

*A discrete, size-structured model of phytoplankton
growth in the chemostat.*

Julien Arino — Jean-Luc Gouzé — Antoine Sciandra

N° 3963

Juin 2000

THÈME 4



*Rapport
de recherche*

A discrete, size-structured model of phytoplankton growth in the chemostat.

Julien Arino , Jean-Luc Gouzé , Antoine Sciandra *

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet Comore

Rapport de recherche n° 3963 — Juin 2000 — 23 pages

Abstract: We introduce inhomogeneous, substrate dependent cell division in a nonlinear matrix model of size-structured population growth in the chemostat, first introduced by Gage *et al.* [7] and later analyzed by Smith [11]. We show that mass conservation is verified, and conclude that our system admits one non zero globally stable equilibrium, which we express explicitly. We then proceed to several numerical simulations, and briefly compare the prediction of the model to data, whose obtention we discuss.

Key-words: chemostat, structured model, discrete model, inhomogeneous division size

* Observatoire océanographique de Villefranche sur Mer

Un modèle discret, structuré en taille, de croissance du phytoplancton en chemostat.

Résumé : Nous introduisons une division cellulaire non homogène et dépendante du substrat dans un modèle matriciel non linéaire de population structurée en taille. Le modèle d'origine, introduit par Gage, Williams et Horton [7], et étudié plus tard analytiquement par Smith [11], décrit la croissance de micro-organismes dans un chemostat. Nous montrons que la conservation de la masse est vérifiée par le nouveau modèle, et par conséquent concluons qu'il admet un équilibre non trivial globalement asymptotiquement stable, que nous exprimons analytiquement. Nous procédons ensuite à quelques simulations numériques, et comparons brièvement les prédictions du modèle à des données, dont nous discutons l'obtention.

Mots-clés : chemostat, modèles structurés, modèle discret, taille de division inégale

Contents

1	Model formulation	5
2	Model behavior	9
3	Numerical results	15
4	Experimental data	20
5	Discussion	21

Introduction

A chemostat is a continuous culture device in which organisms (bacteria, phytoplankton) grow, submitted to a flow of nutrient. Chemostats have been extensively studied mathematically (see [12]), mainly using ordinary differential equations. The global behavior of a chemostat is well known, using models such as the Monod model or the Droop model. But these models describe the behavior of the total biomass of the system.

One of the possible ways to gain further knowledge on the behavior of chemostats is to use structured models, describing the evolution of the population in more detail. Moreover, it is now possible to obtain long time series of structured data from chemostats (*e.g.* the automated device described in [4], which will be described more thoroughly in Section 3). Therefore, comparison of model behavior to the data should lead to refinements of the models [3].

However, structured models of the chemostat are less known. Most works on the latter use partial differential equations. In [9] a size structured model is presented, which is extended to the competition case in [5], while [10] introduces cell cycle structuration. However, the mathematical analysis of these models is often hard (one often has to reduce the model to an ODE system by considering moments of the density), so as their numerical integration.

Another kind of structured modelling is the use of discrete time, discrete structure systems (see [6, chapters 1 and 3] for a review of structured discrete models), While the mathematical analysis of such systems can be as tedious as in the continuous case, they have the advantage that they are easy to simulate.

Such a model, time discrete and structured in biomass (size) classes, was introduced in [7] by Gage *et al.*. They showed that a stable distribution of the biomass is reached, in which the biomass is constant and equal for all size classes. They also studied numerically the influence of various factors (number of size classes, flow rate) on the convergence speed. Later, Smith [11] corrected a mistake in the formulation of the model, showed mathematically that the equilibrium is globally stable, and also introduced competition.

But this model is based on a very strong assumption: all cells are born with the same biomass b , and divide when they reach biomass $2b$. Gage *et al.* obtain biomass spectra which resemble the experimentally observed spectra, but by using the assumption that the biomass in each class is log-normally distributed [13].

The aim of the present work is to extend the model of Gage *et al.* to the case where cell division (and consequently, cell birth) can happen for cells in several biomass classes, distributed around a mean division biomass. Using this approach, the division process is included in the model. This is closely related to the models of [9] and [10], where the division rate is defined using a probability density.

This paper is organized as follows: first, we derive our model from the one of [7], introducing division for cells of different sizes. We then analyze the model; using the fact that it conserves the mass, we are able to follow the analysis of [11] and to conclude to the existence of a globally stable non trivial equilibrium, which we characterize. Sample equilibrium distributions are then shown in Section 3, as well as examples of the transient behavior of the

model. Finally, prior to the discussion, biological data is discussed, and briefly compared to model predictions.

1 Model formulation

We suppose that the system under consideration is a well-stirred chemostat. Since the model will be formulated in discrete time, we define T , the iteration period. Let E ($0 < E < 1$) be the dilution rate per iteration period.

In Gage *et al.*, the following biological assumptions are made.

- (H1) In a constant environment (*ie* if the concentration of limiting nutrient is constant and high), the growth of a cell is exponential.
- (H2) Cells are born with a biomass b , they grow, then divide when they reach a biomass $2b$.
- (H3) When a cell divides, it divides into two daughter cells, whose individual biomass is exactly one half of the biomass of the original cell.

We add the following one, to relax the division size hypothesis.

- (H4) The birth biomass b (accordingly, the division biomass $2b$) is not a constant. There exists an average birth biomass b_d (accordingly, an average division biomass $2b_d$) around which the individual birth biomasses are distributed.

We suppose that there is a total number r of biomass classes, and that the minimal birth biomass of a cell is b_{min} (the smallest biomass that a cell can possibly be born with). We use a Michaelis-Menten function, which we note $f(S) = m \frac{S}{a+S}$ (m is the maximal growth rate per iteration period, and a is called the half-saturation constant), to model nutrient uptake and consequent growth of the cells. But the results hold for all functions $f(S)$ such that $f(0) = 0$, $f'(S) > 0$ and $f''(S) < 0$ (see [11, p. 739]). The state variables are $x_t = (x_1(t), \dots, x_r(t))^T$, where $x_i(t)$ is the total biomass in biomass class i at time t , and S_t , the concentration of substrate in the chemostat chamber at time t .

We suppose that the r biomass classes are divided as follows: there are r_g (*growth*) size classes, during which the dynamics is the same as in [7, 11]: the cells in these classes can either proceed to the next biomass class (if they grow of a sufficient amount) or stay in the class (if the growth is not sufficient). There are no births nor divisions in these classes. Following this stable stage, there are r_d (*division*) classes, during which cell division occurs. Correspondingly, we assume that there are r_b (with $r_b = r_d$) size classes prior to the stable stage, in which the dividing cells “fall” (see Figure 1). Hereafter, we will call the latter classes *birth* classes. Therefore, $r = r_b + r_g + r_d$ is the total number of classes, and the structure of the model can be decomposed as follows:

- $i = 1, \dots, r_d$ birth classes.

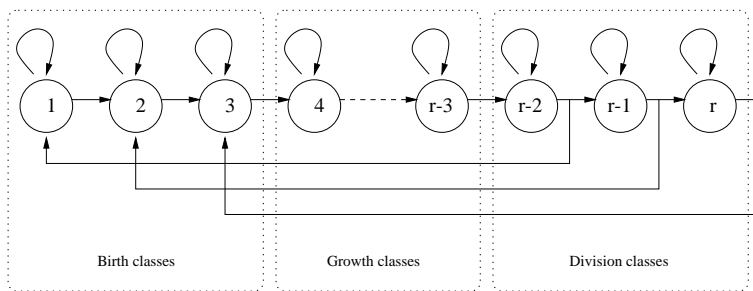


Figure 1: Structure of the model. Example with $r_d = 3 (= r_b)$.

- $i = r_d + 1, \dots, r_d + r_g$ growth classes.
- $i = r_d + r_g + 1, \dots, r$ division classes.

To describe the way a cell progresses along the size classes, we use a constant, M . Since hypothesis (H1) must be fulfilled, M has to account for an exponential growth of the cells. Suppose that a cell has exponential growth between biomasses b and $2b$. If we want to “track” this biomass as it progresses along, say n classes, then using $M = 2^{1/n}$ does the trick. Indeed, defining $M^{i-1}b$ to be the biomass of the cell in class i , we have an exponential progression of the biomass from b to $2b$.

But if we use $M = 2^{1/r}$ as in [7, 11], then the only cells that will double in size are the ones that are born in the first birth class and that divide in the last division class. So in order to account for hypothesis (H2) and to have a one to one correspondence between division classes and birth classes, M has to be chosen differently. Let b_d be the mean birth biomass. This biomass corresponds to one of the birth classes (for example, in Fig.1, we could suppose that class 2 contains the cells which are born with the mean birth biomass b_d). We suppose that cells that are born with biomass b_d undergo division when their biomass has doubled. According to the way we defined the classes, to each division class corresponds one and only one birth class. Therefore M has to be chosen so that the biomass doubles when a cell moves from the class containing b_d to the class containing $2b_d$. Let k be the index of that birth class ($k \in \{1, \dots, r_d\}$). Then k is also the index (division class-wise) of the class which contains biomass $2b_d$, and M is a function of the number of classes that separates these two classes.

Thus $M = 2^{1/(r_b+r_g+1)}$. It has to be noted that M only depends on the number of birth (or division, equivalently) and growth classes. Therefore b_d can be chosen to lie in any birth class.

In order to determine the proportion of cells moving from one class to the next, we proceed as follows. We suppose that there is no washout (*ie* $E = 0$). Class i contains at time t a total biomass $x_i(t)$. Following absorption of nutrient, this biomass becomes $x_i(t)f(S_i)$. Since class i contains cells whose average biomass is $M^{i-1}b_{min}$, and class $i + 1$

contains cells with an average biomass of $M^i b_{min}$, the number of cells from class i moving to class $i + 1$ is $x_i(t)f(S_t)/(M - 1)M^{i-1}b_{min}$. Now this has to be related to the number of cells in class i , which is approximatively given by $n_i(t) = x_i(t)/M^{i-1}b_{min}$.

As a consequence, the proportion of biomass moving from one class to the next is given by

$$P_t = \frac{f(S_t)}{M - 1}$$

Now, in order to follow hypothesis (H4), we suppose that cells that are in a division class and grow sufficiently, can either divide, with a certain proportion, or proceed to the next division class. Furthermore, we suppose that this proportion is a function of the substrate concentration, that can differ in different division classes (*e.g.* one could assume that bigger cells divide even in low substrate concentrations, while smaller cells do not).

Hence let us denote by $D_i(S)$ the proportion of cells in division class i that divide. Then $1 - D_i(S)$ is the proportion of cells of class i that grow instead of dividing. The way these functions depend on the substrate concentration will not be specified, but one can think of Holling type II functional response, or of sigmoidal functions. All that we require is that this function be bounded, that is for $i = 1, \dots, r_d - 1$ and for all $S \in \mathbb{R}_+$, $0 \leq D_i(S) \leq 1$, where the inequality has to be strict for some S . This can be formulated the following way: we suppose that there exists a nonempty subset $S_{int} \subset \mathbb{R}_+$, $S_{int} \neq \{0\}$, defined by

$$S_{int} = \{\bar{S} \in \mathbb{R}^+ : \exists \mathcal{N}(\bar{S}), \forall S \in \mathcal{N}(\bar{S}), \forall i < r_d, D_i(S) \in (0, 1)\} \quad (1)$$

where $\mathcal{N}(\bar{S})$ is a neighborhood of \bar{S} . Note that we do not require that S_{int} be a connected set. This subset will be useful in Theorem 2.3.

In order to constrain the cell sizes, we suppose that in the last division class all cells divide, *ie* that $D_{r_d}(S) = 1$ for all S .

From a biological point of view, these hypotheses mean that there exist substrate concentrations such that: the only size class in which all cells divide is the last one; there are no division classes in which no cells divide. For example, suppose that one observes cell division in cells having a size between 10 and 15 micrometers. There may exist substrate concentrations such that all cells reaching a size of say $12 \mu m$ divide (*e.g.* if subject to starvation). The hypotheses we have just made ensure that this is not always the case, *ie* there also exist substrate concentrations such that some cells grow to reach a size of $15 \mu m$.

Finally, taking into account the washout rate, which applies to all size classes at the same rate ([7]), the model is written as follows.

$$x_{t+1} = (1 - E)A(S_t)x_t, \quad t \geq 0 \quad (2a)$$

$$S_{t+1} = (1 - E)[S_t - f(S_t)U_t] + ES^0 \quad (2b)$$

where $U_t = \mathbb{1}^T x_t$ is the total biomass at time t ($\mathbb{1} = (1, \dots, 1)^T$), and $A(S_t)$ is the following transition matrix (of size $r \times r$).

$$A(S_t) = \begin{bmatrix} \alpha(S_t) & \beta(S_t) \end{bmatrix} \quad (3)$$

with

$$\alpha(S_t) = \begin{bmatrix} 1 - P_t & 0 & \dots & & & & & & & 0 \\ MP_t & & & & & & & & & \\ & & \ddots & \ddots & & & & & & \\ & & & MP_t & 1 - P_t & & & & & \\ \hline & & & MP_t & 1 - P_t & & & & & \\ & & & & & \ddots & \ddots & & & \\ \hline & & & & & & MP_t & 1 - P_t & & \\ \hline 0 & \dots & & & & & \dots & & & MP_t \\ \vdots & & & & & & & & & \vdots \\ 0 & \dots & & & & & \dots & & & 0 \end{bmatrix}$$

a $(r_b + r_g, r)$ -matrix, where the blocks correspond respectively to r_b , r_g and r_d rows, and

$$\beta(S_t) = \begin{bmatrix} MP_t D_1 & & & & & & & & & \\ & MP_t D_2 & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & & & & MP_t D_{r_d-1} & & & \\ \hline & & & & & & & & & MP_t \\ 0 & \dots & & & \dots & & & & & 0 \\ \vdots & & & & & & & & & \vdots \\ 0 & \dots & & & \dots & & & & & 0 \\ \hline 1 - P_t & & & & & & & & & \\ MP_t(1 - D_1) & 1 - P_t & & & & & & & & \\ & & \ddots & \ddots & & & & & & \\ & & & & & & 1 - P_t & & & \\ & & & & & & MP_t(1 - D_{r_d-1}) & 1 - P_t & & \end{bmatrix}$$

is a (r_d, r) -matrix. The first, second and last blocks are r_b , r_g and r_d rows long, respectively. In these matrices, we use the following notation: $P_t = f(S_t)(M - 1)^{-1}$, and $D_i = D_i(S_t)$.

In order to keep P_t in a reasonable interval (*ie* $P_t \in [0, 1]$), some restrictions have to be made.

The minimal doubling time D_{min} and the maximal growth rate μ_{max} are two expressions of the same quantity (since $D_{min} = \ln 2 / \mu_{max}$), and are species constants.

What we must constrain is the iteration period T , since the maximal growth rate per iteration period m is given by $m = \mu_{max} T$.

Now, what happens in general cases? We are given a μ_{max} (or equivalently a D_{min}), we choose a number of classes, and want to study the behavior of the system. Hence the determination of T is what comes last.

The first constraint to be verified is that $m(M-1)^{-1} \leq 1$. Therefore we obtain $T \leq \frac{2^{1/(r_b+r_g)}-1}{\mu_{max}}$. But the constraint $rT \leq D_{min}$ expressed in [11] also has to be satisfied. Written in terms of μ_{max} and adapted to our model, it reads $T \leq \frac{\ln 2}{(r_b+r_g)\mu_{max}}$.

Hence, supposing μ_{max} given and the number of classes chosen, we must require that

$$T \leq \min\left(\frac{2^{1/(r_b+r_g)}-1}{\mu_{max}}, \frac{\ln 2}{(r_b+r_g)\mu_{max}}\right) \quad (4)$$

Now another problem which could arise is that $U_t f(S_t) > S_t$ for some t . This would lead to negative values of S_{t+1} , which must of course be forbidden. In order to avoid this problem, U_t has to be constrained. We proceed as in [11]. We fix an $\eta \in (0, 1)$, and require that $f(S)U/S < \eta$. Defining $W > S^0$ an upper bound on $U + S$ that depends on the range of initial conditions that one wishes to accommodate, we require that

$$\frac{m}{a}W < \eta \quad (5)$$

We observe that these conditions should not be viewed as conditions on m , but only on the iteration period T and the number of classes r .

2 Model behavior

Let us begin by verifying that the *conservation principle* (or *mass conservation principle*) holds. We have

$$\mathbb{1}^T A(S) = [1 + (M-1)P_t]\mathbb{1}^T = (1 + f(S_t))\mathbb{1}^T$$

therefore

$$U_{t+1} = \mathbb{1}^T x_t = \mathbb{1}^T A(S_t)x_t = (1 - E)(1 + f(S_t))U_t$$

As a consequence,

$$U_{t+1} + S_{t+1} = (1 - E)(U_t + S_t) + ES^0 \quad (6)$$

which means that mass conservation is verified. In words all the mass that is present in the chemostat at time t is still present at time $t + 1$, save for what enters (ES^0) and for the quantity $E(U_t + S_t)$ that is washed out.

Hence all the results in the single population case of [11] hold. The method that is used is the following.

- Since mass conservation is verified, it is quite easy to show that the total “mass” contained in the chemostat tends to a fixed quantity, namely S^0 .
- Therefore, the dynamics of the system can be studied on the invariant set $U + S = S^0$, inside of which the system reduces to a one dimensional system.
- On this set, the global behavior of the system can be studied. Smith showed that under certain conditions, there exists a globally stable equilibrium for this simplified system.

- Following this, the global dynamics of the two dimensional system can be deduced.
- Finally, using a result of Golubitsky *et al.* [8], the distribution of biomass in all classes can be deduced.

For the convenience of the reader, we recall here some of the proofs of [11].

First, we can solve (6) easily, and obtain

$$U_{t+1} + S_{t+1} = S^0 - (S^0 - U_0 - S_0)(1 - E)^t, \quad t \geq 1 \quad (7)$$

Let us now define Γ , a positive bounded set, as $\Gamma = \{(x, S) \in \mathbb{R}_+^{r+1}; \mathbb{1}^T x + S \leq W\}$ (W being defined as in Eq. (5)).

Proposition 2.1 ([11]) *If $(x_0, S_0) \in \Gamma$, then $(x_t, S_t) \in \Gamma$ for $t \geq 1$, $S_t - U_t f(S_t) > 0$ for $t \geq 1$ and*

$$S_t + U_t \rightarrow S^0, \quad t \rightarrow \infty \quad (8)$$

Proof As we are in a discrete case, to show the positive invariance of Γ , we need only show that $(x_1, S_1) \in \Gamma$ if $(x_0, S_0) \in \Gamma$. Let $(x_0, S_0) \in \Gamma$. Then $U_0 + S_0 \leq W$. Two cases arise.

First, if $S_0 > 0$, then

$$\frac{f(S_0)}{S_0} U_0 \leq f'(0)W = \frac{m}{a}W < \eta < 1$$

so $S_0 - f(S_0)U_0 > (1 - \eta)S_0 > 0$. Therefore, $S_1 > (1 - E)(1 - \eta)S_0 + ES^0 > 0$.

Now, if $S_0 = 0$ then $S_1 = ES^0 > 0$. Using (6), we have $U_1 + S_1 = (1 - E)(U_0 + S_0) + ES^0 < (1 - E)W + EW = W$ and also $U_1 = (1 - E)U_0 \geq 0$.

Finally, Equation (8) follows immediately from Equation (7), since $1 - E < 1$. \square

Let us now consider the system restricted to the positively invariant set $\{(U, S) \in \mathbb{R}_+^2; U + S = S^0\}$. On this set, we can use a technique standard to those systems that conserve the mass: we replace S by $S^0 - U$ (with $0 \leq U \leq S^0$ since S is positive). Thus, when restricted to this set the system (2a) (2b) becomes

$$U_{t+1} = (1 - E)(1 + f(S^0 - U_t))U_t \quad (9)$$

Proposition 2.2 *If $(1 - E)(1 + f(S^0)) \leq 1$, then $\lim_{t \rightarrow \infty} U_t = 0$, for all solutions of (9) with $U_0 \in [0, S^0]$. If $(1 - E)(1 + f(S^0)) > 1$, then $\lim_{t \rightarrow \infty} U_t = \tilde{U}$, for all solutions of (9) with $U_0 \in [0, S^0]$.*

Proof [[11]] Let $F(U) = (1-E)(1+f(S^0-U))U$. Then $U_{t+1} = F(U_t)$. The global behavior of Eq. (9) will be obtained by studying the fixed points of F . $F : [0, S^0] \rightarrow [0, (1-E)S^0]$, and $F(S^0) = (1-E)S^0$. F is an increasing function of U , since

$$\begin{aligned} F'(U) &= (1-E)[1+f(S^0-U) - Uf'(S^0-U)] \\ &> (1-E)[f(S^0-U) + 1 - W\frac{m}{a}] \\ &> (1-E)[f(S^0-U) + 1 - \eta] \\ &> 0 \end{aligned}$$

by (5). On the other hand, $F'(U)$ is decreasing in U since $f'(S)$ is decreasing in S .

Two cases have to be considered. First, if $F'(0) \leq 1$, 0 is the only fixed point of F in $[0, S^0]$. Indeed, by the Mean Value theorem, there exists $\xi \in (0, U_t)$ such that $U_{t+1} = F(U_t) - F(0) = F'(\xi)U_t$. Since F' is decreasing and that $F'(0) \leq 1$, we have $U_{t+1} < U_t$, and so $\lim_{t \rightarrow \infty} U_t = 0$.

Now if $F'(0) > 1$, we have $F(S^0) < S^0$, and therefore, denoting by F^t the t^{th} iterate of F , $F^t(S^0) \searrow \tilde{U}$ as $t \rightarrow \infty$. On the other hand, for a small $\epsilon > 0$, $F(\epsilon) > \epsilon$, so $F^t(\epsilon) \nearrow \tilde{U}$ as $t \rightarrow \infty$. Let now $U_0 \in (0, S^0]$. For small enough ϵ , we have $\epsilon < U_0 \leq S^0$, and therefore $F^t(\epsilon) < F^t(U_0) \leq F^t(S^0)$, which finally implies that $F^t(U_0) \rightarrow \tilde{U}$ as $t \rightarrow \infty$. \square

To compute the positive fixed point of F , we define λ as the unique solution, when it exists, of $F(U) = U$, i.e

$$f(\lambda) = (1-E)^{-1} - 1 \quad (10)$$

Noting $B = (1-E)^{-1} - 1$, one finds

$$\lambda = \frac{aB}{m-B} \quad (11)$$

if $m > B$, and no solution otherwise. This is the classical chemostat behavior: if the dilution rate E is larger than the maximal growth rate of the organisms, then the population cannot compensate the loss due to the outflow, and it becomes extinct.

Finally, if $\lambda < S^0$, \tilde{U} is given by

$$\tilde{U} = S^0 - \lambda \quad (12)$$

Now the dynamics of the 2 dimensional system

$$U_{t+1} = (1-E)(1+f(S_t))U_t \quad (13a)$$

$$S_{t+1} = (1-E)(S_t - f(S_t)U_t) + ES^0 \quad (13b)$$

can be studied. Let $\Omega = \{(U, S) \in \mathbb{R}_+^2; U + S < W\}$.

Theorem 2.1 ([11]) *If $(1-E)(1+f(S^0)) < 1$, then for all solutions such that $(U_0, S_0) \in \Omega$,*

$$(U_t, S_t) \rightarrow (0, S^0), \quad t \rightarrow \infty$$

If $(1 - E)(1 + f(S^0)) > 1$, then there exists a non zero steady state, and for all solutions such that $(U_0, S_0) \in \Omega$,

$$(U_t, S_t) \rightarrow (S^0 - \lambda, \lambda), \quad t \rightarrow \infty$$

Proof Let us begin by ruling out the trivial case. If $U_0 = 0$, then for all $t \geq 0$, $U_t = 0$ and, since $U_t + S_t \rightarrow S^0$ as $t \rightarrow \infty$, $S_t \rightarrow S^0$ as $t \rightarrow \infty$. Hence in all that follows, we will suppose that $U_0 > 0$.

Solutions of (13a) (13b) starting in Ω approach the invariant line segment $S + U$ in Ω at an exponential rate $(1 - E)^t$. On this invariant line segment the dynamics are given in Proposition 2.2 and depend on whether $(1 - E)(1 + f(S^0))$ is larger than one or less than one.

Let us first consider the case where $(1 - E)(1 + f(S^0)) < 1$. Then we can find $\mu \in (0, 1)$ and $\bar{S} > S^0$ such that $(1 - E)(1 + F(S^0)) < \mu$ for all $S \in [0, \bar{S}]$. Since $U_t + S_t \rightarrow S^0$ as $t \rightarrow \infty$, it follows that $S_t < \bar{S}$ for all large t , say $t \geq \tau$. Consequently, $U_{t+k} \leq \mu^k U_\tau$ for $k \geq 0$ and therefore $U_t \rightarrow 0$ as $t \rightarrow \infty$. From this and $U_t + S_t \rightarrow S^0$ as $t \rightarrow \infty$, we see that $S_t \rightarrow S^0$. Therefore in this case $(0, S^0)$ is globally stable.

Let us now consider the second case, ie $(1 - E)(1 + f(S^0)) > 1$. For convenience, let us define

$$G(U, S) = ((1 - E)(1 + f(S))U, (1 - E)(S - f(S)U) + ES^0)$$

for $(U, S) \in \Omega$. $G : \Omega \rightarrow \Omega$. Then (13a) (13b) becomes

$$(U_{t+1}, S_{t+1}) = G(U_t, S_t)$$

We first study the local stability of the fixed points $(0, S^0)$ and (\tilde{U}, λ) . The Jacobian matrix of G at $(0, S^0)$ is given by

$$J_0 = \begin{bmatrix} (1 - E)(1 + f(S^0)) & 0 \\ -(1 - E)f(S^0) & 1 - E \end{bmatrix}$$

Its eigenvalues are $1 - E \in (0, 1)$ and $(1 - E)(1 + f(S^0)) > 1$, so $(0, S^0)$ is a saddle fixed point for G . The stable manifold of $(0, S^0)$ is the S axis and the unstable manifold is the portion of the segment $U + S = S^0$ joining $(0, S^0)$ to (\tilde{U}, λ) .

The Jacobian matrix of G at (\tilde{U}, λ) is

$$J_1 = \begin{bmatrix} 1 & (1 - E)\tilde{U}f'(S^0 - \tilde{U}) \\ -E & (1 - E)(1 - f'(S^0 - \tilde{U})\tilde{U}) \end{bmatrix}$$

Its eigenvalues are $1 - E$ and $1 - (1 - E)\tilde{U}f'(S^0 - \tilde{U})$. Since $\tilde{U}f'(S^0 - \tilde{U}) \leq Wm/a < \eta < 1$ by (5), both eigenvalues belong to $(0, 1)$ and (\tilde{U}, λ) is locally asymptotically stable.

It is enough then to show that (\tilde{U}, λ) belongs to the omega limit set of every orbit $\{(U_t, S_t)\}_{t \geq 0}$ with $U_0 > 0$. Let Λ be such an omega limit set. Obviously, Λ is a nonempty, compact, invariant subset ($G\Lambda = \Lambda$) belonging to the line segment $U + S = S^0$ in Ω . From

the dynamics of G restricted to this line segment, if Λ contains a point different from $(0, S^0)$ then necessarily Λ also contains (\tilde{U}, λ) since it is closed and invariant. In this case we are done. But Λ must contain such a point since $U_0 > 0$ and therefore (U_0, S_0) does not belong to the stable manifold of $(0, S^0)$. This completes the proof. \square

Now that we know the global behavior of the 2 dimensional system, we can use the following result of Golubitsky *et al.* to derive the equilibrium distribution of the x_i .

Theorem 2.2 ([8]) *Suppose that T_k is a sequence of nonnegative primitive matrices, and that $T_k \rightarrow T$ as $k \rightarrow \infty$, where T is also nonnegative and primitive. If e is the Perron-Frobenius eigenvector of T satisfying $\mathbb{1}^T e = 1$ and $x_{k+1} = T_k x_k$ is a sequence starting with $x_0 \geq 0$ and $x_0 \neq 0$, then*

$$\frac{x_k}{\mathbb{1}^T x_k} \rightarrow e, \quad k \rightarrow \infty$$

Therefore, we have the following result.

Theorem 2.3 *Let e be the Perron-Frobenius eigenvector of $(1-E)A(\lambda)$ satisfying $\mathbb{1}^T e = 1$. If $(1-E)(1+f(S^0)) > 1$, $x_0 \neq 0$ and $\lambda \in S_{int}$, then the system (2a) (2b) admits one globally asymptotically stable non trivial equilibrium (\tilde{x}, \tilde{S}) , where*

$$\frac{\tilde{x}}{\tilde{U}} = e \tag{14}$$

Proof The matrices $A(S_t)$ and $A(\lambda)$ are obviously nonnegative, for all $t > 0$. Under the hypotheses of the theorem, $\lambda \in S_{int}$, so there exists a neighborhood \mathcal{N}_1 of λ such that for all $S \in \mathcal{N}_1(\lambda)$ and all $i < r_d$, $0 < D_i(S) < 1$. Since $(U_t, S_t) \rightarrow (S^0 - \lambda, \lambda)$ as $t \rightarrow \infty$ by Theorem 2.1, it is possible to find $\tau > 0$ such that, for all $t \geq \tau$, $S_t \in \mathcal{N}_2(\lambda)$, where $\mathcal{N}_2(\lambda)$ is a neighborhood of λ that we can make as small as we want. In particular, if τ is such that $\mathcal{N}_2 \subset \mathcal{N}_1$, then $0 < D_i(S_t) < 1$ for all $t \geq \tau$.

Therefore, for all $t \geq \tau$, $A(S_t)$ and $A(\lambda)$ are irreducible, as it is easy to verify that any class can be reached from any other class, in a finite number of steps. Since the trace of $A(S_t)$ and the trace of $A(\lambda)$ are positive, they are also primitive [2, p. 34]. Hence the conditions of Theorem 2.2 are fulfilled. \square

The Perron-Frobenius eigenvalue is $(1 - E)(1 - P + MP)$ ($= (1 - E)(1 + f(\tilde{S}))$), its associated eigenvector has the following form.

$$e = \frac{1}{\tilde{U}} \begin{pmatrix} D_1(\tilde{S}) \\ \vdots \\ 1 - \prod_{k=1}^{r_b-1} (1 - D_k(\tilde{S})) \\ \hline 1 \\ \vdots \\ 1 \\ \hline 1 \\ 1 - D_1(\tilde{S}) \\ \vdots \\ \prod_{k=1}^{r_d-1} (1 - D_k(\tilde{S})) \end{pmatrix} \quad (15)$$

where the blocks correspond to r_b , r_g and r_d rows. It is easily verified that $(1 - E)A(\tilde{S})e = (1 - E)(1 - P + MP)e$.

Let us denote by x_i^b the i^{th} birth class ($i = 1, \dots, r_b$), by x_i^g the i^{th} growth class ($i = 1, \dots, r_g$), and by x_i^d the i^{th} division class ($i = 1, \dots, r_d$). Using (14) and (15), the stable distribution can be computed. This is expressed in the following result.

Proposition 2.3 *The equilibrium biomass distribution is given by the following formulas:*

- for $i = 1, \dots, r_b - 1$ (birth classes except the last one):

$$x_i^b = 1 - \prod_{k=1}^i (1 - D_k(\tilde{S})) \quad (16)$$

- for $i = 1, \dots, r_g$ (growth classes), for the last birth class and the first division class:

$$x_{r_b}^b = x_i^g = x_1^d = 1 \quad (17)$$

- for $i = 2, \dots, r_d$ (division classes except the first):

$$x_i^g = \prod_{k=1}^{i-1} (1 - D_k(\tilde{S})) \quad (18)$$

Now let us consider the case where $\lambda \notin S_{int}$. More precisely, what happens when $D_i(\lambda) = 1$ for some $i < r_d$? This is the object of the following proposition.

Proposition 2.4 *Suppose that $(1 - E)(1 + f(S^0)) > 1$ and $x_0 > 0$. Furthermore, suppose that there exists $k < r_d$ such that $D_k(\lambda) = 1$. Then the system (2a) (2b) admits one globally asymptotically stable non trivial equilibrium (\tilde{x}, \tilde{S}) , with*

$$\frac{\tilde{x}}{\tilde{U}} = e_k$$

where e_k is the Perron-Frobenius eigenvector of the sub-matrix of $A(\lambda)$ comprising its first $r_b + r_g + k$ rows and columns, and such that $\mathbf{1}^T e_k = 1$.

It is easy to see that the equilibrium biomass distribution is the same as the one given by Proposition 2.3, but with a “left shift” of the classes:

- Birth classes: from 1 to k (instead of r_b , where $k < r_b$).
- Growth classes: from $k + 1$ (instead of $r_b + 1$) to $r_b + r_g$.
- Division classes: from $r_b + r_g + 1$ to $r_b + r_g + k$ (instead of $r = r_b + r_g + r_d$).

and where the classes from $r_b + r_g + k + 1$ to r are empty.

The converse proposition also holds: if for some $k < r_d$ we have $D_k(\lambda) = 0$, then there is a “right shift” in the classes, the k first birth classes being empty.

3 Numerical results

Figure 2 shows a comparison of the stable distributions of numbers (computed using the approximation $n(i) = x_i/M^{i-1}b_{min}$) as given by the homogeneous model of Gage *et al.* and by our model. In this example, we have assumed that the mean division (and consequently birth) biomass is located in the middle of the division classes.

Figure 3 shows the transient behavior of the model. We can see that the biomass in each size class has damped oscillations prior to the equilibrium. The biological parameters used in this computation are the ones corresponding to the data set (and to the algal species) that will be shown in the next section.

Figure 4 shows the substrate concentration and total biomass, as well as the normalized number of cells, corresponding to the same simulation. The number of cells oscillates, and this behavior lasts until the biomass distribution has reached its equilibrium.

The transient behavior of the system can be studied as in [7], by defining the wavelength of complex eigenvalue λ_i as

$$\omega_i = 2\pi T / \tan^{-1}\left(\frac{\Im(\lambda_i)}{\Re(\lambda_i)}\right)$$

In Figure 5, two numerical experiments are shown. On the left hand side, the wavelength of the complex eigenvalue of greatest magnitude is plotted, for an increasing number of division classes, while the total number of classes is constant. We can see that as the proportion of classes that are division classes rises, the wavelength of the first complex

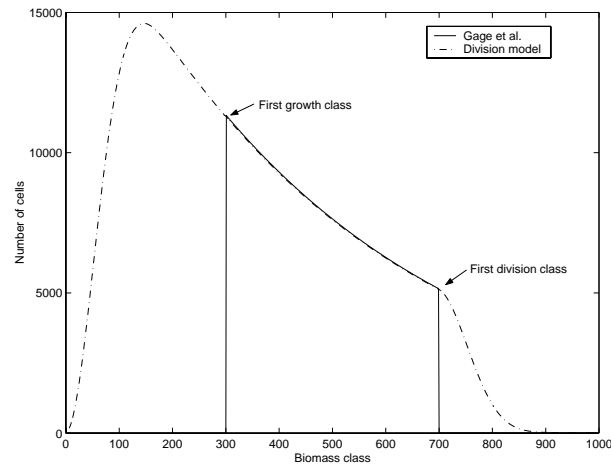


Figure 2: Comparison of the equilibrium distributions (converted to numbers), as given by the two models. This example uses $r_b = r_d = 300$ and $r_g = 400$. The vertical lines show the passage from one type of class to another.

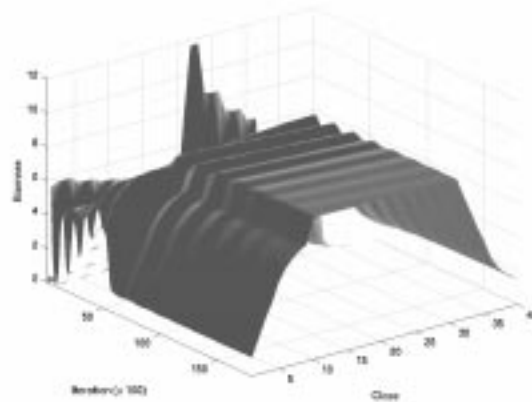


Figure 3: Transient behavior of the model following an initial condition of Dirac type in class 20 (of a total of 40). The parameters are $T = 0.001$, $r_b = r_d = 10$. The figure shows a sampling (every 100 iterations) out of a total 18000 iterations (corresponding to 18 days).

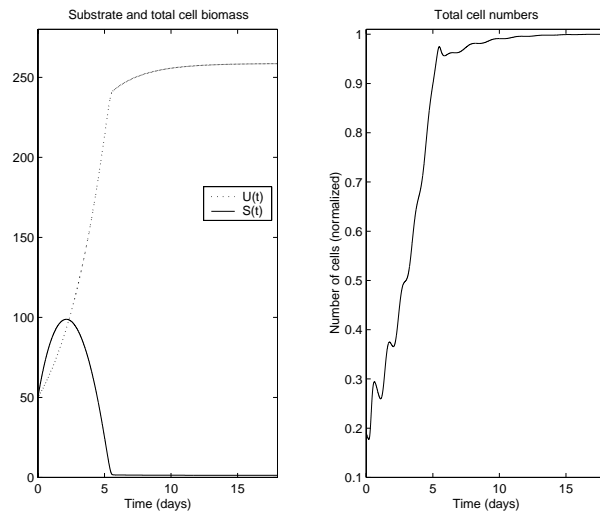


Figure 4: Behavior of the total biomass and of the substrate (left), and of the total number of cells (right), under the same conditions as in Figure 3.

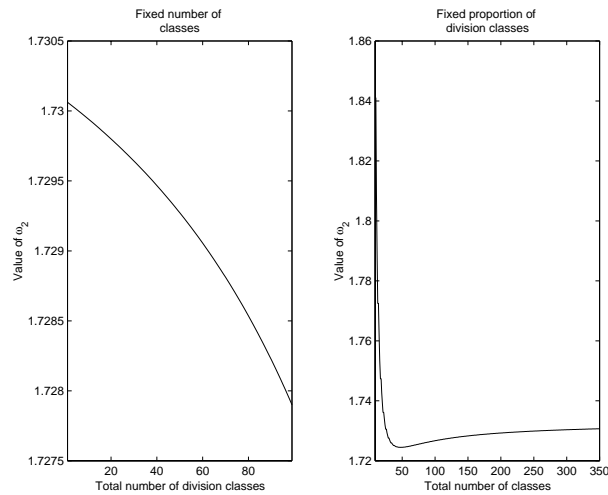


Figure 5: Wavelength of the first complex eigenvalue: (left) for a fixed total number of classes (200), as the number of division classes progresses from 1 to 100; (right) for a fixed proportion (1/4) of division classes, as the total number of classes progresses from 10 to 300.

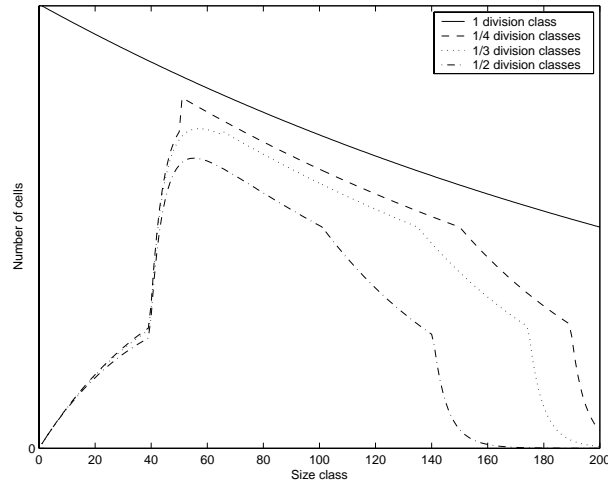


Figure 6: Comparison of distributions obtained for 200 size classes, when the proportion of these classes that are division classes changes: one division class, $1/4 = 50$ division classes, $1/3 = 66$ division classes and $1/2 = 100$ division classes.

eigenvalue decreases. The right hand side of the figure shows that with a given proportion of the classes being division classes, the wavelength of the complex eigenvalue first decreases, then stabilizes.

Hence it seems that the proportion of classes in which cells divide (and accordingly are born) is more determinant for the transient behavior of the system than the total number of classes. Interpretation of this fact is rather straightforward: the more division classes (and accordingly birth classes), the faster the “dispersion” of cells.

The proportion of division classes also influences the equilibrium distribution, as can be seen in Figure 6. In this figure, we show, for a fixed total number of classes, the shape of the distribution (converted to numbers) as the proportion of division classes among the total number of classes rises. The plain curve corresponds to the prediction of the model of Gage *et al.* Then, as the proportion of division classes grows, the distribution becomes more and more narrow.

Figure 7 illustrates the behavior pointed out in Proposition 2.4. The plain curve shows the standard equilibrium distribution. The dashed-dot curve shows what happens when the proportion of cells dividing is one, in a division class that is not the last: there are no more cells in the classes following this class (class 162 on the figure). On the other hand, if for a certain number of division classes, there is no division, we obtain an equilibrium distribution like the one corresponding to the dashed curve (named “late division” distribution in the figure): classes are “right shifted”, and there are no cells prior to the class corresponding to the first non zero division class.

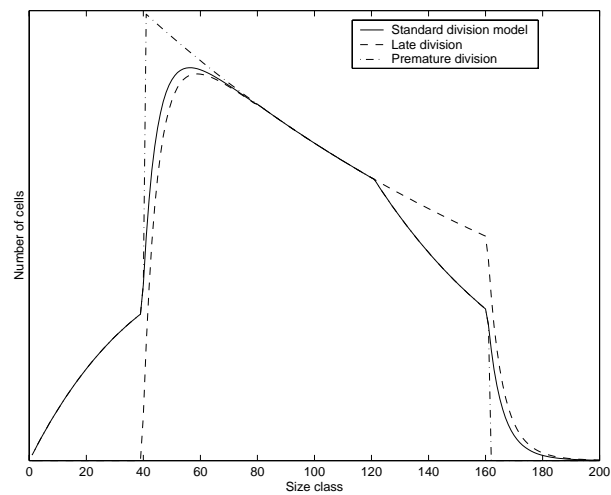


Figure 7: Comparison of a sample distribution with two extreme cases (all converted to cell numbers). The distribution was computed using 200 size classes (80 birth classes, 40 growth classes and 80 division classes). Premature division case: $D_{40}(\lambda) = 1$. Late division case: $D_i(\lambda) = 0$ for $i = 1, \dots, 39$.

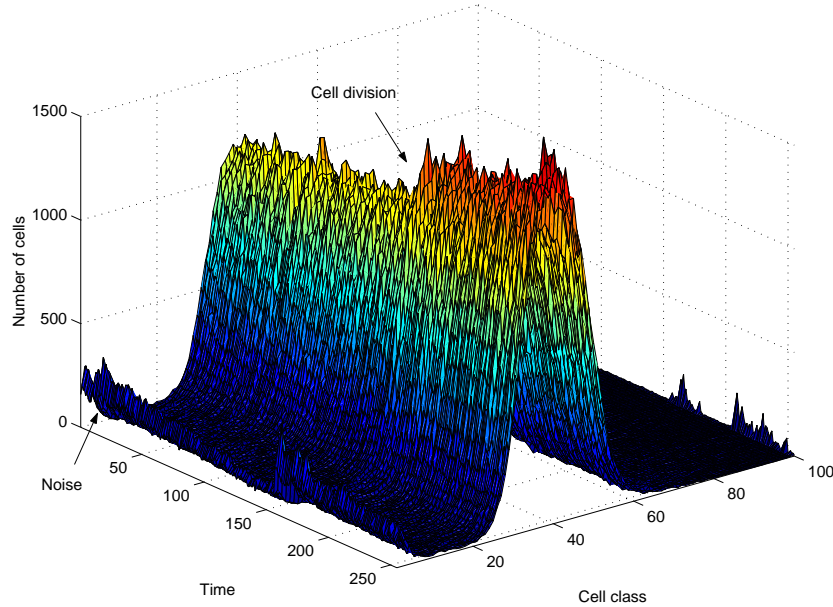


Figure 8: Sample data set of *Cryptomonas sp.*. This series represents three weeks of chemostat culture. There are 100 size classes, from $1.5\mu\text{m}$ to $15\mu\text{m}$.

4 Experimental data

As was mentioned in the introduction, we have access to very long time series of structured data from a chemostat, such as the dataset shown in Figure 8. This dataset corresponds to the algal specie *Cryptomonas*. The maximal growth rate of this specie is $\mu_{max} = 0.7 \text{ day}^{-1}$, corresponding to a minimal doubling time of approximately one day, and the half saturation constant is $1 \mu\text{Mol}$. The dilution rate in this experiment was 0.4 day^{-1} , and the input concentration was $260 \mu\text{Mol}$.

Let us now shortly mention how this data is obtained. The chemostat is fully automated, both on the operating side and on the measurement side [4]. The dilution rate and the input nutrient concentration are computer controlled; this allows fluctuating inputs, as well as long time, nearly unattended functioning.

But the main feature of this device is that it allows nearly continuous monitoring of several variables: substrate concentration (by colorimetric methods, using a Technicon Auto-analyzer), cell size and number distributions (by means of a particle counter (HIAC/ROYCO with laser sensor HRLD 400)), and even chlorophyllia concentration (using a spectrophotometer). All of these measures are computer controlled. Typical experiments last between one and four months, resulting in sets of one thousand to four thousand measures.

We will shortly discuss the comparison of the predictions of the model to this data in the next section.

5 Discussion

We have shown that the introduction of size and substrate dependent cell division in the model of Gage *et al.* does not fundamentally modify its dynamical behavior. Since our model sums to a Monod type model, we are able following the analysis of [11], to prove the existence of a globally stable non trivial equilibrium. This is done by first studying the reduced model of substrate and total biomass, which is shown to have a globally stable equilibrium, and then using a weak ergodic theorem of Golubitsky *et al.* to derive the distribution of biomass in the size classes.

Then, we have presented several numerical simulations. These simulations show that the ratio of division classes to the total number of classes is determinant of the shape of the equilibrium distribution. The proportion of division classes also has an impact on the transient behavior of the model: as it increases, the frequency of the oscillations in individual size classes reduces, as well as their duration.

Regarding the equilibrium distribution, as can be seen in Figure 9, there exists quite a difference between the data and the model prediction. But, by tuning of the division parameters, one can obtain many different shapes for the equilibrium distribution, the main characteristic being that these distributions are leptokurtic. The precise identification of the division functions remains to be treated. However, we do not expect to be able to fit such data very precisely using the predictions of our model. Indeed, it can be inferred from Figure 9 that the doubling of the biomass that is a major hypothesis of this model, should be relaxed. We can see that the diameter of the cells in the sample ranges from $3 \mu\text{m}$ to about $8 \mu\text{m}$. Converting this to biovolume gives a range of 14 to more than $200 \mu\text{m}^3$. To relax the biomass doubling hypothesis, one could consider introducing division into cells of different sizes, as is done in [1, 9], or division into more than two cells, as Williams suggests in [13].

From a modelling point of view, the comparison of Figures 3 and 4 pleads for the use of structured models of the chemostat. Indeed, consider the total biomass curve in Figure 4. It is smooth, while in Figure 3 we observe damped oscillations: the summation hides the more complicated behavior of the individual biomass classes.

To conclude this discussion, let us note that we have not treated here the competition case. But the analysis of [11] should still hold, were we to consider competition.

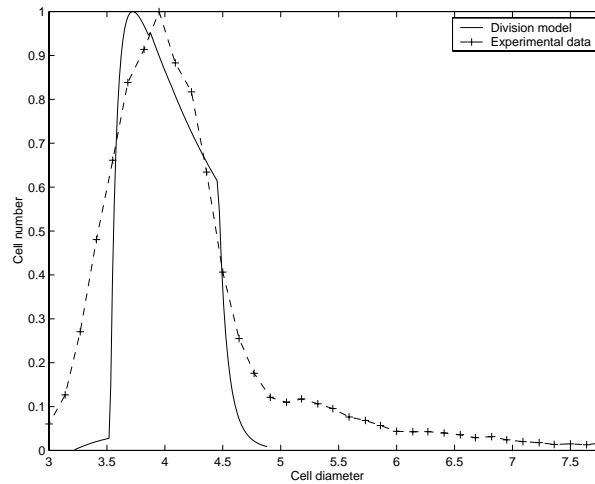


Figure 9: Comparison of a sample spectrum with an equilibrium distribution (converted to numbers). The distribution was computed using 200 size classes (80 birth classes, 40 growth classes and 80 division classes).

References

- [1] O. Arino and M. Kimmel. Asymptotic analysis of a cell cycle model based on unequal division. *SIAM J. Appl. Math.*, 47:128–145, 1987.
- [2] A. Berman and R. Plemmons. *Nonnegative matrices in the mathematical sciences*, volume 9 of *Classics in Applied Mathematics*. SIAM, 1994.
- [3] O. Bernard and J.-L. Gouzé. Non-linear qualitative signal processing for biological systems: application to the algal growth in bioreactors. *Mathematical Biosciences*, 157:357–372, 1999.
- [4] O. Bernard, G. Malara, and A. Sciandra. The effects of a controlled fluctuating nutrient environment on continuous cultures of phytoplankton monitored by computers. *Journal of Experimental Marine Biology and Ecology*, 197:263–278, 1996.
- [5] J. M. Cushing. A competition model for size-structured species. *SIAM J. Appl. Math.*, 49(3):838–858, 1989.
- [6] J. M. Cushing. *An introduction to structured population dynamics*, volume 71 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, 1998.

-
- [7] T. B. Gage, F. Williams, and J. Horton. Division synchrony and the dynamics of microbial populations: A size-specific model. *Theoretical Population Biology*, 26:296–314, 1984.
- [8] M. Golubitsky, E. B. Keeler, and M. Rothschild. Convergence of the age structure: applications of the projective metric. *Theoretical Population Biology*, 7:84–93, 1975.
- [9] J. Metz and O. Diekmann, editors. *The dynamics of physiologically structured populations*, volume 68 of *Lecture Notes in Biomathematics*. Springer-Verlag, 1986.
- [10] M. Pascual and H. Caswell. From the cell cycle to population cycles in phytoplankton-nutrient interactions. *Ecology*, 78(3):897–912, 1997.
- [11] H. L. Smith. A discrete, size-structured model of microbial growth and competition in the chemostat. *J. Math. Biol.*, 34:734–754, 1996.
- [12] H. L. Smith and P. Waltman. *The theory of the chemostat. Dynamics of microbial competition.*, volume 13 of *Cambridge Studies in Mathematical Biology*. Cambridge University Press, 1995.
- [13] F. M. Williams. Dynamics of microbial populations. In B. Patten, editor, *Systems Analysis and Simulation in Ecology*, pages 198–267. Academic Press, 1971.



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
<http://www.inria.fr>
ISSN 0249-6399