Stochastic models for the chemostat

Fabien Campillo, Marc Joannides, Irène Larramendy-Valverde

MAMERN11: 4^{th} International Conference on Approximation Methods
and Numerical Modelling in Environment and Natural Resources
Saidia (Morocco), May 23-26, 2011

Stochastic models for the chemostat

Fabien Campillo[1], **Marc Joannides**[1,2] **and Irène Larramendy-Valverde**[2]

[1] Project–Team MERE, INRIA/INRA, UMR MISTEA, Montpellier
Fabien.Campillo@inria.fr

[2] Université Montpellier 2 / I3M
marc.joannides@univ-montp2.fr and larra@math.univ-montp2.fr

Keywords: stochastic differential equations, chemostat, pure jump process, diffusion approximation, tau-leap method, Monte Carlo method, Gillespie algorithm.

Abstract. *The chemostat is classically represented, at high population scale, as a system of ordinary differential equations. Our goal is to establish a set of stochastic models that are valid at different scales: from the small population scale to the scale immediately preceding the one corresponding to the deterministic model. At a microscopic scale we present a pure jump stochastic model that gives rise, at the macroscopic scale, to the ordinary differential equation model. At an intermediate scale, an approximation diffusion allows us to propose a model in the form of a system of stochastic differential equations.*

1 Introduction

The dynamics of a single species/single substrate chemostat is usually described by a set of ordinary differential equations (ODE) derived from a mass balance principle, see [4, 1]. If $s(t)$ denotes the concentration of nutrient (substrate) and $b(t)$ the concentration of the organism (biomass) at time t (expressed in *g/L*), then the couple $x(t) = (b(t), s(t))$ is the solution of:

$$\dot{b}(t) = [\mu(s(t)) - D]\, b(t)\,, \tag{1a}$$

$$\dot{s}(t) = -k\, \mu(s(t))\, b(t) + D\, [s^{\text{in}} - s(t)] \tag{1b}$$

where $D > 0$ is the dilution rate, $s^{\text{in}} > 0$ the substrate concentration in the influent, and $k > 0$ the stoichiometric coefficient. The initial condition lies in the positive orthant, that is $b(0) \geq 0$ and $s(0) \geq 0$. Equation (1) will also be denoted: $\dot{x}(t) = f(x(t))$. The specific growth rate function $\mu(s)$ is non-negative; we suppose that $\mu(0) = 0$, $\mu(s) > 0$ for $s > 0$, $\mu(s) \leq \mu_{\max} < \infty$ and that it is continuous at 0. Commonly used models are the Monod model (uninhibited growth) and the Haldane model (inhibited growth).

The modeling process that leads to (1) relies on the fact that the stochastic effects can be neglected. This is possible only at macroscopic scale, for high population sizes, and under homogeneity conditions. At all other scales or when the homogeneity conditions is not met, random effects cannot be neglected.

2 Pure jump model $X_t = (B_t, S_t)$

We consider aggregated jumps obtained by adding up small and frequent jumps resulting from individual events. The resulting stochastic process will be a pure jump process $X_t = (B_t, S_t)$, fully determined by its jumps and the corresponding jump rates; the state variable will be denoted $x = (b, s)$. We consider five jumps: ① biology term: biomass increase of size $\nu_1(x)$ at rate $\lambda_1(x)$; ② biology term: substrate decrease of size $\nu_2(x)$ at rate $\lambda_2(x)$; ③ inflow term: substrate inflow of size $\nu_3(x)$ at rate $\lambda_3(x)$; ④ outflow term: biomass outflow of size $\nu_4(x)$ at rate $\lambda_4(x)$; ⑤ outflow term: substrate outflow of size $\nu_5(x)$ at rate $\lambda_5(x)$. We propose:

	① biomass increase	② substrate decrease	③ substrate	④ biomass	⑤ substrate
	biology		inflow	outflow	
rate $\lambda_i(x)$	$K_1\,\mu(s)\,b$	$K_2\,k\,\mu(s)\,b$	$K_3\,D\,s^{\text{in}}$	$K_4\,D\,b$	$K_5\,D\,s$
jump $\nu_i(x)$	$\begin{pmatrix} \frac{1}{K_1} \\ 0 \end{pmatrix}$	$-\begin{pmatrix} 0 \\ \frac{1\wedge K_2\,s}{K_2} \end{pmatrix}$	$\begin{pmatrix} 0 \\ \frac{1}{K_3} \end{pmatrix}$	$-\begin{pmatrix} \frac{1\wedge K_4\,b}{K_4} \\ 0 \end{pmatrix}$	$-\begin{pmatrix} 0 \\ \frac{1\wedge K_5\,s}{K_5} \end{pmatrix}$

This choice comply with the mass balance principle and the stochastic mass action law [6]. Note that the jumps $\nu_i(x)$ essentially do not depend on x except for the negative jumps near the border $\{(b,s) \in \mathbb{R}^2_+; b = 0 \text{ or } s = 0\}$.

3 Stochastic differential representation of X_t

The process X_t can also be described as the solution of a stochastic differential equation. Indeed, X_t is in fact a Markov process that can be represented as the following (jump) SDE:

$$X_t = X_0 + \sum_{i=1}^{5} \int_{(0,t]\times[0,\infty)} \nu_i(X_{u^-})\, 1_{\{v\le\lambda_i(X_{u^-})\}}\, N^i(\mathrm{d}u \times \mathrm{d}v) \tag{2}$$

where N^i are independent random Poisson measures with intensity measure $\mathrm{d}u\times\mathrm{d}v$ (Lebesgue measure). We can easily derived the following semi-martingale decomposition:

$$X_t = X_0 + \int_0^t f_K(X_u)\,\mathrm{d}u + \sum_{i=1}^{5} M_t^i \tag{3}$$

where $f_K(x) \stackrel{\text{def}}{=} \sum_{i=1}^5 \nu_i(x)\,\lambda_i(x)$ and M_t^i are five independent square-integrable martingales with zero mean given by: $M_t^i \stackrel{\text{def}}{=} \int_0^t \int_0^\infty \nu_i(X_{u^-})\, 1_{\{v\le\lambda_i(X_{u^-})\}}\, \tilde{N}^i(\mathrm{d}u \times \mathrm{d}v)$, $i = 1\cdots 5$ and $\tilde{N}^i(\mathrm{d}u \times \mathrm{d}v) \stackrel{\text{def}}{=} N^i(\mathrm{d}u \times \mathrm{d}v) - \mathrm{d}u \times \mathrm{d}v$ are centered random Poisson measures. The infinitesimal generator of the process X_t is:

$$\mathcal{A}\phi(x) = \sum_{i=1}^{5} \lambda_i(x)\,[\phi(x+\nu_i(x)) - \phi(x)] = \lambda(x) \int_{\mathbb{R}^2_+} [\phi(y) - \phi(x)]\,\rho(x, \mathrm{d}y) \tag{4}$$

for all $\phi : \mathbb{R}^2_+ \mapsto \mathbb{R}$ continuous with compact support, where:

$$\lambda(x) \stackrel{\text{def}}{=} \textstyle\sum_{i=1}^5 \lambda_i(x)\,, \quad \rho(x, \mathrm{d}y) \stackrel{\text{def}}{=} \sum_{i=1}^5 \bar{\lambda}_i(x)\,\delta_{x+\nu_i(x)}(\mathrm{d}y)\,, \quad \bar{\lambda}_i(x) \stackrel{\text{def}}{=} \frac{\lambda_i(x)}{\sum_{i'=1}^5 \lambda_{i'}(x)}\,.$$

The non-explosion and existence of moments for the process X_t is detailed in [1].

To describe how the process X_t can be simulated, let $\tau_0 = 0$ and $\tau_n = \inf\{t > \tau_{n-1}\,;\, X_t \neq X_{\tau_{n-1}}\}$ and $Y_n = X_{\tau_n}$. It is well know that (*i*) Y_n is a Markov chain on $\mathbb{R}^2_+$ with transition probability $\rho(x, \mathrm{d}y)$; (*ii*) for all $n \geq 1$, conditionally on $Y_0, \ldots, Y_{n-1}$, the holding times $\tau_1 - \tau_0, \ldots, \tau_n - \tau_{n-1}$ are independent and exponentially distributed of intensity parameters $\lambda(Y_0), \ldots, \lambda(Y_{n-1})$. So the algorithm consists in successively simulating the holding times $t_{n-1} - t_{n-1} \sim \mathrm{Exp}(\lambda(X_{t_{n-1}}))$ and to sample the new state according to $X_{t_n} \sim \rho(X_{t_{n-1}}, \mathrm{d}y)$.

4 Discrete time approximations

For any small $\Delta t > 0$ given, let $t_n = n\,\Delta t$. We propose a discrete time Poisson approximation $(\tilde{X}_{t_n})_{n\geq 0}$ of $(X_t)_{t\geq 0}$: on the interval $[t_n, t_{n+1})$ we froze the rate functions $\lambda_i(X_t)$ to $\lambda_i(X_{t_n})$ so that we get a Poisson distribution. The jumps $\nu_i(X_t)$ are also frozen to $\nu_i(X_{t_n})$. Let $\tilde{X}_0 = X_0$, the approximation is defined by:

$$\tilde{X}_{t_{n+1}} = \tilde{X}_{t_n} + \textstyle\sum_{i=1}^5 \nu_i(\tilde{X}_{t_n})\, \mathcal{P}_n^i(\Delta t\, \lambda_i(\tilde{X}_{t_n})) \tag{5}$$

where $(\mathcal{P}_n^i(\rho))_{n\in\mathbb{N}, i=1\cdots 5}$ are independent Poisson variables with intensities ρ. We have:

$$\mathbb{E}[\tilde{X}_{t_{n+1}} | \tilde{X}_{t_n} = x] = x + f_K(x)\,. \tag{6}$$

In other words, the infinitesimal increments of the conditional mean follow the O.D.E. (1). Also:

$$\mathrm{cov}[\tilde{X}_{t_{n+1}} | \tilde{X}_{t_n} = x] = \textstyle\sum_{i=1}^5 \mathrm{cov}[\nu_i(x)\, \mathcal{P}_n^i(\Delta t\, \lambda_i(\tilde{X}_{t_n})) | \tilde{X}_{t_n} = x] = \begin{pmatrix} \tilde{\Sigma}_1^2 & 0 \\ 0 & \tilde{\Sigma}_2^2 \end{pmatrix} \tag{7}$$

with $\tilde{\Sigma}_1^2 = \Delta t \left\{ \frac{1}{K_1}\, \mu(s)\, b + \frac{1}{K_4}\, (1 \wedge K_4\, b)^2\, D\, b \right\}$ and $\tilde{\Sigma}_2^2 = \Delta t \left\{ \frac{1}{K_2}\, (1 \wedge K_2\, s)^2\, k\, \mu(s)\, b + \frac{1}{K_3}\, D\, s^{\mathrm{in}} + \frac{1}{K_5}\, (1 \wedge K_5\, s)^2\, D\, s \right\}$.

In (5), the variable $\mathcal{P}_n^i(\Delta t\, \lambda_i(x))$ is Poisson distributed with parameter $\Delta t\, \lambda_i(x)$. When this parameter is large (greater than 10 or 20) then this last distribution is very close to the normal distribution of mean $\Delta t\, \lambda_i(x)$ and variance $\Delta t\, \lambda_i(x)$. Hence, we get a (discrete time) normal approximation $(\tilde{\xi}_{t_n})_{n\geq 0}$ of $(X_t)_{t\geq 0}$ by letting $\tilde{\xi}_0 = X_0$ and, conditionally on $\tilde{\xi}_{t_{n-1}} = x$: $\tilde{\xi}_{t_{n+1}} = x + \sum_{i=1}^5 \nu_i(x)\, \mathcal{N}_n^i$ where $\mathcal{N}_n^i$ are 5 independent Gaussian random variables : $\mathcal{N}_n^i \sim \mathcal{N}(\lambda_i(x)\, \Delta t\,,\, \lambda_i(x)\, \Delta t)$. So conditionally on $\tilde{\xi}_{t_n} = x$, $\tilde{\xi}_{t_{n+1}}$ is normal with mean (6) and covariance matrix (7). Let $\tilde{\xi}_{t_n} = (\tilde{\beta}_{t_n}, \tilde{\sigma}_{t_n})$, given $\tilde{\beta}_{t_n} = b$ and $\tilde{\sigma}_{t_n} = s$:

$$\tilde{\beta}_{t_{n+1}} = b + \big[\mu(s) - (1 \wedge K_4\, b)\, D\big]\, b\, \Delta t + \sqrt{\Delta t\, \tfrac{\mu(s)\, b}{K_1}}\; w_n^1 + \sqrt{\Delta t\, \tfrac{(1\wedge K_4\, b)^2\, D\, b}{K_4}}\; w_n^4 \tag{8a}$$

$$\begin{aligned} \tilde{\sigma}_{t_{n+1}} = s &+ \big[-(1 \wedge K_2\, s)\, k\, \mu(s)\, b + D\, s^{\mathrm{in}} - (1 \wedge K_5\, s)\, D\, s \big]\, \Delta t \\ &+ \sqrt{\Delta t\, \tfrac{(1\wedge K_2\, s)^2\, k\, \mu(s)\, b}{K_2}}\; w_n^2 + \sqrt{\Delta t\, \tfrac{D\, s^{\mathrm{in}}}{K_3}}\; w_n^3 + \sqrt{\Delta t\, \tfrac{(1\wedge K_5\, s)^2\, D\, s}{K_5}}\; w_n^5 \end{aligned} \tag{8b}$$

where w_n^i are i.i.d. $\mathcal{N}(0,1)$. In both approximations (5) and (8), a mechanism that forces the processes $\tilde{X}_{t_n}$ or $\tilde{\xi}_{t_n}$ to stay within the positive orthant $\mathbb{R}^2_+$ should be used.

5 Diffusion model $\xi_t = (\beta_t, \sigma_t)$

System (8) is the Euler-Maruyama time discretization of the diffusion process $\xi_t = (\beta_t, \sigma_t)$ solution of the following SDE:

$$\begin{aligned} \mathrm{d}\beta_t &= \big[\mu(\sigma_t) - D\big]\, \beta_t\, \mathrm{d}t + \sqrt{\tfrac{\mu(\sigma_t)\, \beta_t}{K_1}}\; \mathrm{d}W_t^1 + \sqrt{\tfrac{D\, \beta_t}{K_4}}\; \mathrm{d}W_t^4 \\ \mathrm{d}\sigma_t &= \big[-(1 \wedge K_2\, \sigma_t)\, k\, \mu(\sigma_t)\, \beta_t + D\, (s^{\mathrm{in}} - \sigma_t) \big]\, \mathrm{d}t \\ &\qquad + \sqrt{\tfrac{k\, \mu(\sigma_t)\, \beta_t}{K_2}}\; \mathrm{d}W_t^2 + \sqrt{\tfrac{D\, s^{\mathrm{in}}}{K_3}}\; \mathrm{d}W_t^3 + \sqrt{\tfrac{D\, \sigma_t}{K_5}}\; \mathrm{d}W_t^5 \end{aligned}$$

where for large K_i's the terms $(1 \wedge K_4\,\beta_t)$, $(1 \wedge K_2\,\sigma_t)$, $(1 \wedge K_5\,\sigma_t)$ are replaced by 1; and where W_t^i are independent standard Wiener processes. Existence and uniqueness issues for this equation, as well as its behavior near the axes, are detailed in [1].

6 Discussion

In contrast with previous stochastic chemostat models [5, 2, 3] where the stochasticity was introduced according to an ad hoc approach, in the present work we propose a family of models where the structure of the noise emerges from the very dynamics and where the scale parameters can be tuned according to the problem under interest. In particular it allows us to propose hybrid models where the cell population dynamics features stochasticity as the substrate is in fluid dynamics (ODE).

The approach proposed here can be applied to any model of population dynamics especially in cases of difference of scale between the different dynamics (e.g. cell/substrate). The dynamics of interacting populations cannot be modeled by a single model but rather by a family of models whose domain of validity depends on the scale at which the dynamics are considered. For example the normal approximation model or the ODE model are valid in high population levels, hence using such models to infer extinction characteristics like extinction time and extinction probabilities is not valid.

The SDE model is adapted for the confrontation to the data as it allows us to build a statistical model and the associated likelihood function. One of the next important steps, that we will investigate in coming work, will be to propose an adapted statistical procedure to estimate the scale parameters K_i, and in a second step to estimate the parameters (D, s^{in}...). In the future we will also investigate the long-term behavior of these models as well as their optimal command.

REFERENCES

[1] Fabien Campillo, Marc Joannides, and Irène Larramendy. Stochastic models of the chemostat. Rapport de Recherche RR-7458, INRIA, November 2010.

[2] Thomas C. Gard. Asymptotic and transient analysis of stochastic core ecosystem models. In *Fourth Mississippi State Conference on Differential Equations and Computational Simulations, Electronic Journal of Differential Equations, Conference 03*, pages 51–62, 1999.

[3] Lorens Imhof and Sebastian Walcher. Exclusion and persistence in deterministic and stochastic chemostat models. *Journal of Differential Equations*, 217(1):26–53, 2005.

[4] Hal L. Smith and Paul E. Waltman. *The Theory of the Chemostat: Dynamics of Microbial Competition*. Cambridge University Press, 1995.

[5] G. Stephanopoulos, R. Aris, and A.G. Fredrickson. A stochastic analysis of the growth of competing microbial populations in a continuous biochemical reactor. *Mathematical Biosciences*, 45:99–135, 1979.

[6] Darren J. Wilkinson. *Stochastic Modelling for Systems Biology*. Chapman & Hall/CRC, 2006.