pour intégrer ces hétérogénéités dans des modèles prédictifs (Smith et al., 2002 ; Cliff et Haggett, 2006 ; Adjemian et al., 2007 ; Christakos et al., 2007).

2.1.3.2.2 Les modèles à sous-populations

Une deuxième évolution du modèle théorique de Kermack-McKendrick, qui vise à mesurer l'effet de l'hétérogénéité spatiale sur l'expression endémo-épidémique d'une maladie transmissible, consiste à introduire un cadre métapopulationnel par une plus grande compartimentalisation du système d'équations. Le concept de *métapopulation*, déjà évoqué au chapitre 3 (3.2), a eu beaucoup d'influence en écologie et en épidémiologie théoriques (Hanski et al., 1995 ; Grenfell et Harwood, 1997). Il suppose que la distribution spatiale de nombreuses espèces peut être décrite par la séparation d'une *métapopulation* en n *sous-populations*. Chacune d'elle est caractérisée par sa propre dynamique, tandis qu'un transfert d'individus (et d'une infection) est possible de l'une à l'autre, selon une étroitesse de couplage des sous-populations spécifiée. Cela amène à considérer un modèle compartimental dans lequel chaque classe d'état épidémiologique (S, E, I, R par exemple) est déclinée en autant de classes qu'il y a de sous-populations (ou communautés). Dans ce modèle, les individus sont généralement plus à même d'entrer en contact les uns avec les autres à l'intérieur d'une même sous-population qu'avec d'autres individus de l'extérieur, ce qui amène à s'affranchir de l'hypothèse d'un mélange complètement homogène de la population globale.

A partir de là, comme précédemment, deux formulations sont possibles selon que les mobilités individuelles sont, ou non, explicitement représentées. La première est de

considérer une matrice $[\beta_{i,j}]$ pour exprimer les taux de contacts entre les communautés i et j (figure 20). Celle-ci est définie en fonction de différents critères : distance et taille des différentes communautés, par exemple (Bolker et Grenfell, 1995 ; Grenfell et al., 1995 ; Lloyd et May, 1996 ; Hagens et al., 2004 ; Choisy et al., 2007).

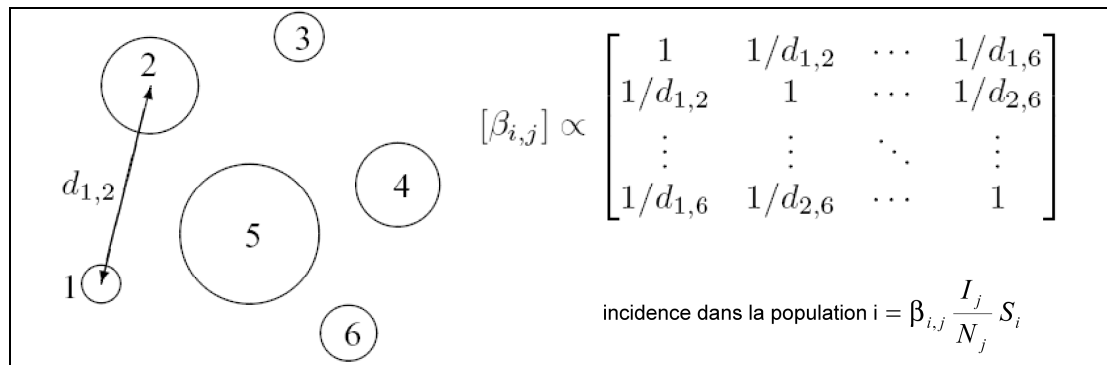


Figure 20. Cadre métapopulationnel dans lequel les taux de la matrice des contacts sont définis de manière inversement proportionnelle à la distance séparant les communautés, représentées par des cercles (d'après Choisy et al., 2007)

Une version plus complexe prendrait en considération le fait qu'un individu infectieux appartenant à la communauté i transmette la maladie à un susceptible de la communauté j , suite à une rencontre dans la localité k (Keeling, 2005a). Dans cette première formulation, les mobilités ne sont pas explicitement représentées et il est supposé que les individus réalisent des visites courtes dans les communautés extérieures. La transmission entre deux individus à l'intérieur d'une même communauté se fait toujours, comme dans le modèle de Kermack-McKendrick, selon la loi de l'action de masse.

Dans une autre formulation, les mobilités sont explicitement représentées ; les individus migrent, plus ou moins durablement, d'une communauté à l'autre, sans que pour autant, une infection puisse avoir lieu durant le processus de migration. La situation est alors celle d'un graphe dirigé où les nœuds représentent les différentes communautés et les liens les possibilités de migration de l'une à l'autre (Murray et

Cliff, 1977 ; Sattenspiel et Dietz, 1995 ; Chalvet-Monfray et al., 1998 ; Artzrouni et Gôûteux, 1999 ; Keeling et Gilligan, 2000a/b ; Rodriguez et Torres-Sorando, 2001 ; Arino et van den Driessche, 2003 ; Brockmann et al., 2005 ; Colizza et al., 2007 ; Auger et al., 2008). En réglant les taux ou probabilités de migration d'un contexte local i à une autre j , on donne une structure particulière au réseau, par exemple hiérarchique, qui affectera la propagation de la maladie (Watts et al., 2005).

Le cadre métapopulationnel ainsi précisé se révèle adéquate pour explorer les conditions de persistance endémo-épidémique d'une maladie transmissible, faisant intervenir la notion de synchronisation des dynamiques locales de l'infection. Lorsqu'en effet les oscillations épidémiques locales deviennent désynchronisées, dans le cas d'une extinction locale de l'infection, un effet de sauvetage peut se produire par son transfert depuis une communauté où la maladie est toujours présente de sorte qu'au niveau global, la maladie se maintient. La désynchronisation est, logiquement, plus susceptible de se produire dans les modèles stochastiques, où chaque communauté vit différentes chaînes d'événements. Elle est également influencée par le degré de couplage, défini par les taux de contact ou de migration entre les différentes sous-populations. Elle caractérisera ainsi davantage les populations composées de communautés de petite taille et modérément isolées. Au contraire, la situation décrite par des communautés de grande taille (moins sensibles à la stochasticité) et fortement interreliées se rapproche davantage de celle d'homogénéité proposée par le modèle de base non structuré (Bonabeau et al., 1998a/b).

Il est important de noter que pour de tels modèles métapopulationnels qui intègrent l'hétérogénéité spatiale et la structuration de la population, il est plus difficile d'obtenir une formule du taux de reproduction de base R_0 que pour le modèle non structuré (Lloyd et May, 1996 ; Diekmann et Heesterbeek, 2000 ; Rodriguez et Torres-Sorando, 2001 ; Hagenaars et al., 2004). Surtout, dans une version stochastique du modèle métapopulationnel et avec un degré de couplage modéré, R_0 peut ne pas avoir de rapport avec la taille finale de l'épidémie, donnée par le cumul des cas d'infection. En effet, si $R_0 > 1$ est une condition nécessaire au démarrage d'une épidémie à l'intérieur d'une sous-population, comme dans le modèle compartimental de base, $R_0 > 1$ ne garantit pas qu'une épidémie se produira ni, dans le cas où elle se déclare, si elle se propagera au-delà du contexte local initial (Watts et al., 2005 ; Cross et al.,

2007). La taille finale, donnée par le cumul des cas, et la durée de l'infection sont largement déterminées par un faible nombre d'événements stochastiques, lorsqu'un individu infectieux introduit la maladie dans une communauté non infectée. Il s'ensuit que durée et taille finale pourront varier d'une simulation à l'autre alors que le taux de reproduction de base reste constant, et inversement : des configurations similaires de la maladie pourront être obtenues à partir de valeurs différentes de R_0 .

Des variables alternatives au taux de reproduction de base ont été définies pour compenser les limites de R_0 que nous venons de souligner. A. D. Cliff et P. Haggett (2006) proposent une version spatiale du taux de reproduction de base : définie au niveau des communautés distinguées dans le modèle, elle mesure la propension, pour une localité touchée par l'infection, à infecter d'autres localités au cours de la période durant laquelle elle reste infectée. Une autre quantité seuil souvent recherchée dans les modèles stochastiques concerne la persistance locale de la maladie : il s'agit de la taille de la communauté requise (*CCS*, ou *critical community size*) pour que la maladie, une fois établie après une période de croissance initiale, ne connaisse pas d'extinction locale due aux fluctuations stochastiques mais se maintienne par elle-même sans apport d'infectieux extérieurs (Bartlett, 1957 ; Grenfell et Harwood, 1997 ; Keeling, 1997 ; Lloyd-Smith et al., 2005 ; Choisy et al., 2007).

2.2 Un apport des modèles à individus discrets : les effets de la structure du réseau de contacts sur le développement épidémique

La progression de la théorie sur l'expression des maladies transmissibles, dans un sens *top-down*, passe par l'introduction de différentes sources d'hétérogénéité inhérentes aux entités individuelles et à la structure de leurs contacts dans les modèles dynamiques. L'approche métapopulationnelle représentait un premier pas dans cette direction, qui a conduit à remettre en cause l'hypothèse de l'action de masse au niveau d'une population globale, en distinguant deux types de contacts : des contacts locaux à l'intérieur d'une même sous-population et d'autres, de longue portée, reliant les différentes sous-populations. Cette approche, qui reporte l'hypothèse de l'action de masse au niveau des contacts intra-communautaires, se révèle adaptée à l'exploration

des effets liés à la structuration en communautés d'une population d'hôtes, mais reste cependant insuffisante pour atteindre le niveau individuel et prendre en compte certains détails à ce niveau qui peuvent profondément affecter l'expression globale de la maladie.

Cela amène à des changements méthodologiques conduisant à formaliser, dans les modèles, le processus de transmission et d'évolution de la maladie au niveau où ils se réalisent : celui de l'individu. En ce qui concerne l'évolution de la maladie, l'intégration de l'individualité dans les modèles épidémiques permet la prise en compte explicite de périodes de latence et d'infectiosité déterminées (Keeling, 1997). On s'affranchit ainsi de l'hypothèse simpliste des modèles compartimentaux d'un taux constant réglant les flux d'individus entre les différents états épidémiologiques, ce qui revenait à considérer une distribution exponentielle des périodes de latence et infectieuses (Choisy et al., 2007). En ce qui concerne la modélisation du processus de transmission, une variété d'approches ont été développées, utilisant un formalisme mathématique ou informatique, visant à s'affranchir de l'hypothèse de l'action de masse et à représenter plus ou moins explicitement le passage de la maladie entre des entités individuelles distribuées dans un réseau d'interaction. Un rapprochement a ainsi eu lieu entre la théorie des épidémies et la théorie des réseaux sociaux, dans l'objectif d'étudier les effets de la topologie du réseau des connexions individuelles sur l'expression endémo-épidémique globale.

2.2.1 Les modèles à réseaux de contacts fixes

Une première catégorie de modèles s'emploie à analyser l'expression de la maladie dans des réseaux théoriques statiques, aux propriétés topologiques spécifiées *a priori* par le modélisateur. Dans ces modèles, les processus de mobilité, de rencontre, de liaison (quel qu'en soit le mode) et de séparation entre les individus ne sont pas explicitement représentés. Le changement de la topologie du réseau est donc supposé intervenir sur une échelle de temps beaucoup plus longue que celle des processus épidémiologiques qui l'affectent (Newman, 2002 ; Koopman, 2005b ; Parham et Ferguson, 2006). L'accent est mis sur l'hétérogénéité des connexions individuelles par lesquelles la maladie peut transiter sans que ne soient considérés, en amont, les

nombreux facteurs sociaux, spatiaux, comportementaux desquels résultent ces connexions.

2.2.1.1 La variété des approches de modélisation employées

Parmi les méthodes employées pour intégrer cette hétérogénéité, une première catégorie regroupe les approches analytiques. Certaines partent du modèle épidémique de base en y intégrant une composante relative à la distribution des individus, pour formuler le taux de reproduction de base notamment (Keeling et Grenfell, 2000 ; Aparicio et Pascual, 2007 ; Bansal et al., 2007).

Une autre approche analytique, fondée sur la théorie de la percolation, prédit, étant données la distribution des connexions individuelles et une probabilité fixée de transmission lors d'un contact contagieux (Susceptible-Infectieux), le sort d'une maladie introduite dans le réseau sans pour autant représenter sa dynamique. Le seuil défini sur le taux de reproduction de base peut en effet être rapproché du seuil de percolation (Moore et Newman, 2000 ; Newman, 2002 ; Meyers et al., 2005).

Au contraire, d'autres approches utilisent un formalisme d'équations différentielles pour représenter cette dynamique (Bansal et al., 2007). Parmi celles-ci, figure la méthode d'approximation par paires, qui considère l'évolution du nombre de paires différentes définies par les différents états épidémiologiques ([SS] [SI] [SR] [II] [IR] [RR], pour un modèle SIR) (Keeling, 1999 ; Keeling et Eames, 2005). Figure également la méthode par laquelle un système d'équations décrit la variation de classes d'individus distinguées selon deux critères : l'état épidémiologique et le degré, ou nombre de connexions partagées avec d'autres individus (Pastor-Satorras et Vespignani, 2001a/b).

En contraste avec les approches analytiques, les approches par la simulation numérique, plus intuitives, représentent explicitement les entités individuelles qui constituent les nœuds du réseau, ainsi que les connexions qui relient ces nœuds, désignant autant de contacts potentiels au cours desquels la maladie peut se transmettre.

Parmi ces approches, les automates cellulaires apportent un cadre de modélisation individu-centré, généralement stochastique, où la probabilité d'infection pour un individu susceptible, représenté par une cellule, est déterminée par le nombre de cellules infectées dans son voisinage. Traditionnellement, les automates cellulaires reposent sur une trame régulière de cellules, où chaque individu est connecté à ses 4 ou 8 plus proches voisins (Mollison et Kuulasmaa, 1985 ; Rhodes et Anderson, 1996 ; Fuentes et Kuperman, 1999). Mais d'autres structures plus irrégulières du réseau de connexion ont été proposées (O'Sullivan, 2001 ; Badaritti et al., 2007) et les effets de ces topologies irrégulières sur la diffusion épidémique et la persistance endémique explorés par le biais de simulations (Read et Keeling, 2003 ; Shirley et Rushton, 2005).

2.2.1.2 Les apports théoriques de ces modèles

Les diverses approches de modélisation à réseaux de contacts évoquées précédemment, représentant plus ou moins explicitement les processus épidémiologiques au niveau individuel, ont contribué à faire progresser la théorie épidémiologique dans le sens d'une meilleure prise en compte des effets de la distribution de la population et de la structure des contacts individuels, orientée en amont par la structure socio-spatiale locale. Différents réseaux particuliers peuvent être distingués selon différentes propriétés topologiques, essentielles pour la dynamique globale de la maladie :

- Le degré moyen K , ou nombre moyen de connexions incidentes à un nœud du réseau, et la distribution de ce degré.
- Le coefficient de *clustering* C , qui traduit l'existence de communautés locales dans le réseau. Calculé pour chaque nœud puis moyenné pour tous les nœuds du réseau, ce coefficient mesure parmi les nœuds connectés à un nœud, leur tendance à être également connectés entre eux. Soit $K(x)$, l'ensemble des nœuds connectés au nœud x . Le coefficient de *clustering* $C(x)$ pour le nœud x est alors le nombre de connexions existant, rapporté au nombre de connexions possibles parmi $K(x)$. Quand C est grand (figure 21a), les connexions dans le réseau sont dites locales et la population est dite *clusterisée* ou communautarisée ; quand C est petit, il y a beaucoup de connexions dites de longue portée et le réseau devient non structuré (figure 21b).

Le modèle épidémique de base représente la situation d'un réseau complet où tous les individus sont connectés entre eux et ont une probabilité égale d'entrer en contact avec n'importe quel autre individu. Le réseau de leurs contacts effectifs peut être approché par un graphe aléatoire, dont les liaisons sont établies aléatoirement entre les individus, mais régulier, dans le sens où chaque individu possède le même nombre de connexions (le degré est invariant).

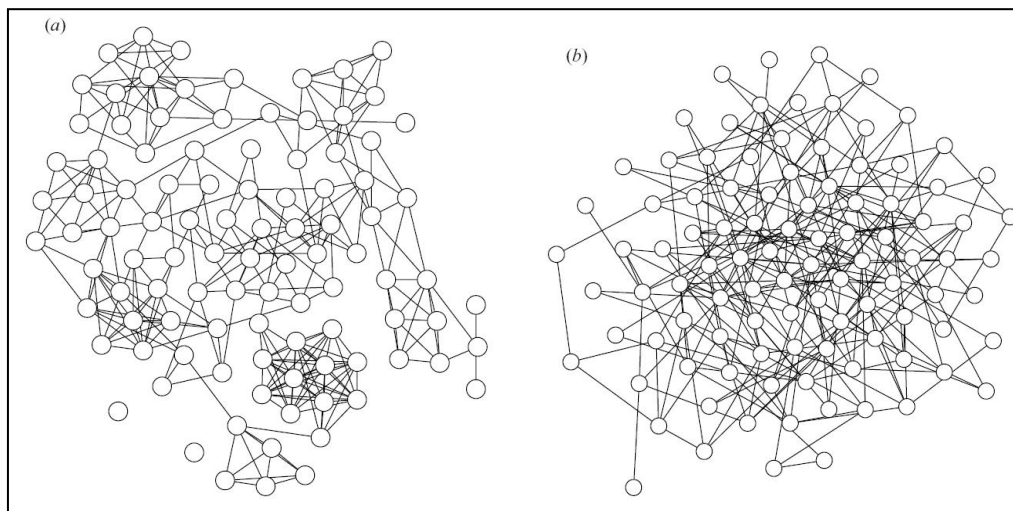


Figure 21. Exemples de réseaux à fort ((a), $C = 0,7$) et faible ((b), $C = 0,2$) coefficient de *clustering* pour un nombre de nœuds identique (100 nœuds) et un degré moyen $K = 5$ (source : Keeling, 1999)

Les résultats épidémiologiques d'un modèle compartimental et d'un modèle à réseau de contacts fixe, aléatoire et régulier, différeront cependant, dans la mesure où le premier, contrairement au second, considère des contacts instantanés et un réseau de contacts effectifs constamment régénéré (Bansal et al., 2007). Ainsi, le second prédira un taux de reproduction de base R_0 plus faible et un développement de la maladie moins important que le premier (Keeling, 1999). Pour des paramètres spécifiques à la maladie identiques, l'écart entre les résultats des deux types de modèles augmentera d'autant plus que celui à réseau de contacts explicite considèrera un réseau présentant un degré moyen faible et un coefficient de *clustering* élevé (Keeling, 1999 ; Keeling, 2005b ; Keeling et Eames, 2005 ; Shirley et Rushton, 2005).

Pour illustrer l'effet du coefficient de *clustering*, considérons un réseau à trame régulière, comme celui d'un automate cellulaire classique à une dimension où les connexions sont strictement locales. On modifie une connexion incidente à un nœud du réseau, pour la réaffecter à un autre nœud pris aléatoirement (figure 22).

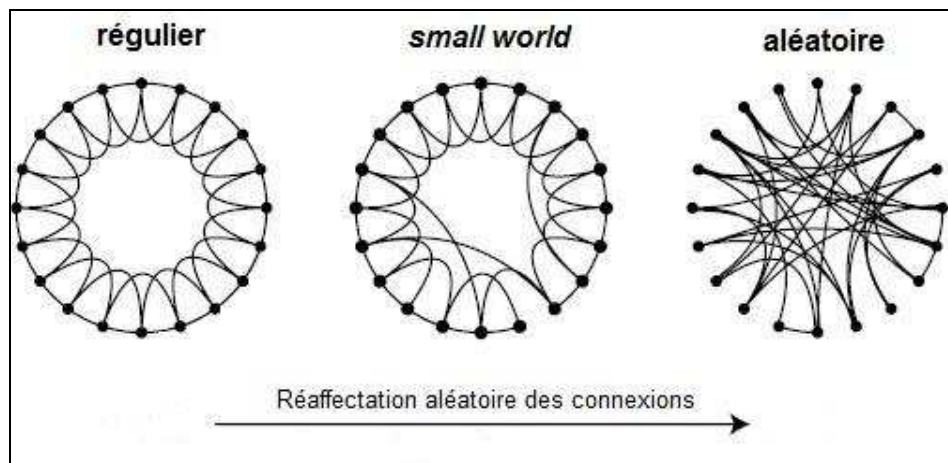


Figure 22. Le réseau *small world*, intermédiaire entre le réseau à trame régulière et le réseau aléatoire (source : Watts et Strogatz, 1998)

La répétition de cette opération, amenant à modifier des connexions locales en connexions de longue portée, fait diminuer le coefficient de *clustering*. Elle introduit un effet petit monde ou *small world* (Watts et Strogatz, 1998), pour se rapprocher progressivement de la situation du graphe aléatoire avec une distribution Poisson du degré. Ce faisant, le phénomène d'agrégation locale des individus infectieux et d'épuisement des susceptibles parmi leurs contacts, qui ralentit la propagation de l'infection, est atténué. Il ressort donc que la probabilité d'occurrence d'une épidémie majeure et la rapidité de sa propagation sont augmentées lorsque le coefficient de *clustering* du réseau de contacts diminue (Moore et Newman, 2000).

Un autre élément caractéristique du réseau, dont l'effet sur la dynamique de l'infection a été étudié ces dernières années, est la distribution du degré. A nombre de connexions égal, une distribution de ces connexions hétérogène facilite la propagation de l'infection dans le réseau, bien que sa dynamique diffère de celle prédite par le

modèle à mélange homogène et degré invariant (Anderson et May, 1991 ; Keeling, 2005a ; Barthélémy et al., 2005 ; Shirley et Rushton, 2005 ; Bansal, 2007). Cet effet est par ailleurs augmenté lorsque le réseau est assortatif, qui fait que les nœuds de degré semblable sont plus amenés à être connectés entre eux. Le passage du réseau régulier au réseau aléatoire fait augmenter l'hétérogénéité du nombre de connexions individuelles, généralement jusqu'à une distribution Poisson. Le réseau qui atteint un niveau extrême d'hétérogénéité des degrés est le réseau invariant d'échelle ou *scale free* (figure 23), dont la distribution du degré suit une loi de puissance (Barabási et Albert, 1999). Dans ce réseau hautement hiérarchisé, une minorité de *super-spreaders* possèdent un grand nombre de connexions, ce qui leur donne un rôle clé dans la propagation et la persistance de la maladie : plus aptes à s'infecter, ils sont également plus aptes à transmettre la maladie à beaucoup d'autres.

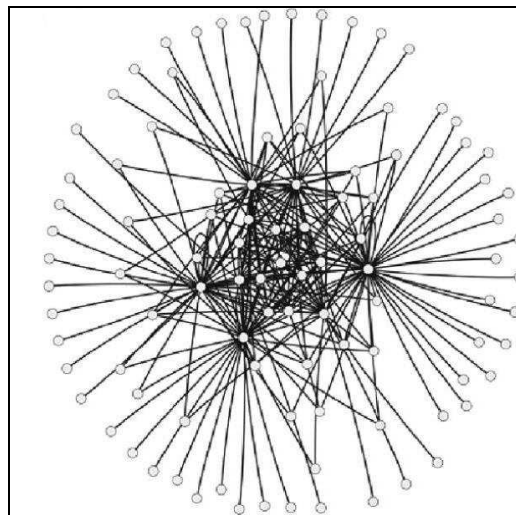


Figure 23. Exemple de réseau aléatoire *scale free* comportant 100 nœuds et au degré moyen $K = 5$ (source : Bansal et al., 2007)

Comme dans le modèle métapopulationnel, lorsque le modèle est stochastique et la distribution du degré hétérogène, le calcul du taux de reproduction de base ne suffit pas à prédire le succès ou non de l'épidémie après introduction d'individus infectieux. Plusieurs études analytiques, en lien avec la théorie de la percolation, ont montré qu'une large gamme de réseaux invariants d'échelle, contrairement aux réseaux plus

homogènes, ne présentent pas ou présentent un seuil épidémique situé très bas, ce qui suggère qu'une épidémie majeure a toujours une probabilité non nulle de se développer et persister même pour des faibles probabilités de transmission (May et Lloyd, 2001 ; Pastor-Satorras et Vespignani, 2001b; Barthélémy et al., 2005). Cependant, $R_0 > 1$ ne signifie pas qu'une épidémie majeure se développera forcément. Une autre étude fondée sur un réseau invariant d'échelle assortatif a précisé que le développement ou non d'une large épidémie est largement fonction du nombre de contacts du patient zéro, qui détermine la possibilité, pour l'infection, d'atteindre ou non à ses débuts, des *super-spreaders* (Meyers et al., 2005).

2.2.2 Les modèles à réseaux de contacts émergents

Le développement de modèles individus-centrés avec une approche constructiviste, allant dans le sens d'une meilleure prise en compte de la diversité biologique des individus et de la structure de leurs interactions socio-spatiales, amène à représenter explicitement dans des modèles à base d'agents les mécanismes comportementaux par lesquels les individus entrent en contact, c'est-à-dire les processus de mobilité, de rencontre, de liaison et de séparation entre les individus (Daudé et Eliot, 2005 ; Eliot et Daudé, 2006 ; Roche et al., 2008 ; Linard et al., 2009).

Ces modèles sont élaborés selon différents degrés de réalisme, allant du modèle théorique représentant des particules se déplaçant aléatoirement dans un espace isotrope et susceptibles d'entrer en *collision* (González et Herrmann, 2004), aux modèles ayant pour objectif de représenter une *réalité virtuelle* sur laquelle appliquer une méthode expérimentale. Ces derniers vont jusqu'à mettre en jeu plusieurs milliers d'agents, distribués dans un environnement hétérogène constitué de plusieurs milliers de localisations, elles-mêmes reliées entre elle par différents modes de transport. Les individus se déplacent de l'une à l'autre selon un programme d'activités journalières précisé en entrée (Eubank et al., 2004 ; Barrett et al., 2005). En lien avec le degré de résolution spatio-temporelle du modèle choisi, l'ensemble des déplacements ou certains types de déplacements peuvent ne pas être explicitement représentés, de sorte que seuls importent les résultats de la mobilité, c'est-à-dire la présence des individus dans tel ou tel lieu (Bian, 2004 ; Huang et al., 2004 ; Müller et al., 2004 ; Dunham, 2005). A moins que le moyen de transport que les individus empruntent ne soit lui-

même considéré dans le modèle comme un lieu où les individus entrent potentiellement en contact.

Un élément essentiel qui distingue cette catégorie de modèles individus-centrés de ceux évoqués dans le paragraphe précédent est que les réseaux de contacts ne sont plus définis *a priori*, mais *a posteriori* des simulations. Dans ces réseaux de contacts émergents, les connexions ne représentent alors plus des contacts potentiels mais les contacts qui se sont effectivement produits au cours de la période simulée. La modélisation explicite des comportements individuels et des structures socio-spatiales qui les orientent représente un moyen d'approcher la réalité des structures d'interactions, en alternative aux méthodes observationnelles très difficiles à mettre en œuvre (Keeling et Eames, 2005), et de les situer par rapport aux réseaux idéalisés présentés précédemment. En règle générale, ces réseaux de contacts du *monde réel*, plus étudiés pour les populations humaines que pour les populations animales (Faust et Skvoretz, 2002 ; Wey et al., 2008), ne sont ni complètement réguliers ni complètement aléatoires (Zhou et al., 2005). Beaucoup font ressortir des effets *petit monde* avec un coefficient de *clustering* relativement élevé (Watts et Strogatz, 1998 ; Eubank et al., 2004), ainsi que des distributions de degrés hétérogènes.

3. Les modèles existants de simulation de l'expression de la peste bubonique

Trois principales expériences de modélisation dynamique sont à remarquer à propos de la peste bubonique. Toutes utilisent un système d'équations différentielles pour formaliser les processus de transmission et d'évolution de la maladie directement au niveau agrégé des populations. Néanmoins, chacune se distingue par des choix de modélisation différents : du point de vue du niveau de résolution spatiale et temporelle considéré d'abord, des populations représentées ensuite, des hypothèses

simplificatrices réalisées enfin, relatives à la structuration de ces populations et aux comportements individuels.

3.1 Le modèle réaction-diffusion de Noble (1974) : conditions de diffusion pandémique de la peste

La première expérience de modélisation réalisée (Noble, 1974) visait à reproduire la diffusion pandémique de la Peste Noire à travers l'espace européen au XIV^{ème} siècle. Le modèle utilisé est un modèle de type réaction-diffusion, adapté du modèle classique de Kermack-McKendrick (1927). A ce niveau d'analyse, le modèle, déterministe, ne considère que la variation des deux classes *Susceptibles* et *Infectées* de la population humaine, sans considération du renouvellement de la population : le mécanisme de propagation *via* les piqûres des puces infectées sur les rats, si tant est qu'il était bien à l'œuvre durant la deuxième pandémie (chapitre 1, 3.2.2 ; chapitre 2, 2.2.2.1), est globalisé dans un paramètre du modèle, le coefficient de transmission.

A travers ce modèle sont explorées les conditions, en termes de densité de population humaine, d'apparition d'une onde de propagation épidémique, dans l'hypothèse où la répartition de la population est uniforme, tandis que les déplacements humains sont de faible portée et se font aléatoirement. Cette hypothèse apparaît bien éloignée de la réalité et génère une diffusion de proche en proche en contradiction avec le schéma classique d'une diffusion hiérarchique observée sur de nombreux foyers (chapitre 1, 3.2.5).

En dépit de cela, pour une densité humaine initiale fixée à 130 km⁻², la vitesse de propagation estimée à travers ce modèle (320-640 km/an) apparaît en adéquation avec celle mesurée rétrospectivement dans le cas de la Peste Noire (Christakos et Oléa, 2005). Par contre, elle est nettement plus élevée que celle de la peste bubonique aux Etats-Unis (Adjemian et al., 2007) ou en Inde (Christakos et al., 2007).

3.2 Le modèle à métapopulations de Keeling et Gilligan (2000) : conditions d'invasion et de persistance de la peste bubonique dans le cas d'un complexe constitué d'une seule espèce de rongeurs et de puces

Une deuxième expérience de modélisation compartimentale réalisée plus récemment (Keeling et Gilligan, 2000 a/b) analyse la dynamique de la maladie à l'intérieur d'un foyer endémique. Le modèle intègre explicitement dans ses équations le cycle de transmission impliquant une espèce de puces et les rats. Il représente explicitement leur dynamique démographique qui obéit, en l'absence de la peste, à une loi logistique, celle des puces étant liée à celle des rats. On reconnaît par là que la peste bubonique est d'abord une zoonose vectorielle transmise occasionnellement à l'homme, ce dernier ne pouvant pas à son tour transmettre la maladie et constituant une impasse dans le cycle de transmission. Le taux d'infection des rats est défini selon la loi d'action de masse : il est fonction du nombre de puces libres infectieuses qui réussissent à se fixer sur un hôte. Ces puces libres sont elles-mêmes libérées par les rats morts naturellement ou de la maladie. L'évolution de la maladie chez le rat suit un schéma SIR : les rats qui guérissent de la maladie sont immunisés à vie, tandis que la période de latence n'est pas distinguée de la période infectieuse. La dynamique de la maladie dans la population humaine, qui n'affecte pas l'expression globale de la maladie, n'a pas besoin d'être représentée explicitement dans le système d'équations. Une variable globale est cependant définie, le potentiel de transmission à l'homme, qui décrit la variation du nombre de puces infectieuses libérées par les rats morts et qui échouent à se fixer sur d'autres hôtes rats. Cette quantité mesure en somme la variation de la source pathogène pour l'homme, qui découle de la densité de la population murine et de la prévalence de l'infection dans cette population.

M. J. Keeling et C. A. Gilligan (2000 a/b) présentent une version métapopulationnelle et stochastique de leur modèle, qui leur permet de considérer une population de rongeurs et leurs puces associées structurée en communautés, avec possibilité de migration d'individus d'une communauté à une autre voisine. Le modèle présente en outre une variation saisonnière de la capacité de charge des puces, destinée à refléter la sensibilité de l'abondance puclidienne à la température (chapitre 2, 2.2.1). A travers

ce cadre de modélisation théorique, les auteurs apportent une contribution aux questions irrésolues des mécanismes et conditions de persistance enzootique de la peste ponctuée d'épizooties, susceptibles de générer des cas humains (chapitre 3, 3.3 ; 3.4). Leurs simulations suggèrent que la peste peut persister pendant plus de 100 ans dans un petit nombre de communautés sans apport extérieur de nouveaux cas d'infection, tandis que dans la plupart des autres communautés, la peste connaît des extinctions suivies de réactivations, révélant le caractère instable de ces poches locales d'infection. Des épizooties murines susceptibles de provoquer des cas humains peuvent se produire à partir de ce fond global endémique. La persistance sur le long terme de l'infection est optimum lorsque la proportion initiale des rats susceptible est modérée (entre 30 et 50 %), ce qui accorde du crédit aux défenseurs de la cause immunitaire, facteur important de cette persistance (chapitre 3, 3.1). Après avoir formulé le taux de reproduction de base R_0 et l'avoir relié à la densité des rats, les auteurs de ce modèle ont étudié les conditions d'une invasion épizootique majeure pouvant générer une épidémie humaine, après introduction d'un faible nombre de rats infectieux dans une population entièrement susceptible. Leurs calculs ont mis en évidence une densité seuil d'approximativement 3000 rats/km², densité seuil généralement dépassée en milieu urbain.

3.3 Le modèle de Foley et al. (2007) : l'expression de la peste bubonique dans le cas d'un complexe constitué de multiples espèces de rongeurs et de puces

Le modèle développé par M. J. Keeling et C. A. Gilligan (2000) convient davantage aux cycles épidémiologiques n'impliquant qu'une seule espèce de rongeurs associée à une seule espèce de puces, comme c'est approximativement le cas sur le foyer indien ou bien encore sur le foyer malgache. Or, nous avons vu précédemment (chapitre 1, 3.2) que la peste peut prendre une grande variété de formes selon les foyers. Elle va jusqu'à impliquer une multitude d'espèces de rongeurs, plus ou moins résistants à l'infection, et de puces, aux spectres d'hôtes et compétences vectorielles variables. Tenant compte de cet aspect essentiel, J. E. Foley et al. (2007) ont récemment proposé

un modèle compartimental dérivé du modèle à transmission vectorielle de Ross-McDonald. La principale évolution est qu'il formalise explicitement la transmission de la peste entre h populations d'hôtes et v populations de vecteurs.

Deux matrices décrivent les taux de contacts contagieux entre les différents couples hôte-vecteur, l'une du vecteur à l'hôte, l'autre de l'hôte au vecteur. A la différence du modèle de Keeling-Gilligan qui représente explicitement le cycle de mobilité des puces (fixation sur un hôte puis libération), l'interaction entre les puces et leur hôte se fait, dans ce modèle, comme dans le modèle de Ross-McDonald. Elle implique trois paramètres : une voracité des puces (ou fréquence des piqûres des puces sur leur hôte), une efficacité de la transmission lors d'une piqûre potentiellement contagieuse, et un coefficient de spécificité d'une espèce de puce pour une espèce de rongeur.

Une autre différence avec le modèle de Keeling-Gilligan tient dans le fait que l'évolution de la maladie chez les rongeurs suit une évolution SIRS, ce qui signifie que ceux qui guérissent de la maladie ont une immunité temporaire. Pour les puces, le schéma est de type SEIRS : la période de latence de l'infection est incluse.

Un point commun avec le modèle précédent concerne la dynamique des différentes populations, régulée en l'absence de l'infection par une croissance logistique et un effet saisonnier. Ce dernier s'applique cette fois à la reproduction des rongeurs mais également au taux de retrait des puces de la classe infectieuse, comme il a été vu que la température affecte le temps de survie de la puce bloquée (chapitre 1, 2.3 ; chapitre 2, 2.2.1).

Les analyses réalisées par les auteurs de ce modèle visent, comme dans le modèle précédent, à préciser les conditions amenant la peste à persister sur le long-terme. Son apport spécifique tient, pour l'essentiel, dans l'estimation des espèces de puces et de rongeurs du foyer nord-américain qui contribuent le plus à la persistance de la maladie, au vu de leurs valeurs spécifiques pour les différents paramètres du modèle. La résistance relative à l'infection des rongeurs, la compétence vectorielle et le spectre d'hôte des puces ressortent comme des facteurs explicatifs importants. Le modèle, appliqué à un foyer enzootique très bien documenté de Californie (Davis et al., 2002), réussit ainsi à simuler l'invasion et la persistance de la maladie pour certaines combinaisons d'espèces, tandis qu'il échoue pour d'autres.

Conclusion du chapitre 4

Le chapitre que nous refermons à présent était destiné à présenter les contributions d'une théorie épidémiologique qui se fonde sur une longue tradition de modélisation des systèmes dynamiques pour étudier les conditions d'invasion épidémique et de persistance endémique d'une maladie transmissible dans une ou plusieurs populations. Cette approche théorique a longtemps évolué dans l'ombre d'une épidémiologie observationnelle qui, par l'application de modèles statistiques, se heurte à la difficulté de prendre en compte d'une manière adéquate les effets de dépendance des états épidémiologiques des individus, liés à leurs contacts, pour préciser le risque d'invasion épidémique et de persistance endémique.

Aujourd'hui, la tendance est à la convergence des deux approches. D'un côté, l'évolution des techniques d'observation de la terre, de suivi des populations à des résolutions toujours plus fines, de caractérisation génétique des pathogènes améliorent la précision des études de risque observationnelles. La modélisation statistique multi-niveaux propose une méthodologie adéquate pour prendre en compte des facteurs structurels au côté des facteurs individuels (Gatrell, 2005). D'un autre côté, les progrès vertigineux de l'informatique, concomitants à la pénétration du paradigme de la complexité, poussent à l'intégration progressive des connaissances issues des approches observationnelles, sources d'hétérogénéité dans les modèles dynamiques, pour faire progresser la théorie de l'expression des maladies transmissibles (Koopman, 2004 ; 2005).

Dans ce sens, les modèles qui formalisent le processus de transmission selon la loi de l'action de masse ont mis l'accent, dans un premier temps, sur l'importance de la densité de population susceptible et sa dynamique démographique, à travers le théorème du seuil épidémique, pour des paramètres spécifiques à la maladie fixés (Anderson et May, 1991). Ce n'est que dans un deuxième temps qu'ont été examinés les effets de l'hétérogénéité spatiale, de la distribution de la population et des structures d'interaction sur l'expression de la maladie, d'abord dans un cadre *métapopulationnel* qui représente une population distribuée en communautés

interreliées, ensuite par l'utilisation de modèles individus-centrés (Keeling, 2005 a/b). L'utilisation de tels modèles ajoute une orientation constructiviste et expérimentale à une géographie des maladies traditionnellement plus tournée vers les études observationnelles, où l'on explore les configurations simulées de la maladie en fonction de nos hypothèses sur la structure spatiale locale, la distribution de la population et les comportements individuels.

Parce qu'ils représentent de manière plus concrète les mécanismes causaux au niveau auquel ils se produisent, les modèles individus-centrés présentent l'avantage de la clarté par rapport à leurs aînés compartimentaux et un paramétrage plus aisé. Surtout, du point de vue de leurs résultats, ces modèles amènent à reconsidérer le théorème du seuil épidémique généré par leurs prédécesseurs. Une partie des différences de sorties entre les deux catégories de modèles provient du fait que les modèles individus-centrés comportent généralement une composante stochastique, due au fait qu'il est impossible de définir exactement le comportement de chaque individu. Ainsi, dans un modèle stochastique à réseau de contacts explicite, après introduction de la maladie, un taux de reproduction de base R_0 supérieur à 1 ne garantira pas que l'infection évoluera en une épidémie et réciproquement. La variation de la taille finale et de la configuration des épidémies simulées pour une même valeur initiale de R_0 augmentera d'autant plus que les contacts entre individus sont hétérogènes. Lorsque la distribution des degrés du réseau de contacts suit une loi puissance, l'existence même d'un seuil épidémique peut disparaître.

Il est des maladies pour lesquelles le modèle compartimental d'action de masse se révèle une bonne approximation du système de transmission. Les maladies humaines à propagation directe ne nécessitant pas un contact très rapproché telles que la grippe en font partie. L'explosion des mobilités humaines va dans le sens de l'hypothèse d'un mélange homogène des individus considérée dans ce modèle (Bonabeau et al., 1998a/b). Il est d'autres maladies qui sont moins bien représentées par le modèle d'action de masse de base. C'est *a priori* le cas de la peste bubonique, qui circule dans des populations de rongeurs communautarisées, à mobilité réduite et dont la transmission nécessite un échange de puces et une fixation, parfois longue, de celles-ci. Les modèles développés jusqu'ici à propos de cette maladie ont permis d'explorer les conditions d'un endémisme à un niveau régional, lié à la diversité des espèces de

puces et de rongeurs impliquées dans le cycle épidémiologique (Foley et al., 2007) ou à la structure *métapopulationnelle* de la population (Keeling et Gilligan, 2000a/b). Ces modèles conservent cependant l'hypothèse de l'action de masse réalisée au niveau de la population entière ou des différentes communautés.

Il ressort donc que jusqu'ici, aucune expérience de modélisation n'a été tentée à une résolution spatiale et temporelle plus fine, se restreignant à explorer le comportement de la maladie à l'intérieur d'une poche locale d'infection, affectant une communauté de rongeurs, et notamment ses conditions de permanence. Dans le chapitre qui suit, nous présenterons le modèle individu-centré à réseau de contacts émergent que nous avons développé dans cette optique, portant sur le foyer endémique de Madagascar. Une telle expérience a impliqué la remise en cause de l'hypothèse de l'action de masse au niveau de la population des rongeurs et de leurs puces. Cela, grâce à la représentation explicite de la distribution des individus et des mécanismes par lesquels ils entrent en contact et se transmettent la maladie.

