

Biomodélisation

Philippe Thieullen, Université Bordeaux 1

26 janvier 2013

2 Modèles continus pour une seule espèce

Dans un modèle en temps discret, on fait l'hypothèse que les changements d'état du système s'effectuent à chaque révolution d'un certain cycle biologique. Dans une représentation continue du modèle, le cycle est très court et semble fluide. En fait, dans les cas mêmes où le cycle de vie est long, comme dans un problème de croissance bactérienne, il est plus rigoureux d'utiliser le modèle continu au détriment du modèle discret, car les différentes phases du cycle ne se produisent pas au mêmes instants pour tous les individus simultanément. Chaque individu, ou bactérie par exemple, évolue indépendamment des autres et les mesures qu'on observe ne sont que des moyennes. Du fait du grand nombre des individus qui constituent la population, les quantités observées apparaissent donc comme des grandeurs continues.

Dans un modèle discret, une loi d'évolution s'écrit

$$\Delta N_t = N_{t+\tau} - N_t = g(N_t),$$

où $\tau = \Delta t$ représente un incrément de temps petit et $g(N_t)$ représente les flux de transfert des individus pendant ce laps de temps. En général, et c'est ce qu'on a toujours fait, $g(N)$ est proportionnel à N . Comme τ n'est plus considéré comme une unité abstraite mais un petit temps, $g(N)$ est aussi proportionnel à τ : la variation du nombre des individus ΔN_t est multipliée par 2 pendant un laps de temps double 2τ . On prend alors plutôt

$$g(N) = \tau N R(N) = \tau f(N)$$

où $R(N)$ est le taux de renouvellement de la population. En prenant $\tau = \Delta t$ très petit, et en se rappelant que

$$\frac{\Delta N}{\Delta t} = \frac{dN}{dt}, \quad (\text{la dérivée de } N(t))$$

une loi d'évolution dans un modèle continu devient une équation différentielle :

$$\frac{dN}{dt} = f(N). \quad (1)$$

Trouver l'évolution du système, c'est résoudre l'équation différentielle (1), c'est-à-dire, trouver $N(t)$ vérifiant (1) sur toute plage de temps $t_0 \leq t \leq t_1$ fixée par l'observateur ; t_0 est appelée *instant initial*. Comme dans les modèles en temps discret, il est nécessaire de fixer la *condition initiale* $N(t_0)$ à l'instant initial t_0 . On admettra que, pour des équations différentielles bien définies et pour tout instant initial t_0 , il existe une unique solution $N(t)$ sur un intervalle de temps maximal $[t_0, t_\infty[$. La solution explicite est cependant souvent impossible à établir et l'on a recours à des simulations numériques sur logiciels. Le rôle du modélisateur consiste à établir la forme correcte de $f(N)$ à chaque situation.

2.1 Les modèles à compartiments

Un modèle à compartiment permet de traiter des problèmes de transfert de matière ou d'individus entre différentes structures biologiques. On peut par exemple s'intéresser à la diffusion d'une molécule marquée ou d'un médicament au travers de différents organes comme le sang, le rein ou le foie. On peut aussi s'intéresser à la répartition d'une population en espace dans différents territoires.

On représente schématiquement un tel modèle par des compartiments reliés entre eux par des flèches. Une variable d'état X, Y, \dots est définie pour chaque compartiment et correspond à une fraction de quantité de matière dans un état donné. Une flèche entre deux compartiments désigne un flux de matière proportionnel à la quantité de matière du compartiment de départ. Les taux de flux a, b, \dots , ou proportion de matière se déplaçant d'un compartiment à l'autre sont affichés près de chaque flèche. Une flèche isolée entrante correspond à un apport extérieur de matière et une flèche isolée sortante, à une perte de matière avec l'extérieur proportionnelle à la quantité de matière de ce compartiment. Un schéma typique de tels modèles est donné à la figure 1. Les constantes de flux a, b, \dots peuvent *a priori* dépendre du temps. Les actions extérieures $U(t), V(t), \dots$ dépendent généralement du temps.

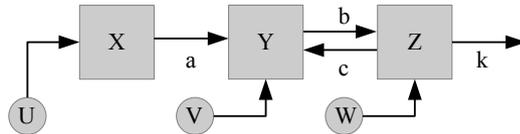


FIGURE 1 – Modèle à compartiment typique

Pendant une variation de temps dt , la variation de quantité de matière de l'état Y dépend du flux entrant $aX + cZ$ en provenance des deux compartiments X et Z , d'un apport extérieur $V(t)$ et du flux sortant $-bY$. En répétant cette analyse pour chaque compartiment, on obtient l'équation différentielle suivante

$$\begin{aligned} X' &= -aX + U(t), \\ Y' &= aX - bY + cZ + V(t), \\ Z' &= bY - (c + k)Z + W(t), \end{aligned} \tag{2}$$

Il est commode de choisir un ordre (X, Y, Z) dans les variables d'état et de respecter cet ordre à droite du signe $=$ dans l'équation précédente. On constate que ce modèle fournit uniquement des *équations différentielles linéaires* tant que les constantes de flux restent indépendantes du temps et des autres variables. On verra plus tard comment modifier ce modèle pour inclure une dépendance par rapport au temps et par rapport aux autres variables d'état. Le système d'équations (2) peut s'écrire encore plus simplement en utilisant la notation matricielle suivant :

$$E = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad A = \begin{bmatrix} -a & 0 & 0 \\ a & -b & c \\ 0 & b & -(c+k) \end{bmatrix}, \quad B = \begin{bmatrix} U(t) \\ V(t) \\ W(t) \end{bmatrix},$$

où E désigne le vecteur colonne des états, A la matrice des taux de flux et B le vecteur colonne des actions extérieures (injection d'un traceur, migration en provenance d'un autre système

écologique, ...) le système (2) s'écrit alors simplement

$$\frac{dE}{dt} = AE + B.$$

On utilise dans l'expression précédente la *règle des produits matriciels*

$$AE = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} aX+bY+cZ \\ dX+eY+fZ \\ gX+hY+iZ \end{bmatrix}$$

où chaque ligne de la première matrice est multipliée terme à terme à chaque colonne de la seconde matrice. La première matrice n'est pas nécessairement carrée, de même que la deuxième matrice n'est pas nécessairement un vecteur colonne. Les *dimensions d'une matrice* $m \times n$ désignent toujours m lignes et n colonnes. Dans un produit de matrices, le nombre de colonnes de la première matrice doit nécessairement coïncider avec le nombre de lignes de la seconde matrice. Par exemple, le calcul suivant désigne une autre forme de produit matriciel

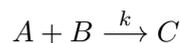
$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} X & U \\ Y & V \\ Z & W \end{bmatrix} = \begin{bmatrix} aX+bY+cZ & aU+bV+cW \\ dX+eY+fZ & dU+eV+fW \end{bmatrix}$$

On notera bien les dimensions 2×2 de la matrice résultant de l'opération de multiplication.

Un *système fermé* est un système sans échange de flux avec l'extérieur. Dans l'exemple précédent, on a donc $u = v = w = k = 0$. Un *système autonome* est un système dans lequel les constantes de flux a, b, \dots ne dépendent pas du temps.

2.2 Les modèles de type réactions chimiques

Dans un modèle de réaction chimique simple, deux molécules A et B interagissent ensemble pour donner une troisième molécule C selon le schéma



Les réactions ne se font pas toutes à la même vitesse et on note par $k(A, B) = \frac{d[C]}{dt}$ la vitesse de la concentration de la molécule C . On suppose qu'on peut appliquer la *loi d'action de masse* qui fait l'hypothèse que les molécules sont des entités indépendantes et que la réaction ne peut se produire que si l'événement

« A et B sont présentes au même endroit »

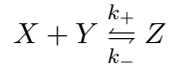
se réalise. La loi d'action de masse fait donc l'hypothèse que les concentrations sont de la forme

$$\frac{d[C]}{dt} = k[A][B], \quad (k \text{ constante}).$$

Pendant qu'une molécule C est créée, deux autres A et B disparaissent ; on a donc aussi

$$\frac{d[A]}{dt} = -k[A][B], \quad \frac{d[B]}{dt} = -k[A][B].$$

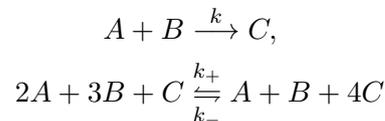
Les réactions sont en général réversibles et il faut alors prévoir deux constantes de vitesse de réaction



On obtient alors un système d'équations différentielles non linéaires

$$\begin{aligned}\frac{d[X]}{dt} &= k_-[Z] - k_+[X][Y], \\ \frac{d[Y]}{dt} &= k_-[Z] - k_+[X][Y] \\ \frac{d[Z]}{dt} &= -k_-[Z] + k_+[X][Y]\end{aligned}$$

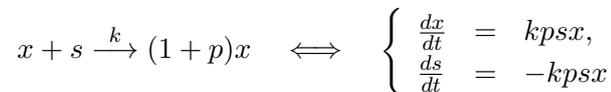
Plusieurs molécules de même identité sont dès fois nécessaires pour réaliser la réaction ; par ailleurs, les molécules ne sont pas transformées. On arrive alors à des équations plus complexes de la forme par exemple



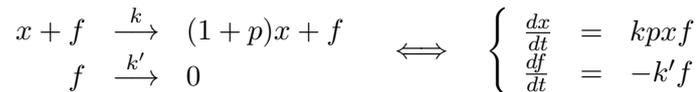
Dans le cas de la deuxième équation, comme il est nécessaire de réaliser simultanément la rencontre de 2 molécules A , de 3 molécules B et d'une molécule C , la vitesse de réaction de gauche à droite doit être proportionnelle à $[A]^2[B]^3[C]$. Par ailleurs, à chaque fois qu'une réaction élémentaire se produit, 1 molécule A disparaît, 2 molécules B disparaissent et 3 molécules C apparaissent. Les équations deviennent

$$\begin{aligned}\frac{d[A]}{dt} &= -k_+[A]^2[B]^3[C] + k_-[A][B][C]^4 - k[A][B], \\ \frac{d[B]}{dt} &= -2k_+[A]^2[B]^3[C] + 2k_-[A][B][C]^4 - k[A][B], \\ \frac{d[C]}{dt} &= 3k_+[A]^2[B]^3[C] - 3k_-[A][B][C]^4 + k[A][B].\end{aligned}$$

L'approche par l'utilisation de la loi d'action de masse permet de modéliser certains mécanismes écologiques comme les mécanismes de croissance ou de compétition entre espèces. Par exemple



représente la croissance d'une certaine biomasse x par consommation d'une ressource s . Ou bien



représente une croissance de x régulée par un facteur f mais pouvant se dégrader. On peut alors se demander quel modèle utilisé à l'avance. Il n'y a pas de réponse évidente à cette question. Une méthode consiste à tester *a posteriori* la validité des résultats et par exemple à

tester les relations d'allométrie que le mod 'ele fait ressortir. Prenons par exemple un problème de croissance simultanée de deux organes sous la dépendance d'un même facteur de croissance

$$\begin{array}{l} x + f \xrightarrow{k_1} 2x + f \\ y + f \xrightarrow{k_2} 2y + f \\ f \xrightarrow{k'} 0 \end{array} \iff \begin{cases} \frac{dx}{dt} = k_1 x f \\ \frac{dy}{dt} = k_2 y f \\ \frac{df}{dt} = -k' f \end{cases}$$

On constate alors la relation

$$\frac{dx}{dy} = \frac{k_1 x}{k_2 y} = \mu \frac{x}{y}$$

qu'on peut intégrer en utilisant la méthode de séparation des variables

$$\frac{dx}{x} = \mu \frac{dy}{y} \iff \ln x = \mu \ln y \iff \frac{x}{y^\mu} = \text{constante.}$$

Cette dernière relation est appelé *relation d'allométrie*.

2.3 Intégration explicite des modèles de base pour une seule espèce

Il y a trois modèles de base à connaître : le modèle exponentiel, le modèle de Gompertz et le modèle logistique.

Exemple 1. Le modèle exponentiel.

$$x \xrightarrow{r/p} (1+p)x \iff x' = rx \quad (3)$$

C'est une équation différentielle du premier ordre (une seule dérivée), linéaire (le second membre dépend linéairement de x), à coefficients constants (le terme kp est indépendant du temps). La solution d'une telle équation est

$$x(t) = x_0 \exp(rt). \quad (4)$$

Exemple 2. Le modèle de Gompertz.

$$\begin{array}{l} x + e \xrightarrow{k} (1+p)x + e \\ e \xrightarrow{r} 0 \end{array} \iff \begin{cases} x' = kpxe \\ e' = -re \end{cases} \quad (5)$$

Pour résoudre cette équation, on cherche à appliquer une méthode de séparation des variables

$$\frac{x'}{x} = pke = -p\frac{k}{r}e' \iff \frac{dx}{x} = -p\frac{k}{r}de$$

On se rappelle maintenant que la primitive $\int \frac{dx}{x} = \ln x$ et par intégration entre 0 et t , on a

$$\ln \frac{x}{x_0} = -p\frac{k}{r}(e - e_0), \quad x = x_0 \exp\left(-p\frac{k}{r}(e - e_0)\right).$$

Par ailleurs $e = e_0 e^{-rt}$ et on obtient finalement la solution

$$x(t) = x_0 \exp \left[p \frac{k}{r} e_0 (1 - e^{-rt}) \right]$$

Le modèle de Gompertz peut aussi s'écrire avec une seule équation en x . On introduit une constant K qui s'interprétera ultérieurement comme une capacité de charge

$$\ln \frac{K}{x_0} = p \frac{k}{r} e_0 \implies \ln \frac{x}{K} = -p \frac{k}{r} e.$$

D'où l'équation différentielle de Gompertz en fonction uniquement de la biomasse

$$x' = -rx \ln \left(\frac{x}{K} \right) \quad (6)$$

et la solution en fonction des paramètres r et K

$$x(t) = x_0 \exp \left[\ln \left(\frac{K}{x_0} \right) (1 - e^{-rt}) \right]. \quad (7)$$

Pour une donnée initiale $x_0 < K$, la biomasse $x(t)$ est une fonction croissante de valeur maximale (en temps infini) K .

Exemple 3. Le modèle logistique.

$$x + s \xrightarrow{k} (1 + p)x \iff \begin{cases} x' = kpxs \\ s' = -kxs \end{cases} \quad (8)$$

On constate que $x' + ps' = 0$ ou bien $x + ps = K = x_0 + p_0 s_0$ est constant dans le temps. En substituant $ps = K - x$ dans la première équation en x et en posant $r = kK$, on obtient l'équation différentielle logistique

$$x' = rx \left(1 - \frac{x}{K} \right) \quad (9)$$

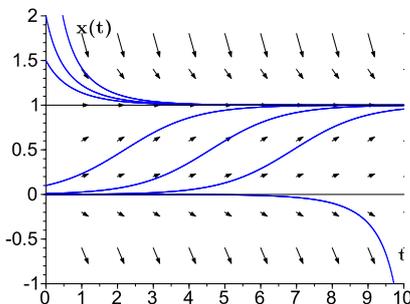


FIGURE 2 – Equation logistique

La figure 2 met en évidence quelques trajectoires $x(t)$ en fonction de t : les trajectoires issues de $x_0 > 0$ convergent vers 1, celles issues de $x_0 < 0$ divergent rapidement et explosent en temps fini (elles n'ont cependant aucune interprétation physique). Les flèches indiquent l'évolution de $x(t)$: une flèche de longueur $(\Delta t, \Delta x)$ est dessinée au dessus de chaque point (t, x) avec $\frac{\Delta x}{\Delta t} = rx(1 - \frac{x}{K})$. Bien qu'en général, il est impossible de résoudre explicitement une équation différentielle, nous allons montrer qu'il existe dans le cadre présent, une méthode simple appelée *méthode de séparation des variables*. La méthode consiste à séparer les variables x et t de chaque côté du signe

égal, puis à rechercher une primitive de chaque fonction

$$\frac{dx}{x(K-x)} = \frac{r}{K} dt \implies \int_{x_0}^x \frac{dx}{x(K-x)} = \int_0^t \frac{r}{K} dt.$$

Pour la première primitive, on utilise une méthode de coefficients indéterminés qui permet de décomposer la fraction de gauche en fractions plus simples. Après avoir ajusté les coefficients, on obtient

$$\frac{1}{x(K-x)} = \frac{A}{x} + \frac{B}{K-x} = \frac{1}{Kx} + \frac{1}{K(K-x)}.$$

On est ramené à trouver des primitives d'expressions plus simple

$$\frac{dx}{Kx} + \frac{dx}{K(K-x)} = \frac{r}{K} dt$$

Après simplification par K

$$\ln\left(\frac{x}{K-x}\right) = \ln\left(\frac{x_0}{K-x_0}\right) + rt \iff \frac{x}{K-x} = \frac{x_0}{K-x_0} e^{rt}$$

et finalement la forme de croissance explicite du modèle logistique

$$\boxed{x(t) = \frac{x_0}{\frac{x_0}{K} + (1 - \frac{x_0}{K}) \exp(-rt)}} \quad (10)$$

2.4 Etats d'équilibre et stabilité

On considère ici un modèle ne faisant intervenir qu'une seule espèce x . On suppose que l'équation d'évolution est donnée par une équation différentielle du premier ordre

$$\frac{dx}{dt} = f(x).$$

Comme dans le chapitre sur les systèmes discrets, un état d'équilibre est une position du système x_* constant dans le temps : $x(t) = x_*$ en tout temps. En particulier $\frac{dx}{dt} = 0$ et donc

$$x_* \text{ est un état d'équilibre} \iff f(x_*) = 0. \quad (11)$$

Par exemple, dans le modèle logistique $f(x) = rx(1 - \frac{x}{K})$, les seuls états d'équilibre sont $x_* = 0$ et $x_* = K$.

Comme dans les modèles à temps discrets, un état d'équilibre est stable si le système revient dans cet état pour toute petite la perturbation. Pour analyser le retour du système vers sa position d'équilibre x_* , nous allons réécrire l'équation différentielle dans de nouvelles variables centrées autour de x_* . On note

$$x(t) = x_* + u(t)$$

où u est censé représenter une petite perturbation du système. Par définition même d'une dérivée,

$$\frac{f(x_* + u) - f(x_*)}{u} \simeq f'(x_*).$$

L'équation d'évolution du système devient alors

$$\begin{aligned}\frac{dx}{dt} &= \frac{du}{dt} = f(x_* + u) \simeq f(x_*) + f'(x_*)u, \\ \frac{du}{dt} &= f'(x_*)u.\end{aligned}$$

On vient d'obtenir l'équation différentielle des petits déplacements u du système autour de sa position d'équilibre, $u' = f'(x_*)u$. En posant $\lambda = f'(x_*)$, on obtient le modèle exponentiel $u' = \lambda u$ qui s'intègre par $u(t) = \exp(\lambda t)$. On appelle λ , la *valeur propre* de l'état d'équilibre. Il est alors facile de comprendre à quelle condition le système retourne vers sa position d'équilibre, c'est-à-dire, à quelle condition $u(t) \rightarrow 0$. Si $f'(x_*) > 0$, le système s'écarte du point d'équilibre x_* ; si $f'(x_*) < 0$, le système s'en rapproche.

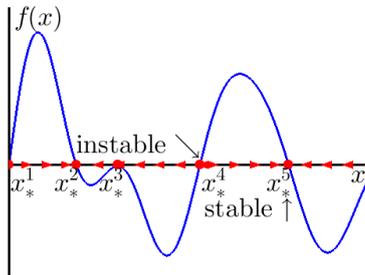


FIGURE 3 – Stabilité des états d'équilibre x_* d'un système décrit par l'équation $\frac{dx}{dt} = f(x)$.

La figure 3 permet de visualiser l'évolution du système dans le temps. Supposons que le système prenne initialement une valeur x_0 située entre les états d'équilibre x_1^* et x_2^* ; comme $f(x) > 0$, $f'(x_1^*) > 0$ et $f'(x_2^*) < 0$, le système va alors progressivement s'éloigner du point instable x_1^* pour se stabiliser ensuite vers le point stable x_2^* . Ce comportement réapparaît pour une valeur x_0 comprise entre x_4^* et x_5^* . L'état d'équilibre x_3^* est un peu différent : comme $f'(x_3^*) = 0$, il est à la fois stable (à droite) en prenant une valeur initiale $x_0 > x_3^*$ et instable (à gauche) en prenant une valeur initiale $x_0 < x_3^*$. Les micro fluctuations d'origine externe des paramètres du système font que cet état d'équilibre peut disparaître à tout moment $f'(x_3^*) < 0$: un point à droite de x_3^* est immédiatement aspiré vers x_2^* . On dira que x_3^* est instable.

On retiendra finalement le critère de stabilité

$$x_* \text{ est un état d'équilibre stable} \iff f'(x_*) < 0. \quad (12)$$

Par exemple, pour le modèle logistique et ces deux états d'équilibre $x_* = 0$ et $x_* = K$, on obtient $f'(x) = r - 2\frac{r}{K}x$ et donc

$x_* = 0$	$f'(0) = r > 0$	x_* est instable
$x_* = K$	$f'(K) = -r < 0$	x_* est stable

2.5 Croissance des vers de sapin au Canada

De nombreuses études écologiques ont montrée l'existence d'un accroissement important récurrents de la population d'une certaine espèce de vers dans les forêts d'épicéa au Canada. La taille de la population est contrôlée essentiellement par deux facteurs : la surface totale du feuillage de la forêt qui détermine la capacité de charge de la population de vers et le nombre

total d'oiseaux qui représente le prédateur naturel. Le modèle utilisé par D. Ludwig, D.D. Jones, C.S. Holling (1978) est le suivant

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K(S)} \right) - B \frac{N^2}{A(S)^2 + N^2},$$

où N désigne le nombre de vers par hectare. Le premier terme du membre de droite est un terme de croissance, le second est un terme de prédation d'origine aviaire. Plus précisément

- r : taux de naissance (1 à 6 larves par an et par individu),
- $K(S)$: capacité de charge dépendant du nombre total S de branches par hectare ; on prend $K(S) = K'S$ avec K' égal au nombre maximum de vers qu'une branche peut supporter (K' à prendre de 100 à 400 vers par branche)
- B : nombre maximum de vers consommés par an et par hectare (B à prendre entre 20 000 et 40 000 vers par an et par hectare). On fait l'hypothèse que la taille de la population des oiseaux est constante.
- $A(S)$: paramètre de saturation de la forme $A(S) = \alpha S$ où α le nombre de vers par branche à partir duquel les oiseaux considèrent la chasse aux vers rentable (α égal à 1 ou 2)