The differential equation counterpart of an individual-based model for yeast population growth

Pablo Gómez-Mourelo\textsuperscript{a,∗}, Marta Ginovart\textsuperscript{b}

\textsuperscript{a} Dpto. Matemática Aplicada, E.T.S. Ingenieros Industriales, Universidad Politécnica de Madrid, c/ José Gutiérrez Abascal, 2, 28006 Madrid, Spain
\textsuperscript{b} Departament de Matemática Aplicada III, Universitat Politècnica de Catalunya, Edifici ESAB, Campus del Baix Llobregat, Avinguda del Canal Olímpic s/n, 08860 Castelldefels, Barcelona, Spain

ARTICLE INFO

Article history:
Received 14 August 2008
Received in revised form 6 January 2009
Accepted 8 May 2009

Keywords:
Individual-based model
PDE Scaling
Yeast population

ABSTRACT

Computer simulations are increasingly used in biological fields. The amazing power, storage ability, and processing speeds available nowadays have enabled the implementation of computer individual-based models (IbMs) to simulate really complex biological populations. Computers can easily keep track of thousands of individuals (often called ‘agents’), whose complex behaviours and large amounts of associated data were daunting only 20 years ago. As such, computer modelling has just entered a field where traditional PDE models used to reign alone. A study of the exchange and non-trivial relationship between these two fields, computer IbMs versus classic partial differential equations (PDEs), is appropriate. The aim of this paper is to compare both approaches through a relevant example, namely the evolution of a yeast population in a batch culture. Thus, this paper deals with the utilization of both classical mathematics and computer science in the solution of problems arising in microbiology. First, an IbM approach to study the evolution of a yeast batch culture is presented. Second, an equivalent PDE model is derived by using some techniques from the interacting particle systems field. Third, a comparison and discussion on the advantages and drawbacks of both modelling tools is given.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Comparison of different scientific models leads to further insights into the behaviour of nature and modelling techniques. The process of modelling biological systems is not always straightforward. In general, the modelling process is a problem in which different routes may be taken, and the modeller must choose which road to take [1–4].

Mathematical and computational approaches provide powerful tools in the study of problems in biological populations and ecosystems [3]. The subject has a rich history intertwined with the development of statistics and dynamical systems theory, but recent analytical advances, coupled with the enhanced potential of high-speed computation, have opened up new vistas and presented new challenges [5].

Individual-based models (IbMs) or “agent-based” models are a bottom-up approach which starts with the ‘parts’ (individuals) of a system and then tries to understand how the system’s properties emerge from the interaction among these ‘parts’ [6].

Four main criteria distinguish an IbM [5,7]:

• The degree to which the complexity of the individual’s life cycle is reflected in the model.
• The extent to which variability among individuals of the same age, size or stage is considered.
• Whether or not the spatial and temporal dynamics of resources used by individuals are explicitly represented.

* Corresponding author.

E-mail addresses: pablo.gomez.mourelo@upm.es (P. Gómez-Mourelo), marta.ginovart@upc.edu (M. Ginovart).

0898-1221/$– see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2009.05.024
Whether real or integer number are used to represent the size of a population (IbMs are built using the mathematics of discrete events).

IbMs have some advantages and drawbacks. Some of the advantages are:

- They make more realistic assumptions than state variable models.
- Individuals are described by attributes and capabilities: they grow, age, develop, acquire resources, reproduce, and interact, changing in many ways over their life cycle and modifying their environment.
- The simulations provide information on the collective behaviour by looking at the behaviour of each element of which it is composed.
- They can address types of questions that are difficult to be addressed with classical models.
- They include highly detailed models with a wide variety of components and mechanisms.
- They are used in controlled simulation experiments to achieve a comprehensive understanding of the key structures and processes.

Some of the disadvantages are:

- The amount of detail is often too high to be supported in terms of what we can measure and parameterize.
- They are mostly applied to pragmatic motivations rather than paradigmatic ones.
- They are more complex than classical and analytically tractable models: many entities, spatial scales, heterogeneities and stochastic events.
- Extensive simulations constitute a brute-force method still lacking an analytical or theoretical framework.
- There is an absence of a common and concise language for communicating.

In this paper we are interested in comparing two different models of the observed macroscopic and microscopic behaviour of yeast populations. On one hand, an IbM called INDISIM-Yeast [8] is presented; on the other hand, a partial differential equation (PDE) for the same system is developed and solved. Lagrangian, individual-based descriptions make attractive graphics and can provide a basis for population features analysis. Eulerian, partial differential equation based, field descriptions capture the essence of population dynamics. We are interested in bridging the gaps between both approaches. Similar works [9] provide an insight into the possibilities of both modelling tools.

We specifically use INDISIM-Yeast to model the biological and metabolic activities of the yeast *Saccharomyces cerevisiae* under batch conditions. INDISIM has already been used to study bacterial growth in yoghurt and on agar plates. The computer code called INDISIM (INDividual DIScrete SIMulations) was developed by Ginovart et al. [10] specifically to study bacterial cultures and it was devised to deal with systems in which bacterial activity is one of the fundamental parts of the system [11,12]. INDISIM is a stochastic model, discrete in space and time, that simulates the behaviour of bacterial populations such that the global properties of the system emerge from the rule-following behaviour of individual microorganisms. The state of each organism is determined by a set of random, time-dependent variables related to spatial location, biomass, cellular cycle and other individual properties.

2. The simulator INDISIM-Yeast

We consider the evolution of a yeast culture over a period of time in a specific environment as occupying a three-dimensional spatial grid. This physical cubic domain is subdivided into cells: spatial cells that are small cubes. Each is identified by Cartesian coordinates. The physical domain is subject to close boundary conditions appropriate to the problem in hand, a batch culture. In batch conditions the medium is not altered by further nutrient addition or removal, so the grid is enclosed by ‘closed’ walls. The time evolution of the system is divided into equal intervals that we identify with program steps.

The individual rules for the yeast cells take into account their motion, uptake of the nutrient (glucose), metabolism of this nutrient to achieve cellular maintenance and new cellular biomass, budding reproduction and cellular viability. The environment where the set of yeast cells evolve is liquid. The composition of the physical lattice affects the environmental conditions.

In the implementation of INDISIM-Yeast we have used random variables for the characterization of individual properties of a yeast cell and a spatial cell, as well as the random updating of individual rules. Hence the simulations are stochastic rather than deterministic. The stochasticity of the simulations implies variability of individual behaviour. As a result we are concerned with heterogeneous systems.

2.1. Modelling the spatial cells and the abiotic components of the medium

We assume the yeast population grows in the bulk of a liquid medium where we consider variables that are space and time dependent. These variables control the amount of abiotic components, and are identified as glucose (the nutrient particles) and ethanol (the metabolites or end-product particles) arising from the yeast cellular activity and excreted to the environment. The environment is continuously changing because glucose particles are consumed and ethanol is accumulating in the medium.

The grid is defined at each time step by

\[
G(t) = \left\{ s_{xyz} \mid s_1(t), s_2(t) \right\}_{x,y,z=1,...,Q}
\]
where $Q$ denotes the size of the spatial domain of the grid; $S_{xyz}$ each of the spatial cells defined within this domain; and $s_1(t), s_2(t)$ the number of glucose and ethanol particles respectively in each cell and at each time step.

At the end of each time step a redistribution of the abiotic particles takes place throughout the domain.

2.2. Modelling a single yeast cell

At each time step, a single organism, an individual yeast cell $i$, is defined by a set of time-dependent variables, which describes and controls its individual properties. For each micro-organism, INDISIM-Yeast implements a set of rules for the following actions: motion, uptake and metabolism of nutrient particles, reproduction and viability, described below. The set of $i$ defines the yeast population with $N = N(t)$ individuals defined at each time step by:

$$P(t) = \{ I_{i}\{ v_1(t), v_2(t), \ldots, v_{10}(t) \} \}_{i=1,...,N}$$

where, for each yeast cell $I_{i}$, $v_1(t), v_2(t), \ldots, v_{10}(t)$ identify its position in the spatial domain; $v_4(t)$, its mass, which, assuming spherical shape for the cell yeast and constant cellular density, enables us to obtain its cellular surface which is proportional to $v_4(t)^{3/2}$; $v_5(t)$, its genealogical age as a number of bud scars on the cellular membrane; $v_6(t)$, the reproduction phase in the cellular cycle where it is, namely the unbudded phase or budding phase; $v_7(t)$, its “start mass”, i.e. the mass to change from the unbudded to budding phase; $v_8(t)$, the minimum growth of its biomass for the budding phase; $v_9(t)$, the minimum time required to complete the budding phase; $v_{10}(t)$, its survival time without satisfying its metabolic requirements.

2.2.1. Motion

The physical position of each yeast cell changes randomly to another neighbouring cell.

2.2.2. Uptake of nutrient particles

At each time step, each yeast cell may take up nutrient particles (glucose) from the medium and be capable of metabolizing them.

The number of nutrient particles entering the yeast cell is assumed to be proportional to the cell’s surface of the yeast and the number of nutrient particles of the spatial cell that it occupies, $S_{xyz}(t_i(t),v_2(t),v_3(t))$. The uptake of nutrients is assumed to be limited by the genealogical age of the cell, say $v_5(t)$, and is defined by the number of bud scars on the cell’s surface [13]. The simulator takes into account the effects arising from bud scars, as these affect the cellular membrane. The maximum number of nutrient particles that one yeast cell may actually absorb is given by

$$U = z_1 c (v_4(t))^{3/2} [1 - K_1 v_5(t)] ,$$

where $z_1$ is a random variable with mean $U_{\text{max}}$ and standard deviation $\sigma = 0.2U_{\text{max}}$, $c$ and $K_1$ constants, and $U_{\text{max}}$ the maximum number of nutrient particles that may be consumed per unit time and unit of cellular surface.

In order to take into account the probability of nutrient particles translocating into the yeast cell through the cellular membrane we consider that $k$ is a given percentage of the amount of nutrient particles that the yeast cell will actually translocate. The final yeast uptake will be

$$\min \left\{ U, k \left( S_{xyz}(t_i(t),v_2(t),v_3(t)) \right) \right\} .$$

2.2.3. Metabolism of nutrient particles

In order to model the metabolism of translocated glucose in a yeast cell, we introduce the following parameters for each yeast and glucose particle:

1. $I$ is a prescribed amount of translocated glucose per unit of biomass that a yeast cell needs to remain viable.
2. $Y$ is a constant which models the metabolic efficiency that accounts for the synthesized biomass units per metabolized glucose particle.
3. $E$ is a residual constant that accounts for the amount of residual product (ethanol) per unit of metabolized glucose particle.

Using the above parameters, and recalling the meaning of $U$, we set the following control rules:

- A maintenance energy $ME$ for the viability of a yeast cell, that depends on its own biomass and the local ethanol concentration where the yeast cell is located (the growth arrest as a consequence of the metabolic final product):

$$ME = Iv_4(t) + K_2 (S_{xyz}(t_i(t),v_2(t),v_3(t))c v_4(t)^{3/2} .$$

- A control relation to check whether the glucose particles absorbed by a yeast cell are enough for its maintenance, $(U - ME) \geq 0$. If $(U - ME) < 0$, evaluate the possibility, at that time step, that the cell remains viable without requiring external energy supply. The longest time that the cell remains viable until the onset of its lysis is defined by the value $t_{ff}$.

- If the viability of the yeast cell is achieved, allow for the increase of its mass from $v_4(t)$ to $v_4(t) + \Delta m$ where $\Delta m = Y(U - ME)$.

- Allow for the excretion of $R$ particles of residual product, ethanol, into the spatial cell where the given yeast cell is located, such that $R = EU$. 


2.2.4. Reproduction

The simulator simplifies the yeast cell cycle by assuming that the model for the cellular cycle, for each yeast cell, involves two clearly differentiated phases or intervals [14], namely:

1. **Phase 1**, or un budded phase. This covers most of phase G1 and a very small fraction of phase S in the traditional division of the cell cycle, and
2. **Phase 2**, or budding phase. This covers a small fraction of G1, most of S and all of G2 and M.

In order to implement phase 1 and phase 2 in the simulator, we observe the following rules:

- **Phase 1**
  
  We assume that, in this phase, the yeast cell is getting ready for budding. The changes into the budding phase takes place only if, at the end of phase 1, the following conditions are satisfied:
  
  - The cell has attained a minimum stochastic cellular start mass \( m_s \), which will be stored in \( t_i \) as \( v_7(t) \).
  - The cell has achieved a minimum growth of its biomass, \( \Delta m_B \).

  When phase 1 begins, a value \( m_5 \) (start) is randomly chosen for each cell in the manner described below. \( m_s \) is the minimum mass the cell must attain during this first part of the process in order to change the cellular phase. Its value is also a function of the individual cell properties. We denote by \( m_{in} \) the value of the mass of an individual cell at the beginning of phase 1.

  The following requirements are checked each time step and for each yeast cell. If \( m_{in} < m_c - \Delta m_B \), then the ‘start mass’ assigned to the cell is \( v_7(t) = m_c + z_2 \), where \( Z_2 \rightarrow N(0, 0.2 m_c) \); otherwise, if \( m_{in} > m_c - \Delta m_B \), then the start mass assigned to the cell is \( v_7(t) = m_{in} + \Delta m_B + z_3 \), where \( Z_3 \rightarrow N(0, 0.2 \Delta m_B) \). The value \( v_7(t) \) remains constant until cellular change takes place.

  Hence, whenever the mass \( v_4(t) \) of a yeast cell satisfies the inequality \( v_4(t) > m_5 \), that cell enters phase 2. Note that, within our model, phase 1 does not need to be completed in a given time interval. The check in our model is whether an individual cell has reached a start mass, irrespective of the original mass value and a specific growth rate. The rate of biomass increments and appearances of the new daughter cells is affected by environmental conditions.

- **Phase 2**

  The budding phase is the least flexible in the cellular cycle as it requires both temporal and growth checks. Within our model, two conditions must be satisfied for cell division to start, a minimum growth of biomass \( \Delta m_B \) and a certain minimum time interval.

  A yeast cell will complete its cellular cycle when (a) it reaches a minimum biomass increase \( \Delta m_B \) (stored in \( v_8(t) \)), given by \( \Delta m_B = \Delta m_{B2} + z_4 \), where \( Z_4 \rightarrow N(0, 0.2 \Delta m_B) \); and (b) it has remained in phase 2 for a minimum time interval given by \( \Delta t = \Delta t_2 + z_5 \) (stored as \( v_9(t) \)), where \( Z_5 \rightarrow N(0, 0.2 \Delta t_2) \).

  The first condition is necessary because a yeast cell must have a minimum number of molecules and satisfy minimum structural requirements in order to function as an independent entity. No time limit is imposed for the cell to grow into this minimum mass. In a culture starved of glucose, or subjected to other inhibitory effects, the growth rate will be slower. On the other hand, the growth rate, even under optimal growth conditions, has to be completed within a minimum time interval; this is the second condition requires.

  The budding phase is completed with cell division into a daughter cell and a parent cell, the daughter carrying all the mass increase in this phase. This physical separation leaves one new scar in the parent cell membrane, increasing its genealogical age. Both parent and daughter yeast cells may remain in the same spatial cell.

We note that the reproduction rules in INDISIM-Yeast are implemented every time a new yeast cell appears. Hence different yeast cells in the culture need not have the same mass when the reproduction process begins. Moreover, the yeast cells involved in the reproduction process remain active; namely, they can change their position, continue to consume nutrient particles and metabolize them, and also dissipate heat and excrete ethanol to the medium. The local environmental conditions, in turn, indirectly affect the overall yeast growth rate.

2.2.5. Viability

The preceding rules are intended for viable yeast cells. We introduce now the rules for when the cells are no longer viable. Whenever a cell, at a given time, does not find enough nutrient particles to satisfy its metabolic requirements, the simulator assigns to this cell a mortality index \( v_{10}(t) \) equal to 1, satisfying the following rules:

- If, after one time step, the cell is still able to satisfy its metabolic requirements, the index increases by one: \( v_{10}(t) = v_{10}(t) + 1 \). Otherwise the index is reduced by one: \( v_{10}(t) = v_{10}(t) - 1 \).

- At each time step, the simulator checks whether \( v_{10}(t) \) exceeds the time \( t_i = t_i + z_7 \), where \( Z_7 \rightarrow N(0, 0.2 t_l) \), and where \( t_l \) denotes an average time interval beyond which the cell cannot survive. Thus, if \( v_{10}(t) > t_i \), the cell is no longer viable.

The following comments are in order. As it stands, the model ignores lysis due to causes external to the yeast culture. In our model the individual cells are no longer viable, either directly or indirectly, for the following reasons: (i) ethanol excess; (ii) low glucose concentration; (iii) diminishing surface to volume ratio; and (iv) the number of scars due to buds (genealogical age). All of these reduce the ability of the individual cell to translocate nutrients. The magnitude of these unfavourable conditions will determine the viability of the individual yeast cell.
3. The system of differential equations counterpart

Normally, a mathematical biology lbM includes two kinds of processes. First, those involving gradual changes (such as motion, biomass growth), and second, those involving abrupt changes (such as births and deaths). The former are usually associated with continuous processes, so we will call them type-C processes in what follows, whilst the latter are normally discontinuous, so they will be identified as type-D processes.

Type-C processes usually include not only deterministic phenomena but stochastic factors as well. Type-C phenomena will be described in what follows by stochastic differential equations (SDEs) and later on by advection equations. Type-D phenomena will be described by Poisson processes with certain intensity rates which will lead to reaction terms. These models were started by Okubo [4]. Taxis-induced movements are an especially relevant case [15]. For a detailed explanation on SDEs and Poisson processes we refer the reader to the works by Oksendal [16] and Arnold [17].

In our case, we are assuming a population made up of $R$ subpopulations of different classes or species of individuals. $N_j, \ j = 1, \ldots, R$ stands for the number of individuals in class $j$. Hence, the total population is $N = \sum_{j=1}^{R} N_j$. We tag each individual with an index $k \in \{1, \ldots, N\}$. Each individual is characterized by:

- the class $c_k$ it belongs to. Hence, if individual $k$ belongs to class $j$, we have $c_k = j$;
- $d$ attributes, which include its position, size, age and whatever relevant features are required for the model. As such, we will describe the state of the $k$-th individual at time $t$ as the $d$-dimensional vector

$$p^k(t) = (p^k_1(t), \ldots, p^k_d(t)).$$

From now on we will call $p^k(t)$ the generalized position of the $k$-th individual at time $t$.

3.1. Mathematical description of type-C processes

The continuous evolution of the $k$-th individual’s generalized position is described by the following SDE:

$$dp^k = a(x_1, \ldots, x_d, c_k, t)dt + b(x_1, \ldots, x_d, c_k, t)dW^k(t)$$

where $a$ represents the speed of the deterministic part of the type-C process, and $b$ is proportional to its uncertainty. $W^k, k = 1, \ldots, N$ are independent $d$-dimensional Brownian motions used for technical reasons [18].

At any time $t$, all the information needed about the system is mathematically comprised in the so-called empirical measure, defined as:

$$\Xi_N(t) = \frac{1}{N} \sum_{k=1}^{N} \delta_{p^k(t)},$$

where $\delta_c$ is the Dirac delta centred at point $c$. The empirical measure is similar to the density of individuals in a certain region. When deriving a PDE to describe the system evolution, we assume that, as $N \to \infty$, the empirical measure $P_N(t)$ tends to a deterministic process with density $\rho(x_1, \ldots, x_d, t)$, i.e.,

$$\lim_{N \to \infty} (P_N(t), f(\cdot, t)) = \int_{\mathbb{R}^d} f(x_1, \ldots, x_d, t) \rho(x_1, \ldots, x_d, t) \, dx_1 \cdots dx_d,$$

for any function $f$ smooth enough.

The density $\rho(x_1, \ldots, x_d, t)$ is a solution [19,20] of the PDE

$$\frac{\partial \rho}{\partial t} = \frac{1}{2} b^2 \Delta \rho - \nabla (a \rho), \quad (1)$$

where

$$\nabla \rho = \sum_{m=1}^{d} \frac{\partial \rho}{\partial x_m} \quad \text{and} \quad \Delta \rho = \sum_{m=1}^{d} \frac{\partial^2 \rho}{\partial x_m^2}.$$ 

The $\frac{1}{2} b^2 \Delta \rho$ term stands for the diffusion, which is consequence of the individuals’ Brownian motion, and the term $\nabla (a \rho)$ accounts for the deterministic part of the individuals’ transport. Gómez-Mourelo [21] successfully used this procedure in the study of fish populations.

3.2. Mathematical description of type-D processes

At any instant $t$, each individual $k$ is assumed to be able to:

1. **die**, at a rate given by $d_{c_k}(p^k(t), t)$;
2. **give birth** to a new individual of a (perhaps different) class $q$, at a rate given by $b_{c_k,q}(p^k(t), t)$;
3. turn into a new individual belonging to a different class \( q \), at a rate given by \( t_{q|q} (P^k(t), t) \). Transition rates from one class into the same one are null.

The rates \( d_{q|}, b_{q|q}, t_{q|q} \) represent the intensity of Poisson random processes, and can be understood as probabilities of death, birth and transition in small time steps.

In this case there are \( R \) empirical measures:

\[
\Xi_j(t) = \frac{1}{N_j} \sum_{q=1}^{R} \rho_j(t),
\]

one for each class of individuals. Again, the empirical measures \( \Xi_j \) are assumed to converge to deterministic processes with density \( \tilde{\rho}_j \) (see, e.g., [20]), provided the numbers of individuals \( N_j \) in each family tend to infinity. Then, see [20], the densities \( \rho_j \) satisfy the following system of PDEs:

\[
\frac{\partial \rho_j}{\partial t} = -d_j \rho_j + \sum_{l=1}^{R} b_{lj} \rho_l + \sum_{l \neq j} t_{lj} \rho_l - \sum_{l \neq j} t_{lj} \rho_j, \quad j = 1, \ldots, R.
\] (2)

3.3. Some remarks

- Every SDE for a type-C process involving an individual is only suitable during the lifetime of that individual, i.e., the time interval between its birth and death (transition into another type of individual would also make an individual disappear).
- When obtaining a counterpart system of differential equations, appropriate initial and boundary conditions must be added. We choose our initial and boundary conditions keeping coherence with the model.
- When the phenomenon under study includes both type-C and type-D phenomena (such as our case), the counterpart system of differential equation is obtained by summing up all type-C and type-D terms of the right-hand side of (1) and (2).
- Phenomena such as movement and mass increase must be treated as a type-C process.

3.4. The differential equations for INDISIM-Yeast

According to the biological description, we have two kinds of individuals: phase 1 yeast cells and phase 2 yeast cells; and two types of particles: glucose and ethanol. Every cell is always either in phase 1 or in phase 2. We describe below the type-C and type-D phenomena for all of them.

- **Phase 1 yeast cells**
  - A phase 1 yeast cell is described by a position \((x, y)\) in \(\mathbb{R}^2\), a biomass \(m\), the mass increased since last bud \(m_t\), and the genealogical age \(a\). The rules are:
    - **Type-C**
      * There is no deterministic spatial movement in \(\mathbb{R}^2\), but yeast cells experience a Brownian motion. Thus, the SDE for them is \(dP = dW\).
      * The biomass \(m\) increases at certain speed \(g\). The same speed applies to \(m_t\), so we have \(dm = dm_t = gm dt\).
      * The age \(a\) is not increased while in phase 1, so at this stage \(da = 0\).
    - **Type-D**
      * Death: we call \(Y\) the death rate.
      * Transitions: we call \(\tau_{12}\) the rate of transition from phase 1 to phase 2. For a phase 1 cell to turn into a phase 2 cell, two conditions must be observed: to have reached a critical biomass \(m_C\) and to have had a mass increase \(\Delta m_{B1}\) since the last scar.

- **Phase 2 yeast cells**
  - A phase 2 yeast cell is described by a position \((x, y)\) in \(\mathbb{R}^2\), a biomass \(m\), the bud mass \(m_B\), the genealogical age \(a\) and the time since the start of bud formation \(a_B\). The rules are:
    - **Type-C**
      * The spatial position behaves as in phase 1.
      * The biomass \(m\) remains constant, and \(m_B\) increases at speed \(g\), so we have \(dm = 0\) and \(dm_B = g dt\).
    - **Type-D**
      * Death follows the same notion and rules as in phase 1.
      * Transitions: we call \(\tau_{21}\) the rate of transition from phase 2 cells to phase 1 cells. Once a cell is in phase 2, for a new daughter to be born, two requirements must be satisfied: a minimum time \(\Delta T_2\) must have elapsed and the bud must have reached a minimum biomass \(\Delta m_{B2}\). Once these requirements are satisfied, the daughter is born, and the mother turns back to phase 1 with a new scar on its membrane (the genealogical age has increased by one). The newborn starts its cellular cycle in phase 1.

- **Glucose and ethanol particles**
  - Glucose particles are eaten from the medium by cells at a rate \(u\) (uptake) and ethanol is produced and released to the medium by cells at a rate \(p\) (production). Furthermore, we assume that glucose and ethanol have a very fast diffusion as compared to the time scale of the phenomenon, so that a homogeneous redistribution of both these particles is considered.
3.5. System of differential equations

- **Phase 1 yeast cells.**
  We note phase 1 yeast cells by the density \( \rho_1(x, y, t; m, m_1, a) \), which stands for the density of phase 1 cells at position \((x, y)\), time \(t\), with mass \(m\), age \(a\), which have increased their biomass by \(m_1\) since the last scar. According to previous sections, the evolution equation will be
  \[
  \frac{\partial \rho_1}{\partial t} = \frac{1}{2} D^2 \Delta \rho_1 - \frac{\partial}{\partial m} (\rho_1 \gamma) - \frac{\partial}{\partial m_1} (\rho_1 \gamma) - \gamma \rho_1 - \tau_{12} \rho_1.
  \]

- **Phase 2 yeast cells.**
  We model phase 2 yeast cells by the density \( \rho_2(x, y, t; m, m_2, a, a_2) \), which stands for the density of phase 2 cells at position \((x, y)\), time \(t\), with mass \(m\), age \(a\), with bud of mass \(m_2\) and age \(a_2\). The corresponding PDE is
  \[
  \frac{\partial \rho_2}{\partial t} = \frac{1}{2} D^2 \Delta \rho_2 - \frac{\partial}{\partial m_2} (\rho_2 \gamma) - \frac{\partial}{\partial a_2} (\rho_2) - \gamma \rho_2 - \tau_{21} \rho_2.
  \]

- **Glucose.** We denote the density of glucose by \( \rho_{Gl} = \rho_{Gl}(t) \), with the equation
  \[
  \frac{d\rho_{Gl}}{dt} = -\frac{1}{\Omega} \int_{\Omega} U(x, y)dx dy,
  \]
  where
  \[
  U(x, y) = \int u(\rho_1 + \rho_2)dmda.
  \]

- **Ethanol.**
  We denote the density of ethanol by \( \rho_{Et} = \rho_{Et}(t) \). The ethanol production is proportional to glucose uptake under the stoichiometric ratio 2:1. Hence,
  \[
  \frac{d\rho_{Et}}{dt} \propto \frac{d\rho_{Gl}}{dt}.
  \]

3.5.1. Some general remarks on domain and boundary conditions

A detailed explanation of the boundary conditions would be too long and beyond the scope of this work. We give only some general guidelines as to how we developed the simulation.

The spatial domain is a square \([0, L] \times [0, L]\), whilst the mass and age variables are restricted to
\[
m \in [0, L_1], m_1 \in [0, L_2], m_2 \in [0, L_3], a \in [0, L_4], a_2 \in [0, L_5].
\]
The values of \(L, L_1, \ldots, L_5\) are chosen according to those used in the IbM.

The boundary conditions are, roughly, as follows:

- **Spatial boundaries**
  The system is closed, so we introduce a null flux condition in the spatial boundaries for \(\rho_1, \rho_2, \rho_{Gl}, \rho_{Et}\). We have written them all for \(\rho_1\). The others can be written in an identical manner.
  \[
  \frac{\partial \rho_1}{\partial x}(x = 0, y, t; m, m_1, a) = 0
  \]
  \[
  \frac{\partial \rho_1}{\partial x}(x = L, y, t; m, m_1, a) = 0
  \]
  \[
  \frac{\partial \rho_1}{\partial y}(x, y = 0, t; m, m_1, a) = 0
  \]
  \[
  \frac{\partial \rho_1}{\partial y}(x, y = L, t; m, m_1, a) = 0.
  \]

- **Main mass and age boundaries**
  \[
  \frac{\partial \rho_1}{\partial m}(x, y, t; m = 0, m_1, a) = 0
  \]
  \[
  \rho_1(x, y, t; m = L_1, m_1, a) = 0
  \]
  \[
  \frac{\partial \rho_2}{\partial m}(x, y, t; m = 0, m_2, a, a_2) = 0
  \]
  \[
  \frac{\partial \rho_2}{\partial m}(x, y, t; m = L_1, m_2, a, a_2) = 0
  \]
\[
\rho_1(x, y, t; m = m_{B2}, m = 0, a = 0) = \int \tau_{21} \rho_2(\cdot, m_{B2}, \cdot)
\]
\[
\rho_1(x, y, t; m = m, m_l = 0, a) = \int \tau_{21} \rho_2(\cdot, m, \cdot)
\]
\[
\rho_1(x, y, t; m, m_l = L_2, a, \cdot) = 0
\]
\[
\rho_2(x, y, t; m = 0, a, a_B = 0) = \int \tau_{12} \rho_1
\]
\[
\rho_2(x, y, t; m = L_3, a, a_B = 0) = 0.
\]

All integrals above are computed over their corresponding domains.

3.5.2. Shape of some functions

We give below some inexhaustive information about the shape of the main functions involved in the PDE. These functions are in strict correspondence with those used in INDISIM-Yeast.

- Uptake rate:
  \[
  \min \left\{ \rho_{Gl}, \max \left\{ 0, \frac{cm^{2/3}}{3} \left( 1 - \frac{a}{2a_{\max}} \right) \right\} \right\}
  \]

- Maintenance energy (ME):
  \[
  Im + K_2 m^{2/3} \rho_{Gl}.
  \]

- Yeast growth rate:
  \[
  \max \{0, Y(\text{uptake} - \text{ME})\}.
  \]

- Mortality:
  \[
  \gamma = H(\rho_{Et} - \rho_{Et:max}).
  \]

  where \(H(t)\) stands for the standard Heaviside function, that is:

  \[
  H(t) = \begin{cases} 
  1 & \text{if } t > 0, \\
  0 & \text{otherwise.}
  \end{cases}
  \]

- Transition rate from phase 1 to phase 2:
  \[
  \tau_{12}(m, m_l) = H(m - M_{c12}) H(m_l - M_{12}).
  \]

- Transition rate from phase 2 to phase 1:
  \[
  \tau_{21}(m_g, a_g) = H(m_g - M_c) H(a_g - T).
  \]

All previous functions have been chosen to correspond to those used in the lbM, following the guidelines in Section 3. Smooth versions of the Heaviside function were used in the numerical resolution, in order to avoid numerical instabilities. An explicit Lax–Wendroff scheme was used in the numerical resolution.

4. Results and discussion

The values of the parameters in both types of simulation were chosen in order to generate comparable developments of the yeast system, rather than based on experimental measures. Also, the values chosen in the PDE and lbM do match, according to the relationship explained in Section 3.

In Fig. 1 we present the results of the simulations of both approaches. The main magnitudes of the yeast culture are represented in the graphs: nutrient (glucose), residual product (ethanol), total number of both cells in phase 1 and cells in phase 2, and genealogical age (number of scars) distribution at a time frame (chosen to be representative of the exponential growth of the culture).

All the simulations were carried out in both two- and three-dimensional grids. Both simulations assume that:

- Glucose fermentation takes place in a batch culture, i.e. neither cells nor nutrients nor end-particles can either enter or exit.
- The glucose is homogeneously spread all over the spatial grid at the beginning of the simulation. There is no ethanol at simulation start; it gradually appears over time.
- The simulation starts with one single daughter yeast cell.

A formal study of the equivalence of both models, such as a statistical or analytical study, was not the goal of this work. However, the visual overall aspect of representations in Fig. 1 suggests a fairly qualitative agreement; extensive numerical
Fig. 1. a1, a2: Glucose and ethanol evolution; b1, b2: Evolution of the number of yeast cells; c1, c2: A snapshot of genealogical age distribution at a representative time frame. The numbers 1 and 2 are associated to the IbM and PDE simulations, respectively, in all figures.

Simulations were carried out, strongly indicating that both approaches show very similar results under different initial conditions and parameter values.

Also, there is a reasonable explanation for those remaining slight differences between the outcomes of both approaches; in fact, the limiting process we used to derive a PDE approach assumes that the number of particles involved in the IbM is sufficiently large. This assumption is not sufficiently satisfied for the initial stages of the simulation, where few particles are handled. We did expect some discrepancies since the beginning of our first draft. The level to which these disparities would arrive was another motivation for this work.

Another source of dissimilarities is the stochasticity and discreteness associated with the IbM. Many of the features of the IbM are naturally represented by integer values, leading to discontinuous jumps along the simulation. Also, random variables are used in the IbM in order to introduce variability in the system’s evolution and composition. Meanwhile, the PDE approach cannot reproduce these behaviours, as it yields continuous, deterministic, smooth results and diagrams.

We were interested in comparing the numerical cost of both modelling approaches. The numerical resolution of a PDE proves daunting beyond 3 spatial dimensions (plus time), not to say 7, as in the present case. When studying biological systems, many individual features are usually considered. At least at the beginning, the scientist/modeller will keep all the individual characteristics in the model, just in case they are relevant for the overall population dynamics. As roughly each individual feature yields another spatial dimension in the PDE counterpart, the resulting PDE soon becomes intractable. While it is quite easy to add new attributes into the IbM computer code, the same strategy proves unfeasible in the PDE counterpart. Nonetheless, it is also true that IbMs can carry high computation costs when dealing with many particles.
which interact; in particular, the computation of interactions between \( N \) particles increases (roughly) with \( N^\alpha \), with \( \alpha \geq 2 \), making such simulations hardly feasible for very large values of \( N \). If this were the case and there were few dimensions, the PDE resolution speed could overcome that of the lbM.

The ease of implementation of some individual actions also differs from one approach to the other. While they may be easily included in the lbM, they can correspond to twisted, hard-to-analyze reaction–diffusion terms in the PDE. For instance, it would have been quite natural to consider in our yeast study an individual energy reservoir, whose magnitude would evolve depending upon the environmental conditions. While we thought of including a corresponding term in the PDE, it increased the complexity far too much. We finally decided to use a simplified version of INDISIM-Yeast, in order to keep the PDE counterpart reasonable and tractable.

lbM computational simulations are accessible and intuitive. They do not require deep mathematical knowledge and can be very helpful for non-experts. However, they are less standardized than PDEs as a scientific tool. The game-like appearance of lbM simulations is considered by some scientists a signal of the absence of scientific rigour, something never discussed on any PDE approach. We believe lbM modelling to be valuable, but its limitations must be known and appreciated.

As a final remark, after working with both two approaches, the lbM has seemed to us better in the case of yeast cultures handled in this work, for the following reasons.

- There is a complexity inherent in the yeast biology (two different phases for cellular reproduction, mass- and age-structured populations, etc.).
- Many of the yeast attributes have an integer nature, which is difficult to handle in PDEs (whose solutions are assumed to be smooth).
- The number \( N \) of biotic and abiotic elements is low and their interaction is moderate, yielding fast lbM simulations.
- If a modification/update of our study was considered, the lbM would require, in general, shorter modifications than the PDE resolution. The former would require only some lines of code to be changed, whilst a complete equation rearrangement would be required in the latter.

Acknowledgements

Gómez-Mourelo is supported by Ministerio de Educación y Ciencia (Spain), proyecto MTM2005-00423, and FEDER. Ginovart is supported by Ministerio de Educación y Ciencia, proyecto CGL2007-65142/BOS.

References