# Pantoea Model description

To describe the model, we use the ODD protocol (“Overview, Design concepts, and Details”) which helps to ensure that the model explanation is complete (Grimm et al., 2010). This document should be read in conjunction with the paper “Application of Machine Learning Techniques to an Agent-Based Model of Pantoea” and the code available at <https://github.com/miguel-fc/ABM-Pantoea>.

## Purpose

To develop a computational model to study the growth of *Pantoea* sp. YR343 due to several plant growth-promoting characteristics that it exhibits. *Pantoea* will grow in a batch culture under aerobic growing conditions, which can be used to reproduce the agar plate experimental protocol and explore the consequence of different priorities in the individual use of nutrients on the system dynamics.

## Model Overview

Interfaz de usuario gráfica, Diagrama

Descripción generada automáticamente

Fig 1. ABM Pantoea, Model Overview. Schematic shows entities and a brief description of model structure and processes.

Figure 1 provides an overview of the system we are modelling. The left side of the figure represents the real-world system whilst the right side represents the entities and rules that are applied to these entities.

## Entities, State Variables, and Scales

Our model has two types of entities: individuals (green ovals in Figure 1) and square grid cells (patches), black patches in Figure 1, of culture medium. Each individual represents a unique bacterium of *Pantoea* and has the variables: unique identification number, size, location (X,Y grid cell coordinates of where it is), mass, reproduction mass, counter for each reproduction cycle, individual variables for each nutrient uptake, density, diameter, viability and reproduction time. Table 1 provides more details on these variables along with their values (if constant across all bacteria) or range if variable for each bacterium.

Table 1. Model state variables by entity.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Entity | State Variable | Description | Value/Range | Units |
| Individuals  (constant) | density\_pa | Density | 1.1e-9 | g.cm-3 |
| width\_big\_pa | Maximum width | 0.5 | μm |
| width\_small\_pa | Minimum width | 0.25 | μm |
| length\_big\_pa | Maximum length | 1.1 | μm |
| length\_small\_pa | Minimum length | 0.55 | μm |
| mmol\_small\_pa | Minimum mass | 2.97E-13 | mmol |
| mmol\_big\_pa | Maximum mass | 2.37E-12 | mmol |
| ini\_biomass\_pa | Initial biomass | 1.78E-12 | mmol |
| Diameter | Diameter for each bacterium | 0.55 to 1.1 | μm |
| Individuals  (variable) | Unique id | Unique id of bacterium | 0 – inf | N/A |
| Location | Patch X, Y values | 0,0 to -125, -125 | N/A |
| Mass | Mass of the bacterium | Random distribution between minimum and maximum biomass | μg |
| Reproduction mass | Mass relative to reproduction | 75% of initial biomass | μg |
| Glucose-uptake | Rate of glucose uptake | According to μmax and stoichiometric coefficients of global reaction obtained using TEEM | molNutrient\*molcell-1\*h-1 |
| Oxygen-uptake | Rate of oxygen uptake |
| Ammonium-uptake | Rate of ammonium uptake |
| Bicarbonate-uptake | Rate of bicarbonate uptake |
| Patches  (constants) | glucose-medium | Glucose concentration | User defined | mM |
| oxygen-medium | Oxygen concentration | User defined | mM |
| ammonium-medium | Ammonium concentration | User defined | mM |
| bicarbonate-medium | Bicarbonate concentration | User defined | mM |
| dioxide | Carbon dioxide concentration | User defined | mM |
| water | Water concentration | User defined | mM |
| Patches (variable) | All nutrients concentration | Diffuse Coefficient | User defined | mM |

The model assumes that the smallest individual represents a bacterium with a diameter of ~ 0.55 μm and the largest a bacterium with a diameter of ~ 1.1 μm. To characterize the composition of the microbial cells, the model uses the empirical formula C4.17H8O1.75N so that each bacterium is assumed to have this elementary cell composition.

A two-dimensional lattice of 251 x 251 grid cells (patches) represents the agar plate (see agar plate and black squares (patches) in Figure 1) that contains the culture medium; each grid-cell represents a volume that can be tuned by changing the world dimensions: depth, and length. A volume was calculated using a portion of the wells in the experimental setup (see main paper) with dimensions of 605 μm × 605 μm × depth, where depth ranges from 10 μm to 50 μm. The grid cell variables are: unique position identifier in X,Y coordinates, total amount of each nutrient: glucose, oxygen, ammonium, bicarbonate, carbon dioxide, and water, along with counters to determine the number of bacteria within it. For details see Table 1.

All microbial and culture medium processes are discretized into time steps. One time step represents 30 seconds; for the current work the simulations were run for at least 4000-time steps (32 hours), but the user may choose the max time to run the model, for example, using High Performance Computing Facilities the model could run until all the bacteria die. With these units, graphical and numerical model outputs are the molar concentration of nutrients and metabolic products expressed in millimolar (mM), and the growth curve of *Pantoea*.

## Process Overview and Scheduling

The initial configuration of our model has two parts (See Figure 1): the first one for the system and the second one for the entities (culture medium (grid cells) and bacteria (individuals)). The initial system setup sets the world size and topology, and the time scale factor (time step). The topology of the world is programmed using a box mode (world with vertical and horizontal limits). The initial culture medium concentrations (within grid cells) and the initial bacterial population (via the number of individuals) are established using random variables, normal probability distributions with mean values which are provided by the user following experimental procedures.

At each time step a group of individuals are controlled using a set of time-dependent variables for each bacterium. Every individual performs the following processes: nutrient uptake, cellular maintenance, biomass synthesis, and bipartition. Culture medium processes are different depending on the management agar plate protocol and experimental conditions, but in any case, the culture medium is randomly diffused to simulate the mass transfer process in solid phase.

For each time step the time-dependent variables of microorganisms (individuals) and culture medium (grid cells) are calculated, updating the graphics and digital outputs according to the time scale proposed. The model also controls the whole carbon and nitrogen mass inside and outside of the system to ensure the carbon and nitrogen are balanced. During the simulation processes the bacteria (individuals) are called in a different random order in each time step and the state variables changes are immediately assigned generating an asynchronous update (Figure 1).

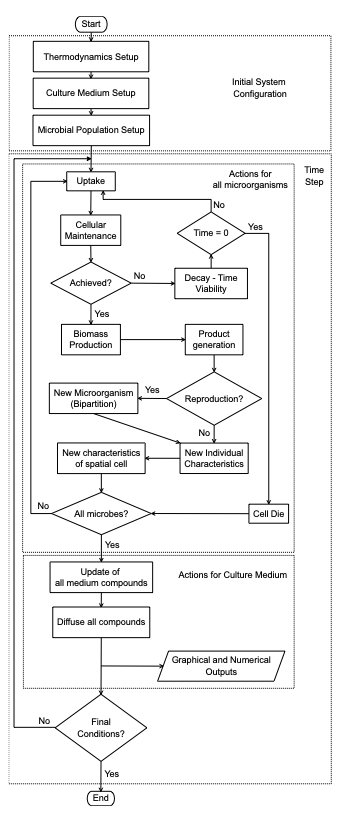


Fig 1. Flow Chart of the Pantoea Model

## Design Concepts

## Basic Principles

The model has two kinds of behavior-rules, rules for the individuals (bacteria) and different rules for the environment (culture medium – grid cells). The set of individuals and the environment are called the system (agar plate). All the rules are used at the level of the sub-models, and they are explained in the corresponding sections below. The individual rules are: nutrient uptake, cellular maintenance, biomass synthesis, and bipartition. The system rules are those that mimic the general agar plate procedures (nutrient diffusion).

## Emergence

The system dynamics emerge as the result of the interaction between bacteria (individuals) and the culture medium (grid cells) that they find inside of the agar plate (system). The model outputs are simulated volume (nl), simulated time (hours), the biomass over the time evolution, nutrient uptake rate consumption and all nutrients’ concentrations on the agar plate. All the model outputs appear at the system level and from the individual bacterial activity.

## Adaptation

All the individuals (bacteria) are programmed with the same rules; some of these rules will be executed and others not, depending on the internal states of the individuals and/or the characteristics of their local environments. Individuals act one after another, not in parallel. Hence, after one individual carries out all of its actions the composition of the spatial cell where it lives changes, and the next individual meets a different medium composition (assuming they are within the same grid cell) in relation to the previous acting or post-acting individuals.

To execute cellular maintenance each individual follows the energy reaction written using TEEM (McCarty, 2007; Rittmann and McCarty, 2001). Each bacterium, according to its biomass, calculates the nutrient quantity to accomplish its maintenance if the cellular maintenance is not fulfilled the bacteria cannot execute the subsequent individual rules. The third individual rule is performing biomass synthesis for growth. This rule is only executed when the amounts of nutrients, taken in during the uptake rule are enough to accomplish the maintenance requirement and after the updated nutrient amounts allow execution of the corresponding synthesis reaction in the aerobic phase. The next individual rule is whether to divide or not, depending on whether it has reached the minimum reproduction mass. The reproduction mass (mR) is the mass the bacterium must reach to start bipartition (mR is obtained from a normal random distribution with a mean value of 75% of the largest bacterium size). Additionally, the model includes an individual rule to control and avoid the overlay of bacterium on its surrounding culture medium (grid cell).

## Interaction

*Pantoea* is the only bacteria species in the virtual culture chamber. The microorganisms interact with the culture medium; therefore, there is an indirect interaction in which nutrient competition takes place among the bacteria that share the same spatial cell.

## Collective

The simulated bacteria (individual) do not form aggregates, each individual acts uniquely.

## Stochasticity

Several processes are modelled on criteria of randomness: i) the reproduction rule, ii) a portion of the uptake rule, iii) some parameters involved in the culture chamber management or operating protocol, and iv) a part of the initial system configuration.

Inside the reproduction rule, we consider that the reproduction threshold biomass for each bacterium is determined using a value from a normal distribution. For the physical separation of the two bacteria the original mass is separated into two new bacteria with masses according to a value from the normal distribution with mean value 0.5 and standard deviation 0.075. Thus, the mass of the original bacterium does not divide exactly in the proportion 50-50.

Inside the uptake rule, we consider that at each time step, each individual nutrient uptake-rate for each nutrient is established from a normal distribution with the mean value calculated from stoichiometric coefficients using TEEM and the individual mass of each bacterium with a standard deviation of 10% of this value.

Regarding the agar plate (gird cell) management: At the beginning of each simulation the model randomly distributes each bacterium (individual) into the system (setup procedure). For the initial system configuration, we consider that the initial culture medium composition, the initial population biomasses are established from normal distributions with mean values determined by the experimental procedure and standard deviations of 10% of these values. To represent the small reactor with constant agitation, we introduce a redistribution of nutrients and metabolic products in random time steps.

## Observation

The graphical and numerical outputs of the model are the concentration (mmol·l-1) of each culture medium component (glucose, oxygen, ammonium, bicarbonate, carbon dioxide and water) and the time evolution of the microbial biomass (mg·ml-1). The user can obtain all simulated data in the output file with the extension “.txt”).

## Initialization

The user can adjust: i) the culture medium composition (mmol·l-1) of glucose and ammonium, ii) initial amount of viable micro-organisms (bacteria), iii) time step (min), iv) the size of the culture chamber fixing the values of depth (um) and total world length (um), v) the individual max time availability (min), vi) the maintenance energy requirement for aerobic phase (gCdonor·gCmic-1·h-1), vii) thermodynamic efficiency to simulate microbial metabolic reactions, viii) the maximum growth rate for the bacteria population (h-1), ix) the diffusion coefficient for the nutrients, and x) the maximum number of bacterium on each culture (grid) cell.

## Input Data

There is no data input needed to start and run our model.

## Sub models

The uptake, cellular maintenance and biomass synthesis sub models are obtained from INDISIM-Paracoccus (Araujo Granda et al., 2016) and INDISIM-Denitrification (Araujo-Granda et al., 2020).

## Uptake

Each nutrient uptake depends on the individual’s capacity to capture nutrients through the cell membrane-associated proteins and on the nutrient availability in the medium. In our model, to determine the amount of each nutrient captured (absorbed) by each bacterium at each time step, two values are compared, the maximum uptake capacities (Ui, molnutrient·h-1) of the bacterium and the nutrient available in the culture medium (Ai, molnutrient·h-1), and the lowest value is chosen. Ui is assumed to be proportional to the individual mass and to the uptake-rate (ui) i being the nutrient, so:

Ui = ui \* individual-mass,

ui is a model parameter which represents the amount of nutrient that could be absorbed per unit of time and mass. Its units are molnutrient·molmass-1·h-1, where molmass denotes the moles of microbial mass (the microbial mass equals C4.17H8O1.75N). Ai is assumed to be proportional to the nutrient amount in each spatial (grid) cell and to the availability coefficient (ai), i being the nutrient, so:

Ai = ai \* nutrient-amount,

ai is a model parameter directly related to the nutrient characteristics and not to the types of micro-organisms involved, which represents the fraction of each nutrient in a spatial (grid) cell that is accessible per unit of time and for the individual, its units are h-1.

Following the INDISIM (Ginovart et al., 2002b)framework the maximum population growth rate (μmax) has been used to estimate the individual maximum uptake-rates (ui). Using this value and performing calculations with the stoichiometric coefficients of each metabolic reaction adjusted by TEEM we obtained the maximum uptake-rate for each nutrient. *Steptime* is the model variable used to give values to the availability coefficient (ai).

## Cellular Maintenance

Before biomass synthesis, it is necessary that each bacterium achieve some energetic requirements to ensure its viability. The cellular maintenance rule has two main components, the maintenance requirement, and the energy reaction (Re) (0.0417 C6H12O6 + 0.25 O2 --> 0.25 CO2 + 0.25 H2O) which is written using TEEM. The maintenance requirements are proportional to the individual’s mass. The coefficients determine an amount of nutrients per time step for cellular maintenance. We consider an appropriate maintenance requirement for soil heterotrophic microorganisms of 0.0015 gCdonor·gCmic-1·h-1, which was assumed in the model for the aerobic phase.

The energy reaction (Re) indicates the stoichiometry that the nutrients follow when each bacterium executes this action or rule. Each bacterium achieves its maintenance when the amount of nutrients taken in is enough to accomplish the maintenance requirement and these amounts also allow it to execute the corresponding energy reaction. Performing calculations with the energy reaction, we establish the maintenance requirements for the aerobic phase.

When the individual carries out its maintenance, the CO2 and the reduced electron acceptors are expelled to the culture. If the bacterium cannot reach its maintenance requirements, the availability time is penalized by one time step. After the maintenance, if the remaining glucose taken up and the quantity of electron acceptor are higher than zero, the individual can perform biomass synthesis.

## Individual biomass synthesis

With the nutrient intakes updated and using the stoichiometric coefficients for the metabolic reaction adjusted by TEEM (0.0417 C6H12O6 + 0.1287 O2 + 0.0267 NH4+ + 0.0267 HCO3- --> 0.0267 C4.17H8O1.75N1 + 0.1654 CO2 + 0.21 H2O), each bacterium divides the amount of each nutrient taken up by its respective stoichiometric coefficient and selects the smallest value (the limiting nutrient). This information provides for the demands of each one of the nutrients and drives the creation of new mass and metabolic products generation. After executing the metabolic reaction, the CO2 produced is released to the culture medium. The execution of the metabolic reaction is limited to the existence of sufficient quantities of electron donors and acceptors. After this, if there are any remaining intakes, the syntheses finish, and the remaining intakes are returned to the culture medium.

## Bipartition

The bipartition reproduction process is a rule which is taken from INDISIM (Ginovart et al., 2002a). Table 2 provides more details of the variables related with this procedure along with their values (if constant across all bacteria) or range if variable for each bacterium.

Table 2. Variables used in reproduction procedure

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Simulation variable** | **Units** |
| Pantoea average initial biomass | *ini\_biomass\_pa* | mmol |
| Pantoea biomass | *pa\_biomass* | mmol |
| Minimum biomass needed to reproduce | *pa-biomass-reproduction* | mmol |
| Reproduction time counter | *repr\_time\_pa* | min |
| Reproduction time threshold | *rep\_pa* | min |
| Time step in simulation | *min/steptime* | ---- |
| Standard deviation 10% | *DST* | ---- |
| Maximum number of bacteria allowed per patch | *pmax* | units |
| Number of bacteria in the actual patch (grid cell) | *count turtles-here* | units |

If *pa\_biomass ≥ pa-biomass-reproduction*, the individual can accumulate one count (min/steptime) in *repr\_time\_pa* until *repr\_time\_pa > abs (random-normal rep\_pa (0.1 \* rep\_pa)).*

The biomass division is made randomly at 50% ± 10 % of the actual biomass of bacterium and zeroing the *rep\_pa* counter in both new individuals. The *pa-biomass-reproduction* is updated in both cells (individuals) which also follow the equation:

*pa-biomass-reproduction = abs random-normal (ini\_biomass\_pa) (DST \* ini\_biomass\_pa).*

If *count turtles-here ≤ pmax - 1* the bacterium divides and the reproduction is successful otherwise it looks for empty space in patches (grid cell) adjacent to the current patch. The bipartition procedure only proceeds if there is an adjacent patch with *count turtles-here ≤ pmax.* If the number of bacteria in the patch of the system is higher than a maximum threshold the bacterium cannot execute the bipartition procedure until there exists some empty space (*count turtles-here ≤ pmax)* either in the current cell or one of its eight neighbors.

## Cell death

Table 3. Variables used in cell death procedure

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Simulation variable** | **Units** |
| Cell viability counter | *time-viability\_pa* | min |
| Maximum cell viability | *max-time-viability\_pa* | min |
| Time step in simulation | *min/steptime* | ---- |
| Standard deviation 10% | *DST* | ---- |

Each bacteria has a time viability (*time-viability\_pa),* that represents the maximum time in which bacteria could survive without having enough nutrients to supply its maintenance requirements.

*time-viability\_pa = abs random-normal (max-time-viability\_pa) (DST \* max-time-viability\_pa)*

At each simulation time step the bacteria could maintain its initial *time-viability\_pa,* while getting enough nutrients or could decrease its *time-viability\_pa* following the next rule:

*time-viability\_pa = time-viability\_pa - min/steptime*

If *time-viability\_pa* is equal to *zero* the cell dies.

## Agar Plate operation

The rules related to the culture medium on the agar plate are: Nutrient diffusion. We use the Netlogo primitive called “diffuse” to simulate the mass diffusion process in solid-phase linked with an input parameter called *diffusion-coefficient*. This procedure executes to update the nutrient (glucose, oxygen, ammonium, bicarbonate, carbon dioxide and water) concentration in patches (grid cells). The nutrient amounts are distributed equally among the 8 neighbor patches following the expression: *1/8 \* diffusion-coefficient.*

## References

Araujo-Granda, P., Gras, A., Ginovart, M., and Moulton, V. (2020). INDISIM-Denitrification, an individual-based model for study the denitrification process. *J. Ind. Microbiol. Biotechnol.* 47, 1–20. doi:10.1007/s10295-019-02245-8.

Araujo Granda, P., Gras, A., Ginovart, M., and Moulton, V. (2016). INDISIM-Paracoccus, an individual-based and thermodynamic model for a denitrifying bacterium. *J. Theor. Biol.* 403, 45–58. doi:10.1016/j.jtbi.2016.05.017.

Ginovart, M., López, D., and Valls, J. (2002a). INDISIM, an individual-based discrete simulation model to study bacterial cultures. *J. Theor. Biol.* 214, 305–19. doi:10.1006/jtbi.2001.2466.

Ginovart, M., López, D., Valls, J., and Silbert, M. (2002b). Individual based simulations of bacterial growth on agar plates. *Phys. A Stat. Mech. its Appl.* 305, 604–618. doi:10.1016/S0378-4371(01)00581-7.

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., and Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecol. Modell.* 221, 2760–2768. doi:10.1016/j.ecolmodel.2010.08.019.

McCarty, P. L. (2007). Thermodynamic electron equivalents model for bacterial yield prediction: modifications and comparative evaluations. *Biotechnol. Bioeng.* 97, 377–388. doi:10.1002/bit.

Rittmann, B., and McCarty, P. L. (2001). *Environmental Biotechnology: Principles and Applications*. First edit. , ed. McGraw-Hill McGraw-Hill Available at: https://www.mheducation.com/highered/product/environmental-biotechnology-principles-applications-rittmann-mccarty/9781260440591.html.