Ecological Modelling 299 (2015) 1–13

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

A modeling approach of the chemostat

Coralie Fritsch[a,b,*], Jérôme Harmand[b,c], Fabien Campillo[b]

[a] Université Montpellier 2/I3M, 2 Place Eugène Bataillon, 34095 Montpellier Cedex 5, France
[b] MODEMIC Project-Team, INRA/INRIA, UMR MISTEA, 2 Place Pierre Viala, 34060 Montpellier Cedex 01, France
[c] Laboratoire de Biotechnologies de l'Environnement, UR0050, INRA, Avenue des Étangs, 11100 Narbonne, France

ARTICLE INFO

Article history:
Received 10 June 2014
Received in revised form 28 October 2014
Accepted 20 November 2014

Keywords:
Chemostat model
Stochastic chemostat model
Mass structured chemostat model
Individually-based model (IBM)
Monte Carlo

ABSTRACT

Population dynamics and in particular microbial population dynamics, though intrinsically discrete and random, are conventionally represented as deterministic differential equations systems. In these type of models, populations are represented by continuous population sizes or densities usually with deterministic dynamics.

Over the last decades, alternate individual-based models have been proposed where population is explicitly represented as a set of individuals. These may include stochastic dynamics or stochastic rules. With reference to the last class of models we can also associate pure jump processes where the population is described as a discrete population size with stochastic discrete event evolutions.

In the first class of models the population dynamics and its representation may be viewed respectively as deterministic and continuous, in the second class they may be viewed respectively as stochastic and discrete.

In this present work, we present a modeling approach that bridges the two representations. This link can be mathematically described as a functional law of large numbers in high population size asymptotics. These results suggest new strategies of modeling and simulation. We illustrate this approach on the modeling of the chemostat.

1. Introduction

Biological continuous cultures in chemostat play an important role in microbiology as well as in biotechnology. Different formulations are used to represent these processes. The mechanisms of growth and cell division may indeed be described at the cell level or at the population level.

In the case of the former, the population is represented as a discrete population size or as a finite set of individuals. The mechanisms acting on these individuals could be discrete or continuous, stochastic or deterministic. Among other possibilities, we may think of stochastic birth and death processes (BDP) or individual-based models (IBM). Even if IBM are not intrinsically stochastic, here we consider stochastic IBM. Note that most IBMs integrate stochastic entries or stochastic rules and can be considered as stochastic. Even if these kinds of models can also integrate continuous components (space localization, nutrients, etc.) and partially deterministic dynamics, we will refer to them as stochastic and discrete (in their representation of the population).

In the latter case, the population is described in terms of continuous population size or continuous population densities and the dynamics are classically represented as systems of ordinary differential equations (ODE) or as integro-differential equations (IDE) or partial derivative equations. With reference to this point of view, we will refer to these models as deterministic and continuous.

Our purpose is not to propose a full review of these types of models and of their classification; these questions go far beyond the scope of this work and we suggest the reader to consult works like Ferrer et al. (2009) and Hellweger and Bucci (2009).

Our purpose is to present a modeling approach through a specific example of chemostat dynamics. The proposed approach bridges discrete/random to continuous/deterministic within the framework of a "large population size" asymptotic, which allows us to prove, under certain assumptions, the convergence in distribution of the former models toward the latter ones (Campillo and Fritsch, 2014), and to propose new strategies in terms of modeling and simulation.

* Corresponding author at: Université Montpellier 2/I3M, 2 Place Eugène Bataillon, 34095 Montpellier Cedex 5, France. Tel.: +33 4 99 61 23 80.
E-mail addresses: coralie.fritsch@supagro.inra.fr (C. Fritsch), Jerome.Harmand@supagro.inra.fr (J. Harmand), Fabien.Campillo@inria.fr (F. Campillo).

http://dx.doi.org/10.1016/j.ecolmodel.2014.11.021
0304-3800/

Hence in large population sizes, a simulation of the discrete/random model would be similar to that of the continuous/deterministic one. Of course this is true under certain assumptions, and especially in an asymptotic framework: in practice it is difficult to a priori know what large population size means. Beyond the mathematical analysis, it is possible to rely on numerical simulations to get an idea of the convergence of these former models to the latter ones. It is also interesting to understand how the former models behave when they are not close to the latter, that is to say when the continuous/deterministic models are no longer valid.

Beyond the antagonism discrete/random vs. continuous/deterministic, it seems appropriate to propose a new modeling approach where the "model" is not a specific computer or mathematical representation defined once and for all but rather a set of representations. To infer the links between these representations, the scope of validity of each different representation, as well as the capabilities of the associated simulation and control tools, are required.

As we will see the relation between IBM's and integro-differential equations is rather straightforward. There are also natural links between IBM-like representations to PDE representations. One may think of meshfree particle methods for the approximation of PDE fluid dynamics. In the context of IBMs, we can consult the following modeling examples: Faugeras and Maury (2007) for fish population movements, Gómez-Mourelo and Ginovart (2009) for yeast population growth in batch culture and Campillo and Champagnat (2012) for clonal plant dynamics.

The first model of the chemostat appeared in the 1950s (Monod, 1950; Novick and Szilard, 1950). This first model has always retained its relevance in particular because of its simplicity (Smith and Waltman, 1995). Several other models appeared later such as the so-called population balance models proposed by Fredrickson et al. (1967) which rely on a representation of the population structured in mass (Ramkrishna, 1979).

More recently several stochastic models in (unstructured) population size appeared in order to account for the demographic or environmental sources of randomness (Crump and O'Young, 1979; Stephanopoulos et al., 1979; Imhof and Walcher, 2005; Grasman et al., 2005; Campillo et al., 2011; Deygout et al., 2013). In particular, for the demographic noise, according to a now classic approach, the model is described at the level of the individual as a discrete stochastic birth and death process that can be approximated at a meso-scale by a continuous diffusion process when population sizes are large enough, and that reduces at a macro-scale to the solution of the classic chemostat ODE when these population sizes are very large.

The individual-based model which we propose in this paper has been studied mathematically in Campillo and Fritsch (2014) where we proved in particular its convergence in distribution to the solution of an IDE similar to that proposed in Fredrickson et al. (1967).

We propose in this article to illustrate this approach on the model of chemostat: starting from the classical ordinary differential equation model in dimension 2, we propose other representations in the form of an integro-differential equation (continuous and deterministic) or as an individual-based model (discrete and random) both structured in mass. By model reduction, these representations can be reduced to the classical model (continuous and deterministic) or as a birth and death process (discrete and random). We explain the links between these different representations of the same model, as well as their respective advantages and limitations, specifically in terms of simulation.

In Section 2 we introduce the different models, we detail in particular the proposed IBM. In Section 3 we describe the (almost) exact simulation algorithm of the IBM. Using simulations, in Section 4 we highlight the differences between each of these representations. The paper ends with a discussion in Section 5.

2. The models

2.1. The ODE model

The classic chemostat model reads:

$$\dot{S}_t = D(\mathbf{s}_{\text{in}} - S_t) - k\mu(S_t)Y_t \tag{1}$$

$$\dot{Y}_t = (\mu(S_t) - D)Y_t, \tag{2}$$

where S_t and Y_t are respectively the *substrate concentration* and the *bacterial concentration* (mg/l) which are assumed to be uniform in the vessel; D is the dilution rate (1/h), $\mathbf{s}_{\text{in}}$ is the substrate input concentration (mg/l), k is the (inverse of) yield constant. The specific growth rate μ could for example be the classic Monod kinetics:

$$\mu(s) = \mu_{\max}\frac{s}{K_s + s} \tag{3}$$

with maximum specific growth rate $\mu_{\max}$ and half-velocity constant K_s.

In biochemical engineering, System (1)–(2) corresponds to the classic continuous stirred-tank reactor (CSTR) under well-mixing conditions (Smith and Waltman, 1995).

2.2. The IDE model

Instead of representing the dynamics of the bacterial population inside the chemostat through the aggregated state variable Y_t, one may wish to represent the state of the bacterial population structured in mass, that is to consider the density of population $p_t(x)$ w.r.t. their mass in a reference volume V. Hence $\int_{m_0}^{m_1} p_t(x)\mathrm{d}x$ is the number of cells which mass is between m_0 and m_1 and the link with the bacterial concentration is:

$$Y_t \overset{\text{def}}{=} \frac{1}{V}\int_0^{m_{\max}} x p_t(x)\mathrm{d}x$$

where $0 < m_{\max} < \infty$ is an upper bound for the mass of a bacterium. The evolution equation for the couple $(S_t, p_t(x))$ has been established by Fredrickson et al. (1967) as the population balance equations for growth-fragmentation models (see also Ramkrishna, 1979), they read:

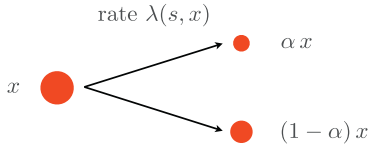
$$\dot{S}_t = D(\mathbf{s}_{\text{in}} - S_t) - \frac{k}{V}\int_0^{m_{\max}} \rho(S_t, x)p_t(x)\mathrm{d}x, \tag{4}$$

$$\frac{\partial}{\partial t}p_t(x) + \frac{\partial}{\partial x}\left(\rho(S_t, x)p_t(x)\right) + \left(\lambda(S_t, x) + D\right)p_t(x) = 2\int_0^{m_{\max}} \frac{\lambda(S_t, z)}{z} q\left(\frac{x}{z}\right) p_t(z)dz \tag{5}$$

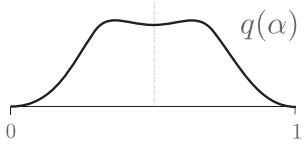
for $x \in [0, m_{\max}]$. Here, like in the previous model S_t is the substrate concentration (mg/l) which is assumed to be uniform in the vessel.

In (4) and (5), $\rho(s, x)$ and $\lambda(s, x)$ are respectively the growth function and the division rate of a bacterium of mass x with a substrate concentration s, the mass distribution of the daughter cells is represented by the probability density function $q(\alpha)$ on [0]. We detail these functions now:

(i) **Cell division** – Each individual of mass x divides itself at a rate $\lambda(s, x)$ into two individuals with respective masses αx and $(1-\alpha)x$:



where α is distributed according to a given probability density function $q(\alpha)$ on [0], and s is the substrate concentration. We suppose that the p.d.f. $q(\alpha)$ is symmetric with respect to 1/2, i.e. $q(\alpha) = q(1-\alpha)$:



Hence, the p.d.f. of the division kernel of a cell of mass x is $q(y/x)$ with support $[0, x]$. In the case of perfect mitosis, a cell of mass x is divided into two cells of masses $x/2$ so that $q(\alpha) = \delta_{1/2}(\alpha)$. We suppose that q is smooth (which is not the case for the perfect mitosis) and that $q(0) = q(1) = 0$. Thus, relatively to their mass, the division kernel is the same for all individuals. This allows us to reduce the model to a single division kernel but more complex possibilities can also be investigated.

(ii) **Mass growth** – The growth function $\rho : \mathbb{R}_+ \times [0, m_{\max}] \mapsto \mathbb{R}_+$ describes the evolution of the mass of an individual cell within the chemostat, i.e. in the model (4) and (5) the mass of an individual cell starting from the mass m_0 at a given time t_0 will evolve according to:

$$\dot{x}_t = \rho(S_t, x_t), \quad t \geq t_0, \quad x_0 = m_0$$

until the time of division or washout. To ensure the existence and uniqueness of the solution of (4) and (5) and of this last EDO, we assume that application $\rho(s,x)$ is Lipschitz continuous w.r.t. s uniformly in x. To ensure a coherence to this equations we also suppose that $0 \leq \rho(s,x) \leq \overline{\rho}$ for all $(s,x) \in \mathbb{R}_+ \times [0, m_{\max}]$, and that in the absence of substrate the bacteria do not grow, i.e. $\rho(0,x) = 0$ for all $x \in [0, m_{\max}]$. To ensure that the mass of a bacterium stays between 0 and $m_{\max}$, it is finally assumed that $\rho(s, m_{\max}) = 0$ for any $s \geq 0$.

In a relaxed context where $\rho(s,x)$ does not satisfy the previous hypothesis, it is easy to link the model (4) and (5) to the classic chemostat model (1) and (2). Indeed suppose that:

$$\frac{1}{V}\int_{\mathcal{X}} \rho(S_t, x) p_t(x)\mathrm{d}x = \mu(S_t) Y_t$$

which is the case when the growth function $x \mapsto \rho(s,x)$ is proportional to x, i.e. $\rho(s,x) = \mu(s)x$. First (4) reduces to (1) and then we can check that Y_t is solution of (2) (see details in Campillo and Fritsch, 2014).

2.3. The BDP model

We consider an hybrid model, where the substrate concentration S_t follows the same continuous/deterministic dynamic (1):

$$\dot{S}_t = D(\mathbf{s}_{\text{in}} - S_t) - k\mu(S_t)\frac{m}{V}\mathcal{Y}_t \tag{6}$$

but now m is the mean mass of an individual cell and $\mathcal{Y}_t$ is the number of cells in the chemostat. The dynamic of $\mathcal{Y}_t$ is discrete/stochastic, namely a birth and death stochastic process where at time t and conditionally to $\mathcal{Y}_t = n$, the process jumps from n to $n+1$ with rate $\mu(S_t)$ and jumps from n to $n-1$ with rate D, that is:

$$\mathcal{Y}_{t+h} = n + \begin{cases} 1 & \text{with probability } \mu(S_t)nh + o(h), \\ -1 & \text{with probability } Dnh + o(h), \\ 0 & \text{with probability } 1 - \mu(S_t)nh - Dnh + o(h), \\ i & \text{with probability } o(h) \text{ for all} i \neq 0, -1 \end{cases} \tag{7}$$

for infinitesimally small $h > 0$.

2.4. The IBM model

In the *individual-based model (IBM) structured in mass* the bacterial population is represented as a set of individuals growing in a perfectly mixed vessel of volume V (l). In this IBM each individual is characterized, not by its position, but only by its mass $x \in [0, m_{\max}]$. At time t the state of the system is defined by:

$$(S_t, \nu_t) \tag{8}$$

where S_t is the *substrate concentration* (mg/l) which is supposed to be uniform in the vessel; and ν_t is the state of the *bacterial population*, that is N_t individuals represented only by their mass: x_t^i (mg) will denote the mass of the individual number i for $i = 1, \ldots, N_t$.

It will be convenient to represent the population $\{x_t^i\}_{i=1,\ldots,N_t}$ at time t as the following sum of Dirac delta functions:

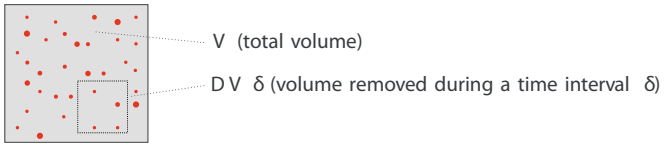
$$\nu_t(x) = \sum_{i=1}^{N_t} \delta_{x_t^i}(x). \tag{9}$$

where $\delta_{x_t^i}(x)$ is the Dirac delta function in x_t^i: $\int \phi(x)\delta_{x_t^i}(x)\mathrm{d}x = \phi(x_t^i)$ for any test function ϕ. For example $\int_{m_0}^{m_1} \nu_t(x)\mathrm{d}x$ is the number of cells with mass between m_0 and m_1 at time t; and $\int_{m_0}^{m_1} x\nu_t(x)\mathrm{d}x$ is the cumulated mass of cells with mass between m_0 and m_1 at time t (see Dieckmann and Law, 2000, for more details on this representation).

The IBM dynamics combine discrete evolutions (cell division and bacterial washout) as well as continuous evolutions (the growth of each individual and the dynamics of the substrate). We now describe the four components of the dynamics, first the discrete ones and then the continuous ones which occur between the discrete.

(i) **Cell division** – *Each individual of mass x divides at rate* $\lambda(s,x)$ *into two individuals of respective masses* αx *and* $(1-\alpha)x$ *where* α *is distributed according to the given p.d.f.* $q(\alpha)$ *on* [0, 1], *and s is the substrate concentration.*

(ii) **Washout** – *Each individual is withdrawn from the chemostat at rate D.* This mechanism is equivalent to a death process. In *perfect mixing* conditions, individuals are uniformly distributed in the volume V independently from their mass. During a time step δ, a total volume of $DV\delta$ is withdrawn from the chemostat:



and therefore, if we assume that all individuals have the same volume considered as negligible, during this time interval δ, an individual has a probability $D\delta$ to be withdrawn from the chemostat, D is the dilution rate. This rate could possibly depend on the mass of the individual.

At any time t, when the division of an individual occurs, the size of the population instantaneously jumps from N_t to $N_t + 1$; when an individual is withdrawn from the vessel, the

size of the population jumps instantaneously from N_t to $N_t - 1$; between each discrete event the size N_t remains constant and the chemostat evolves according to the following two continuous mechanisms:

(iii) **Growth of each cell** – *Each cell of mass x growths at speed* $\rho(S_t, x)$:

$$\dot{x}_t^i = \rho(S_t, x_t^i), \quad i = 1, \ldots, N_t \tag{10}$$

where $\rho : \mathbb{R}_+^2 \mapsto \mathbb{R}_+$ is given.

(iV) **Dynamics of the substrate concentration** – The substrate concentration evolves according to the ODE:

$$\dot{S}_t = D(\mathbf{s}_{\text{in}} - S_t) - k\tilde{\mu}(S_t, \nu_t) \tag{11}$$

where

$$\tilde{\mu}(s, \nu) \overset{\text{def}}{=} \frac{1}{V}\int_{\mathcal{X}} \rho(s, x)\nu(\mathrm{d}x) = \frac{1}{V}\sum_{i=1}^{N} \rho(s, x^i)$$

with $\nu = \sum_{i=1}^{N} \delta_{x^i}$. Mass balance leads to Eq. (11) and the initial condition S_0 may be random.

The IBM integrate the function $\lambda(s, x)$, $q(\alpha)$ and $\rho(s, x)$ already defined in the IDE but in a different way: the IDE uses them in an "average way" at the population level in contrast with the IBM that uses them for the explicit dynamics of each individual cell.

2.4.1. Convergence in distribution of the individual-based model

Campillo and Fritsch (2014) proved a result that we will now comment on from the application point of view. This result states a "functional" law of large numbers: in large population size the density of population given by the IBM is close to the density of population $p_t(x)$ given by (5). The population size should increase to infinity at any time t, for that purpose we replace the reference volume V by nV (or simply by n), let:

$$V_n = nV$$

We also suppose that the initial population size converges toward infinity with n:

$$\frac{1}{n}\nu_0^n \underset{[n\to\infty]}{\to} \xi_0 \text{weakly}$$

that is $\int_0^{m_{\max}} \phi(x)\nu_0^n(x)\mathrm{d}x = (1/n)\sum_{i=1}^{N_0^n} \phi(x_0^{i,n}) \to \int_0^{m_{\max}} \phi(x)\xi_0(x)\mathrm{d}x$ as $n \to \infty$, and we suppose that $\int_0^{m_{\max}} \xi_0(x)\mathrm{d}x > 0$. We suppose that the initial substrate concentration does not depend on n. Then define $(S_t^n, \nu_t^n(x))$ the IBM process where V is replaced by V_n and the rescaled process:

$$\overline{\nu}_t^n \overset{\text{def}}{=} \frac{1}{n}\nu_t^n.$$

Under these conditions Campillo and Fritsch (2014) stated that the process $(S_t^n, \overline{\nu}_t^n)_{0\le t\le T}$ given by the IBM converges toward the solution $(S_t, \overline{p}_t)_{0\le t\le T}$ of the IDE model (4) and (5) in a suitable sense with initial condition (S_0, ξ_0).

3. Simulation of the models

The simulation of the ODE system (1) and (2) is straightforward; for the simulation of the IDE system (4) and (5) we make use of an explicit Euler time-scheme coupled with an uncentered upwind finite difference space-scheme (see details in Appendix A).

3.1. Simulation of BDP model

The simulation of the system (6) and (7) is achieved with an adaptation of the classic "stochastic simulation algorithm" (SSA) also called "Gillespie algorithm" (Gillespie, 1977). It is an exact simulation algorithm, up to the approximation scheme for the ODE (6), in the sense that it simulates a realization of the exact distribution of the stochastic process $(S_t, \mathcal{Y}_t)$ given by (6) and (7). To apply the algorithm we need to suppose that there exists $\overline{\mu} < \infty$ such that:

$$\mu(s) \le \overline{\mu}, \quad \forall s \ge 0.$$

Then the SSA is given by the Algorithm 1.

Algorithm 1. Stochastic simulation algorithm (SSA) or Gillespie algorithm for the Monte Carlo simulation of the BDP model (6) and (7).

sample $(S_0, \mathcal{Y}_0)$
$\mathcal{Y} \leftarrow \mathcal{Y}_0$
$t \leftarrow 0$
while $t \le t_{\max}$ **do**
  $\tau \leftarrow (\overline{\mu} + D)\mathcal{Y}$
  $\Delta t \sim \text{Exp}(\tau)$
  integrate the equation for substrate (6) over $[t, t + \Delta t]$
  $t \leftarrow t + \Delta t$
  $u \sim U[0, 1]$
  if $u \le \mu(S_t)/(\overline{\mu} + D)$ **then**
    $\mathcal{Y} \leftarrow \mathcal{Y} + 1$% division
  else if $u \le (\mu(S_t) + D)/(\overline{\mu} + D)$ **then**
    $\mathcal{Y} \leftarrow \mathcal{Y} - 1$% washout
  end if
end while

3.2. Simulation of the IBM

We now detail the simulation procedure of the IBM. The division rate $\lambda(s, x)$ depends on the concentration of substrate s and on the mass x of each individual cell which continuously evolves according to the system (10) and (11), so to simulate the division of the cell we make use of a rejection sampling technique. It is assumed that there exists $\overline{\lambda} < \infty$ such that:

$$\lambda(s, x) \le \overline{\lambda}$$

hence an upper bound for the rate of event, division and washout combined, at the population level is given by:

$$\tau \overset{\text{def}}{=} (\overline{\lambda} + D)N.$$

Algorithm 2. "Exact" Monte Carlo simulation of the individual-based model: approximations only lie in the numerical integration of the ODEs and in the pseudo-random numbers generators.

sample $(S_0, \nu_0 = \sum_{i=1}^{N_0} \delta_{x_t^i})$ initial substrate concentration and population
$t \leftarrow 0$
$N \leftarrow N_0$ % initial population size
while $t \le t_{\max}$ **do**
  $\tau \leftarrow (\overline{\lambda} + D)N$
  $\Delta t \sim \text{Exp}(\tau)$
  integrate the equations for the mass (10) and the substrate (11) over $[t, t + \Delta t]$
  $t \leftarrow t + \Delta t$
  draw x uniformly in $\{x_t^i; i = 1, \ldots, N_t\}$
  $u \sim U[0, 1]$
  if $u \le \lambda(S_t, x)/(\overline{\lambda} + D)$ **then**
    $\alpha \sim q$
    $\nu_t \leftarrow \nu_t - \delta_x + \delta_{\alpha x} + \delta_{(1-\alpha)x}$ % division
    $N \leftarrow N + 1$
  else if $u \le (\lambda(S_t, x) + D)/(\overline{\lambda} + D)$ **then**
    $\nu_t \leftarrow \nu_t - \delta_x$ % washout
    $N \leftarrow N - 1$
  end if
end while

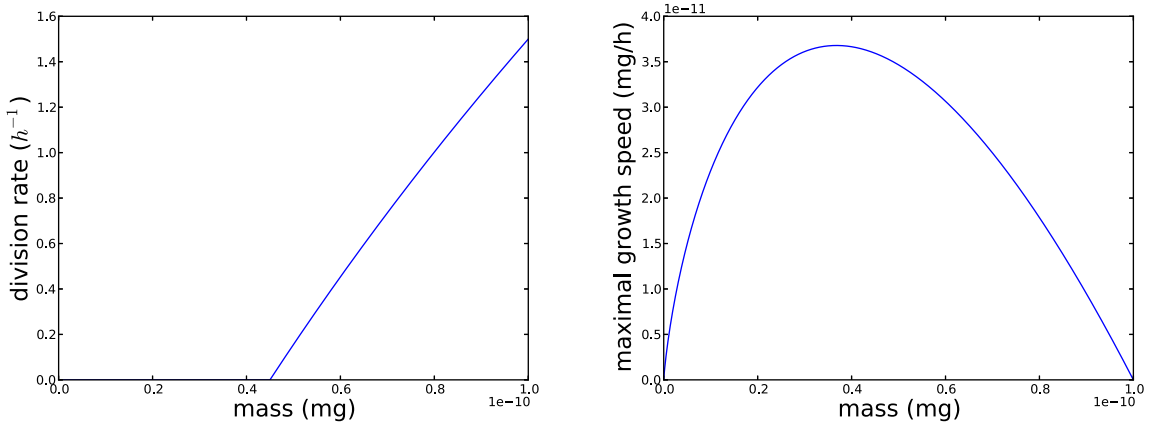


Fig. 1. (Left) Division rate function $\lambda(x)$ defined by (15) with $\overline{\lambda} = 1.5\,\mathrm{h}^{-1}$, $m_{\mathrm{div}} = 0.45 \times 10^{-10}$ mg and $p_\lambda = 6 \times 10^9$. (Right) Maximal growth speed with the Gompertz growth speed function (14) with $r_{\max} = 1.0\,\mathrm{h}^{-1}$, $m_{\max} = 1.0 \times 10^{-10}$ mg (namely the RHS of the inequality (14)).

At time $t + \Delta t$ with $\Delta t \sim \mathrm{Exp}(\tau)$, we determine if an event has occurred and what is its type by acceptance/rejection. To this end, the masses of the N individuals and the substrate concentration evolve according to the coupled ODEs (10) and (11). Then we choose uniformly at random an individual within the population $\nu_{(t+\Delta t)^-}$, that is the population at time $t + \Delta t$ before any possible event, let $x_{(t+\Delta t)^-}$ denotes its mass, then:

(i) With probability:

$$\frac{\overline{\lambda}}{(\overline{\lambda} + D)}$$

we determine if there has been division by acceptance/rejection:

- division occurs, that is:

$$\nu_{t+\Delta t} = \nu_{(t+\Delta t)^-} - \delta_{x_{(t+\Delta t)^-}} + \delta_{\alpha x_{(t+\Delta t)^-}} + \delta_{(1-\alpha)x_{(t+\Delta t)^-}} \quad \text{with} \quad \alpha \sim q \tag{12}$$

  with probability $\lambda(S_t, x_{(t+\Delta t)^-})/\overline{\lambda}$;
- no event occurs with probability $1 - \lambda(S_t, x_{(t+\Delta t)^-})/\overline{\lambda}$.

In conclusion, the event (12) occurs with probability:

$$\frac{\lambda\left(S_t, x_{(t+\Delta t)^-}\right)}{\overline{\lambda}} \frac{\overline{\lambda}}{(\overline{\lambda} + D)} = \frac{\lambda\left(S_t, x_{(t+\Delta t)^-}\right)}{(\overline{\lambda} + D)}.$$

(ii) With probability:

$$\frac{D}{(\overline{\lambda} + D)} = 1 - \frac{\overline{\lambda}}{(\overline{\lambda} + D)}$$

the individual is withdrawn, that is:

$$\nu_{t+\Delta t} = \nu_{(t+\Delta t)^-} - \delta_{x_{(t+\Delta t)^-}} \tag{13}$$

Finally, the events and the associated probabilities are:

- division (12) with probability $\lambda(S_t, x_{(t+\Delta t)^-})/(\overline{\lambda} + D)$,
- washout (13) with probability $D/(\overline{\lambda} + D)$

and no event (rejection) with the remaining probability. The details are given in Algorithm 2.

Technically, the numbering of individuals is as follows: at the initial time individual are numbered from 1 to N, in case division the daughter cell αx keeps the index of the parent cell and the daughter cell $(1 - \alpha)x$ takes the index $N + 1$; in the case of the washout, the individual N acquires the index of the withdrawn cell.

4. Simulation tests

We present simulations of four different models: the individual based-model (IBM), the integro-differential equation (IDE) (4) and (5), the classic chemostat model represented by the system of ordinary differential equations (ODE) (1) and (2), and the birth and death process (BDP) (6) and (7). These models can have similar or

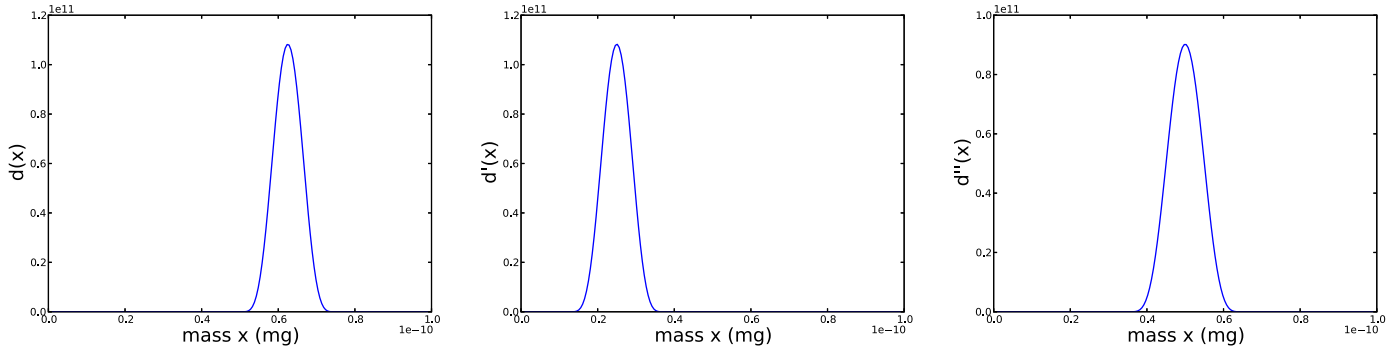


Fig. 2. Initial distributions d (left), d' (center) and d'' (right) respectively defined by Eqs. (16)–(18).

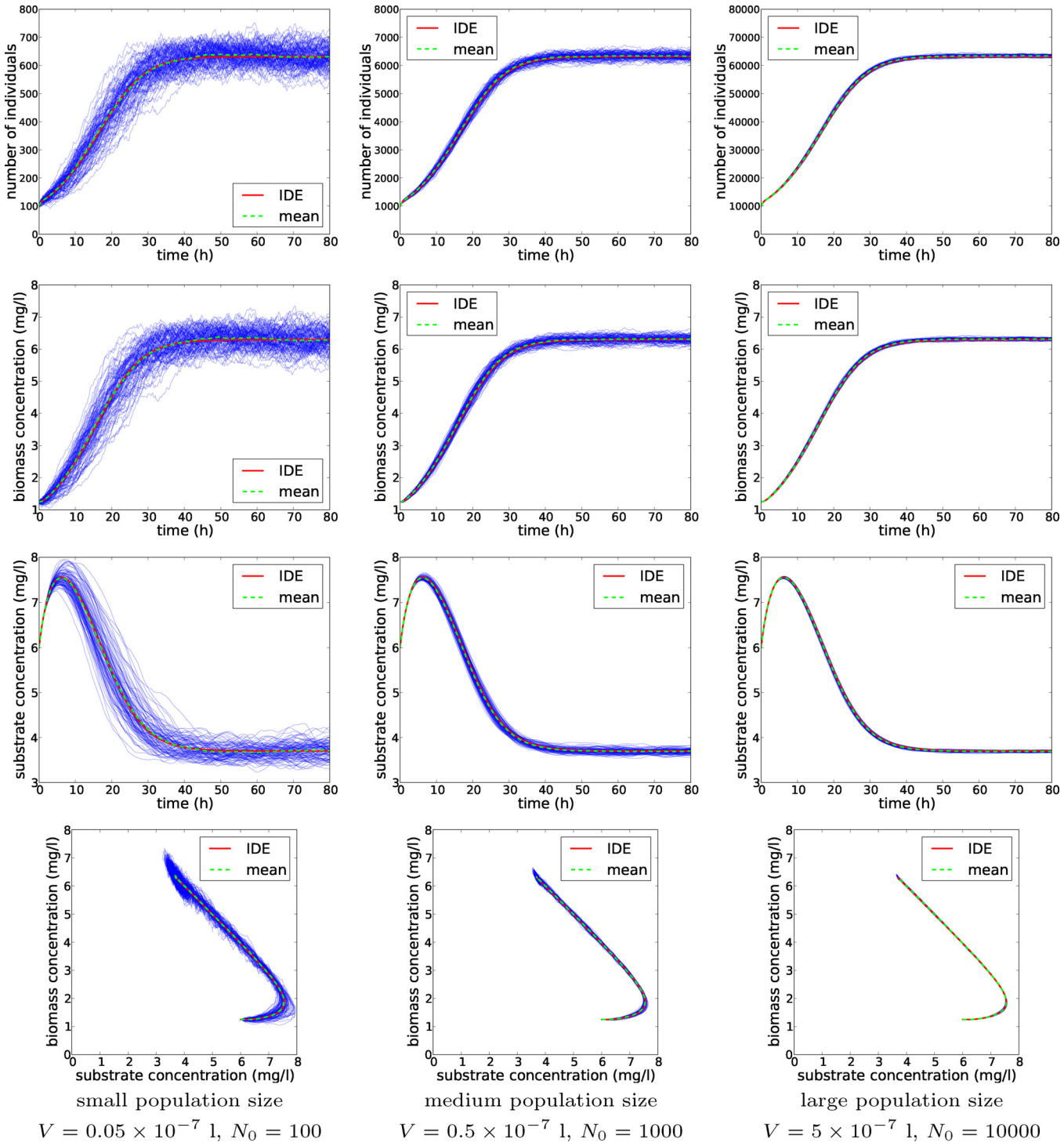


Fig. 3. From top to bottom: time evolutions of the population size, the biomass concentration, the concentration substrate and the concentrations phase portrait for the three levels of population sizes (from left to right: small, medium and large). The blue curves are the trajectories of 100 independent runs of IBM. The green curve is the mean value of these runs. The red curve is the solution of the IDE. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

different behaviors, depending on the model parameters and initial conditions.

Simulations of the BDP and of the IBM were performed respectively by Algorithms 1 and 2. The resolution of the integro-differential equation was made following the numerical scheme given in Appendix A, with a discretization step in the mass space of $\Delta x = 5 \times 10^{-14}$ mg and a discretization step in time of $\Delta t = 0.00125$ h. The numerical integration of the ODE (1) and (2) presents no difficulties and is performed by the function `odeint` of the module `scipy.integrate` of `Python` with the default parameters.

4.1. Simulation functions

We propose ad hoc functions ρ, q, d for simulation purposes. At fixed substrate concentration, individual growth is supposed to

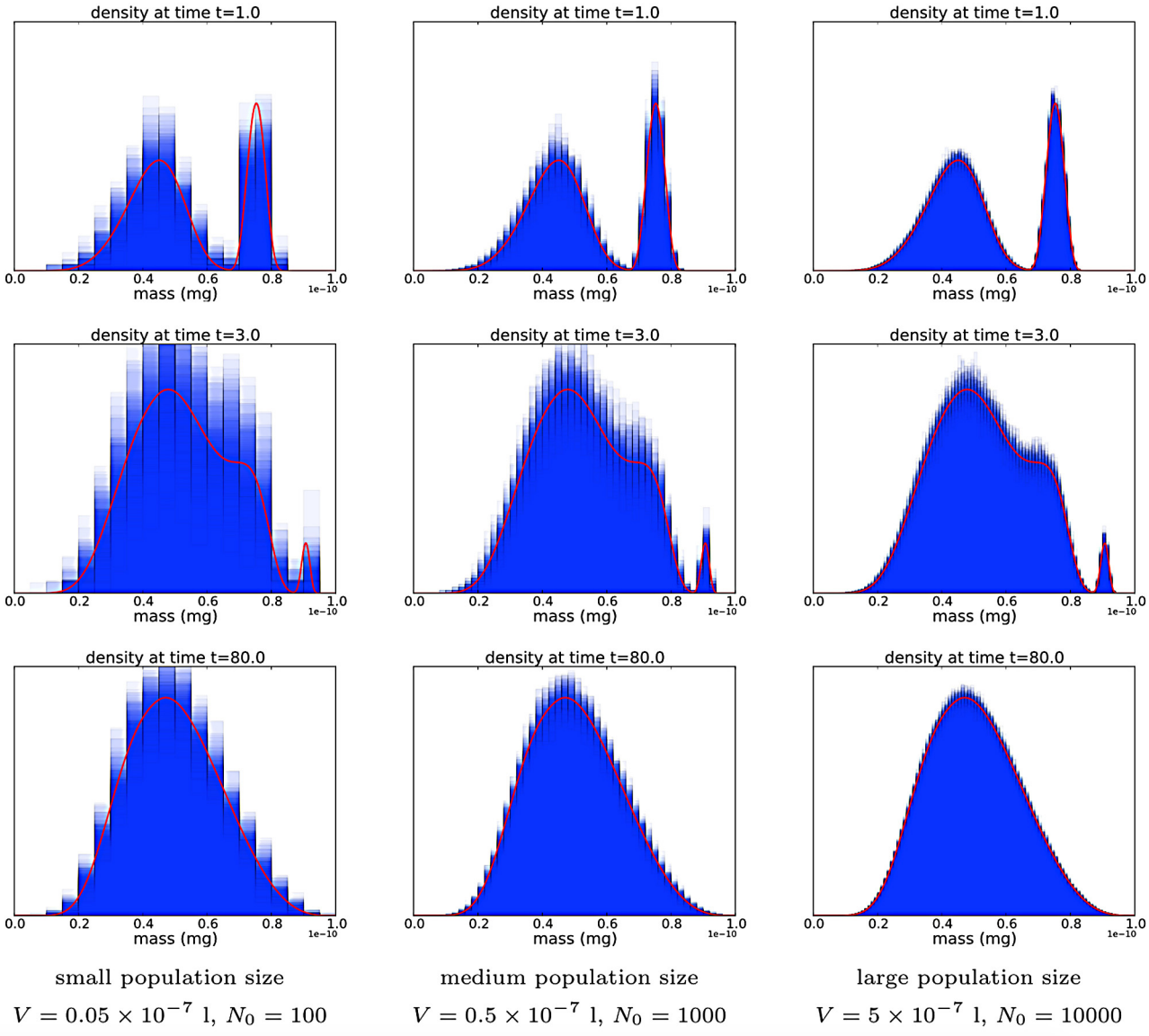


Fig. 4. Mass distribution represented as probability density functions for the time $t = 1$ h (top), $t = 3$ h (middle) and $t = 80$ h (bottom) in small (left), medium (middle) and large (right) population size. For each graph, the blue histograms represent the empirical mass distributions of individuals for the 100 independent runs of IBM. In order to plot the histogram we have adapted the number of bins according to the population size. The red curve represents the mass distribution given by the IDE. The dilution rate D is $0.25\,\mathrm{h}^{-1}$. Again we observe the convergence of the IBM solution to the IDE in large population limit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

be given by a Gompertz function. Moreover, we assume that the specific growth rate of the population depends on the substrate concentration and follows a Monod law:

$$\rho(s,x) = r_{\max}\frac{s}{k_r+s}\log\left(\frac{m_{\max}}{x}\right)x \le r_{\max}\log\left(\frac{m_{\max}}{x}\right)x \tag{14}$$

where $r_{\max}$ is the maximum specific growth rate of the population, k_r is the half-saturation constant and $m_{\max}$ is the maximal size of an individual.

We assume that one individual cannot divide below a mass m_{div}. For the simulations, we choose the following increasing division rate function:

$$\lambda(s,x) = \lambda(x) = \frac{\overline{\lambda}}{\log((m_{\max}-m_{\mathrm{div}})p_\lambda+1)}\log((x-m_{\mathrm{div}})p_\lambda+1)1_{\{x\ge m_{\mathrm{div}}\}} \tag{15}$$

where $p_\lambda > 0$ is a parameter of curvature of the function, see Fig. 1. This "ad hoc" function has been chosen as it meets the desired conditions.

The proportion α of the division kernel $q(\alpha)$ will be computed by a symmetric beta distribution:

$$q(\alpha) = \frac{1}{B(p_\beta)}(\alpha(1-\alpha))^{p_\beta-1}$$

where $B(p_\beta) = \int_0^1 (\alpha(1-\alpha))^{p_\beta-1}\mathrm{d}\alpha$ is a normalizing constant.

The initial distribution of individual masses is given by the following probability density function:

$$d(x) = \frac{1}{C_d}\left(\frac{x-0.5\times10^{-10}}{0.25\times10^{-10}}\left(1-\frac{x-0.5\times10^{-10}}{0.25\times10^{-10}}\right)\right)^5 1_{\{0.5\times10^{-10}<x<0.75\times10^{-10}\}} \tag{16}$$

where C_d is a normalizing constant. This initial distribution, represented in Fig. 2, is a beta distribution which charges big masses.

Table 1
Simulation parameters.

Parameters	Values
S_0	6 mg/l
$\mathbf{s}_{\text{in}}$	10 mg/l
D	0.25 h^{-1}
$m_{\max}$	1.0×10^{-10} mg
m_{div}	0.45×10^{-10} mg
$\overline{\lambda}$	1.5 h^{-1}
p_λ	6×10^{9}
p_β	10
$r_{\max}$	1 h^{-1}
k_r	6 mg/l
k	1

Simulations starting from this distribution will show a transient phenomenon that cannot be reproduced by the classic chemostat model described in terms of ordinary differential equations (1) and (2), see Fig. 6.

4.2. Comparison of the IBM and the IDE

In this section we illustrate the convergence of the IBM to the IDE. In order to do this, we increase the volume of the chemostat and the initial number of individuals in a proportional way. We perform simulations at three levels of population size. The small size level is performed with $V = 0.05 \times 10^{-7}$ l and $N_0 = 100$, the medium one with $V = 0.5 \times 10^{-7}$ l and $N_0 = 1000$ and the large one with $V = 5 \times 10^{-7}$ l and $N_0 = 10,000$. The initial distributions of individual masses are the same, so that the initial biomass concentration is the same for the three sets of parameters.

For each level of population size, we simulate 100 independent runs of the IBM in order to observe the reduction of variance when we increase the number of bacteria. The IDE is numerically approximated using the finite difference schemes detailed in Appendix A. The parameters, which are in the same order of magnitude than the one used by Henson (2003), are given in Table 1. Note that the stoechiometric coefficient k equals 1. This may seem biologically questionable but, by a change of variable, it is equivalent to $k > 1$ with a different half-saturation constant k_r and a different substrate input $\mathbf{s}_{\text{in}}$ (or equivalently by changing the unit of substrate concentration).

Figs. 3 and 4 illustrate the convergence of IBM to IDE. The variances in the evolution of the biomass concentration and of the substrate concentration, as well as the relative variance of the number of individuals, decrease when we increase the number of individuals, see Fig. 3. The normalized size distributions at times $t = 1$, 3 and 80 (h) are represented in Fig. 4 for the IDE (red curve) and 100 independent runs of the IBM (blue histograms) for the small, the medium and the large population. Note that the number of bins was adapted according to the scale of the population in order to obtain clear graphics.

The normalized solution of the IDE (5) is represented in Fig. 5. It corresponds to the time evolution of the normalized mass distribution. At the initial instant this distribution is given by the function (16). It then becomes bimodal. The lower mode corresponds to the bacteria from the division. The upper mode represents bacteria of the initial distribution before their division or washout. We observe the same phenomenon in the realization of IBMs, see Fig. 4. In contrast, the classic chemostat model presented below, see Eqs. (1) and (2), cannot account for this phenomenon. After this transient phenomenon, the normalized mass distribution converges to a stationary state.

As the IDE is the limit of the IBM in large population size, the behavior of the IDE gives informations on the behavior of the IBM. But there is no reason that the IDE corresponds to the mean value of the IBM, because of the correlation between the individuals behaviors.

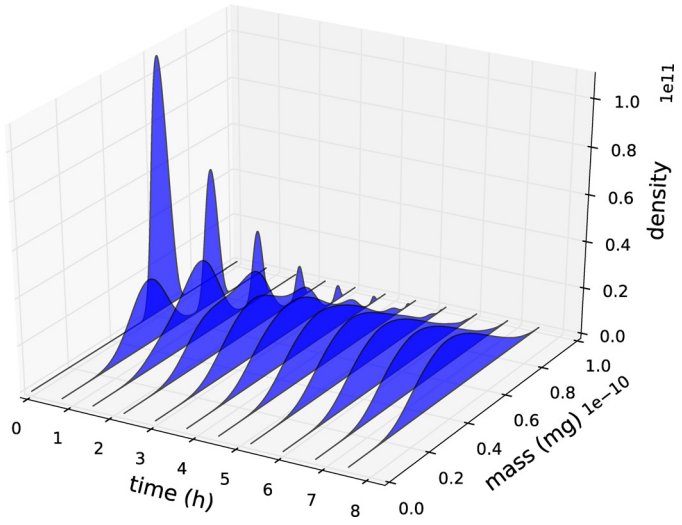


Fig. 5. Time evolution of the mass distributions represented as probability density functions for the IDE (5): we represent the simulation until time $T = 8$ (h) only to illustrate the transient phenomenon caused by the choice of the initial distribution (16). After a few iterations in time this distribution is bimodal, the upper mode growths in mass and disappears before $T = 8$ (h).

4.3. Comparison of the IBM, the IDE and the ODE

We now compare the IBM and the IDE to the classic chemostat model described by the system of ODE's (1) and (2). The growth model in both the IBM and the IDE is of Monod type, so for the ODE model we also consider the classic Monod kinetics (3). The parameters of this Monod law are not given in the initial model and we use a least squares method to determine the value of the parameters $\mu_{\max}$ and K_s which minimize the weighted quadratic distance between $(S_t, X_t)_{t\le T}$ given by (1) and (2) and $(\overline{S}_t, \overline{X}_t)_{t\le T}$, where $\overline{S}_t$ and $\overline{X}_t$ are the means of the variable S_t and $X_t = V^{-1}\int_\chi x\nu_t(x)dx$ given by the IBM (8). This quadratic distance is weighted by the variance of the IBM.

Fig. 6 represents the evolution of the number of individuals, the biomass concentration, the substrate concentration and the trajectories in the phase space for 60 independent runs of the IBM and for the IDE with parameters of Table 2 and with different initial density. The initial number N_0 is adapted so that the average initial biomass concentration is the same in all three cases.

First we consider a simulation based on the initial mass density $d(x)$ defined by (16). With this initial density both the IDE and the IBM feature a transient phenomenon described in the previous section and illustrated in Figs. 5 and 4. Fig. 6 (left) shows a significant difference between the IBM and the IDE on the one hand and

Table 2
Simulation parameters.

Parameters	Values
S_0	5 mg/l
$\mathbf{s}_{\text{in}}$	10 mg/l
D	0.2 h^{-1}
$m_{\max}$	1.0×10^{-10} mg
m_{div}	0.4×10^{-10} mg
$\overline{\lambda}$	1 h^{-1}
p_λ	10×10^{9}
p_β	7
$r_{\max}$	1 h^{-1}
k_r	10 mg/l
k	1

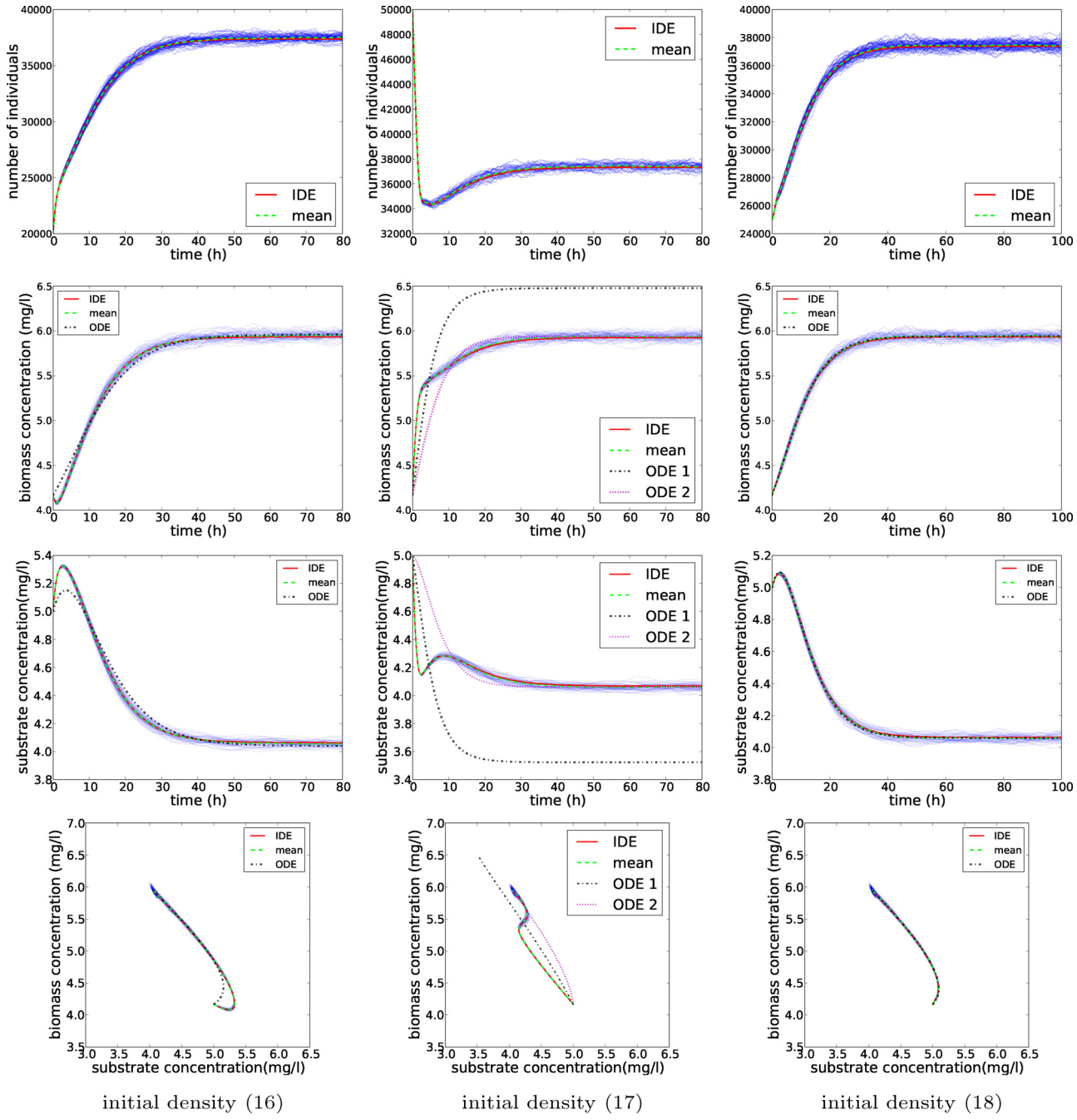


Fig. 6. Top to bottom: Time evolution of the number of individuals, the biomass concentration, the substrate concentration and the concentration trajectories in the phase space according to the initial mass distributions (16) (left), (17) (middle) and (18) (right). In blue, the trajectories of 60 independent runs of the IBM simulated with $V = 3 \times 10^{-7}$ l and $N_0 = 20000$ (left), $N_0 = 50000$ and $N_0 = 25,000$ (right); in green, the mean of the IBM runs; in red, the solution of IDE (4) and (5); in black, the solution of the system (1) and (2). The latter is fitted by the least squares method on the IBM, the parameters of the Monod law (3) are $\mu_{max} = 0.329$ and $K_s = 2.603$ (left), $\mu_{max} = 11.556$ and $K_s = 200.0$ for EDO1 and $\mu_{max} = 9.219$ and $K_s = 183.065$ for EDO2 (middle), $\mu_{max} = 0.397$ and $K_s = 3.991$ (right). Note that in the case (17) (middle) it is unrealistic to fit a classic EDO chemostat to the IBM, indeed it is not possible for that model to fit the transitory behavior of the IBM leading to unrealistic values for the parameters of the chemostat. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the ODE on the other hand, the latter model cannot account for the transient phenomenon. With the first two models, the individual bacteria are withdrawn uniformly and independently of their mass (large mass bacteria have the same probability of withdrawal as small mass bacteria). As the initial state $d(x)$ has a substantial proportion of bacteria with large mass, we have an important division rate at the population scale and a relatively low growth of individuals (see Fig. 1). Therefore at the beginning of the simulation there is an important increase of the number of individuals whereas the biomass decreases. The ODE is naturally not able to account for this transient phenomenon.

Conversely, if we choose an initial density which charges the low masses (see Fig. 2), as the following:

$$d'(x) = \frac{1}{C_{d'}} \left(\frac{x - 0.125 \times 10^{-10}}{0.25 \times 10^{-10}} \left(1 - \frac{x - 0.125 \times 10^{-10}}{0.25 \times 10^{-10}} \right) \right)^5 \times 1_{\{0.125 \times 10^{-10} < x < 0.375 \times 10^{-10}\}} \quad (17)$$

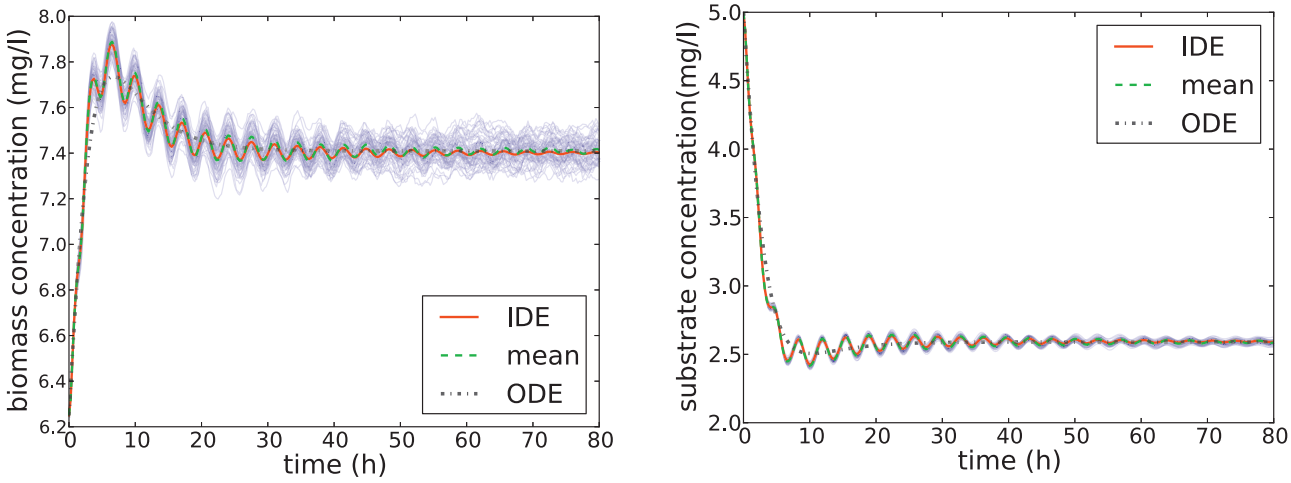


Fig. 7. Evolutions of the biomass (left) and the substrate (right) concentrations of 60 independent runs of the IBM (blue), the mean of the IBM (green), the IDE (red), the ODE (black) fitted by the least squares method on the IBM. The parameters of the Monod law (3) of the ODE are $\mu_{\max} = 0.537$ and $K_s = 4.363$. The division rate function is given by Eq. (19). $\overline{\lambda} = 5\,\mathrm{h}^{-1}$, $m_{\mathrm{div}} = 0.5 \times 10^{-10}$ mg, $p_\beta = 100$, $V = 1.0 \times 10^{-7}$ l, $N_0 = 10,000$. Other parameters are given in Table 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where $C_{d'}$ is a normalizing constant, we observe an important increase in the biomass at the beginning of the simulation for the IBM and the IDE whereas the number of individuals decrease (see Fig. 6 (middle)), which is due to fact that at the beginning of the simulation individuals have masses too low to divide, but with a high "speed of growth" (see Fig. 1). As the randomness is low at the beginning of the simulation of IBMs, the least squares method, weighted by the variance of IBMs, give an ODE which have a strong increase of the biomass concentration and a strong decrease of the substrate concentration near the initial instant, but the stationary state of the ODE (black curves) does not match to the quasi-stationary state of the IBM or the stationary state of the IDE. If we give a high weight to the quasi-stationary state (between $t = 40$ h and $t = 80$ h), we obtain an ODE (magenta curves) with a stationary state which matches the quasi-stationary state of the IBM, but with a significant difference during the transitory state.

This phenomena no longer appear if we use the following density (see Fig. 2):

$$d''(x) = \frac{1}{C_{d''}} \left(\frac{x - 0.35 \times 10^{-10}}{0.3 \times 10^{-10}} \left(1 - \frac{x - 0.35 \times 10^{-10}}{0.3 \times 10^{-10}} \right) \right)^5 1_{\{0.35\times10^{-10} < x < 0.65\times10^{-10}\}} \tag{18}$$

where $C_{d''}$ is a normalizing constant. Indeed, from Fig. 6 (right), the different simulations are comparable, the ODE and the IDE match substantially.

Fig. 7 shows simulations with the following division rate function:

$$\lambda(s, x) = \overline{\lambda} 1_{\{x \geq m_{\mathrm{div}}\}}, \tag{19}$$

with $\overline{\lambda} = 5\,\mathrm{h}^{-1}$, $m_{\mathrm{div}} = 0.5 \times 10^{-10}$ mg and the parameter of the division kernel is $p_\beta = 100$.

Another interesting phenomenon is that we can observe oscillations in the evolution of the biomass and substrate concentrations for the IBM and the IDE, which can not be accounted for by the ODE. These oscillations are due to the distribution which remain bimodal with alternation of the higher density between the lower and the upper mode (see Fig. 8). When the lower mode has a higher density than the upper mode, there are a lot of individuals which grow rapidly, the biomass concentration increases and the substrate concentration decreases. When the upper mode has a higher density than the lower mode, there are more individuals with low growth, the biomass concentration subsequently decreases and the substrate concentration increases.

This oscillation phenomenon is artificial and not, to the best of our knowledge, connected to a real biological case. Our purpose is to mention a rather obvious but nevertheless important point, that IBM and IDE models account for phenomena which cannot be illustrated with the classical ODE model.

4.4. Study of the total washout

One of the main differences between deterministic and stochastic models lies in their way of accounting for the total washout phenomenon (or extinction phenomenon in the case of an ecosystem). With a sufficiently small dilution rate D, the solutions of the system (1) and (2) and of the IDE (4) and (5) converge to an equilibrium point with strictly positive biomass. In fact, the total washout is an unstable equilibrium point and apart from the line corresponding to the null biomass, the complete phase space corresponds to a basin of attraction leading to a solution with a strictly positive biomass asymptotic point. However, from Fig. 9, among the 1000 independent runs of the IBM, 111 converge to the total washout before time $t = 1000$ h; so the probability of total washout at this instant is approximately 11%. The ODE 1 (dot-dashed black line) is fitted to the 1000 IBMs. We can observe that it corresponds to the mean. The ODE 2 (dotted cyan line) is fitted on the non-extinct IBM and corresponds to the mean conditionally to the non extinction. It may be noted that the IDE and the ODE do not correspond to the average value of the IBM since only the latter may reflect the total washout in a finite time horizon.

Now we consider a sufficiently large dilution rate, $D = 0.5\,\mathrm{h}^{-1}$, corresponding to the total washout conditions. Fig. 10 (top) presents the evolution of the biomass concentration in the different models. The runs of the IBM converge to the total washout in finite time whereas both deterministic ODE and IDE models converge exponentially to total washout without ever reaching it in finite time. Fig. 10 (bottom) shows the empirical distribution of the total washout time calculated from 7000 independent runs of the IBM (red curve). This total washout time features a relatively large variance.

It is known that for a birth-death process with constant rates $\tilde{\lambda}$ and D which corresponds respectively to the rates of birth and

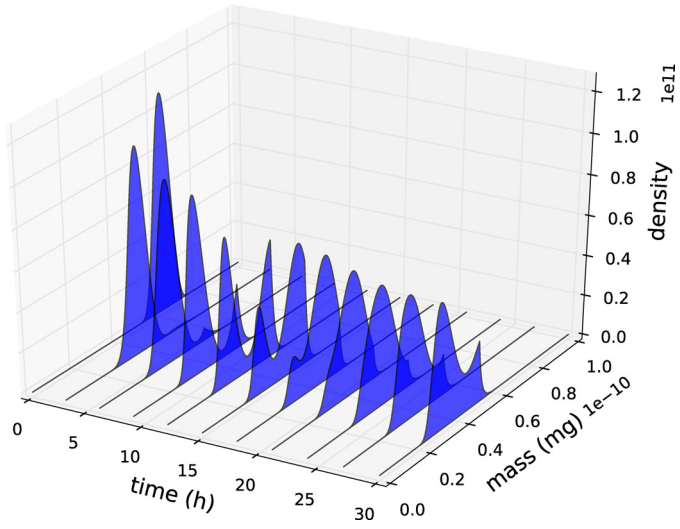


Fig. 8. Time evolution of the normalized mass distribution for the IDE (5) with the division rate function (19), $\bar{\lambda} = 5\,\mathrm{h}^{-1}$, $m_{\mathrm{div}} = 0.0005$ mg, $p_\beta = 100$, $V = 1.0 \times 10^{-7}$ l, $N_0 = 10,000$. Other parameters are given in Table 2.

death and with $\tilde{\lambda} < D$, the probability density function of the time of extinction T is

$$d(t) = N_0 D \frac{(\tilde{\lambda} - D)^2 e^{(\tilde{\lambda}-D)t}}{(\tilde{\lambda} e^{(\tilde{\lambda}-D)t} - D)^2} \left(\frac{D e^{(\tilde{\lambda}-D)t} - D}{\tilde{\lambda} e^{(\tilde{\lambda}-D)t} - D} \right)^{N_0 - 1}. \tag{20}$$

When the birth rate is not constant, we can expect that the probability density function of the time of extinction is of the form (20) where $\tilde{\lambda}$ is the average birth rate of the population. Fig. 10 shows the probability density function (20) (green dotted curve) where $\tilde{\lambda}$ is computed by a least squared method in order to be fitted to the empirical distribution of the total washout time (red solid curve). This constant $\tilde{\lambda}$ depends on the model parameters, in particular on the initial number of individuals N_0 and on the initial distribution of individuals. In our example the initial distribution contains bacteria with higher mass than the quasi-stationary distribution, then in this case the effective division rate near the time $t = 0$ is higher than the quasi-stationary effective division rate and therefore, the constant $\tilde{\lambda}$ will also be higher. Moreover, the higher the initial number of individuals N_0 is, the more negligible is the time taken to reach the quasi-stationary distribution. The dashed blue curve represents the empirical law of the extinction time of the BDP, calculated from 7000 independent runs of the BDP (performed with Algorithm 1), where the function μ in Eqs. (6) and (7) is a Monod function (3) with the same parameters as the ODE fitted on the IBM.

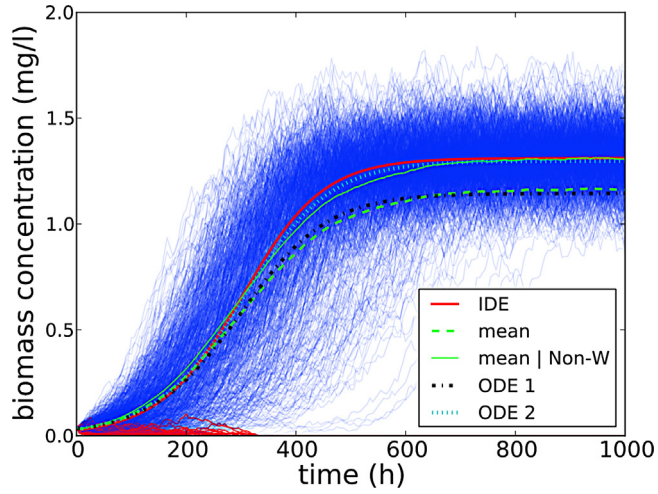
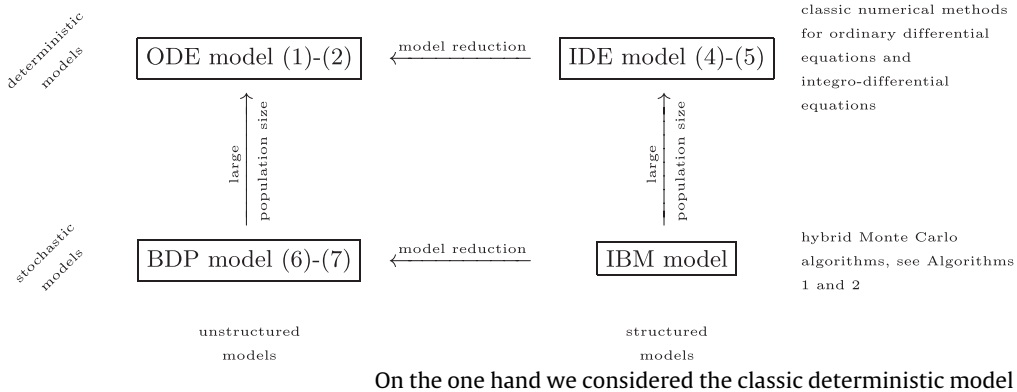


Fig. 9. Time evolution of the biomass concentration. In blue, 1000 independent realizations of the IBM simulated with $V = 0.5 \times 10^{-7}$ l and $N_0 = 30$; in green, the mean of these runs; in red, the solution of the IDE; in black, the solution of the ODE 1 with parameter values $\mu_{\max} = 0.432$ and $K_s = 5.050$, fitted on the IBM, weighted by the variance. In cyan, the solution of the ODE 2 with parameters values $\mu_{\max} = 0.406$ and $K_s = 4.142$, fitted on the IBM given by the non extinction of the population, weighted by the variance of non-extinct populations. Parameters are given in Table 2. The dilution rate D is $0.275\,\mathrm{h}^{-1}$. Among the 1000 independent runs of the IBM, 111 lead to total washout while the deterministic models converge to an equilibrium with strictly positive biomass. The mean value of the 1000 runs of the IBM gives account for the total washout probability while IDE and ODE models do not account for this question. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Discussion

In this work we presented four models of the chemostat together with the analytical and algorithmic gateways bridging one to the other:



On the one hand we considered the classic deterministic model of chemostat as a system of ODE's, and also a birth and death stochastic process hybridized with an ODE describing the substrate dynamics; on the other hand their mass-structured counterparts, a deterministic IDE and also a stochastic IBM hybridized with an ODE. In all cases the evolution of the substrate is represented as an ODE meaning that this part of the model is reasonably represented as a fluid dynamic limit deterministic process. The stochastic model are Markov processes with values in $\mathbb{R}_+ \times \mathbb{N}$ for the unstructured model and with values in $\mathbb{R}_+ \times \mathcal{M}([0, m_{\max}])$ for the mass-structured model. The Markov property allows us to analytically prove the convergence of the stochastic models toward their deterministic counterpart in large population size limits. Moreover, the reduction from the mass-structured models to unstructured ones is obtained by a simplification of the growth function.

The numerical simulations of deterministic models are straightforward and are done using classic integration schemes. The

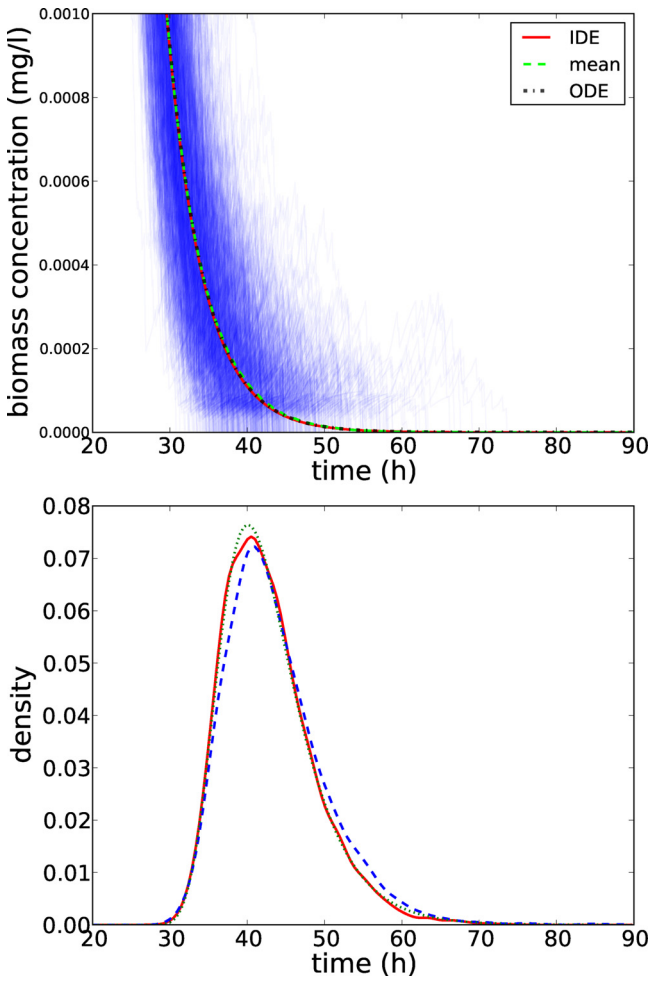


Fig. 10. (Top) Evolution of biomass concentration between $t = 20$ and $t = 90$ h: blue, 1000 independent runs of the IBM; in green, the mean value of these runs; in red the solution of the IDE; in black, the solution to the ODE with parameters $\mu_{max} = 0.578$ and $K_s = 10.0$. The parameters are $V = 10 \times 10^{-7}$ l and $N_0 = 10,000$, the dilution rate D is 0.5 h^{-1}, others parameters are the same as those in Table 2. For both deterministic models, the size of the population decreases exponentially rapidly to 0 but remains strictly positive for any finite time. However, all the runs of the IBM reach total washout in finite time. (Bottom) The continuous red line is empirical distribution of the total washout time calculated from 7000 independent runs of the IBM and plotted using a time kernel regularization. The dashed blue line is the empirical distribution of the total washout time calculated from 7000 independent runs of the birth-death process with the same parameters as the ODE matched on the IBMs. The distribution is also plotted using a time kernel regularization. The green dotted line is the p.d.f. (20) with $N_0 = 10000$, $D = 0.5$ et $\tilde{\lambda} = 0.2922$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

numerical simulation of random models uses almost exact Monte Carlo algorithms, indeed the models are hybrid and the integration of the ODE part of the model is achieved through approximation schemes. These latter algorithms are not computationally realistic in large populations as all events, cell division and cell washout, are explicitly simulated; but it is precisely at this level that the simulation of the deterministic models took over, the whole framework being perfectly consistent.

Still discrete/stochastic models allow the practitioner to explore concepts that are not within the scope of continuous/deterministic approaches. For example, questions of diversity and dynamics within species reduced to small number of individuals can be apprehended only through the former models. The same remark is valid for the question of extinction.

It is important to evaluate the complexity of the models in terms of analysis as well as simulation. For example, it is difficult to determine an *optimal* control law for the IBM while this task is relatively easy in the case of the classic ODE model. In this latter case there is already a large number of results, while in the former case the criteria to optimize are still not well established. However, it is pertinent to test a control law developed on the ODE model (1) and (2) not on the same model but on simulated data generated from the IBM.

Despite their relative complexity, stochastic discrete models are essential in more than one aspect for population dynamics (Hellweger and Bucci, 2009). On the one hand, they allow us to explore situations where deterministic models are totally blind, this is particularly the case for situations close to extinction conditions or near wash-out conditions in the case of chemostat. This question may also be relevant in larger population size (Campillo and Lobry, 2012). On the other hand, they offer a non-reproducible simulation tool close to conditions encountered in practice. As the biologist Georgy Gause pointed out in 1934: "When the microcosm approaches the natural conditions [...] the struggle for existence begins to be controlled by such a multiplicity of causes that we are unable to predict exactly the course of development of each individual microcosm. From the language of rational differential equations we are compelled to pass on to the language of probabilities, and there is no doubt that the corresponding mathematical theory of the struggle for existence may be developed in these terms" (Gause, 1934).

However, stochastic and discrete modeling is essentially devoted to evolutionary population dynamics. It is only more recently that this approach has been extended to all areas of population dynamics with a similar concern to encourage cooperation between different representations of a model (Andrews et al., 2009). It is interesting to note that the same approach is now also adopted in epidemiology where considerations of discrete and random aspects of population dynamics in small sizes are essential (Allen and Lahodny, 2012; Allen and van den Driessche, 2013).

The IBM proposed here is certainly not the most efficient in terms of computational speed: it is asynchronous and requires the simulation of each individual event. There are strategies that accelerate this IBM thanks to approximation techniques. The proposed IBM has the advantage of being an exact Monte Carlo simulation, up to approximation schemes of the ODE, of the very stochastic process which we can analyze and prove the weak convergence in large population toward the ID model. This important property is due to the fact that all the models considered here, including the deterministic ones, are Markovian and that the study of weak convergence of these processes is an important tool in terms of mathematics but also on a practical level in simulation terms.

Finally, this work advocates for the development of hybrid models relevant when the size of a given population fluctuates between large and small values, or when multiple populations are involved some in large sizes, others in small sizes.

Acknowledgements

The authors are grateful to Claude Lobry for discussions on the model, to Pierre Pudlo and Pascal Neveu for their help concerning the programming of the IBM. This work is partially supported by the project "Modèles Numériques pour les écosystèmes Microbiens" of the French National Network of Complex Systems (RNSC call 2012). The work of Coralie Fritsch is partially supported by the Meta-omics of Microbial Ecosystems (MEM) metaprogram of INRA.

Appendix A. Numerical integration scheme for the IDE

To numerically solve the system of integer-differential equations (4) and (5), we make use of finite difference schemes.

Given a time step Δt and a mass step $\Delta x = L/I$, with $I \in \mathbb{N}^*$, we discretize the time and mass space with:

$$t_n = n\Delta t, \quad x_i = i\Delta x.$$

We introduce the following approximations:

$$p_{n,i} \simeq p_{t_n}(x_i), \quad s_n \simeq S_{t_n}.$$

We also suppose first that at the initial time step there is no individual with null mass in the vessel, i.e. $p_{0,0} = 0$; and second that individual with null mass cannot be generated during the cell division step, i.e. q is regular with $q(0) = 0$. This assumption was not necessary in the mathematical development presented in the previous sections but is naturally required to obtain reasonable mass of individuals in the simulation.

For time integration we use an explicit Euler scheme, for space integration, an uncentered upwind difference scheme, which leads to the coupled integration scheme:

$$\frac{p_{n+1,i} - p_{n,i}}{\Delta t} = -\rho(s_n, x_i)\frac{p_{n,i} - p_{n,i-1}}{\Delta x} - \frac{\partial}{\partial x}\rho(s_n, x_i)p_{n,i} - (\lambda(s_n, x_i) + D)\, p_{n,i} + 2\Delta x \sum_{j=1}^{I} \frac{\lambda(s_n, x_j)}{x_j} q\left(\frac{x_i}{x_j}\right) p_{n,j},$$

$$\frac{s_{n+1} - s_n}{\Delta t} = D(\mathbf{s}_{\text{in}} - s_n) - \frac{k}{V}\Delta x \sum_{j=1}^{I} \rho(s_n, x_j)p_{n,j}$$

for $n \in \mathbb{N}$ and $i = 1, \cdots I$, with the boundary condition:

$$p_{n+1,0} = 0$$

and given initial conditions $p_{0,i}$ and s_0.

We finally get:

$$p_{n+1,i} = p_{n,i} + \Delta t \left\{ -\rho(s_n, x_i)\frac{p_{n,i} - p_{n,i-1}}{\Delta x} - \frac{\partial}{\partial x}\rho(s_n, x_i)p_{n,i} - (\lambda(s_n, x_i) + D)\, p_{n,i} + 2\Delta x, \sum_{j=1}^{I} \frac{\lambda(s_n, x_j)}{x_j} q\left(\frac{x_i}{x_j}\right) p_{n,j} \right\}$$

$$s_{n+1} = s_n + \Delta t \left\{ D(\mathbf{s}_{\text{in}} - s_n) - \frac{k}{V}\Delta x \sum_{j=1}^{I} \rho(s_n, x_j)p_{n,j} \right\}$$

for $n \in \mathbb{N}$ and $i = 1, \ldots I$ with boundary condition $p_{n+1,0} = 0$ and given initial conditions $p_{0,i}$ and s_0.

References

Allen, L., van den Driessche, P., 2013. Relations between deterministic and stochastic thresholds for disease extinction in continuous- and discrete-time infectious disease models. Math. Biosci. 243 (1), 99–108.

Allen, L.J.S., Lahodny, G.E., 2012. Extinction thresholds in deterministic and stochastic epidemic models. J. Biol. Dyn. 6 (2), 590–611.

Andrews, S.S., Dinh, T., Arkin, A.P., 2009. Stochastic models of biological processes. In: Meyers, R. (Ed.), Encyclopedia of Complexity and System Science, vol. 9. Springer, New York, pp. 8730–8749.

Campillo, F., Champagnat, N., 2012. Simulation and analysis of an individual-based model for graph-structured plant dynamics. Ecol. Model. 234, 93–105.

Campillo, F., Fritsch, C., 2014. Weak convergence of a mass-structured individual-based model. Appl. Math. Optim., 1–37.

Campillo, F., Joannides, M., Larramendy-Valverde, I., 2011. Stochastic modeling of the chemostat. Ecol. Model. 222 (15), 2676–2689.

Campillo, F., Lobry, C., 2012. Effect of population size in a predator–prey model. Ecol. Model. 246, 1–10.

Crump, K.S., O'Young, W.-S.C., 1979. Some stochastic features of bacterial constant growth apparatus. Bull. Math. Biol. 41 (1), 53–66.

Deygout, C., Lesne, A., Campillo, F., Rapaport, A., 2013. Homogenised model linking microscopic and macroscopic dynamics of a biofilm: application to growth in a plug flow reactor. Ecol. Model. 250, 15–24.

Dieckmann, U., Law, R., 2000. Relaxation projections and the method of moments. In: Dieckmann, U., Law, R., Metz, J.A.J. (Eds.), The Geometry of Ecological Interactions: Simplifying Spatial Complexity. Cambridge University Press, Cambridge, pp. 412–455.

Faugeras, B., Maury, O., 2007. Modeling fish population movements: from an individual-based representation to an advection–diffusion equation. J. Theor. Biol. 247 (4), 837–848.

Ferrer, J., Prats, C., López, D., Vives-Rego, J., 2009. Mathematical modelling methodologies in predictive food microbiology: a SWOT analysis. Int. J. Food Microbiol. 134 (1–2), 2–8.

Fredrickson, A.G., Ramkrishna, D., Tsuchiya, H.M., 1967. Statistics and dynamics of procaryotic cell populations. Math. Biosci. 1 (3), 327–374.

Gause, G.F., 1934. The Struggle for Existence. Willians and Willians, Baltimore.

Gillespie, D., 1977. Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. 81 (25), 2340–2361.

Gómez-Mourelo, P., Ginovart, M., 10, 2009. The differential equation counterpart of an individual-based model for yeast population growth. Comput. Math. Appl. 58 (7), 1360–1369.

Grasman, J., De Gee, M., Herwaarden, O.A.V., 2005. Breakdown of a chemostat exposed to stochastic noise volume. J. Eng. Math. 53 (3), 291–300.

Hellweger, F.L., Bucci, V., 2009. A bunch of tiny individuals – individual-based modeling for microbes. Ecol. Model. 220 (1), 8–22.

Henson, M.A., 2003. Dynamic modeling and control of yeast cell populations in continuous biochemical reactors. Comput. Chem. Eng. 27 (8-9), 1185–1199.

Imhof, L., Walcher, S., 2005. Exclusion and persistence in deterministic and stochastic chemostat models. J. Diff. Equ. 217 (1), 26–53.

Monod, J., 1950. La technique de culture continue, théorie et applications. Ann. Inst. Pasteur 79 (4), 390–410.

Novick, A., Szilard, L., 1950. Description of the chemostat. Science 112 (2920), 715–716.

Ramkrishna, D.,1979. Statistical models of cell populations. In: Advances in Biochemical Engineering, vol. 11. Springer, Berlin Heidelberg, pp. 1–47.

Smith, H.L., Waltman, P.E., 1995. The Theory of the Chemostat: Dynamics of Microbial Competition. Cambridge University Press, Cambridge.

Stephanopoulos, G., Aris, R., Fredrickson, A.G., 1979. A stochastic analysis of the growth of competing microbial populations in a continuous biochemical reactor. Math. Biosci. 45, 99–135.