

Modèles logistiques déterministes et stochastiques

Fabien Campillo, Marc Joannides

► **To cite this version:**

Fabien Campillo, Marc Joannides. Modèles logistiques déterministes et stochastiques. CARI 2010: 10^e Colloque Africain sur la Recherche en Informatique et en Mathématiques Appliquées, Oct 2010, Yamoussoukro, Côte d'Ivoire. pp.110-118. hal-02752476v2

HAL Id: hal-02752476

<https://hal.inria.fr/hal-02752476v2>

Submitted on 15 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1. Modèle déterministe

Le modèle logistique classique s'exprime sous la forme de l'EDO

$$\dot{x}_t = r x_t \left(1 - \frac{x_t}{K}\right) \quad (1)$$

où x_t est la densité de population à l'instant t , $r > 0$ le taux de croissance per capita, $K > 0$ la capacité du milieu ; $x_0 > 0$ désignera la condition initiale. Les deux points d'équilibre de (1) sont 0 et $= K$, le premier est instable et le second stable. La taille n_t de la population s'écrit

$$n_t \stackrel{\text{déf}}{=} N \times x_t$$

où le coefficient de normalisation N désigne la taille initiale de la population, la surface ou le volume du domaine de l'écosystème. L'EDO associée à la taille de la population est $\dot{n}_t = r n_t \left(1 - \frac{n_t/N}{K}\right)$.

2. Modèles stochastiques

2.1. Modèle microscopique de saut pur

Ce modèle [2,1] décrit la dynamique de population sous la forme d'un processus $(N_t)_{t \geq 0}$ de Markov de naissance et mort, i.e. un processus de Markov en temps continu à valeurs dans \mathbb{N} et ne sautant que de $+1$ (naissance) ou -1 (mort). Il s'agit naturellement d'une idéalisation où la population est homogène (tous les individus ont les mêmes taux de naissance et de mort). Le processus N_t est défini par

$$\mathbb{P}(N_{t+h} = j | N_t = i) \stackrel{h \downarrow 0}{\simeq} \begin{cases} \lambda_i h & \text{si } j = i + 1 \text{ (naissance),} \\ \mu_i h & \text{si } j = i - 1 \text{ (mort),} \\ 1 - (\lambda_i + \mu_i) h & \text{si } j = i \end{cases} \quad (2a)$$

avec $\lambda_i \geq 0$, $\mu_i \geq 0$. Le plus souvent on suppose que $\lambda_0 = 0$ (i.e. 0 est un état absorbant). La loi initiale de ce processus est δ_{x_0} . On suppose que les taux de naissance et mort *per capita* sont linéaires en la densité de population :

$$\lambda_i = i \left(b_1 + b_2 \frac{i}{N}\right), \quad \mu_i = i \left(d_1 + d_2 \frac{i}{N}\right), \quad (2b)$$

avec $b_j > 0$ et $d_j > 0$. L'hypothèse "logistique" revient à supposer que

$$\lambda_i - \mu_i = r i \left(1 - \frac{i/N}{K}\right) \quad (2c)$$

avec $r = b_1 - d_1 > 0$ et $K = -\frac{b_1 - d_1}{b_2 - d_2} > 0$. Le générateur infinitésimal du processus $(N_t)_{t \geq 0}$ est défini par

$$Qf(i) = \lim_{t \rightarrow 0} \frac{\mathbb{E}[f(N_t) | N_0 = i] - f(i)}{t} = \tau(i) \int_{\mathbb{R}} [f(y) - f(i)] \rho(i, \mathbf{d}y), \quad i \in \mathbb{N} \quad (3)$$

pour toute fonction $f : \mathbb{N} \rightarrow \mathbb{R}$ mesurable et bornée, où l'intensité $\tau(i)$ d'événements (de naissance ou de mort) est défini par

$$\tau(i) \stackrel{\text{déf}}{=} \lambda_i + \mu_i \quad (4)$$

et le noyau $\rho(i, \mathbf{d}y)$ du processus de saut est défini par

$$\rho(i, \mathbf{d}y) \stackrel{\text{déf}}{=} \frac{\lambda_i}{\lambda_i + \mu_i} \delta_{i+1}(\mathbf{d}y) + \frac{\mu_i}{\lambda_i + \mu_i} \delta_{i-1}(\mathbf{d}y). \quad (5)$$

Lorsque $\lambda_i > 0$ et $\mu_i > 0$ pour tout $i > 0$, alors 0 est un état absorbant et tous les états communiquent. Dans ce cas on peut en fait montrer [1, Ch. 6] [6, p. 149] que la seule mesure de probabilité invariante est définie par $\pi_0 = 1$ et $\pi_i = 0$ pour tout $i > 0$. La probabilité d'extinction se comporte de façon asymptotique comme suit : Si $\sum_{i=1}^{\infty} \frac{\mu_1 \cdots \mu_i}{\lambda_1 \cdots \lambda_i} = \infty$ alors $\mathbb{P}(N_t = 0) \rightarrow 1$ lorsque $t \rightarrow \infty$; sinon

$$\mathbb{P}(N_t = 0) \xrightarrow{t \rightarrow \infty} \frac{\sum_{i=x_0}^{\infty} \frac{\mu_1 \cdots \mu_i}{\lambda_1 \cdots \lambda_i}}{1 + \sum_{i=1}^{\infty} \frac{\mu_1 \cdots \mu_i}{\lambda_1 \cdots \lambda_i}}.$$

On introduit le processus de saut pour la *densité* de population

$$X_t^N \stackrel{\text{déf}}{=} \frac{N_t}{N},$$

On déduit facilement de Q , le générateur Q_N du processus X_t^N :

$$Q_N f(x) = \tau(Nx) \int_{\mathbb{R}} [f(y) - f(x)] \rho^N(x, \mathbf{d}y), \quad x \in \mathbb{N}/N$$

avec $\rho^N(x, \mathbf{d}y) \stackrel{\text{déf}}{=} \frac{\lambda_{Nx}}{\tau(Nx)} \delta_{x + \frac{1}{N}}(\mathbf{d}y) + \frac{\mu_{Nx}}{\tau(Nx)} \delta_{x - \frac{1}{N}}(\mathbf{d}y)$.

2.2. Représentation semimartingale

Dans le but de comparer les modèles déterministes et stochastiques, il est utile d'établir une EDS satisfaite par le processus de saut. Pour f à support compact, on a déjà la représentation de $f(X_t^N)$ comme

$$f(X_t^N) = f(X_0^N) + \int_0^t Q_N f(X_s^N) \mathbf{d}s + M_t^f \quad (6)$$

où M_t^f est une martingale. Sous des conditions de régularité (non explosion), cette décomposition reste valable pour $f = \text{Id}$.

Proposition 2.1 Si le processus X_t^N est régulier et admet un moment d'ordre 2 à tout instant $t \geq 0$, on a la décomposition

$$X_t^N = X_0^N + \int_0^t r X_s^N \left(1 - \frac{X_s^N}{K}\right) ds + M_t \quad (7)$$

où M_t est une martingale nulle en 0, de carré intégrable et de variation quadratique prévisible

$$\langle M \rangle_t = \frac{1}{N} \int_0^t r' X_s^N \left(1 + \frac{X_s^N}{K'}\right) ds \quad (8)$$

avec $r' = b_1 + d_1$ et $K' = \frac{r'}{b_2 + d_2}$.

La preuve s'appuie sur une représentation du processus N_t

$$\begin{aligned} N_t = N_0 &+ \int_0^t \sum_{i \leq N_s^-} d\Pi_s^{b_1}(i) + \int_0^t \sum_{i \leq N_s^-} \sum_{j \leq N_s^-} d\Pi_s^{b_2/N}(i, j) \\ &- \int_0^t \sum_{i \leq N_s^-} d\Pi_s^{d_1}(i) - \int_0^t \sum_{i \leq N_s^-} \sum_{j \leq N_s^-} d\Pi_s^{d_2/N}(i, j) \end{aligned}$$

où $\Pi_s^{b_1}(i)$, $\Pi_s^{b_2/N}(i, j)$, $\Pi_s^{d_1}(i)$ et $\Pi_s^{d_2/N}(i, j)$ sont 4 familles de processus de Poisson indépendantes d'intensités respectives b_1 , $\frac{b_2}{N}$, d_1 et $\frac{d_2}{N}$.

L'équation (7) est l'analogie stochastique naturel du modèle logistique déterministe (1). C'est une EDS dirigée par une martingale à sauts. On observe que son terme de dérive (moyenne instantanée) a la même forme que dans le modèle déterministe. Cependant, à cause de la non-linéarité, l'espérance de X_t^N n'est pas la solution x_t de (1). En revanche, d'après (8), le modèle déterministe (1) peut être vu comme une limite fluide du modèle stochastique (7) [7,8].

2.3. Approximation diffusion

Lorsque la taille de la population est grande, les sauts du processus X_t^N sont fréquents et petits. Il est alors légitime de décrire l'évolution de la densité de population par un processus continu. L'approximation-diffusion [4] de X_t^N se base sur un développement de Taylor de f dans l'expression du générateur Q_N , en négligeant les termes d'ordre supérieur à 2. En effet

$$\begin{aligned} Q_N f(x) &= \lambda_{Nx} [f(x + \frac{1}{N}) - f(x)] + \mu_{Nx} [f(x - \frac{1}{N}) - f(x)] \\ &\simeq \frac{1}{N} (\lambda_{Nx} - \mu_{Nx}) f'(x) + \frac{1}{2N^2} (\lambda_{Nx} + \mu_{Nx}) f''(x) \\ &= r x \left(1 - \frac{x}{K}\right) f'(x) + \frac{1}{2N} r' x \left(1 + \frac{x}{K'}\right) f''(x) \stackrel{\text{déf}}{=} A^N f(x) \end{aligned}$$

avec A^N le générateur du processus de diffusion solution de

$$d\tilde{X}_t^N = r \tilde{X}_t^N \left(1 - \frac{\tilde{X}_t^N}{K}\right) dt + \frac{1}{\sqrt{N}} \sqrt{r' \tilde{X}_t^N \left(1 + \frac{\tilde{X}_t^N}{K'}\right)} dW_t \quad (9)$$

où W_t est un processus de Wiener standard. Cette approximation–diffusion consiste en fait, pour N grand à approcher X_t^N par un processus de diffusion de même moyenne et variance instantanées. Il existe une justification heuristique de cette approximation [5] : pour $\delta > 0$ et conditionnellement à $X_t^N = x$

$$X_{t+\delta}^N = x + \frac{1}{N} (B_\delta - D_\delta)$$

où B_δ et D_δ sont respectivement le nombre de naissances et de morts survenues dans l'intervalle $[t, t + \delta]$. On fait l'hypothèse (i) : δ est suffisamment petit pour que les taux d'occurrence de naissance et de mort ne varient pas significativement dans $[t, t + \delta]$. B_δ et D_δ sont donc les accroissements de processus de Poisson d'intensités respectives λ_{Nx} et μ_{Nx} et sont de loi de Poisson de paramètres respectifs $\lambda_{Nx} \delta$ et $\mu_{Nx} \delta$. Or, pour un nombre suffisamment grand d'événements, ces lois de Poisson seront correctement approchées par des lois normales de mêmes moyennes et variances. On rajoute l'hypothèse (ii) : δ est suffisamment grand pour que l'approximation normale s'applique. On obtient alors

$$\begin{aligned} X_{t+h}^N &\simeq x + \frac{1}{N} (\mathcal{P}(\lambda_{Nx} \delta) - \mathcal{P}(\mu_{Nx} \delta)) && \text{(approx. poissonienne)} \\ &\simeq x + \frac{1}{N} (\mathcal{N}(\lambda_{Nx} \delta, \lambda_{Nx} \delta) - \mathcal{N}(\mu_{Nx} \delta, \mu_{Nx} \delta)) && \text{(approx. normale)} \\ &= x + \frac{\lambda_{Nx}}{N} \delta + \sqrt{\frac{\lambda_{Nx} \delta}{N^2}} \mathcal{N}(0, 1) - \frac{\mu_{Nx}}{N} \delta + \sqrt{\frac{\mu_{Nx} \delta}{N^2}} \mathcal{N}(0, 1) \\ &= x + r x \left(1 - \frac{x}{K}\right) \delta + \frac{\delta}{\sqrt{N}} \sqrt{r' x \left(1 + \frac{x}{K'}\right)} \mathcal{N}(0, 1). \end{aligned}$$

Cette dernière équation est l'écriture du schéma d'Euler pour l'intégration numérique de l'EDS (7). On en déduit que s'il existe δ tel que les hypothèses (i) et (ii) soient satisfaites, alors les réalisations de X^N seront proches (en loi) des trajectoires discrétisées de son approximation–diffusion \tilde{X}_t^N . Remarquons que les hypothèses (i) et (ii) semblent contradictoires et exigent des tailles de populations très grandes, comme pour les populations de microorganismes. En particulier, il est déraisonnable d'utiliser une approximation–diffusion pour estimer ou calculer des instants d'extinction.

3. Fermeture des moments

La loi du processus est entièrement caractérisée par son générateur Q , en particulier, posons :

$$m_p(t) \stackrel{\text{déf}}{=} \mathbb{E}[N_t^p] \qquad u_p(t) \stackrel{\text{déf}}{=} \mathbb{E}[X_t^p] = \frac{m_p(t)}{N^p}$$

en résolvant l'équation de Kolmogorov progressive on peut montrer [1] que :

$$\frac{d}{dt} m_1(t) = (b_1 - d_1) m_1(t) + \frac{(b_2 - d_2)}{N} m_2(t), \quad (10)$$

$$\frac{d}{dt} m_2(t) = (b_1 + d_1) m_1(t) + [2(b_1 - d_1) + \frac{(b_2 + d_2)}{N}] m_2(t) + 2 \frac{(b_2 - d_2)}{N} m_3(t). \quad (11)$$

Ces équations ne sont pas fermées, mais on peut utiliser une technique de fermeture des moments. Si on suppose N_t déterministe alors $m_2(t) = m_1^2(t)$, ce qui permet de retrouver à l'aide de (10) l'équation logistique classique (1). Supposons maintenant que N_t est gaussienne alors $\mathbb{E}(N_t^3) = 3 m_1(t) m_2(t) - 2 m_1^3(t)$ qui plongé dans (11) conduit à un modèle proposant une approximation des deux premiers moments :

$$\frac{d}{dt} \bar{u}_1(t) = r \bar{u}_1(t) - \frac{r}{K} \bar{u}_2(t), \quad (12a)$$

$$\frac{d}{dt} \bar{u}_2(t) = \frac{r'}{N} \bar{u}_1(t) + [2r + \frac{r'}{NK'}] \bar{u}_2(t) - 2 \frac{r}{K} (3 \bar{u}_1(t) \bar{u}_2(t) - 2 \bar{u}_1^3(t)). \quad (12b)$$

Pour identifier les paramètres de ce modèle, il est nécessaire de disposer de plusieurs trajectoires de X_t afin d'estimer le moment d'ordre 2. Pour faire mieux (i.e. travailler avec une seule trajectoire) : il faut utiliser un modèle d'EDS et identifier la variation quadratique (c'est possible sur une seule trajectoire).

À l'aide d'un théorème de comparaison on peut en fait montrer que : $m_1(t) \leq x_t$, $\forall t \geq 0$, i.e. le modèle logistique stochastique (2) est en moyenne inférieur au modèle logistique déterministe (1).

4. Simulations et commentaires

L'algorithme de Gillespie, ou SSA (stochastic simulation algorithm) permet de simuler des trajectoires d'un processus de Markov de saut pur de générateur connu. La Figure 1 montre 3 réalisations du processus X_t^N pour de faibles valeurs de N , ainsi que la trajectoire du modèle déterministe correspondant. Le modèle stochastique n'est pas monotone et la probabilité d'extinction n'est pas nulle.

Lorsque N est grand, les trajectoires de X_t^N (Figure 2) et de son approximation-diffusion \tilde{X}_t^N (Figure 3) présentent les mêmes caractéristiques statistiques.

La Figure 4 appelle plusieurs commentaires. Comme on le sait déjà le modèle déterministe (1) surévalue la moyenne du modèle stochastique. Par ailleurs le transitoire est très différent : le modèle par fermeture des moments (12) converge nettement plus rapidement vers son équilibre, le modèle déterministe (1) correspond mieux au modèle stochastique. Le comportement asymptotique est aussi très différent : la capacité du milieu est

surévaluée par le modèle déterministe (1), le modèle par fermeture des moments correspond au modèle stochastique.

Il existe un autre type de modèle stochastique consistant à perturber directement le taux r par un bruit blanc, c'est l'optique adoptée dans [3] et qui conduit à une EDS différente de (9). Il serait donc nécessaire dans une version étendue de ce travail de comparer les différentes lois asymptotiques des modèles stochastiques.

Il est intéressant de noter que le modèle stochastique de saut se met sous la forme différentielle (7) qui se décompose en deux termes : la dynamique déterministe (1) plus un terme stochastique dont on connaît les caractéristiques.

En conclusion, l'équation logistique classique (1) ne peut vraiment être utilisée que dans un cadre "grande population", dans tous les autres cas il est nécessaire d'adopter une technique de moment (corrigeant certains biais), soit de faire appel à des modèle de type EDS, soit – pour les populations en taille très faible et notamment dans des conditions proches de l'extinction – de faire appel à des modèles microscopique de saut pur.

5. Bibliographie

- [1] L. J. Allen. *An Introduction to Stochastic Processes with Biology Applications*. Prentice Hall, 2003.
- [2] M. S. Bartlett, J. C. Gower, and P. H. Leslie. A comparison of theoretical and empirical results for some stochastic populations models. *Biometrika*, 47(1/2) :1–11, 1960.
- [3] F. Comets and T. Meyre. *Calcul stochastique et modèles de diffusions*. Dunod, 2006.
- [4] S. N. Ethier and T. G. Kurtz. *Markov Processes – Characterization and Convergence*. John Wiley & Sons, 1986.
- [5] D. T. Gillespie. The chemical Langevin equation. *The Journal of Chemical Physics*, 113(1) :297–306, 2000.
- [6] S. Karlin and H. M. Taylor. *A First Course in Stochastic Processes, Second Edition*. Academic Press, 2nd edition, 1975.
- [7] T. G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *Journal of Applied Probability*, 7(1) :49–58, 1970.
- [8] T. G. Kurtz. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *Journal of Applied Probability*, 8 :344–356, 1971.
- [9] J. H. Matis and T. R. Kiffe. On approximating the moments of the equilibrium distribution of a stochastic logistic model. *Biometrics*, 52(3) :980–991, 1996.
- [10] H. Schurz. Modeling, analysis and discretization of stochastic logisticoic equations. *Internation Journal of Numerical Analysis and Modeling*, 4(2) :178–197, 2007.

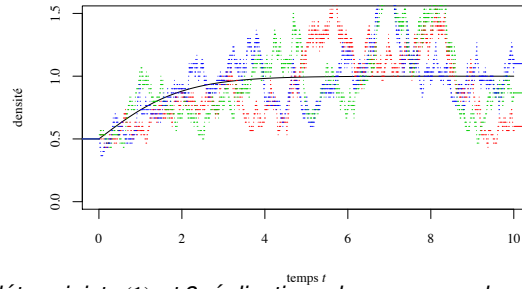


Figure 1. *Modèle déterministe (1) et 3 réalisations du processus de sauts (7) pour $N = 30$.*

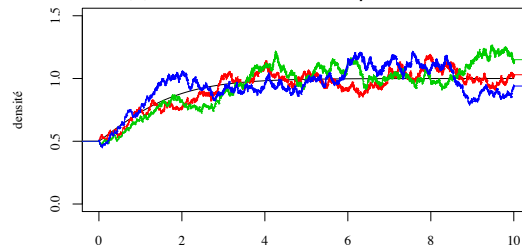


Figure 2. *Modèle déterministe (1) et 3 réalisations du processus de sauts (7) pour $N = 300$.*

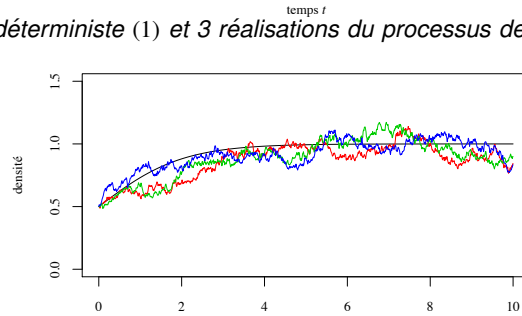


Figure 3. *Modèle déterministe (1) et 3 réalisations du processus de diffusion (9) pour $N = 300$.*

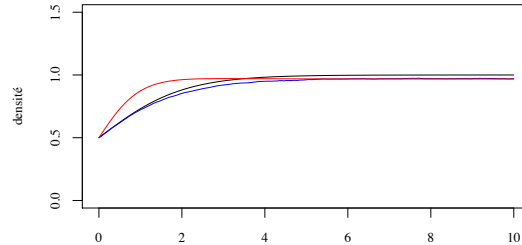


Figure 4. *Modèle déterministe (1) [noir], modèle déterministe par fermeture des moments (12) [rouge] et moyenne de 5000 réalisations du modèle stochastique (7) pour $N = 300$ [bleu]. Ces deux derniers ont la même moyenne asymptotique qui est inférieure à celle du modèle déterministe. Le modèle par fermeture des moments présente un transitoire beaucoup plus court que les deux autres.*
Revue