CEMRACS 2016. Numerical challenges in parallel scientific computing

The CEMRACS project HPC-IIBios: scientific background, motivation and overview

Xavier Portell, Olga Assainova, Valérie Pot, Philippe Baveye, Patricia Garnier

UMR ECOSYS, AgroParisTech, INRA, Université Paris-Saclay
F-78850 Thiverval-Grignon, France.
E-mail: xavier.portell-canal@grignon.inra.fr

August 19, 2016, Marseille

SUMMARY

1. Introduction and background
2. The ANR project Soil µ3D: towards more accurate CO_2 and N_2O gas emissions predictions
3. The model IlBioS: Coupling a lattice-Boltzmann approach to a biological individual-based model
4. Project HPC-IlBioS: understanding the best HPC strategy for the IlBioS approach

SUMMARY

1. Introduction and background
2. The ANR project Soil µ3D: towards more accurate CO_2 and N_2O gas emissions predictions
3. The model IlBioS: Coupling a lattice-Boltzmann approach to a biological individual-based model
4. Project HPC-IlBioS: understanding the best HPC strategy for the IlBioS approach

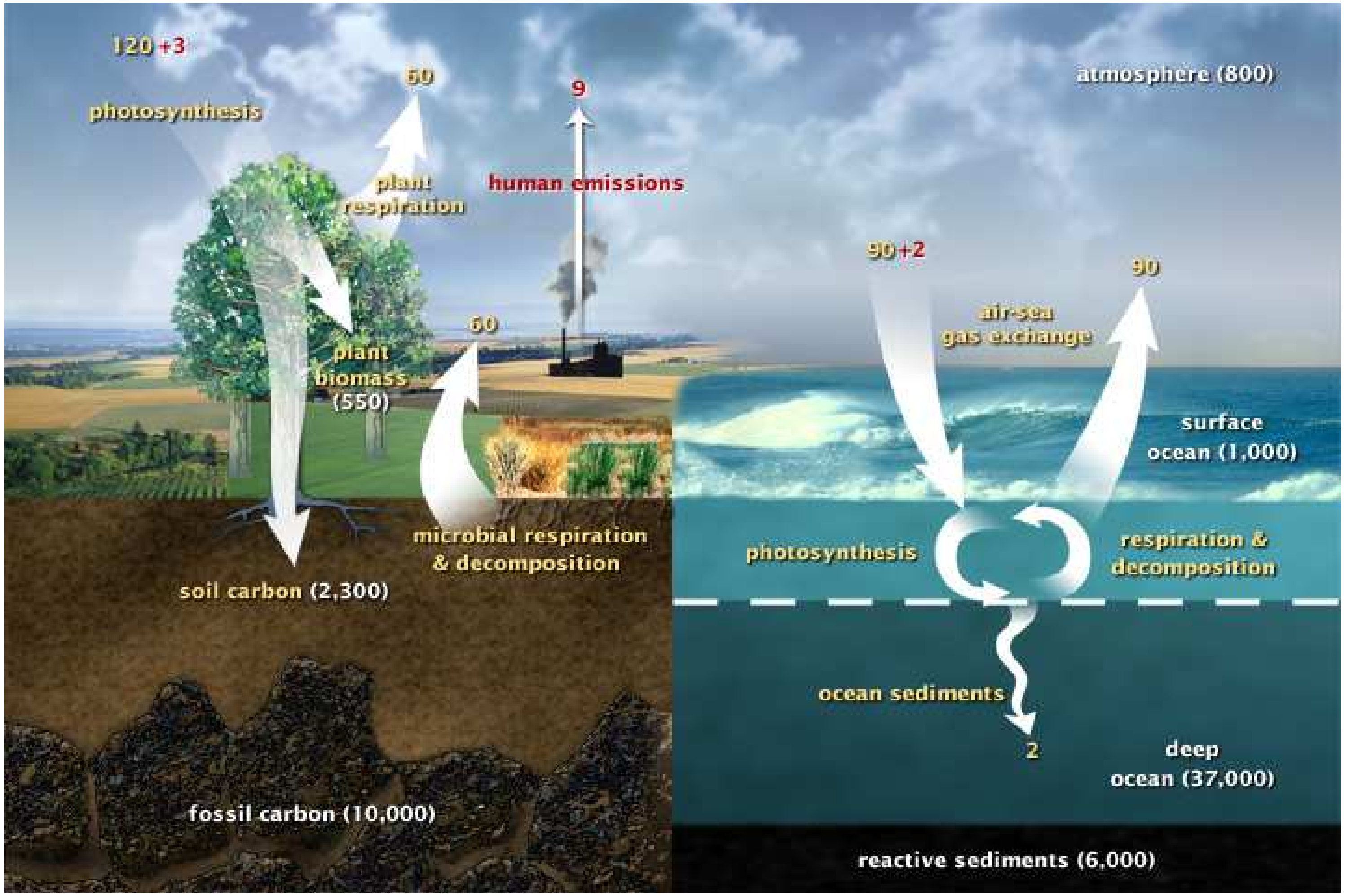
_01

Introduction and background

INRA
SCIENCE & IMPACT

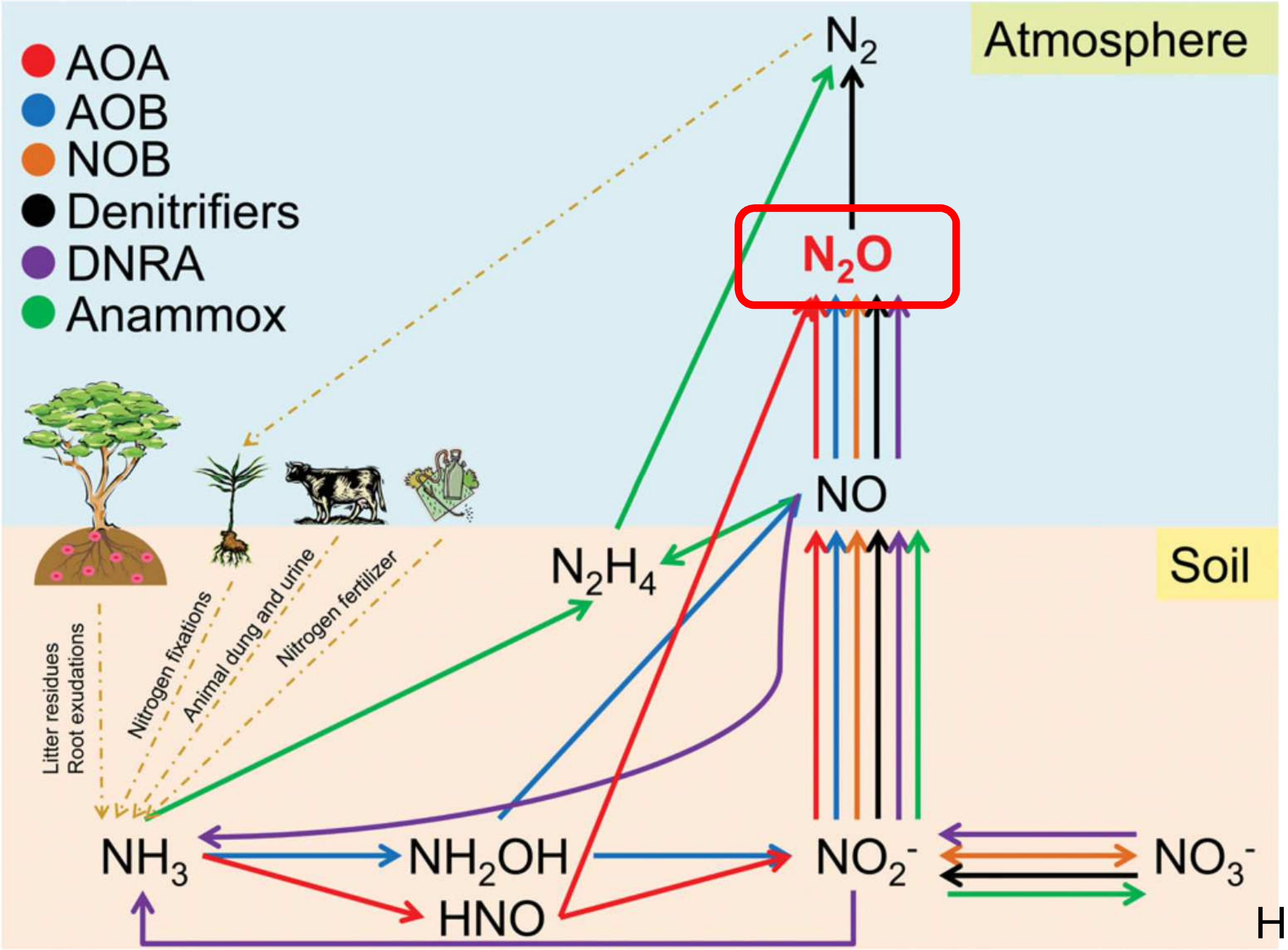
The CO2 and N2O as main GHG

An important part of the CO_2 emissions are due to the activity of the soil microorganisms



Riebbek (2011)

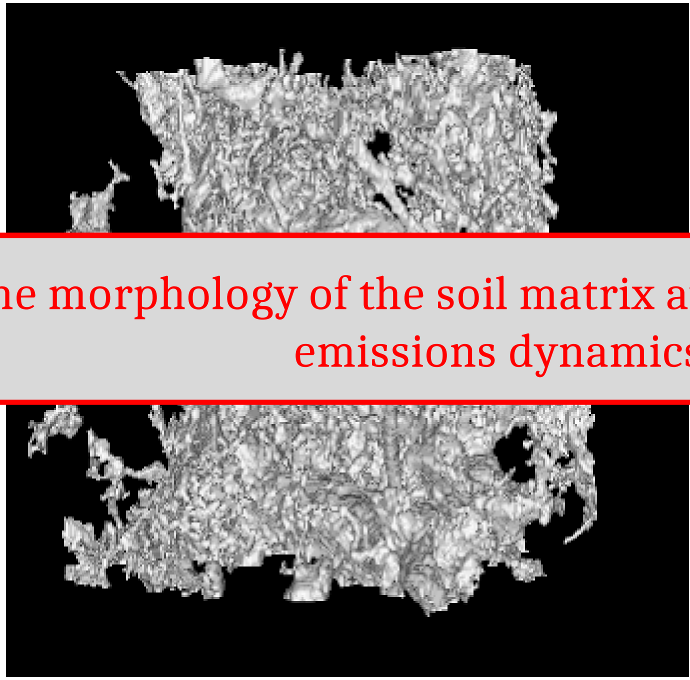
0,03 % of the total GHG emissions but with a 300-fold greater potential for global warming (Thomson et al 2012)



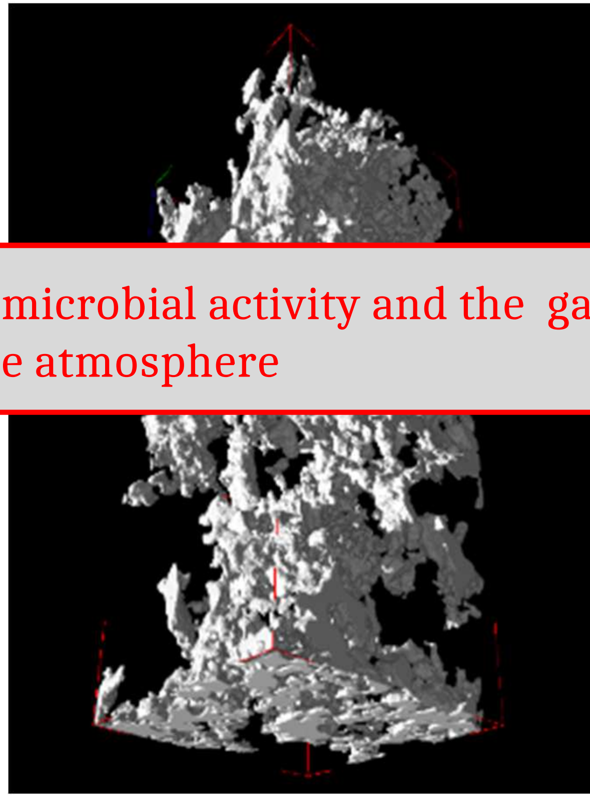
Hu et al (2015)

The soil matrix

The complex geometry of the pore space can now be characterized to a high level of detail and quantify its connectivity and topology characterized.



X-ray Computed Tomography



The morphology of the soil matrix affects microbial activity and the gas emissions dynamics to the atmosphere

Microorganisms in soil tend to be found in microcolonies, this means that "identical organisms" will face different microhabitat conditions

The individuality of the microorganisms plays a role

Bacteria in the top soil of a Podsol.
Eickhorst & Tippkötter, 2008 (SBB)

[1] Compute-aided Detection - Fluorescence In Situ Hybridization

Mass transport processes modelling in soil

To take into account mass transport processes in the soil matrix we can use two main group of methods:

Volume of Fluid (VoF); Finite Element (FE) Methods

$$\frac{\partial c}{dt} = \nabla \cdot (D \nabla c) - \nabla \cdot (\vec{v} c) + R$$ ⇨ **Solving for t and space**

Advantage

Relatively low computing time required

Issues

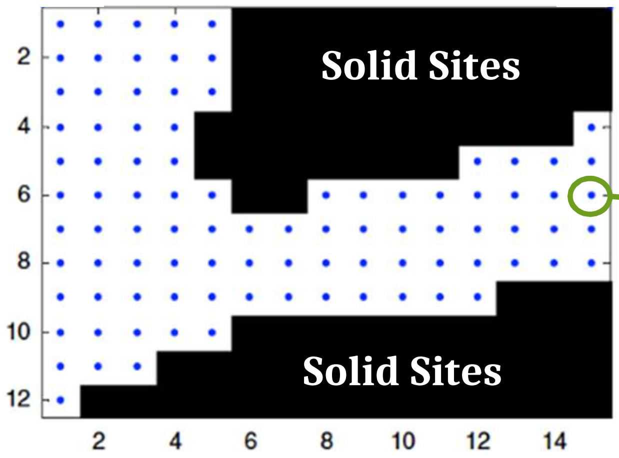
The complex pore space geometry hard to apply

Mathematical resolution difficult and highly specific of the system

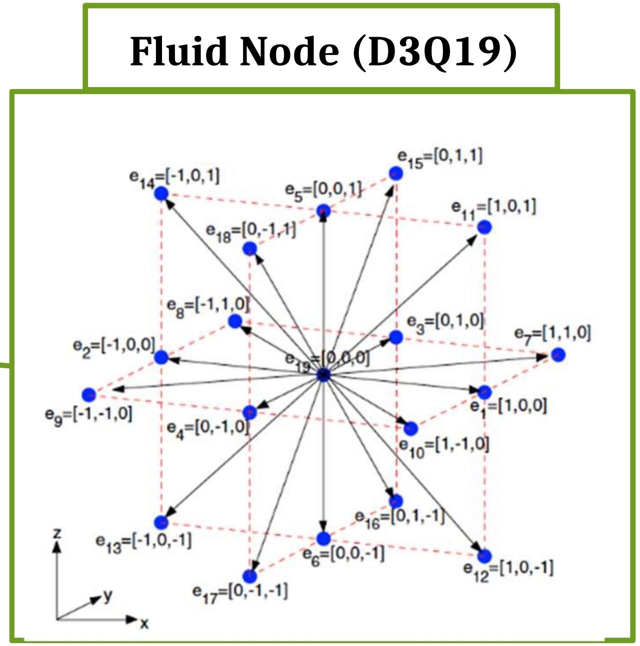
Volume of Fluid (VoF); Finite Element (FE) Methods

$$\frac{\partial c}{dt} = \nabla \cdot (D\nabla c) - \nabla \cdot (\vec{v}c) + R$$ ⟹ **Solving for t and space**

Lattice-Boltzmann modelling



From: www.egr.msu.edu/ ~kutay/Lbsite/



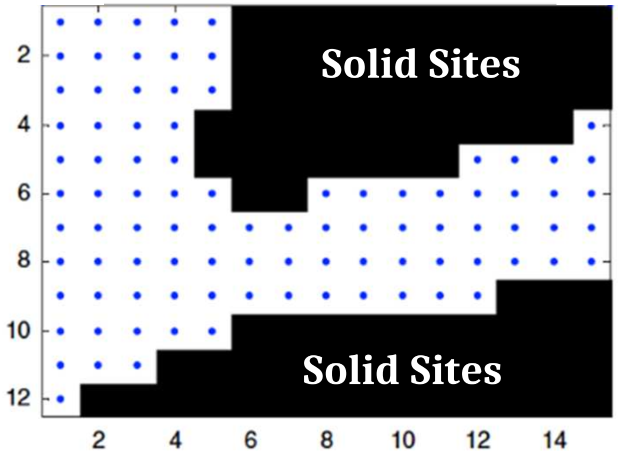
III. Mass transport processes

Volume of Fluid (VoF); Finite Element (FE) Methods

$$\frac{\partial c}{dt} = \nabla \cdot (D \nabla c) - \nabla \cdot (\vec{v} c) + R$$

⇨ **Solving for t and space**

Lattice-Boltzmann modelling



From: www.egr.msu.edu/ ~kutay/Lbsite/

Advantage

Complex pore space geometry easy to handle

Issues

Computationally demanding, but easy to parallelize (GPU or accelerators)

Individual-based modelling

IbMs in microbiology: a definition

"Simulation models that treat individuals as unique and discrete entities which have at least two independent properties."
(Hellweger and Bucci, 2009) *

* Extending a previous definition by Grimm (1999)

IbMs are becoming consolidated in the field. Examples are the µIbMs of: Kreft *et al.* (1998), Dens *et al.* (2005), Bucci *et al.* (2012), Tack *et al.* (2014), and Hellweger *et al.* (2014)

A very simple example: The prey-predator model

Population-based

- Description of the **population density** using diferential equations

$$\frac{dV}{dt} = rV - aVP$$

$$\frac{dP}{dt} = -qP + baVP$$

V =ares density; P =lynks density; *r*, *a*, *b* and *q* are model parameters

Individual-based

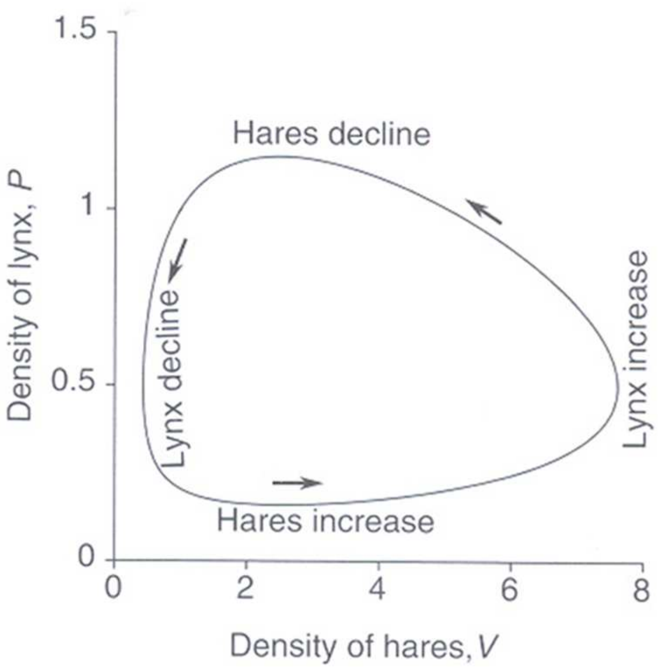
- **Individual properties** and the rules defining the **behaviour** of the individual are modelled and the population evolution ***emerges*** from the interactions into the system

Behaviour:

- Wolfs eat sheeps
- Wolfs create new wolfs
- Sheeps create new sheeps
- Starving individuals may day

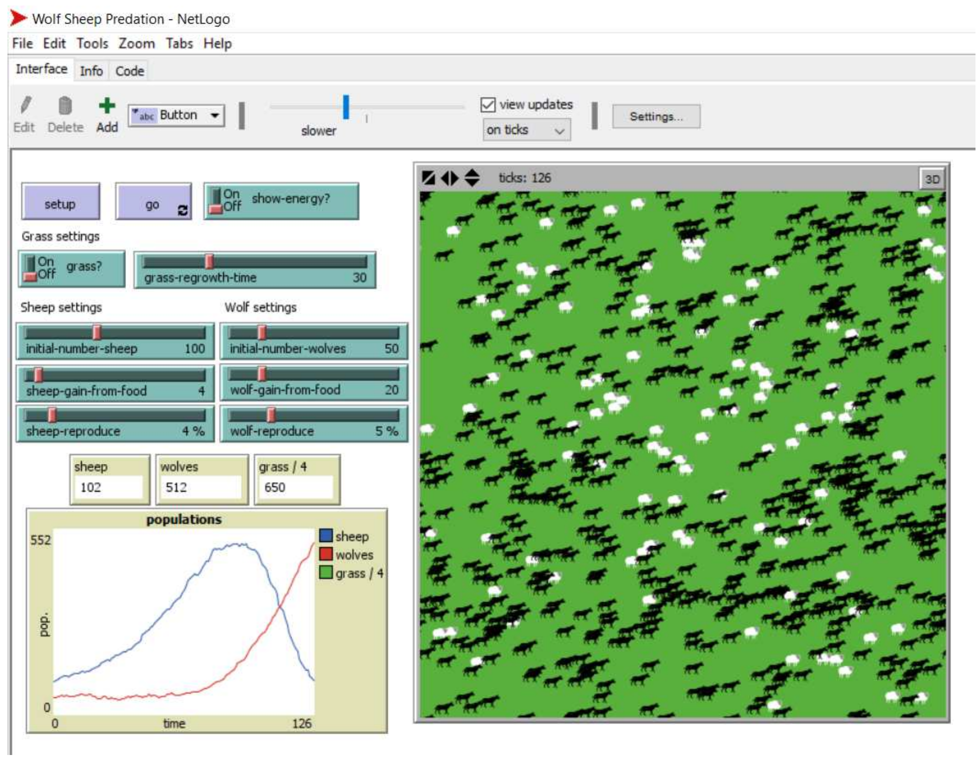
A very simple example: The prey-predator model

Population-based



Source: Kokko (2007)

Individual-based



Source: Netlogo home page (2016)

SUMMARY

1. Introduction and background
2. The ANR project Soil µ3D: towards more accurate CO_2 and N_2O gas emissions predictions
3. The model IlBioS: Coupling a lattice-Boltzmann approach to a biological individual-based model
4. Project HPC-IlBioS: understanding the best HPC strategy for the IlBioS approach

_02

The ANR project Soil µ3D: towards more accurate CO_2 and N_2O gas emissions predictions

INRA
SCIENCE & IMPACT

Emergent properties of soil microbial functions: upscaling from 3D modeling and spatial descriptors of pore scale heterogeneity (Soilµ-3D)

- Funding agency: The French National Research Agency (ANR)
- Starting and ending date: 01/11/2016 to 01/11/2020.
- Quantity: 250 000€
- Participants:
  - 29 permanent researchers
  - 1 post-doctoral position
  - 4 PhD thesis
  - 3 Master 2 stages

INRA SCIENCE & IMPACT

AgroParisTech INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

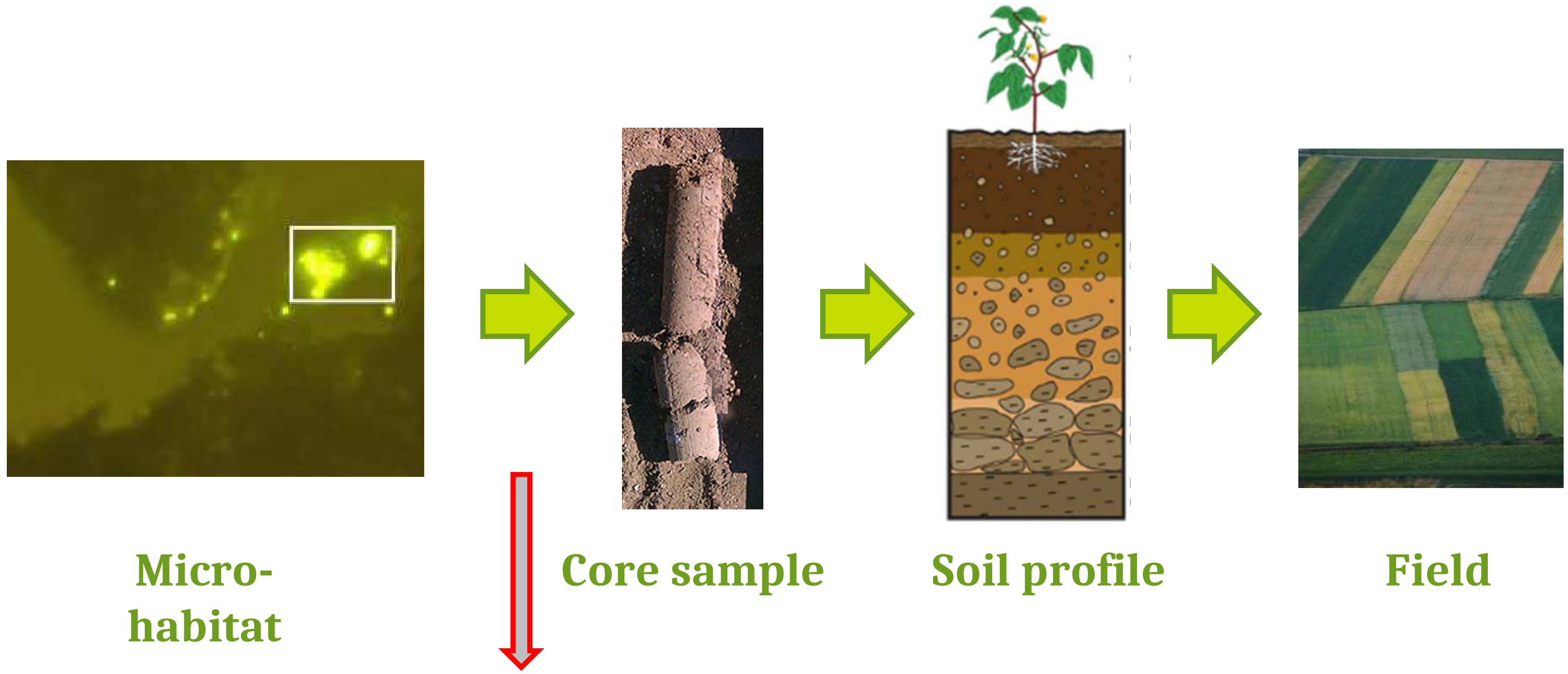
cnrs

UPMC SORBONNE UNIVERSITÉS

Montpellier SupAgro

Aim of the project Soil µ3D

The main goal of the project is to upscale heterogeneities identified at the scale of microhabitats to the soil profile scale.



What do we have to measure to scale up to higher scales?

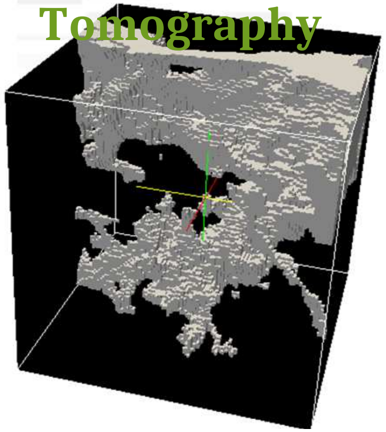
Use of the model IlBioS in the project Soil µ3D

Use of the model IlBioS in the project Soil µ3D

The model ILBioS is built coupling an **Individual-based Model** of the soil bacteria to a **lattice-Boltzmann model** simulating the fluid dynamics and mass transport processes of soluble substrates

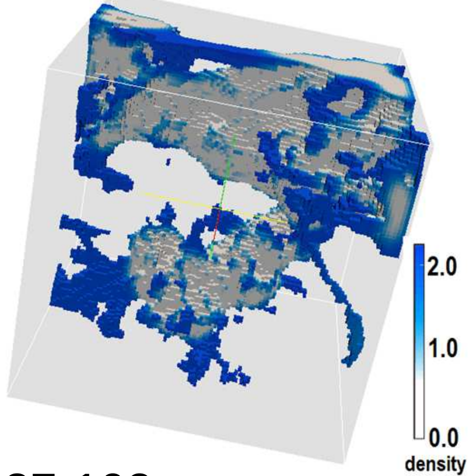
The model ILBios and the Work Package 3

3D Computed Tomography

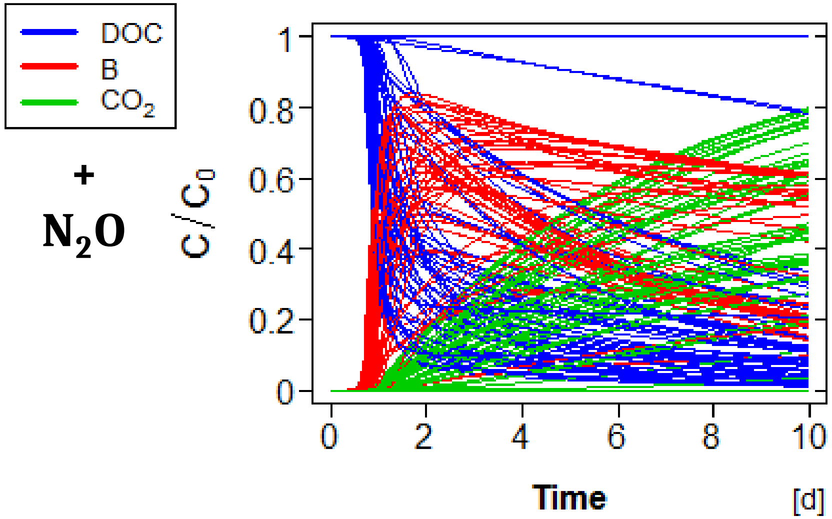


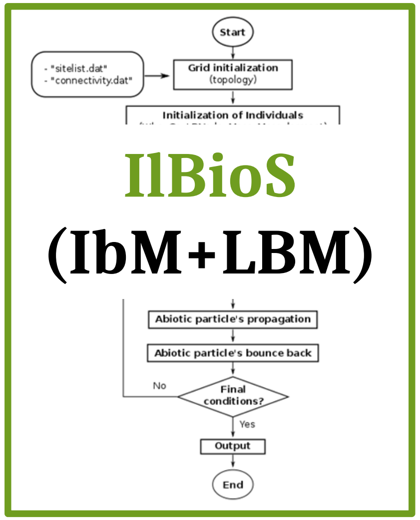
2-P TRT Lattice-Boltzmann *

Solid-Liquid-Gas phases



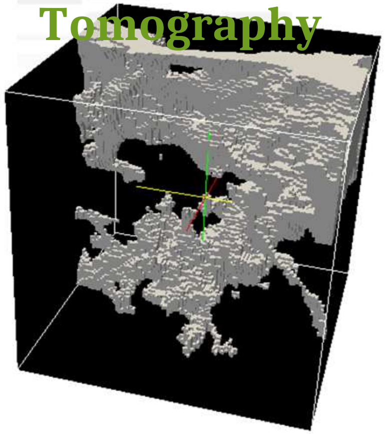
*** Pot et al. (2015)** *Advances in Water Resources* 84, 87-102.



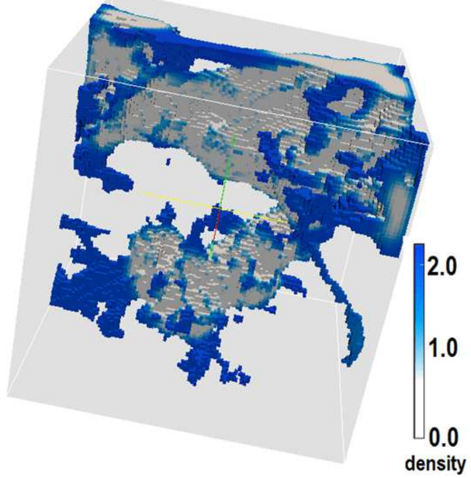


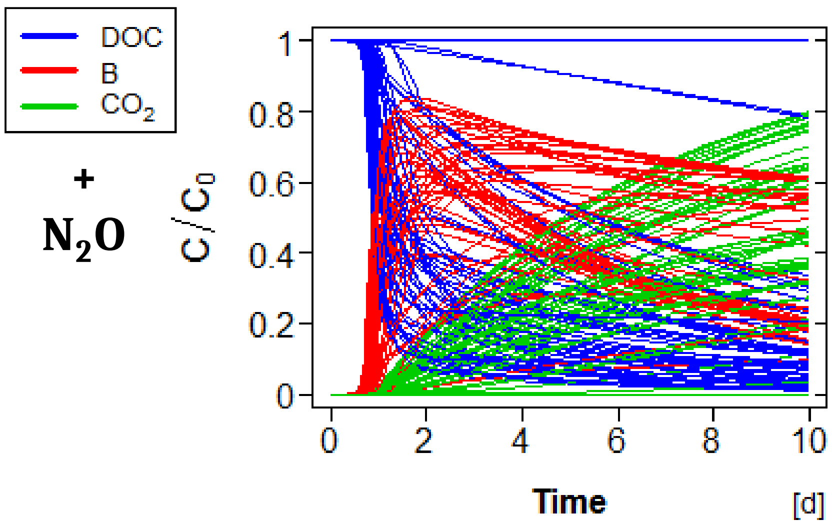
The model ILBios and the Work Package 3

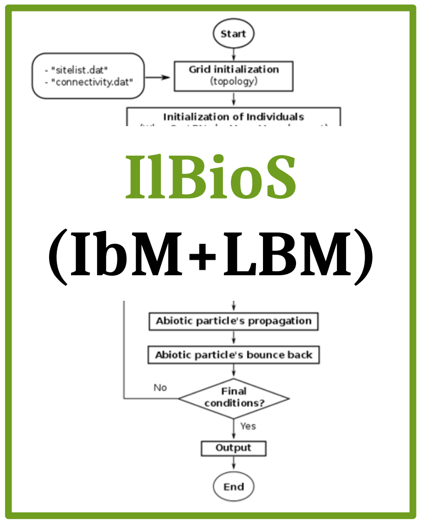
3D Computed Tomography



Solid-Liquid-Gas phases



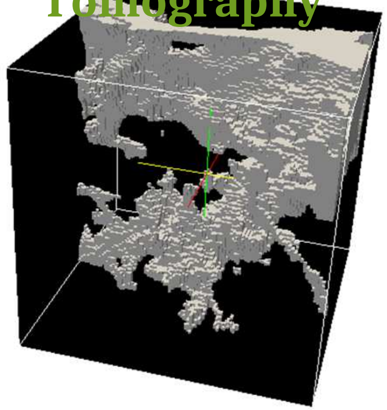


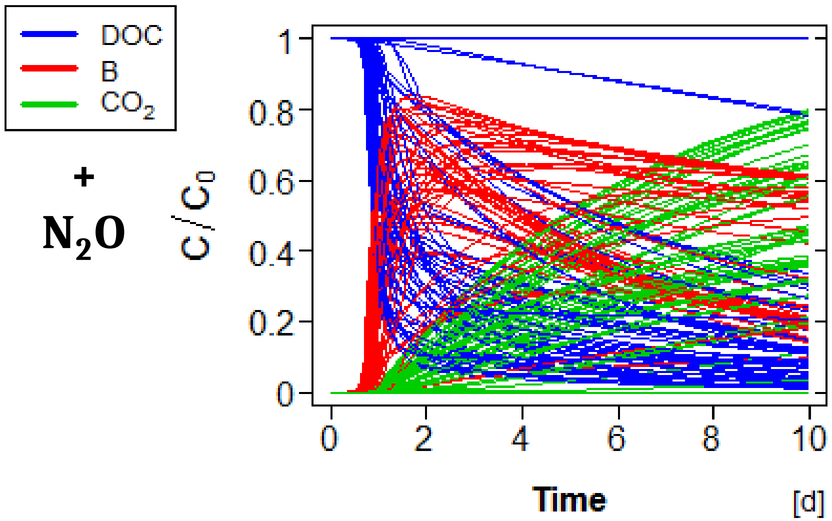


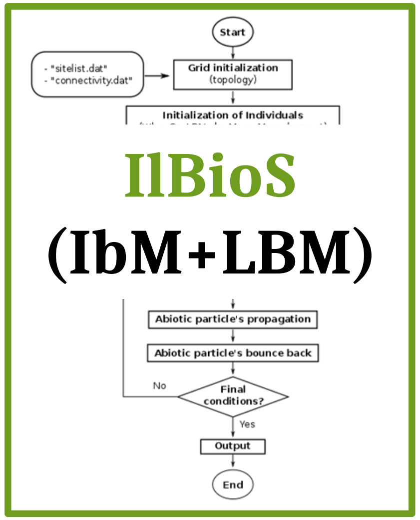
Integration of descriptors in soil profile models

Set of µ-organisation Descriptors

3D Computed Tomography







SUMMARY

1. Introduction and background
2. The ANR project Soil µ3D: towards more accurate CO_2 and N_2O gas emissions predictions
3. The model IlBioS: Coupling a lattice-Boltzmann approach to a biological individual-based model
4. Project HPC-IlBioS: understanding the best HPC strategy for the IlBioS approach

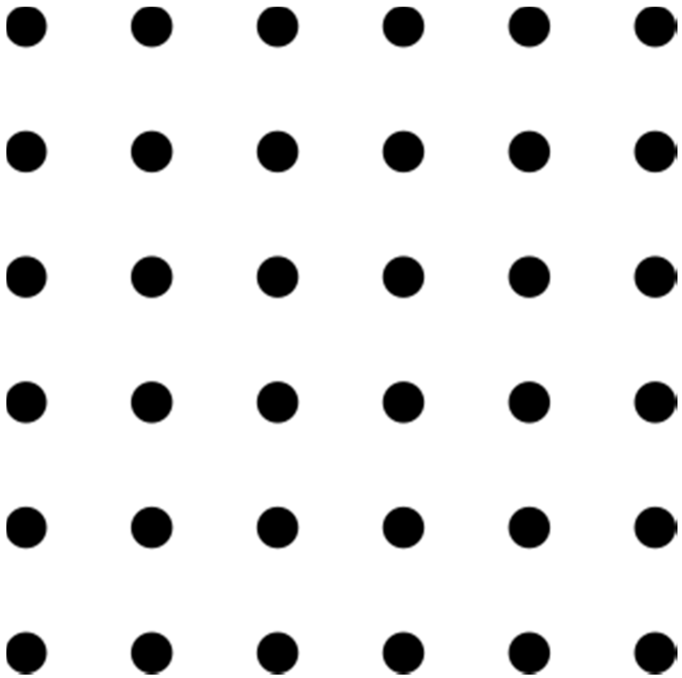
_03

The model IIBioS: Coupling a lattice-Boltzmann approach to a biological individual-based model

INRA
SCIENCE & IMPACT

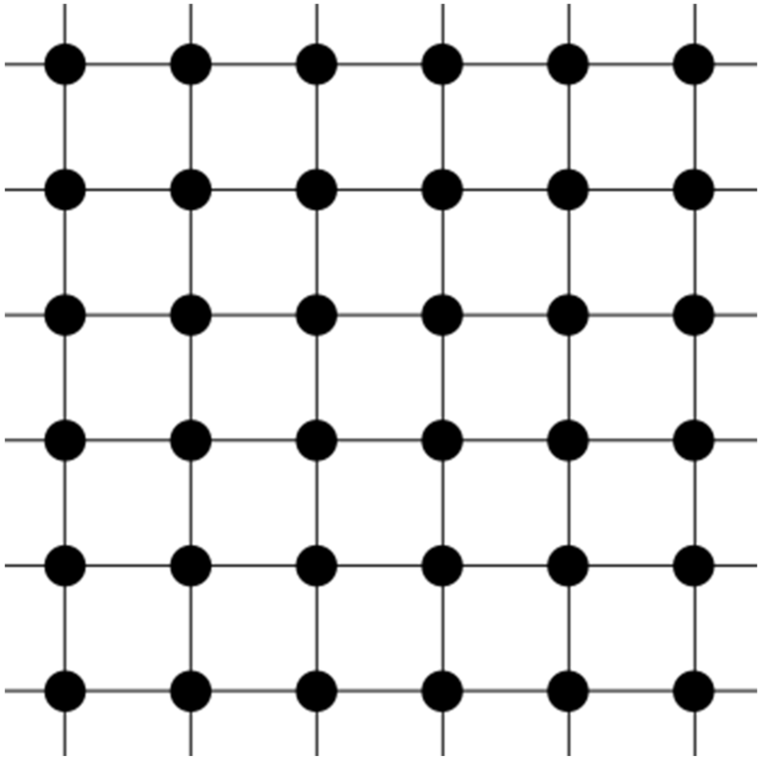
Model conceptualization and description

Three dimensional lattice-Boltzmann (D3Q7)



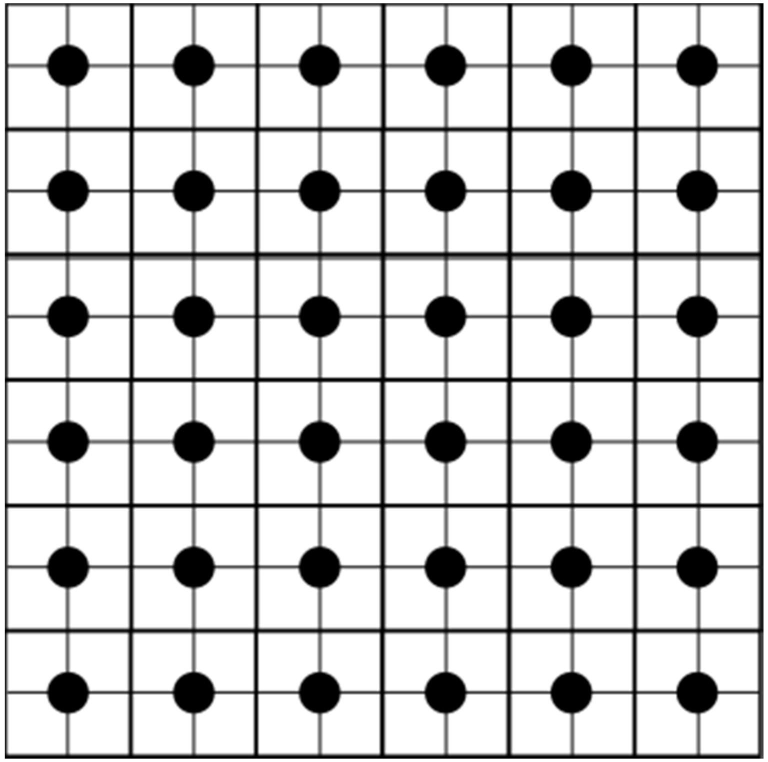
Lattice-Boltzmann Nodes

Three dimensional lattice-Boltzmann (D3Q7)

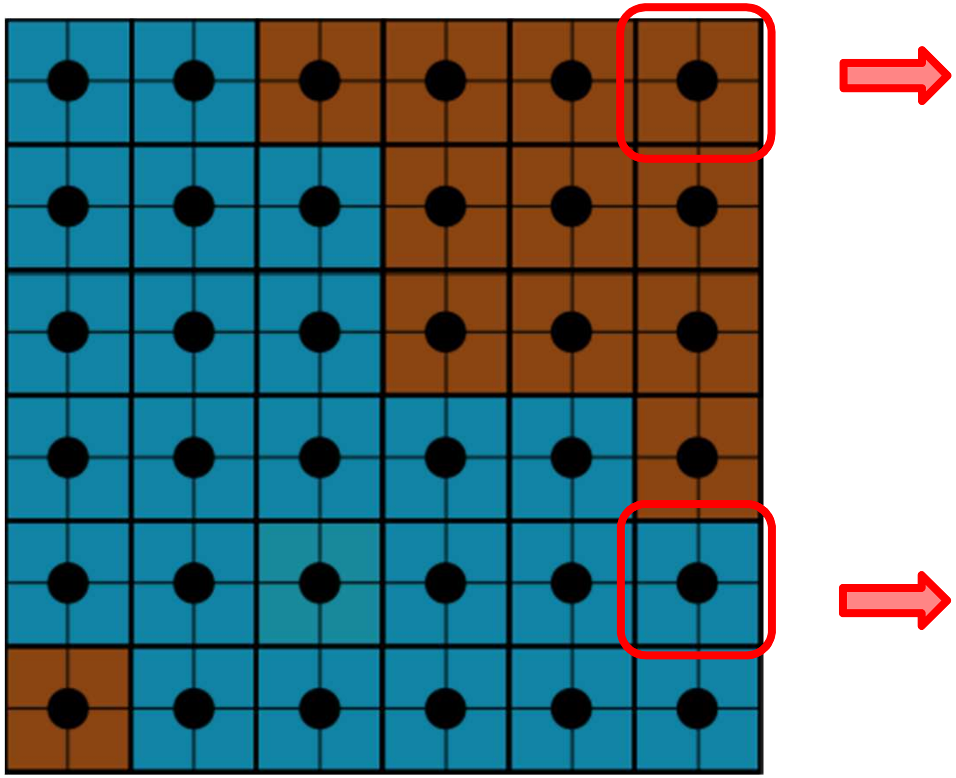


Connected Lattice-Boltzmann Nodes

Three dimensional lattice-Boltzmann (D3Q7)



Three dimensional lattice-Boltzmann (D3Q7)

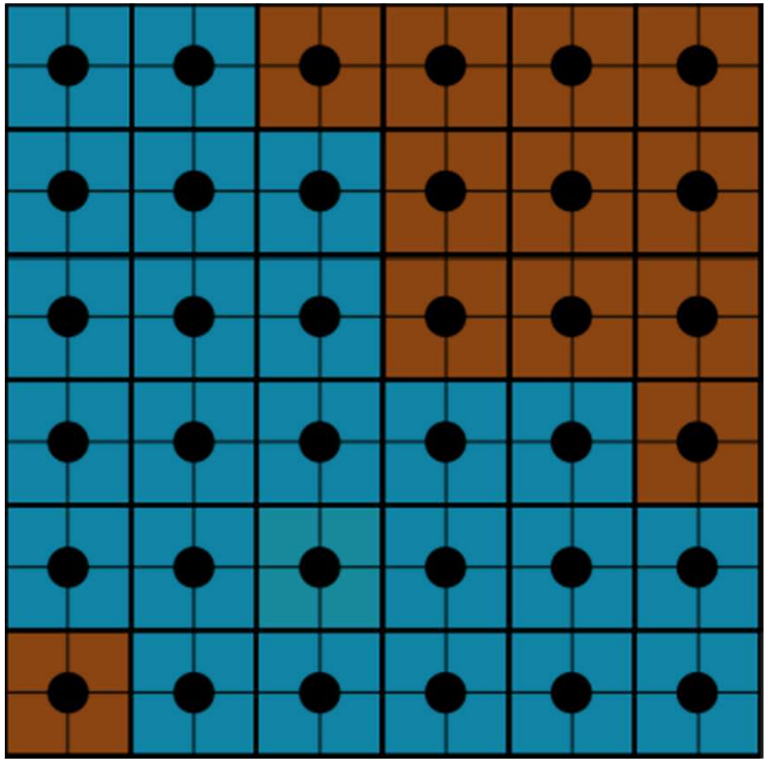


Solid nodes

Fluid nodes

Dissolved Organic Carbon (DOC)

Three dimensional lattice-Boltzmann (D3Q7)



$u(x, y, z, t)$

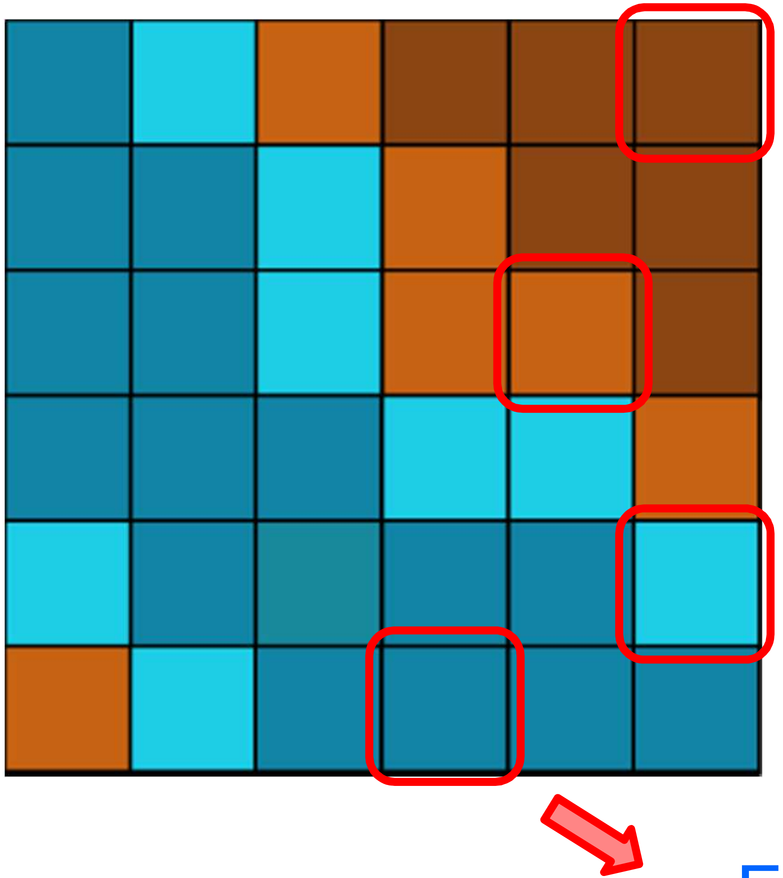
$v(x, y, z, t)$

$w(x, y, z, t)$

$\rho(x, y, z, t)$

$$dt\ \rho + dx\ (u\ \rho) + dy\ (v\ \rho) + dz\ (w\ \rho) = 0$$

Three dimensional lattice-Boltzmann (D3Q7)



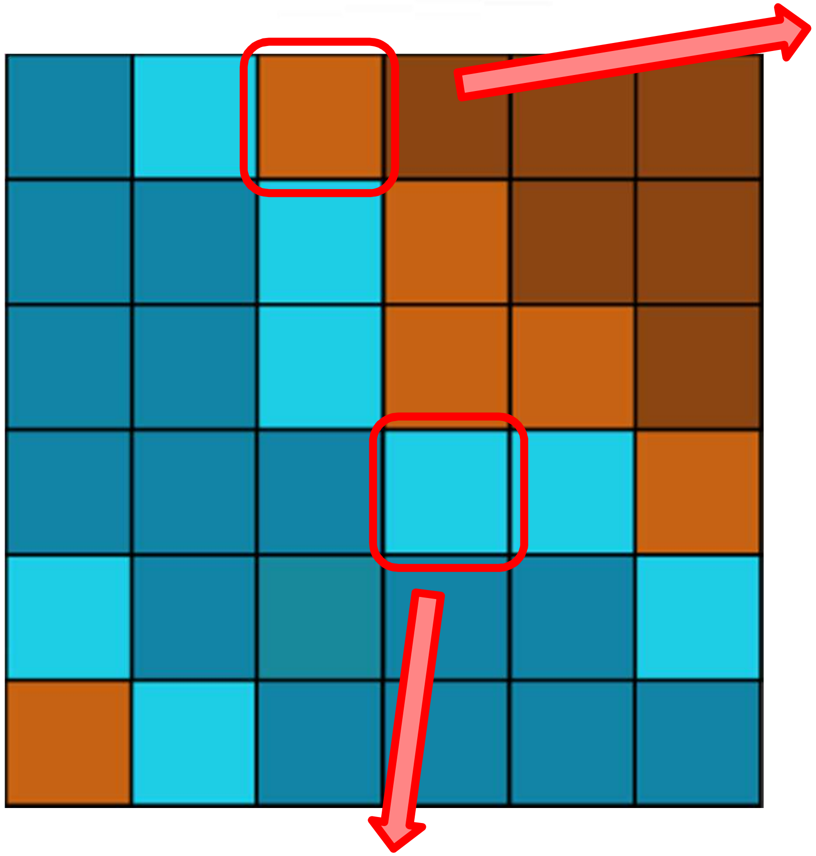
Solid nodes

Boundary
Solid nodes

Boundary Fluid
nodes

Fluid nodes

Three dimensional lattice-Boltzmann (D3Q7)



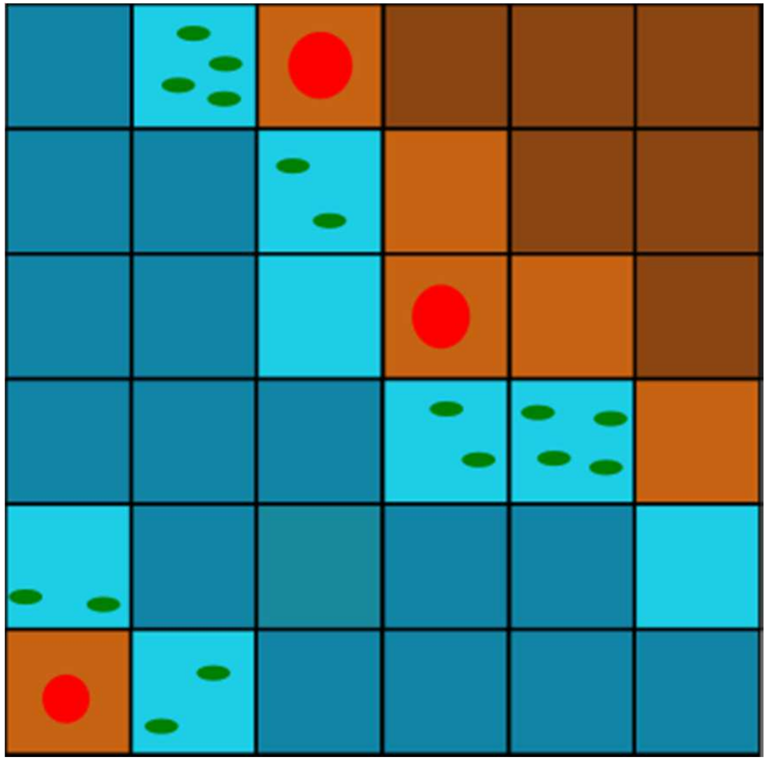
Particulate Organic Matter:

Release DOC to the boundary fluid nodes.

$$dt\ \rho + dx\ (u\ \rho) + dy\ (v\ \rho) + dz\ (w\ \rho) = S_A - S_B$$

Soil bacteria: Uptake DOC from the boundary fluid nodes.

Three dimensional lattice-Boltzmann (D3Q7)

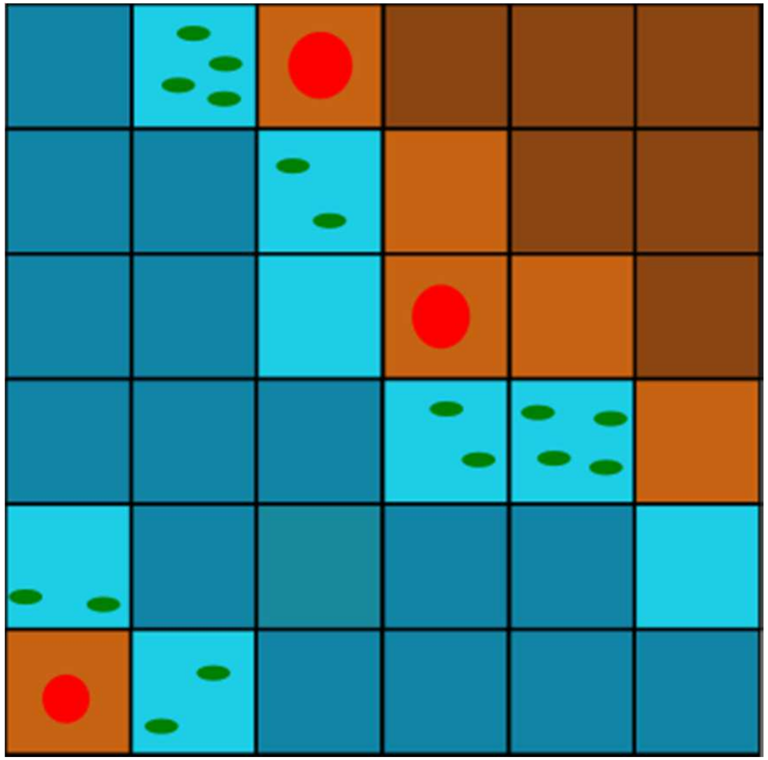


$$dt\,\rho + dx\,(u\,\rho) + dy\,(v\,\rho) + dz\,(w\,\rho) = S_A - S_B$$

$$S_A = \left\{ \begin{array}{ll} 0 & \text{; if solid or bulk} \\ \dfrac{k_{POM}\, m_j}{n} & \text{; if boundary liquid} \end{array} \right\}$$

m_j mass of the POM agent.

Three dimensional lattice-Boltzmann (D3Q7)

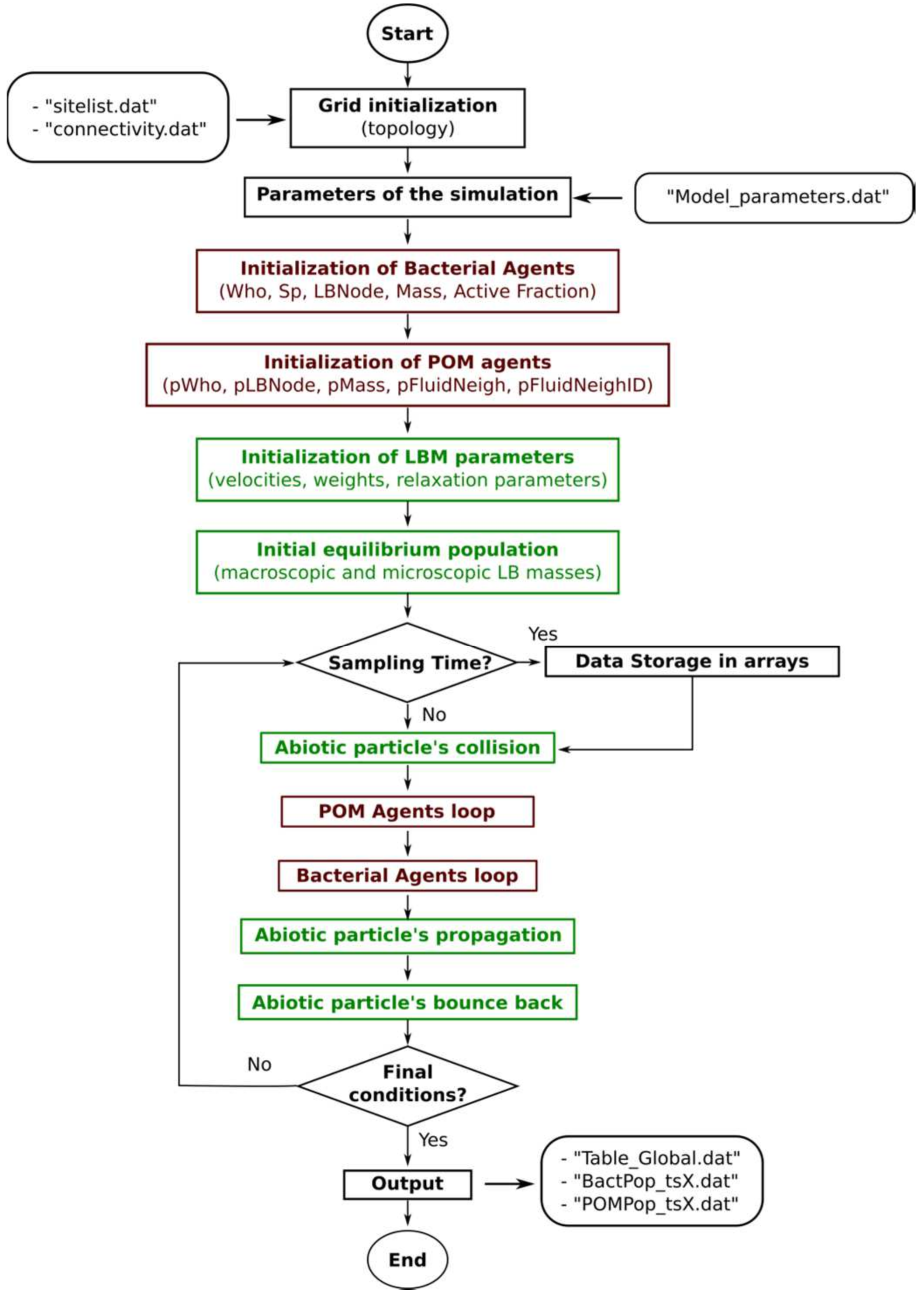


$$dt\,\rho + dx\,(u\,\rho) + dy\,(v\,\rho) + dz\,(w\,\rho) = S_A - S_B$$

$$S_B = \left\{ \begin{matrix} 0 & ;\text{if solid or bulk} \\ \sum \left(\frac{k_{DOC}\,\rho}{\rho + k_{DOC}}\, m_i \right) & ;\text{if boundary liquid} \end{matrix} \right\}$$

m_i mass of the bacterium

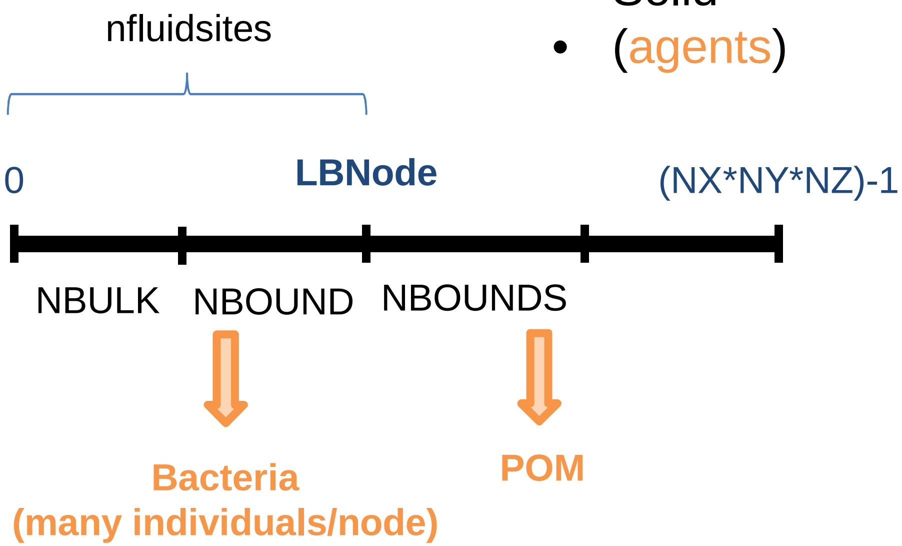
General workflow of the model IlBioS

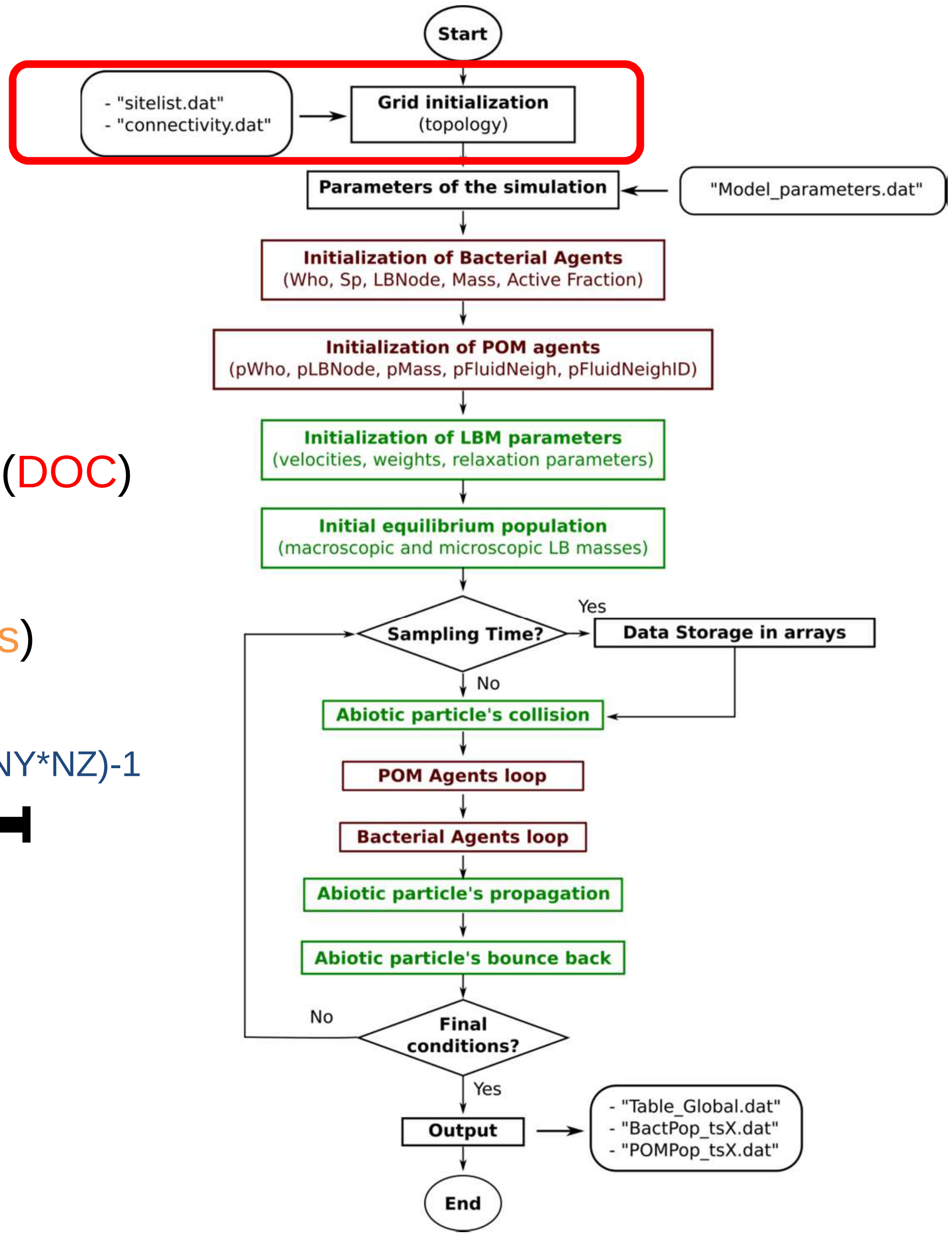


General workflow of the model IlBioS

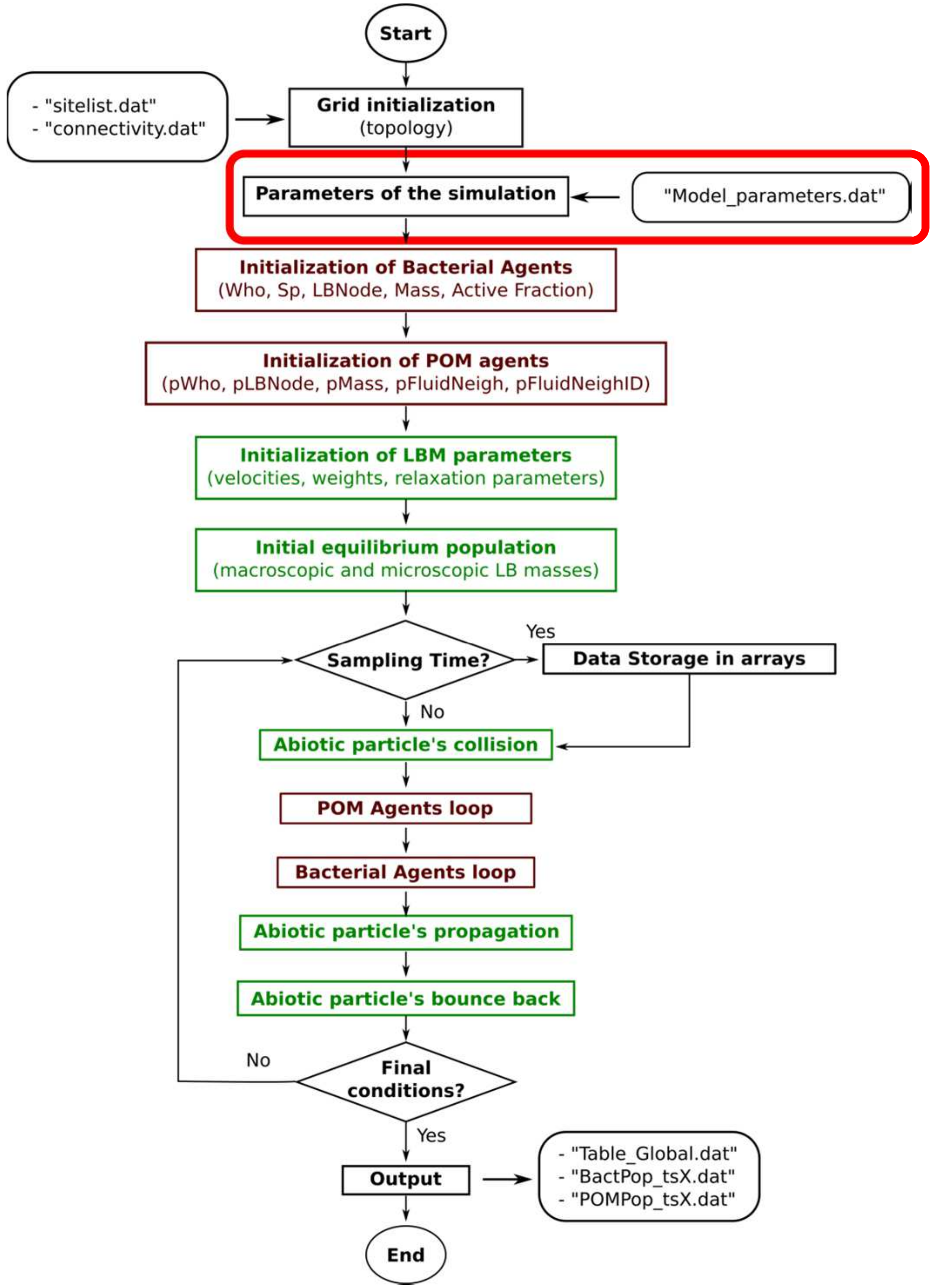
Spatial cells (LB Nodes)

- Water (DOC)
- Air
- Solid
- (agents)





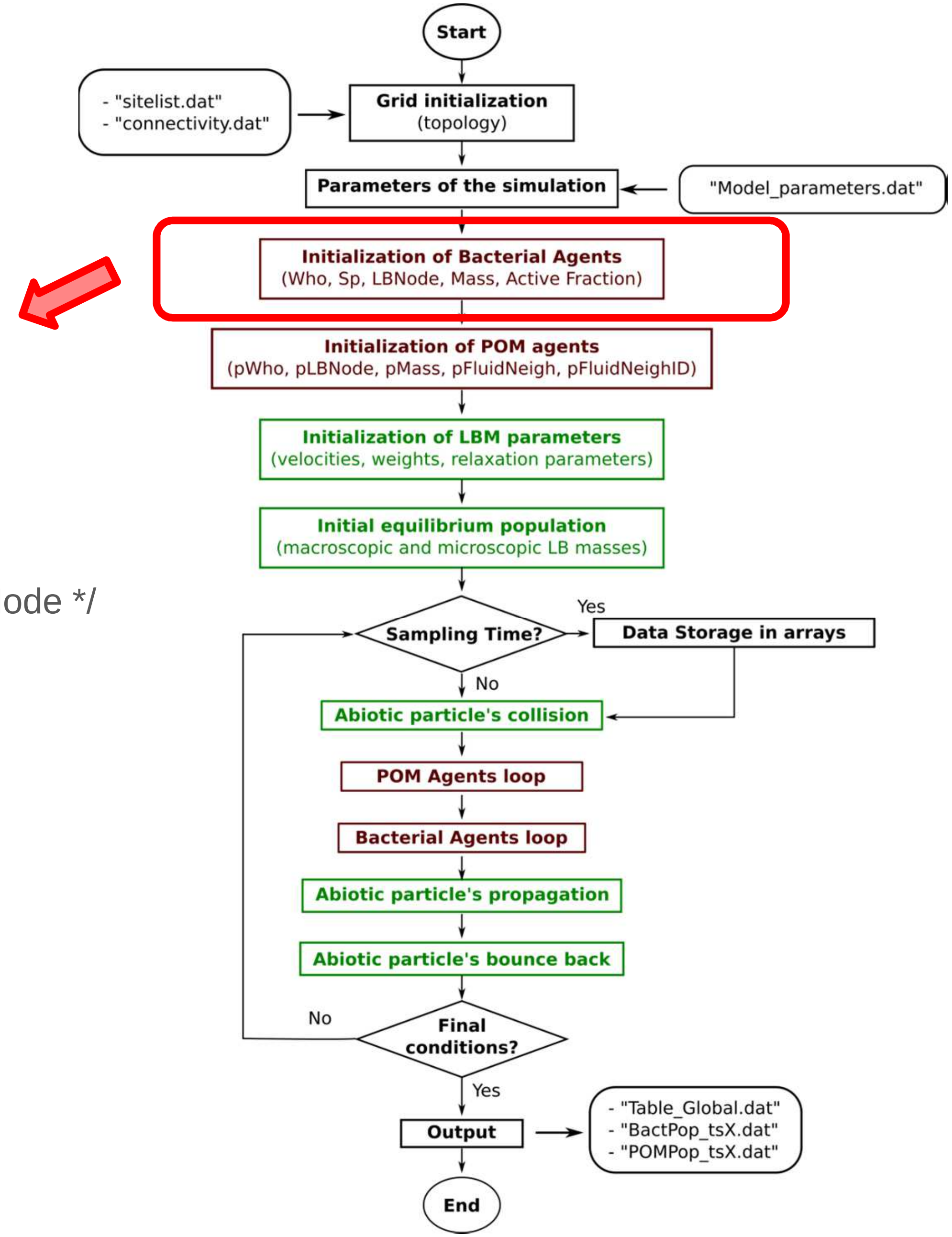
General workflow of the model IlBioS



General workflow of the model IlBioS

Vector holding the individual bacteria properties

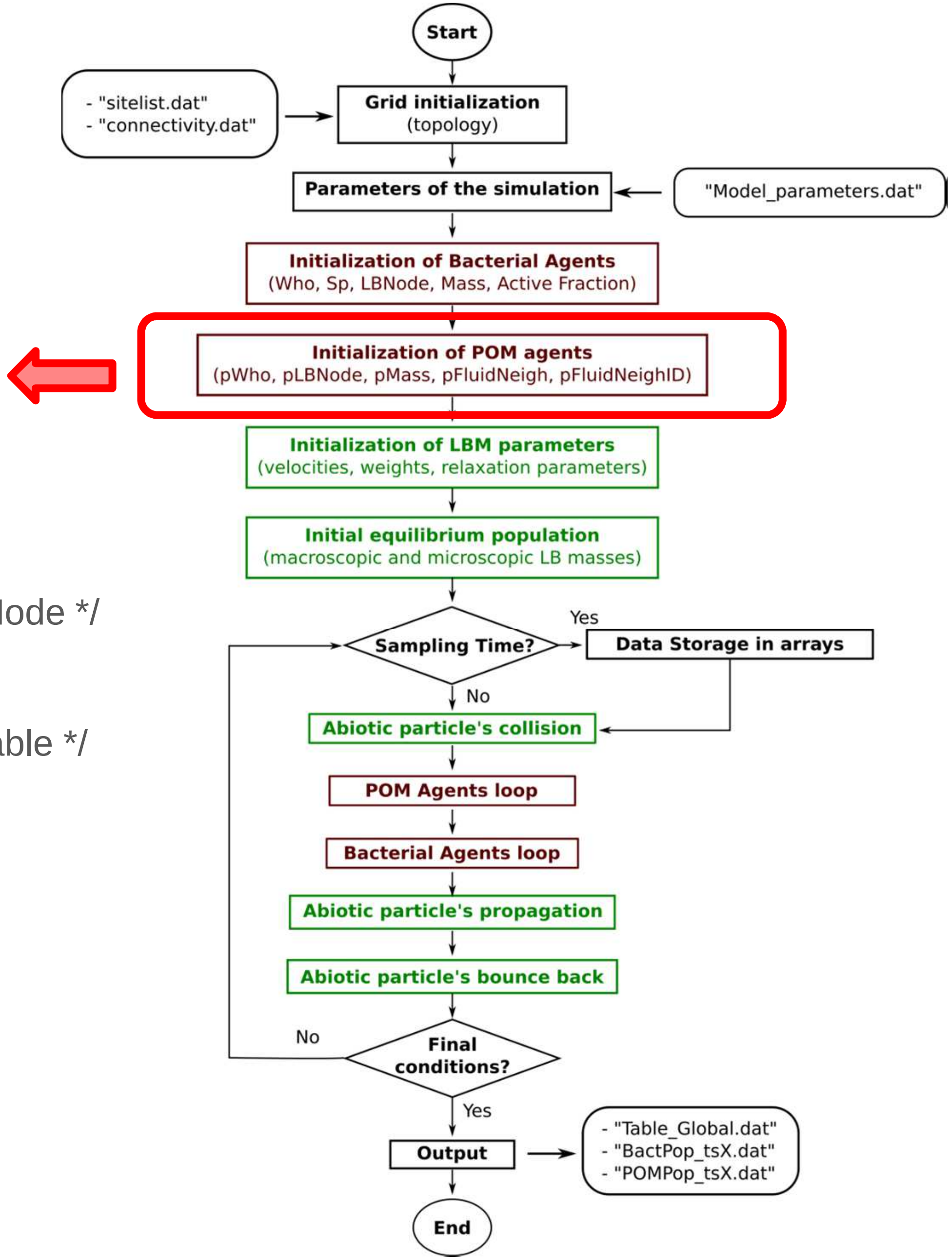
```
Typedef struct BiologicalAgent {
    unsigned long Who;  /* ID of the individual */
    unsigned long  LBNode; /* lattice-Boltzmann Node */
    double Mass;
    double Active_Fraction;
    double Respired_CO2; /* technical variable */
} type_bacteria;
```



General workflow of the model IlBioS

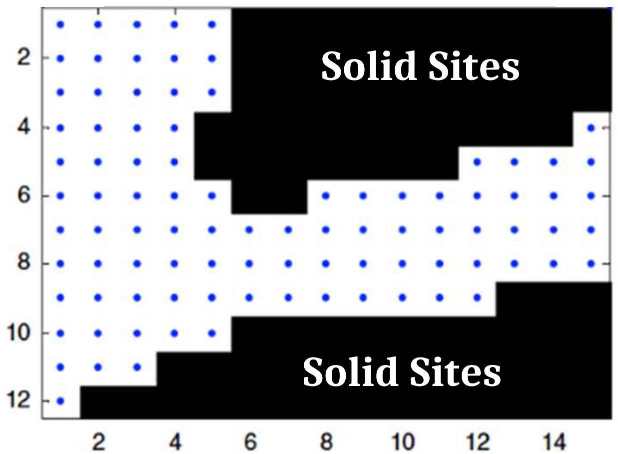
Vector holding the POM properties

```
Typedef struct POMAgent {
      unsigned long pWho;  /* ID */
      unsigned long  pLBNode; /* lattice-Boltzmann Node */
      double pMass;
      int pFluidNeigh; /* Number of fluid neighbors */
      unsigned long *pFluidNeighID; /* technical variable */
} type_POM;
```



General workflow of the model IlBioS

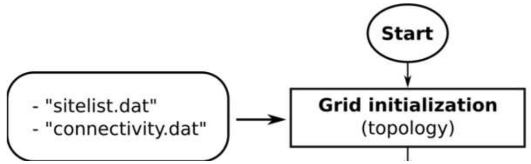
Initialization of the DOC = population Boltzmann



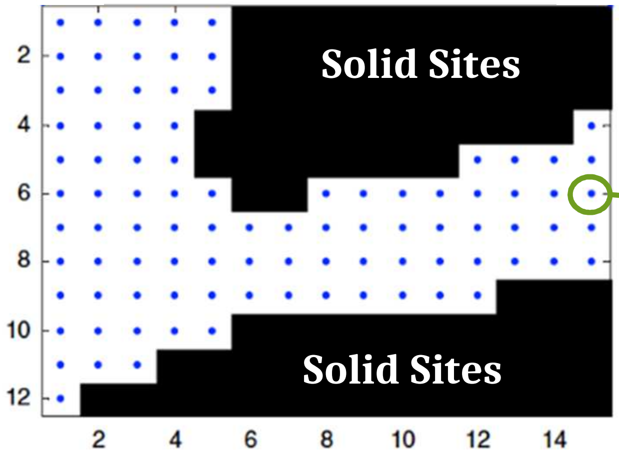
From: www.egr.msu.edu/ ~kutay/Lbsite/

INRA SCIENCE & IMPACT

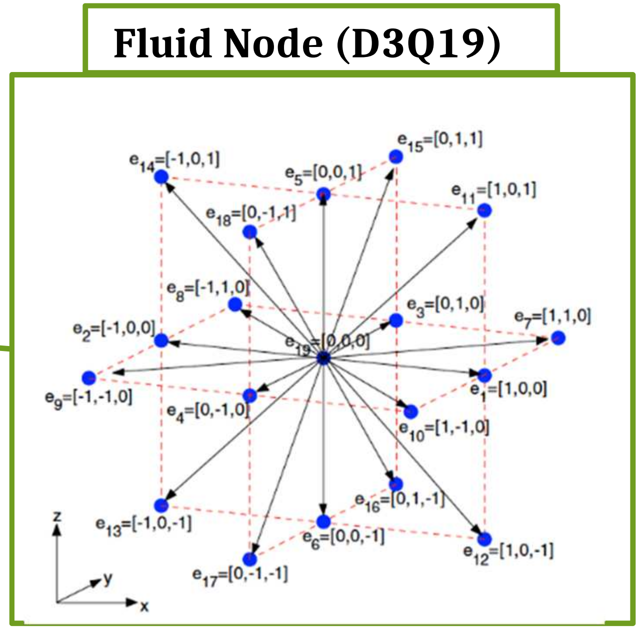
General workflow of the model IIBioS

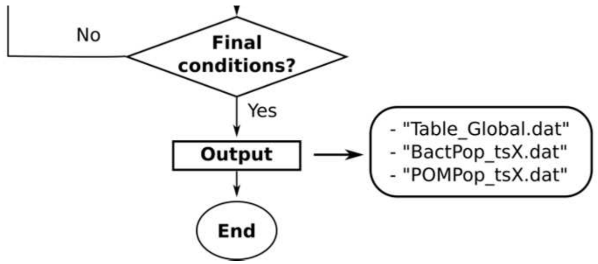


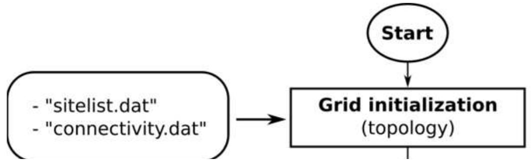
Initialization of the DOC = population Boltzmann



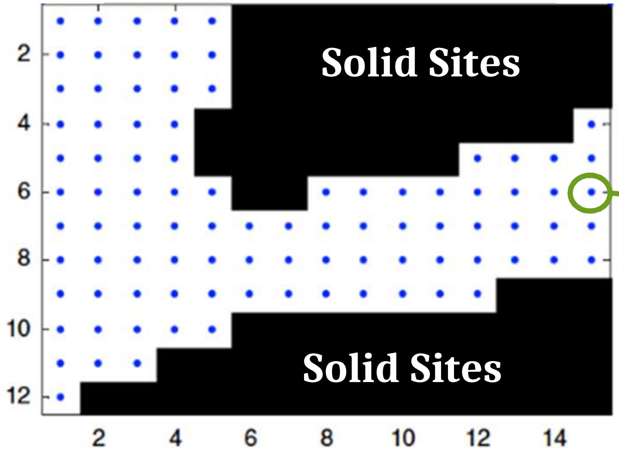
From: www.egr.msu.edu/ ~kutay/Lbsite/



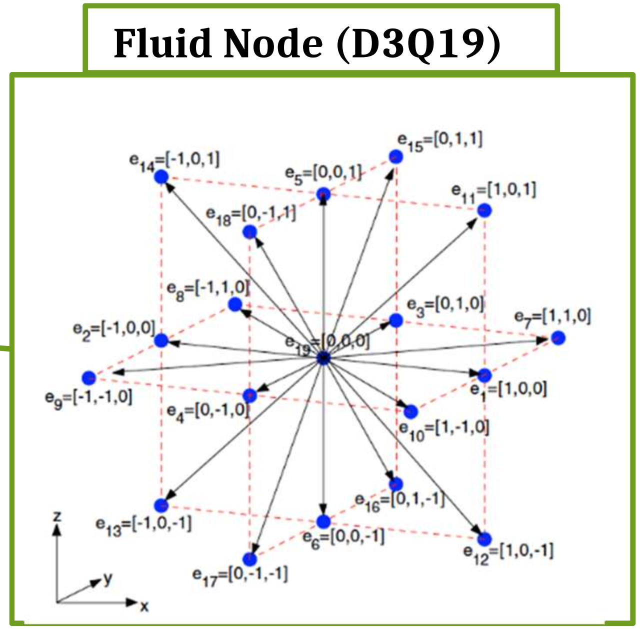




Initialization of the DOC = population Boltzmann

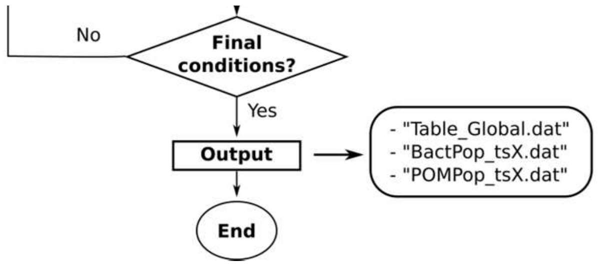


From: www.egr.msu.edu/ ~kutay/Lbsite/

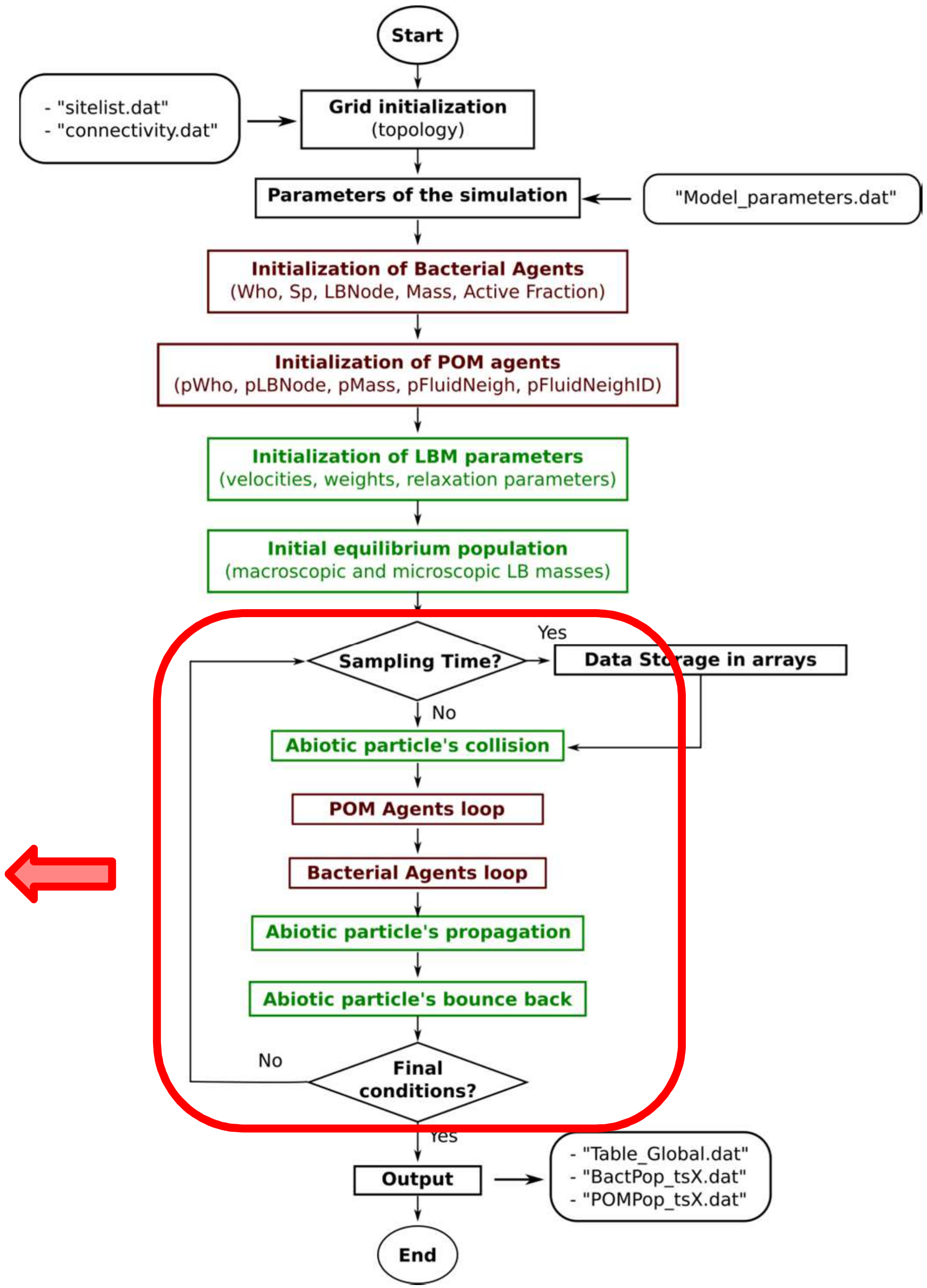



```
double **f = (double **) calloc(Q, sizeof(double*));
for (q = 0; q < Q; q++) {
    f[q] = (double *) calloc(nsite, sizeof (double));
}

double *rho = (double *) calloc(nsite, sizeof(double));
```

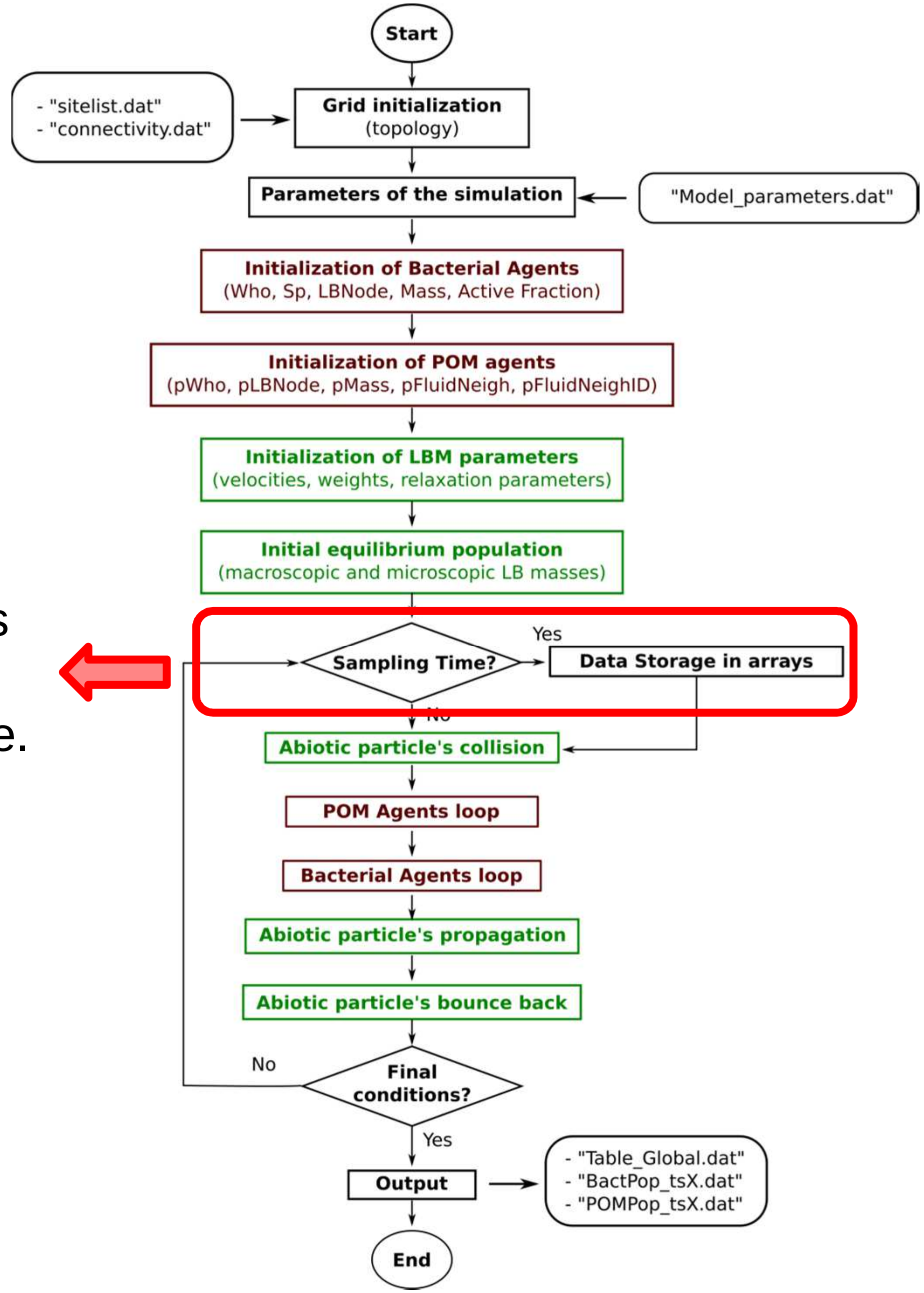


Time step (TS) or main loop

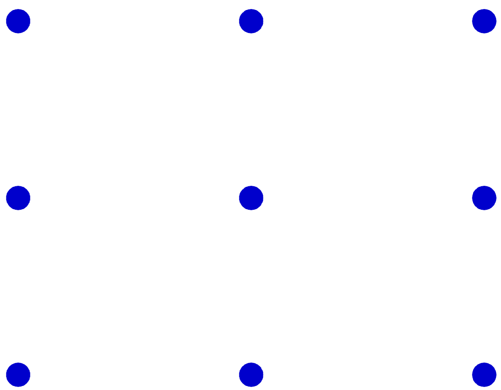


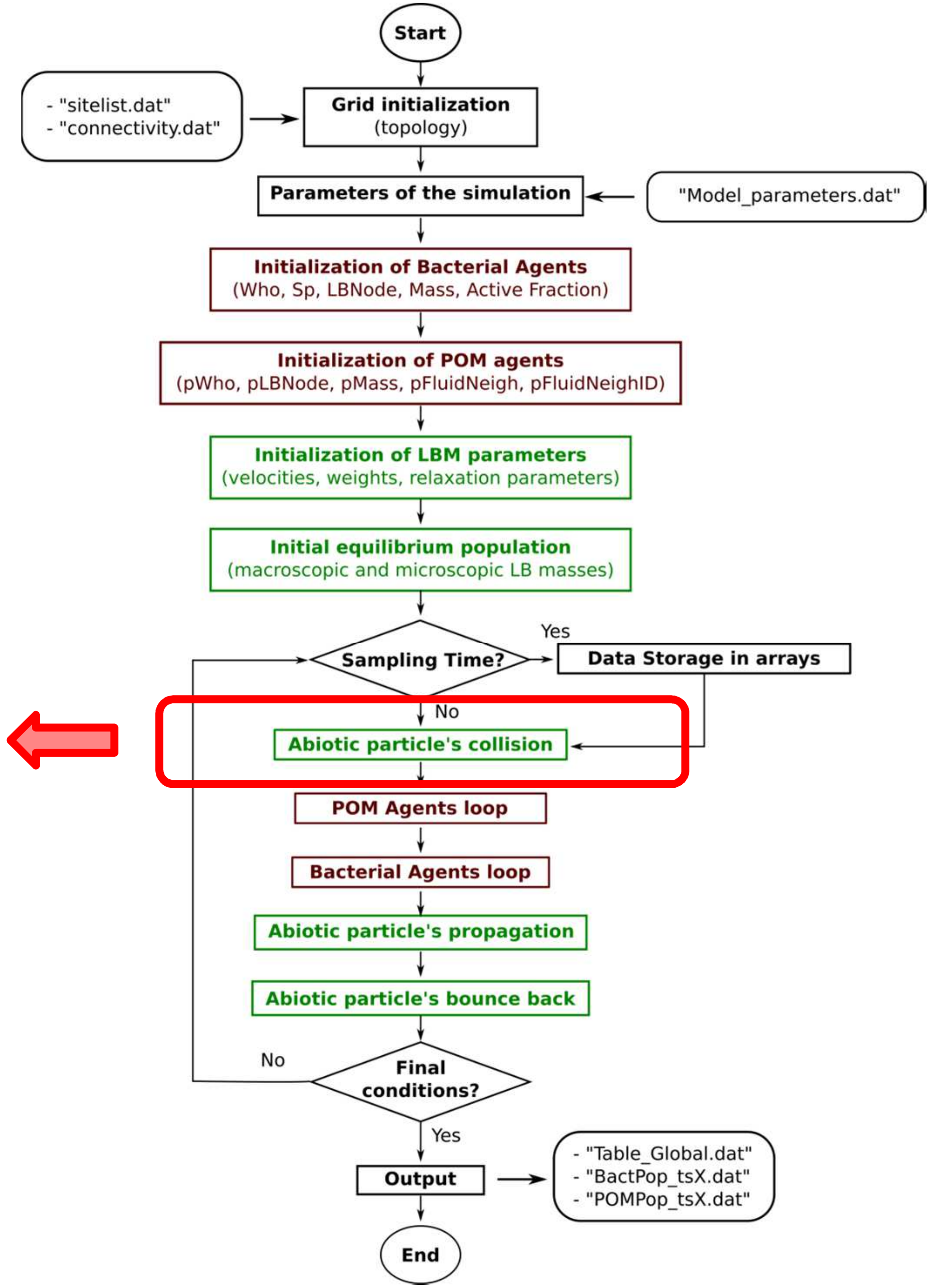
General workflow of the model IlBioS

The State of the bacterial and POM agents and the global out-fluxes are stored in appropriate data arrays every sampling time.



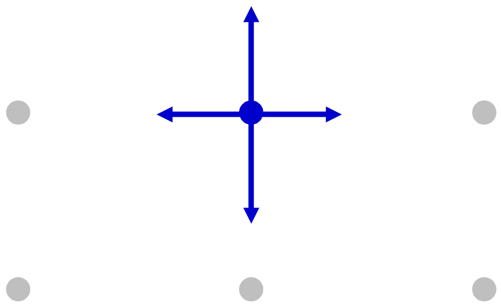
General workflow of the model IlBioS



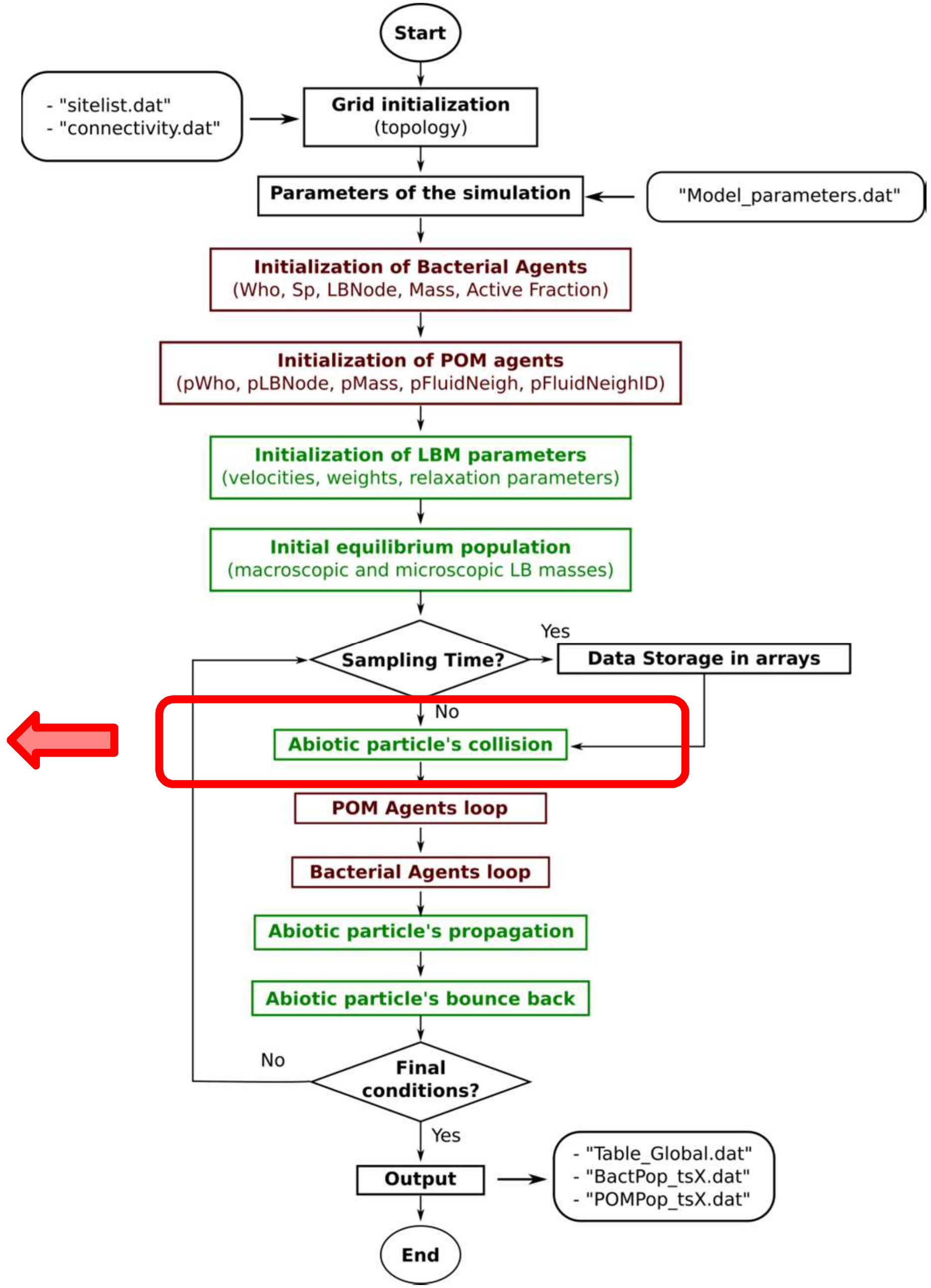


General workflow of the model IlBioS

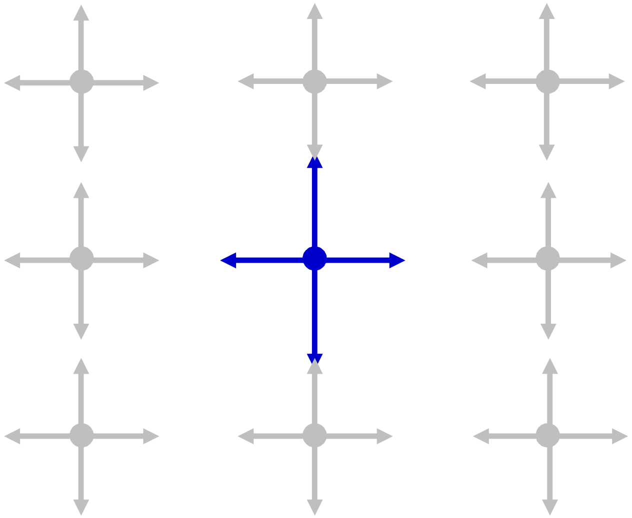
D2Q5

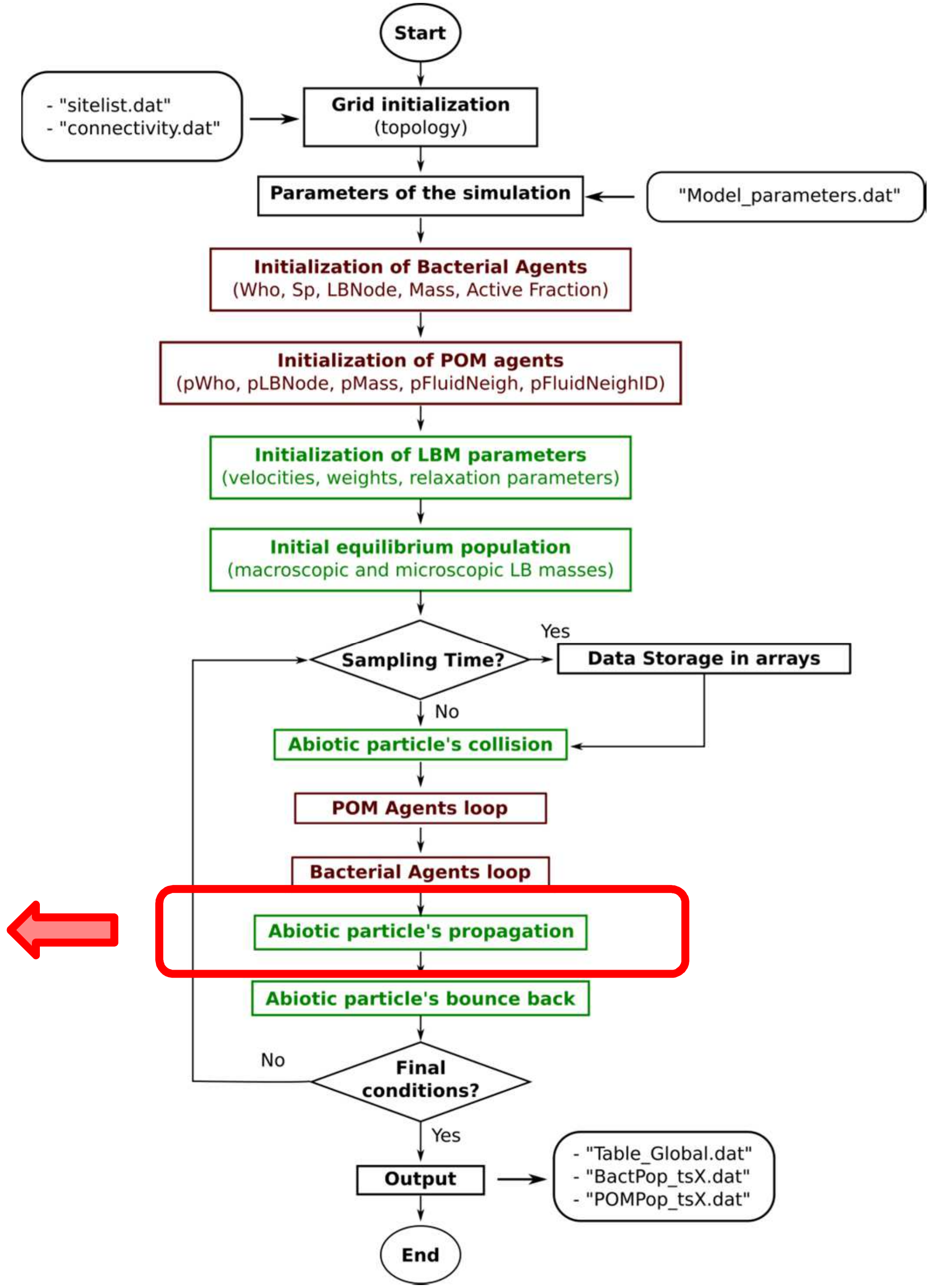


Mass of DOC particles moving
in the unitary directions



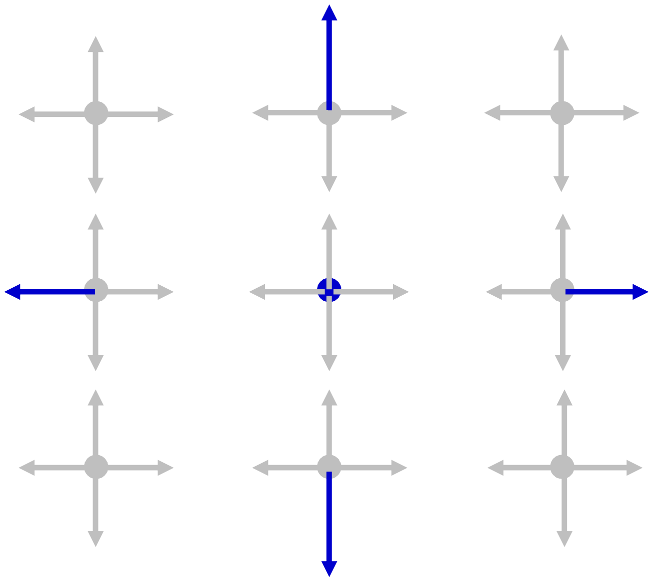
General workflow of the model IlBioS



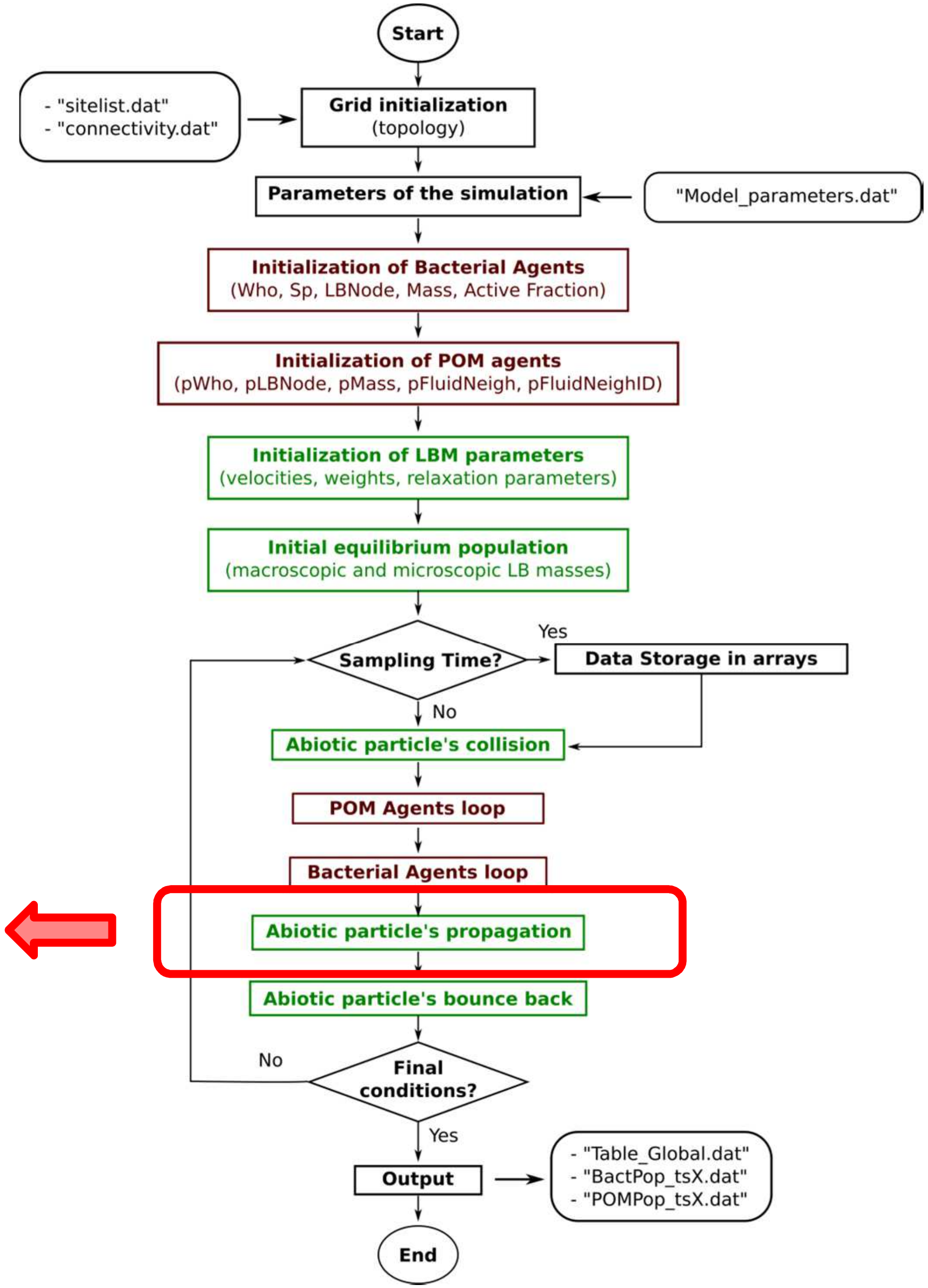


General workflow of the model IlBioS

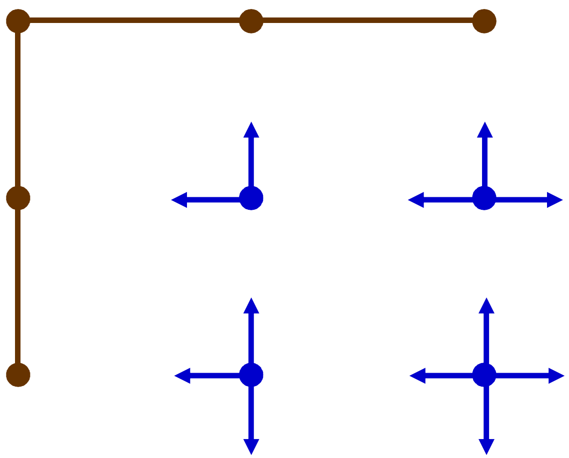
General workflow of the model IlBioS

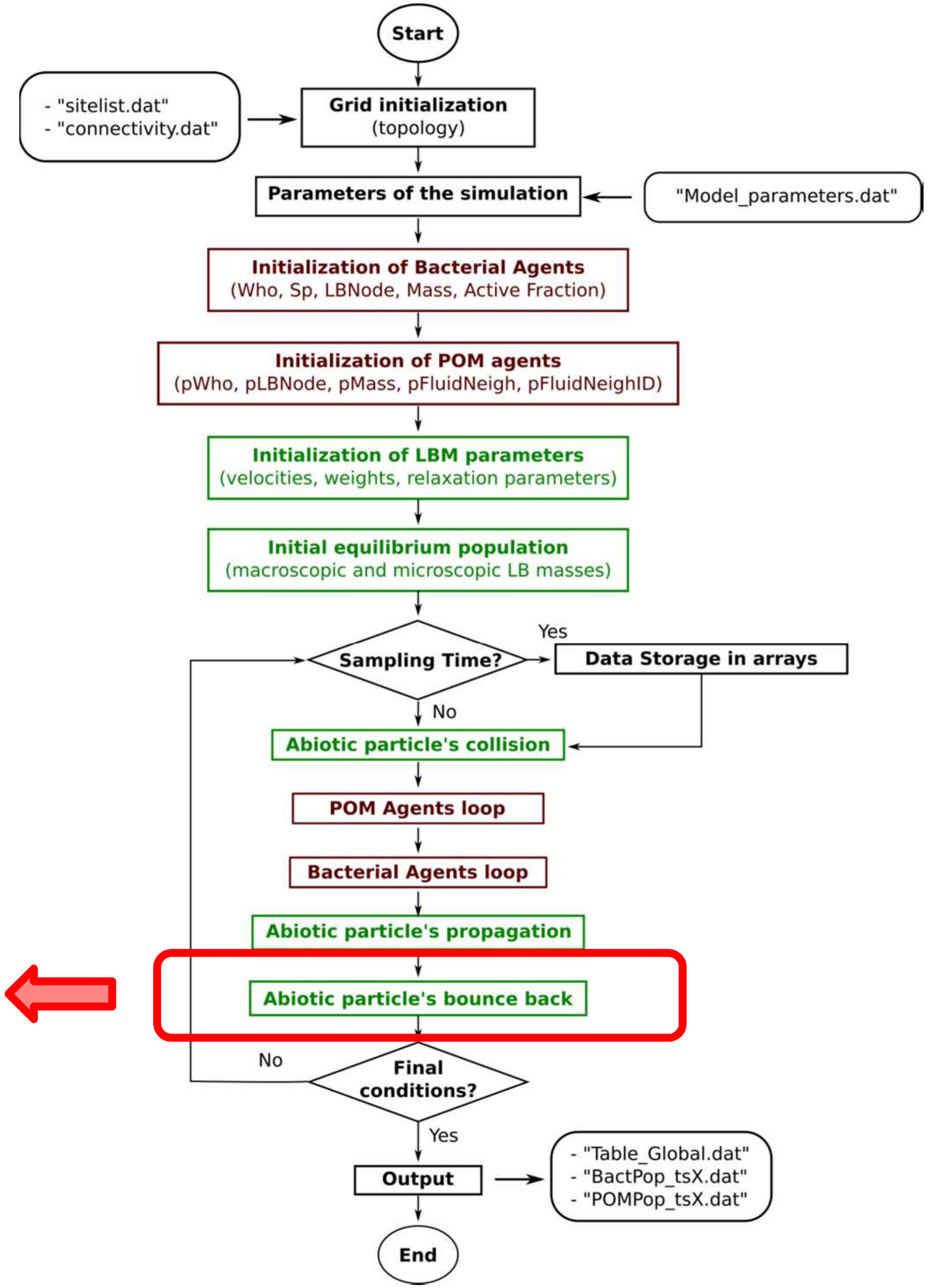


The mass of DOC particles is propagated to the neighbors



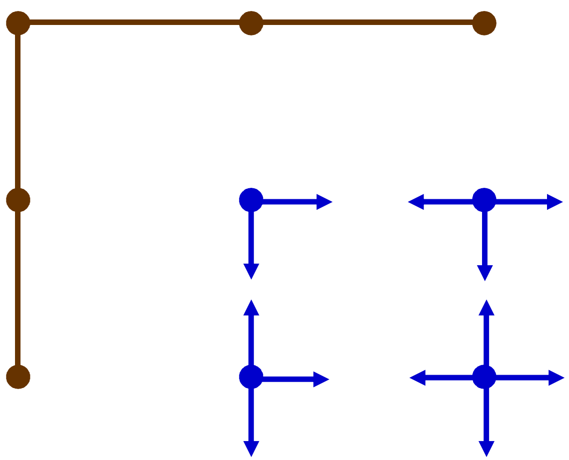
Solid nodes



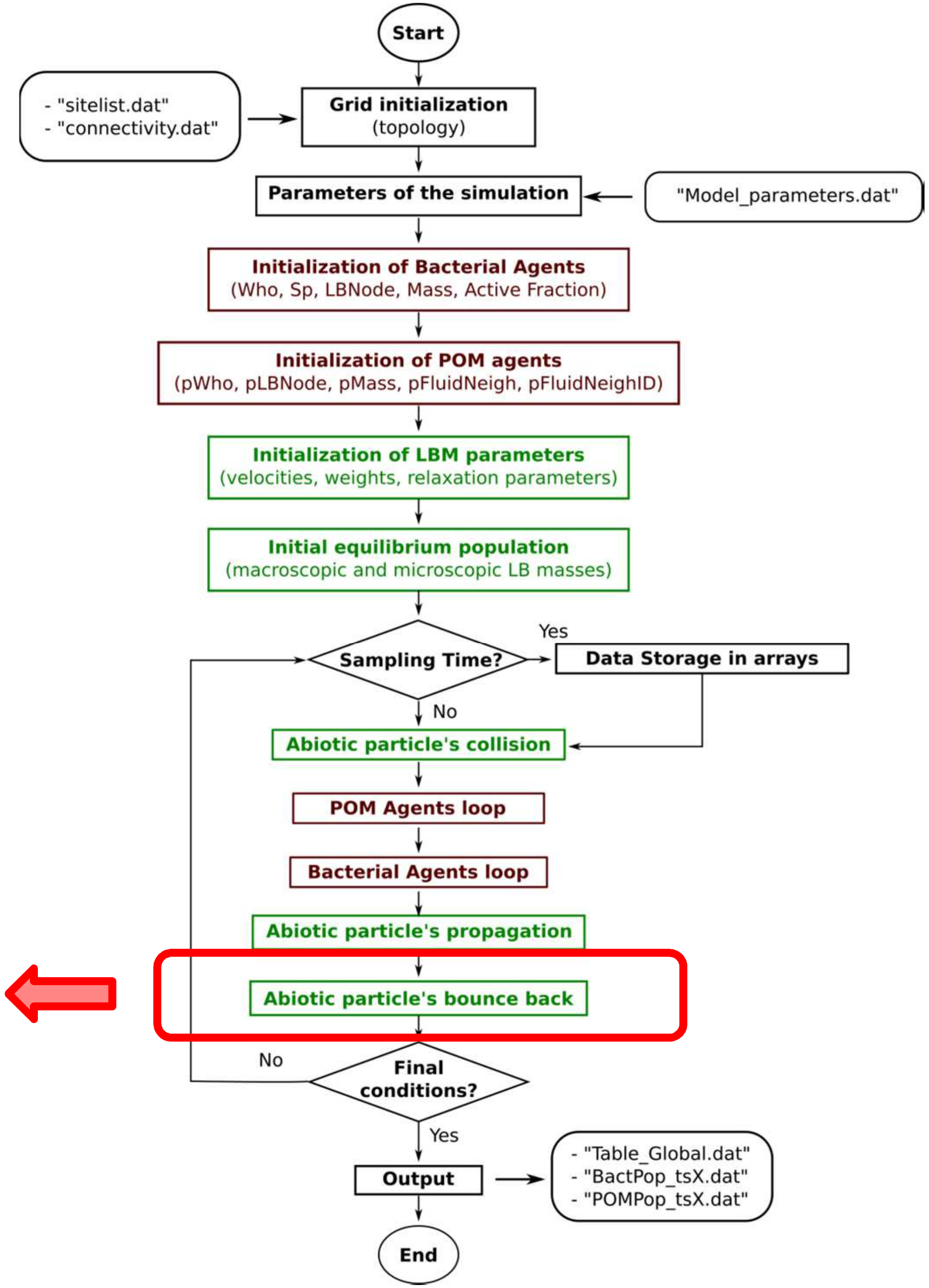


General workflow of the model IlBioS

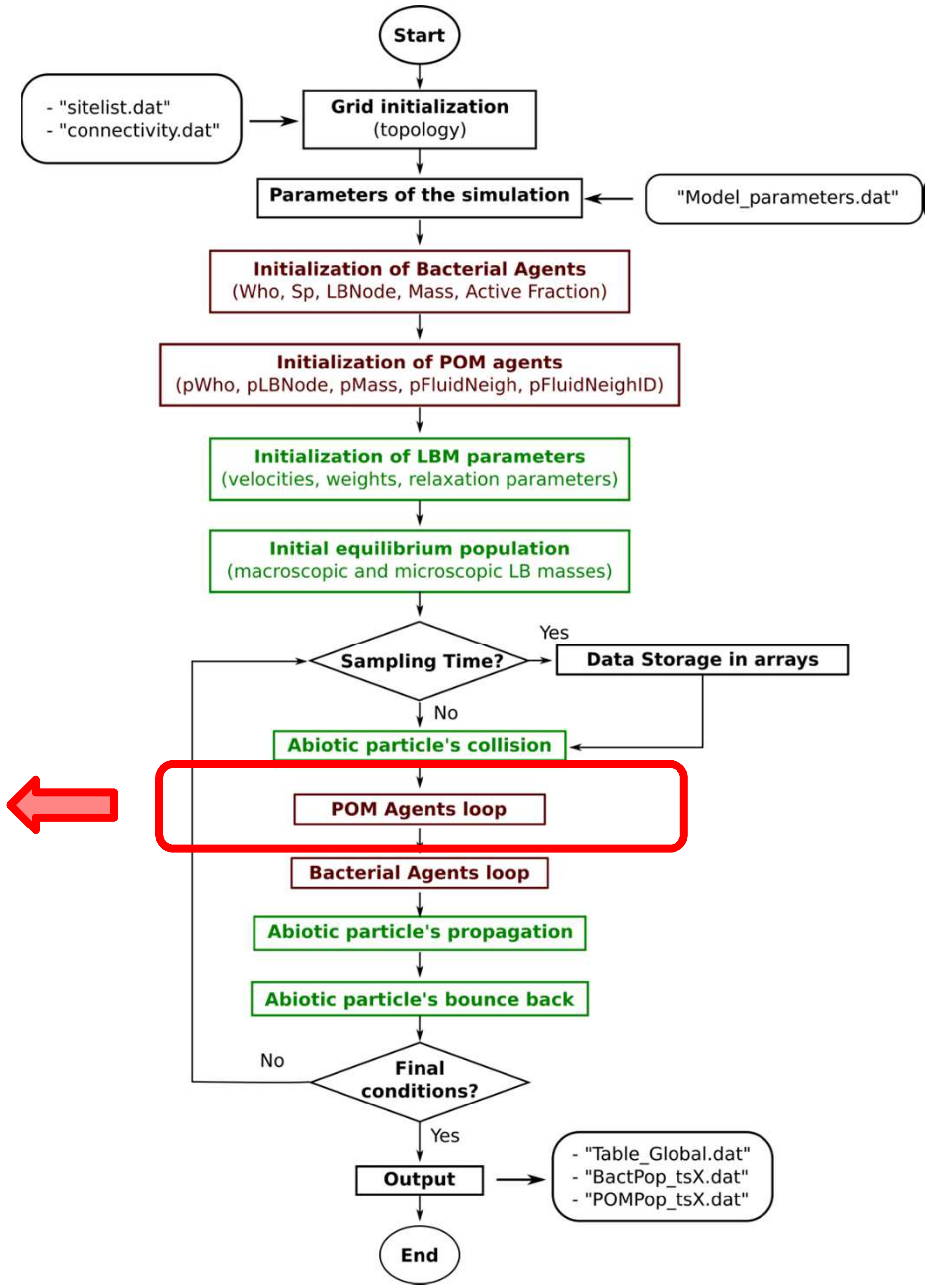
Solid nodes



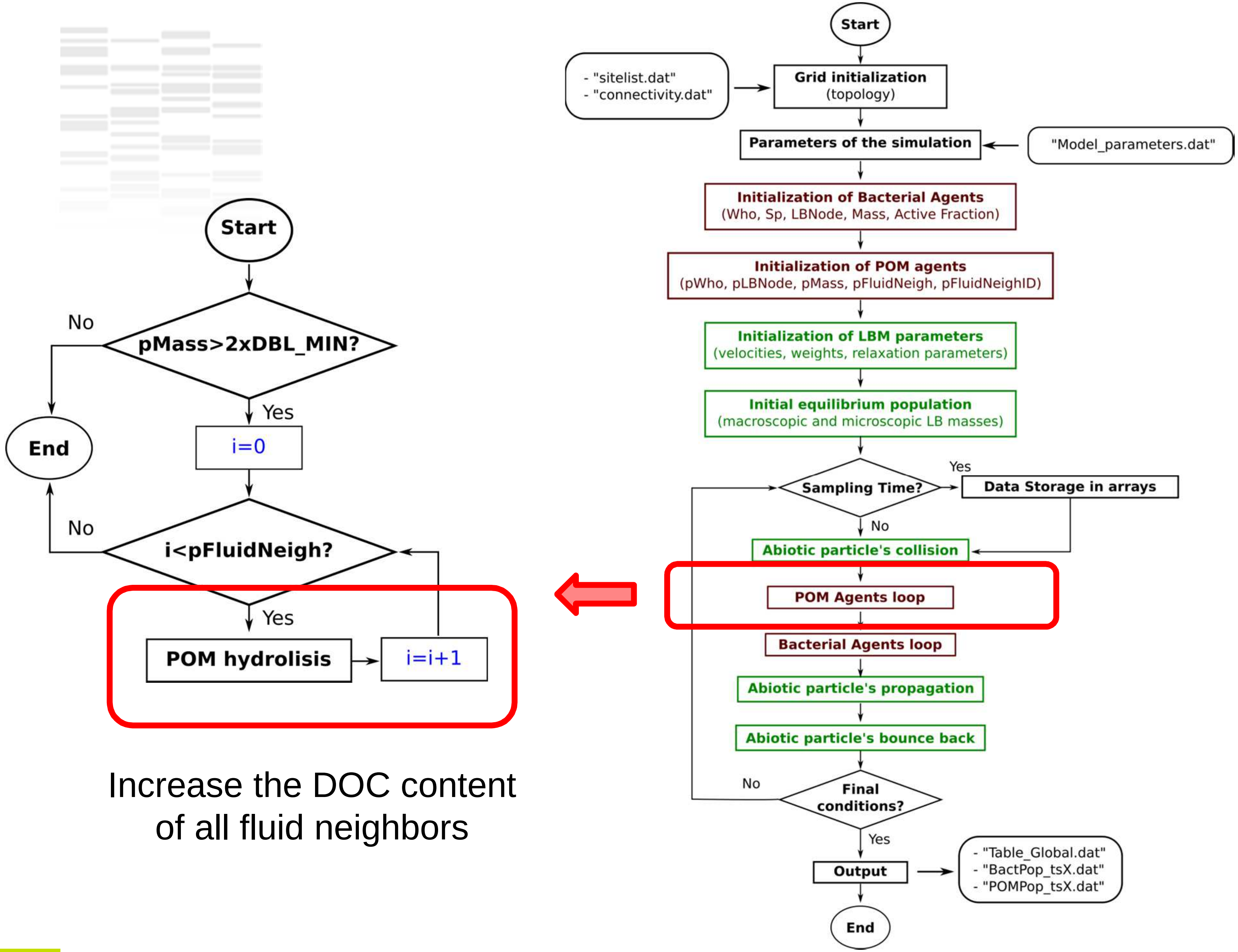
The DOC particles bounce back on the solid nodes directions



General workflow of the model IlBioS

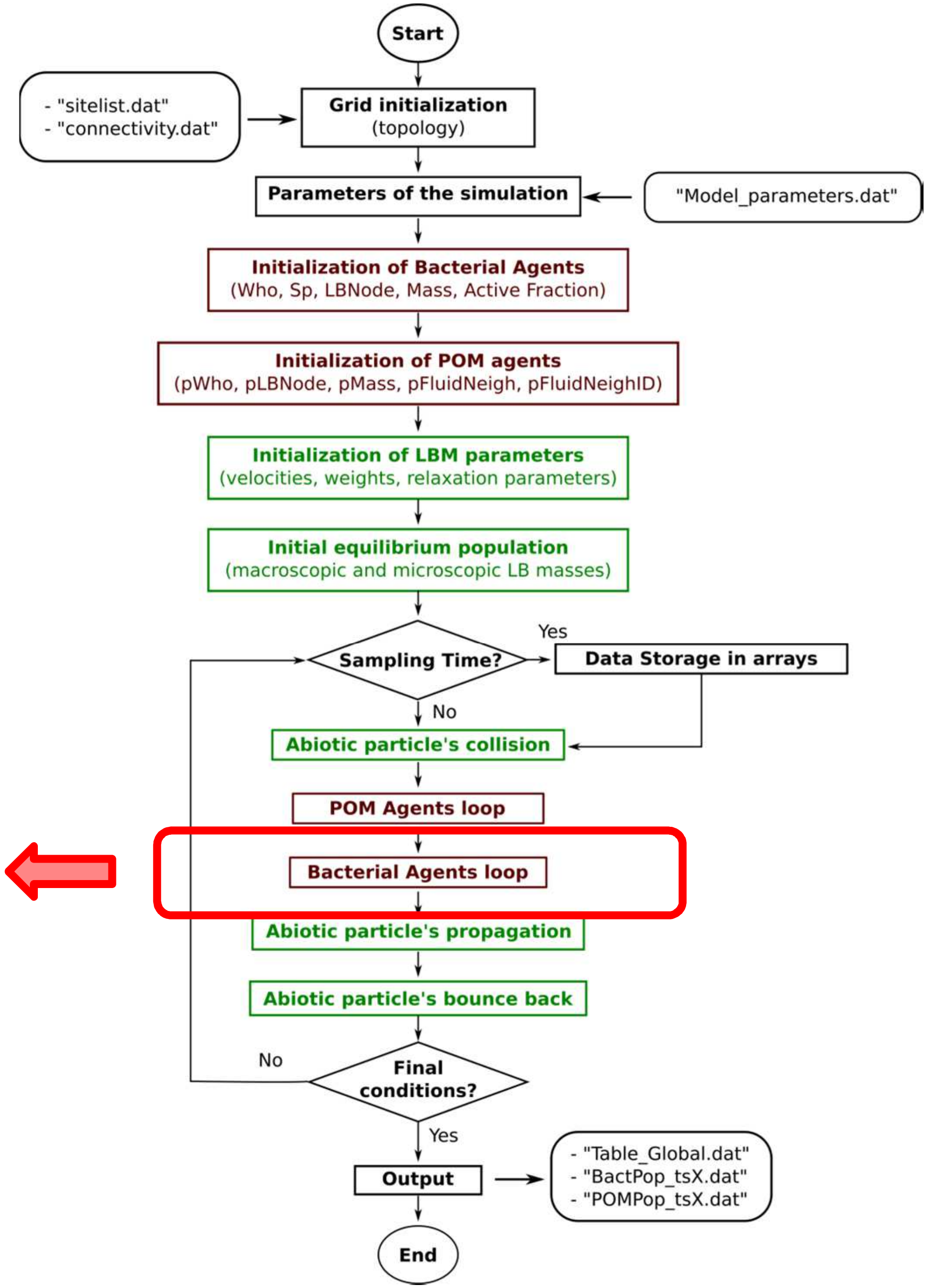


General workflow of the model IlBioS



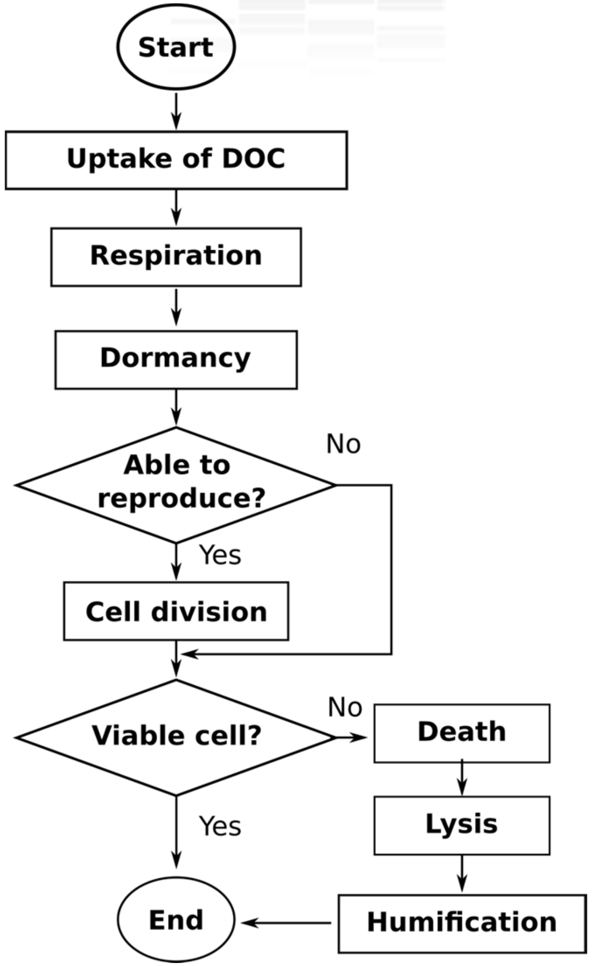
Increase the DOC content
of all fluid neighbors

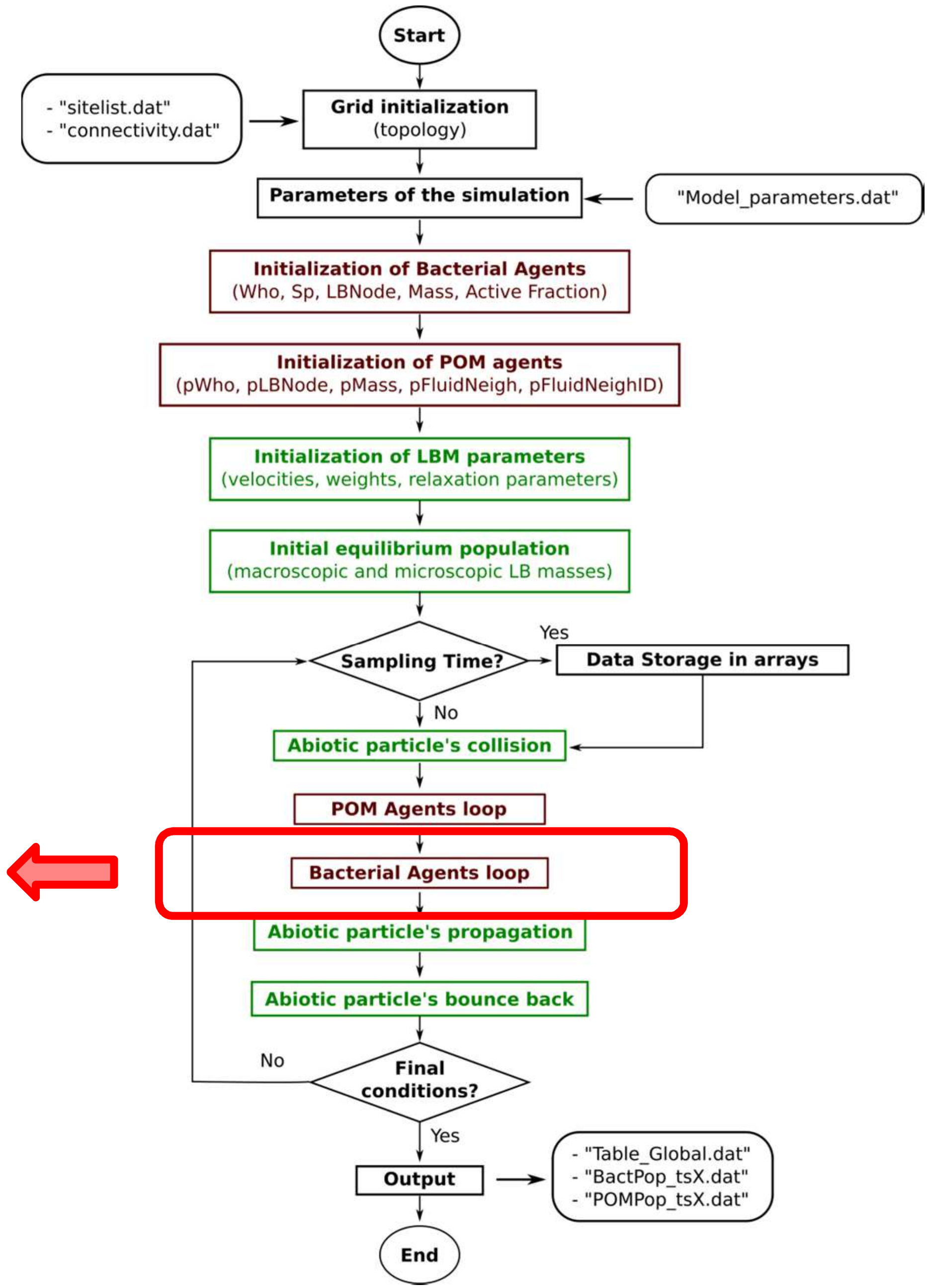
General workflow of the model IlBioS



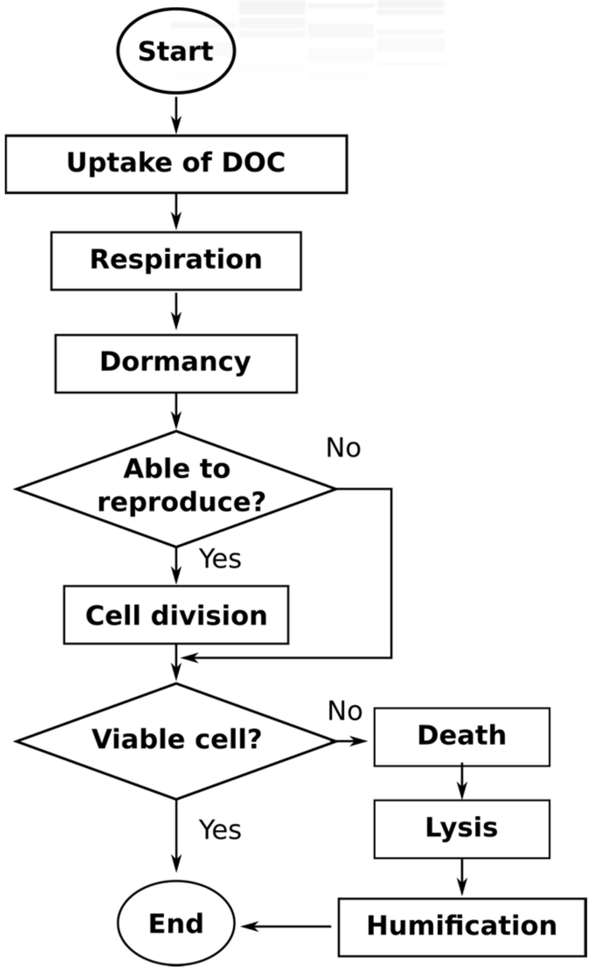
The acting order is randomly changed every time step to avoid privileging bacteria.

General workflow of the model IlBioS

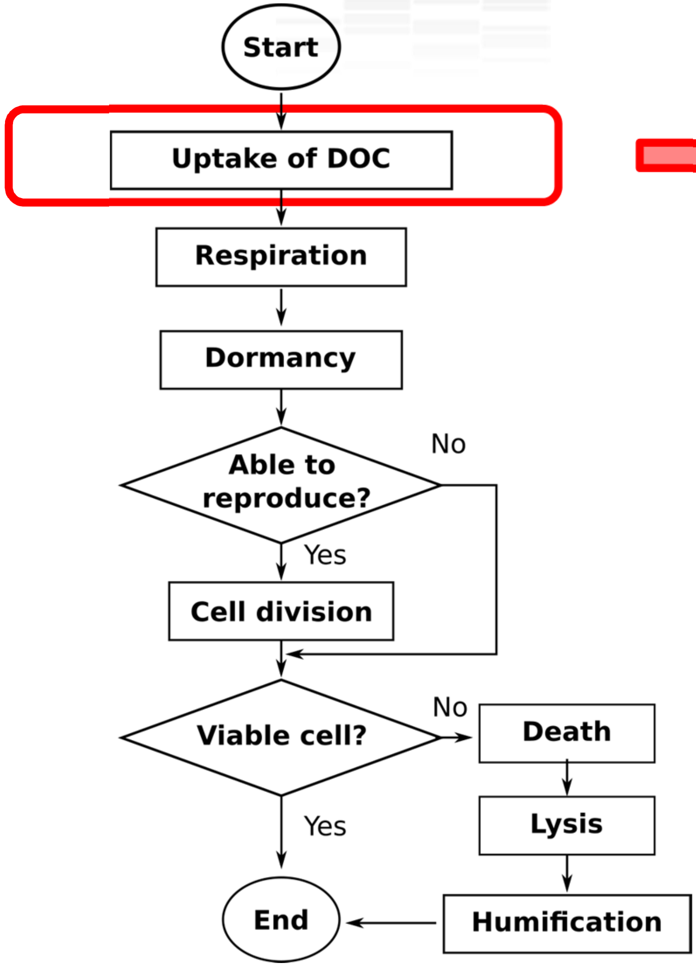




General workflow of the model IlBioS

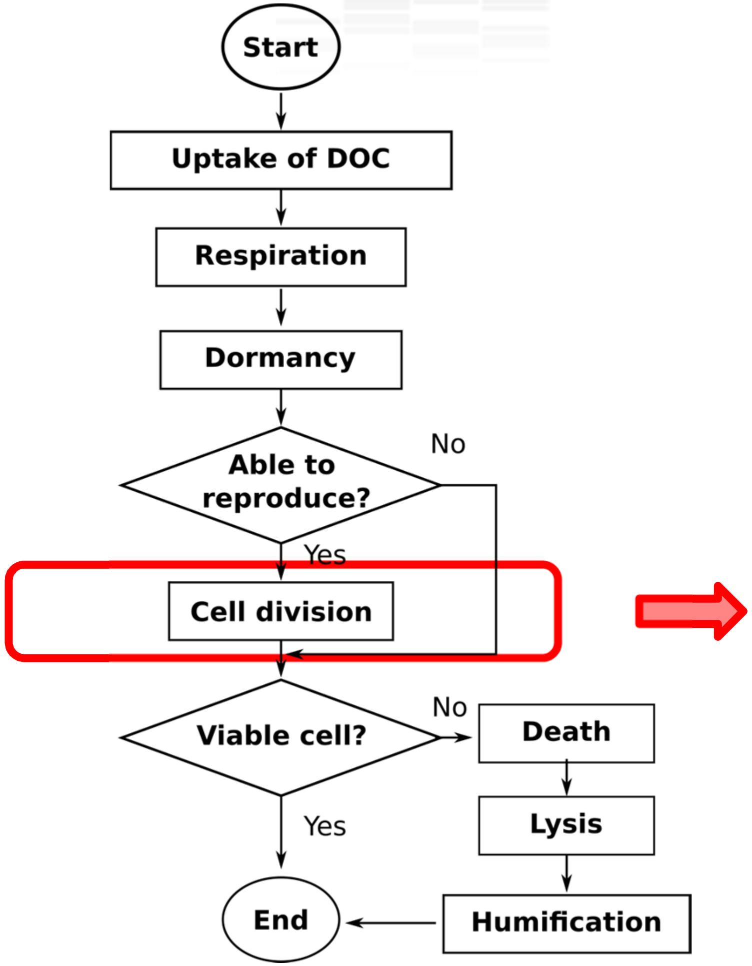


General workflow of the model IlBioS



Decreases the DOC content of the LBNode

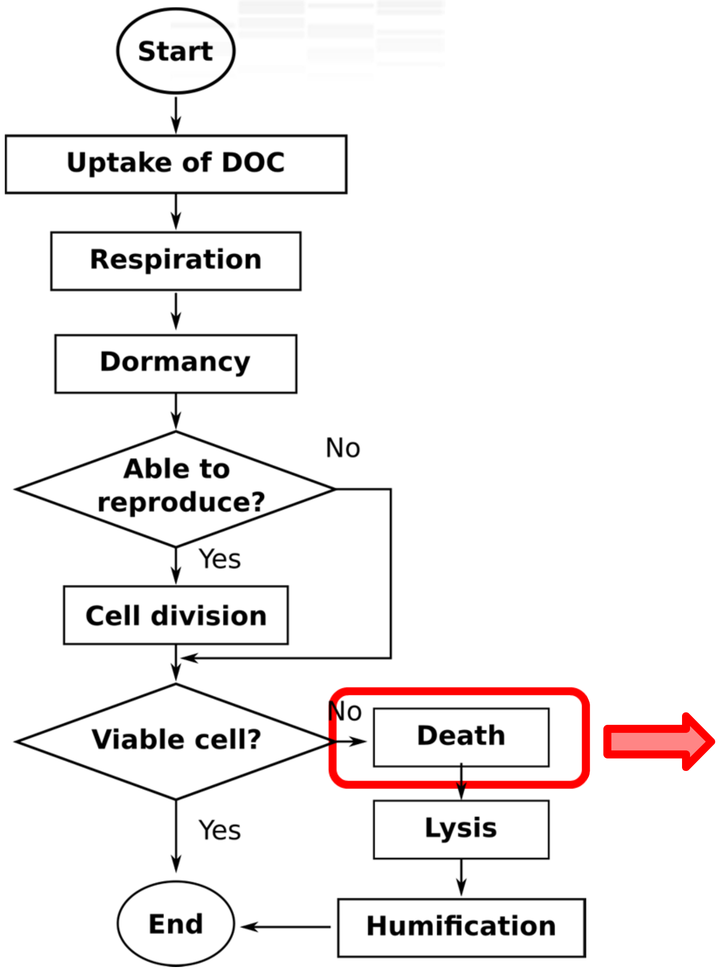
General workflow of the model IlBioS



Is the struct vector completely saturated?

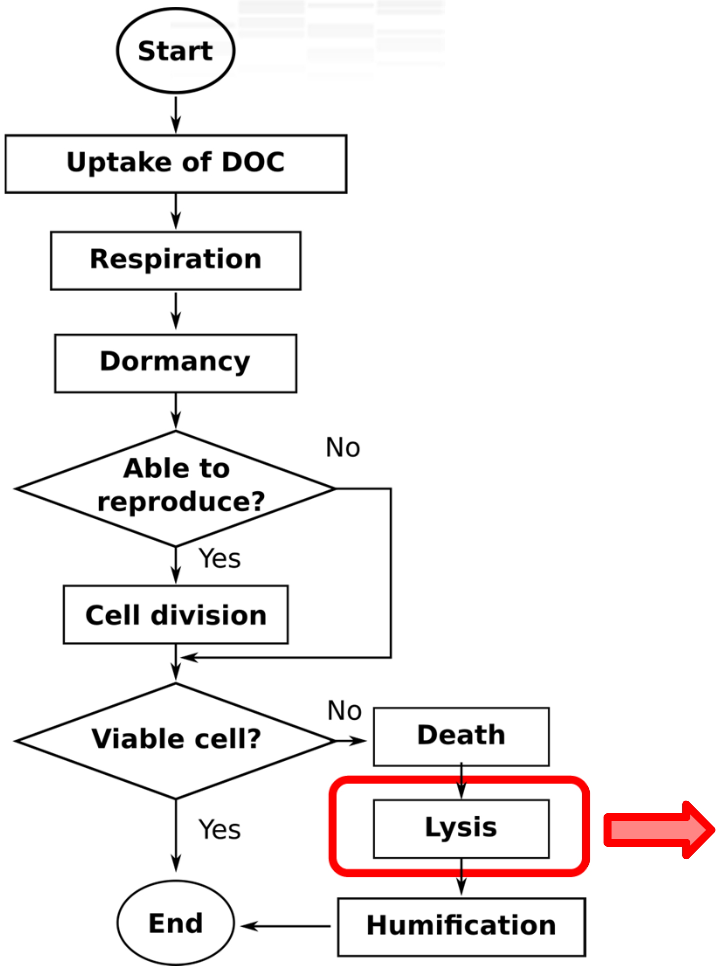
Which one is the first empty position of the vector of struct?

General workflow of the model IlBioS



New empty position in struct vector

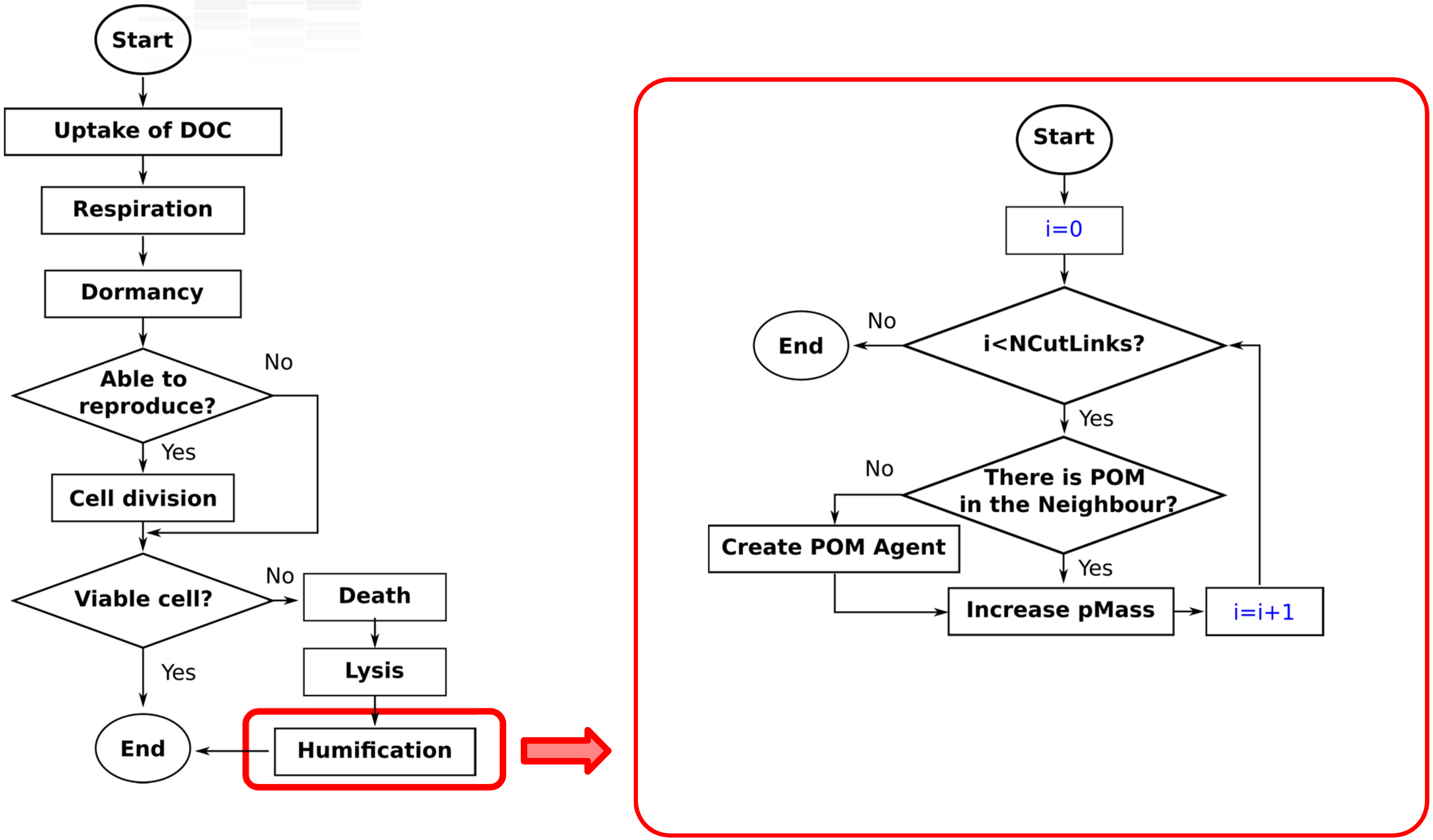
General workflow of the model IlBioS



Increases the local DOC concentration

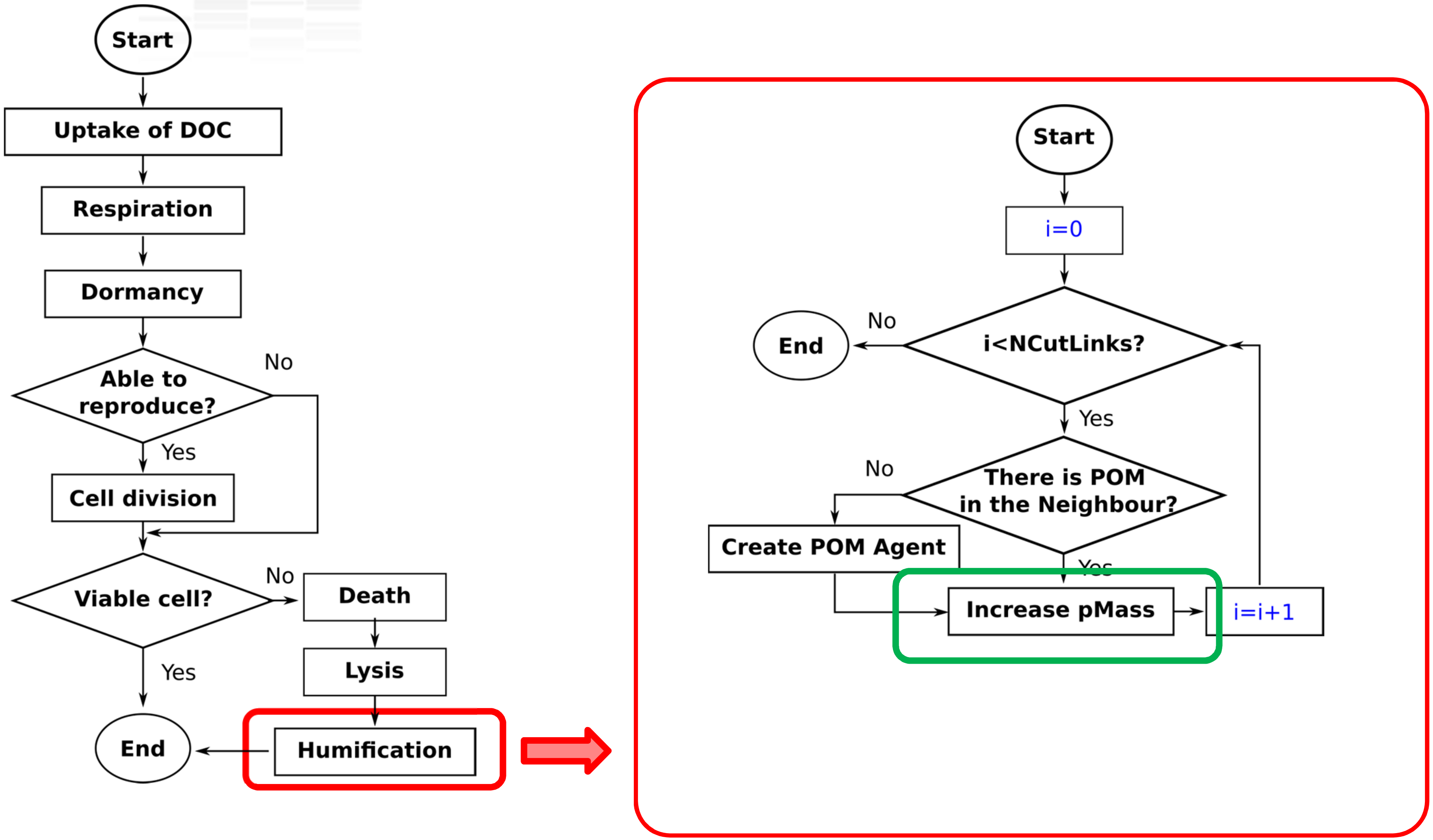
General workflow of the model IlBioS

Every TS, all Bacterial agent perform:

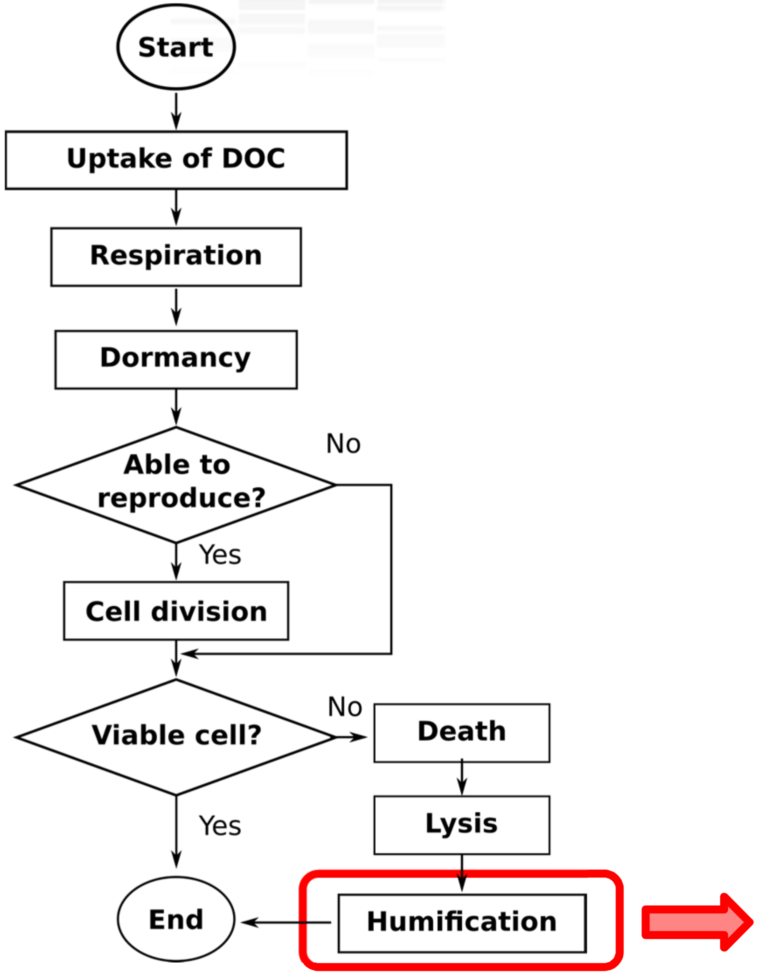


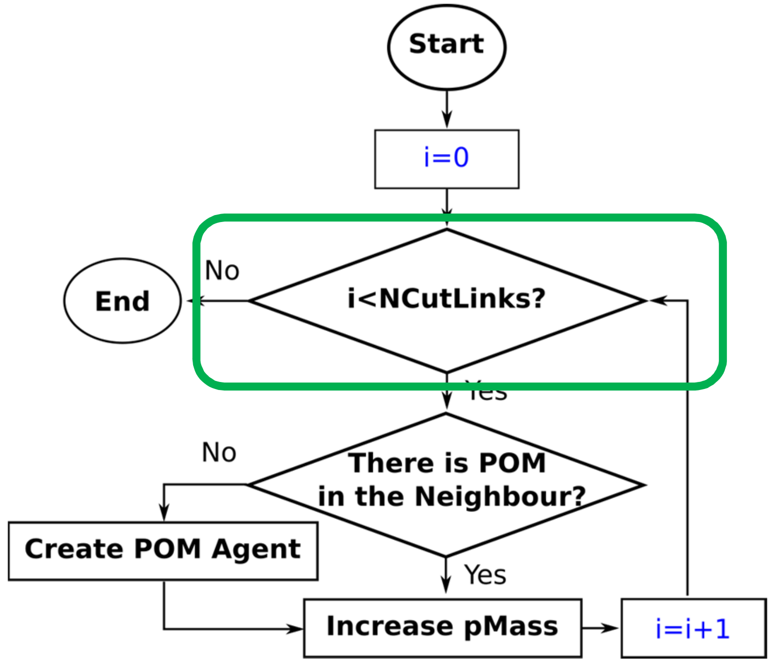
General workflow of the model IlBioS

Every TS, all Bacterial agent perform:



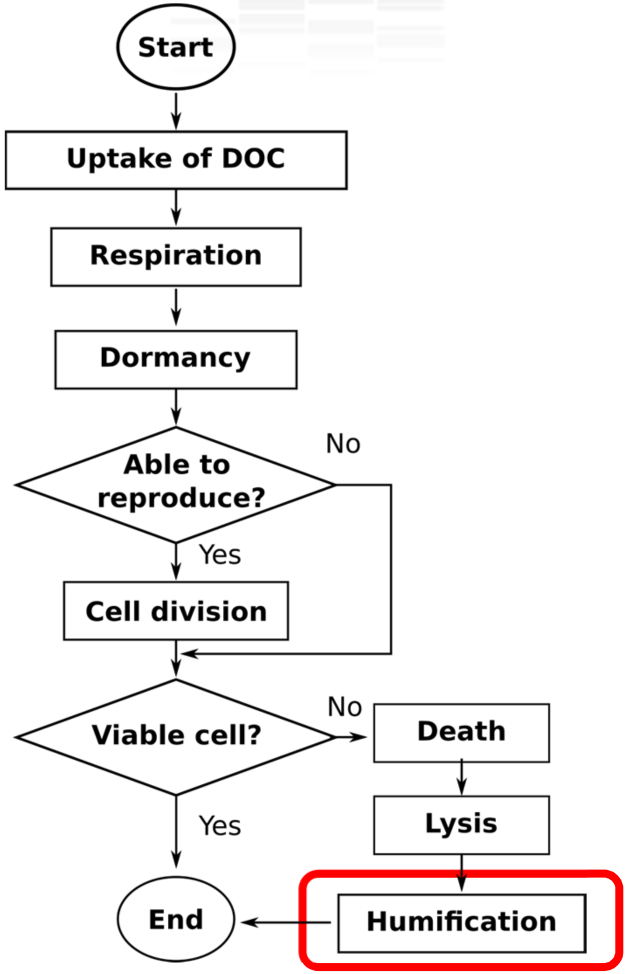
Every TS, all Bacterial agent perform:





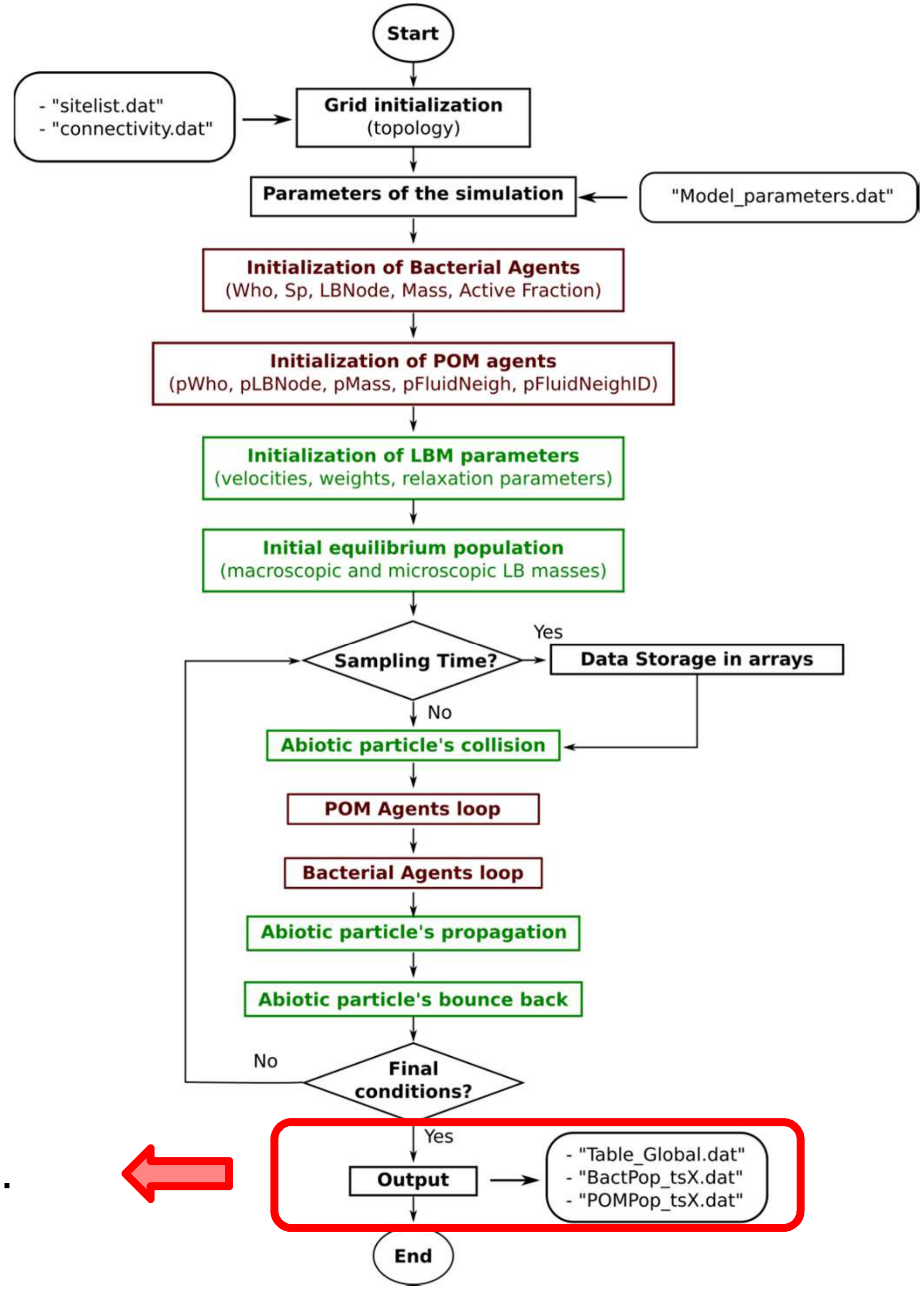
Affects the solid neighbors of the LBNode

Every TS, all Bacterial agent perform:



General workflow of the model IlBioS

Writing the sampled output in files.



General workflow of the model IlBioS

SUMMARY

1. Introduction and background
2. The ANR project Soil µ3D: towards more accurate CO_2 and N_2O gas emissions predictions
3. The model IlBioS: Coupling a lattice-Boltzmann approach to a biological individual-based model
4. Project HPC-IlBioS: understanding the best HPC strategy for the IlBioS approach

_04

Project HPC-IIBioS: understanding the best HPC strategy for the IIBioS approach

Computing time requirements: antecedents

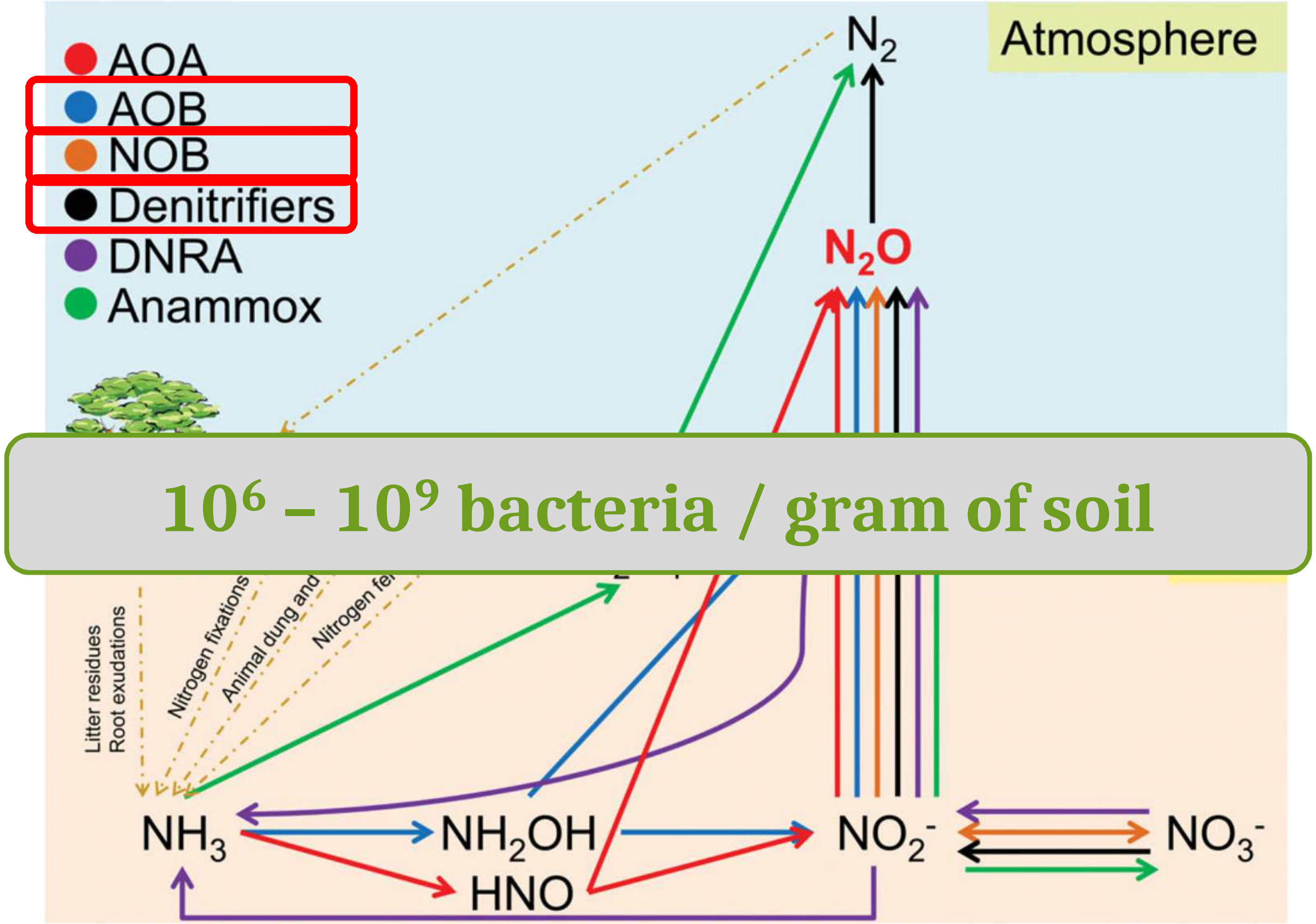
Computing time requirements of IBioS

- 3D CT image of 200^3 nodes.
- 10% of porosity.
- 10 bacterial nodes.
- 5 to 15 POM nodes.
- DOC as a single soluble lattice-Boltzmann substrate.
- Single desktop computer.

25 hours of processing time.

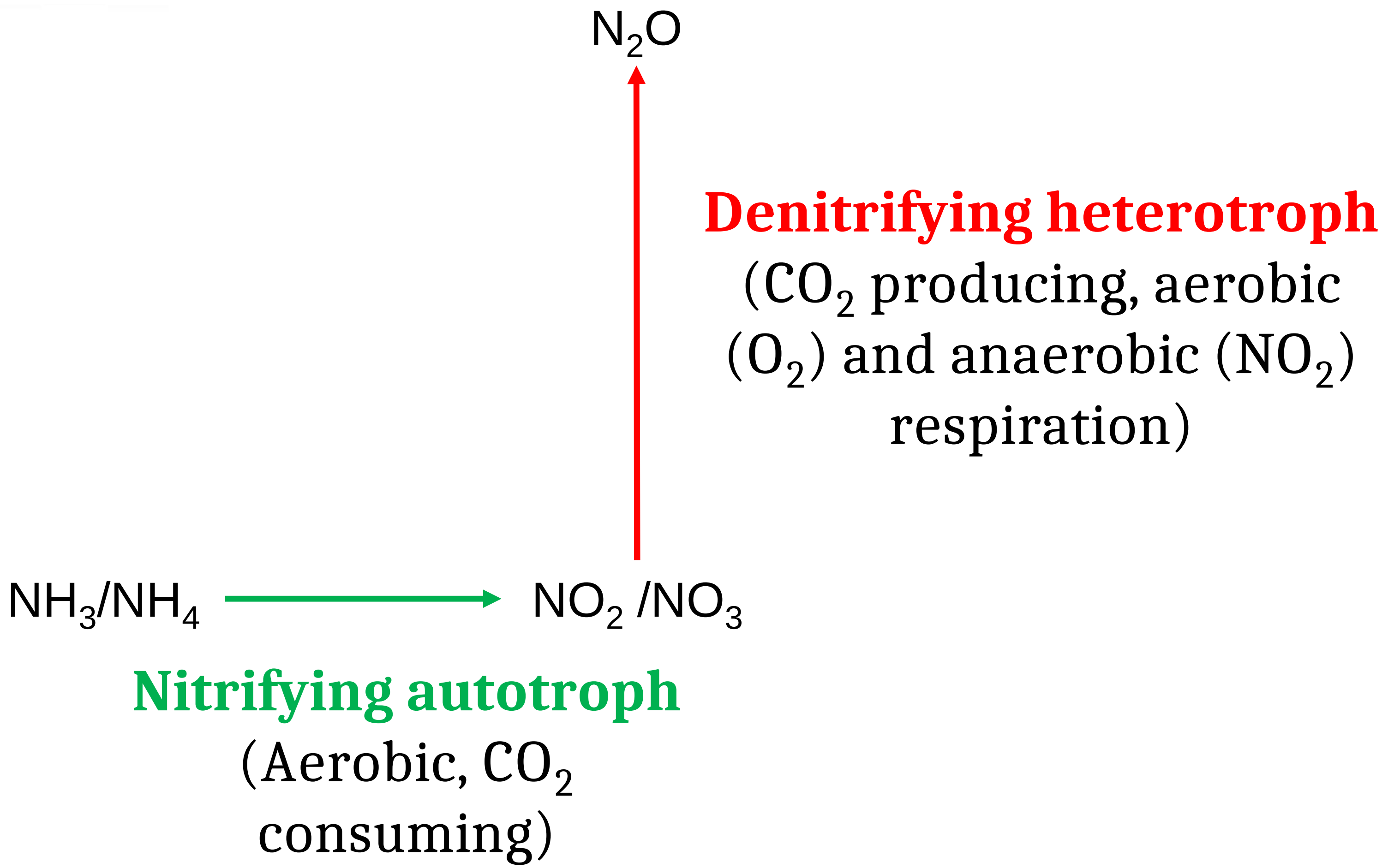
Model complexity of upcoming IlBioS models

Microbial regulation of the terrestrial Nitrous oxide formation (Hu et al., 2015)



10^6 – 10^9 bacteria / gram of soil

Initial (minimal) conceptualization of a multispecies IbM to reproduce CO_2 and N_2O flow in natural samples



Initial (minimal) conceptualization of a multispecies IbM to reproduce CO_2 and N_2O flow in natural samples

Required lattice-Boltzmann substrates in the minimal system:

CO_2, O_2, DOC, NH_4, NO_2/ NO_3, N_2O

Future computing time requirements

Computing time requirements of IlBioS

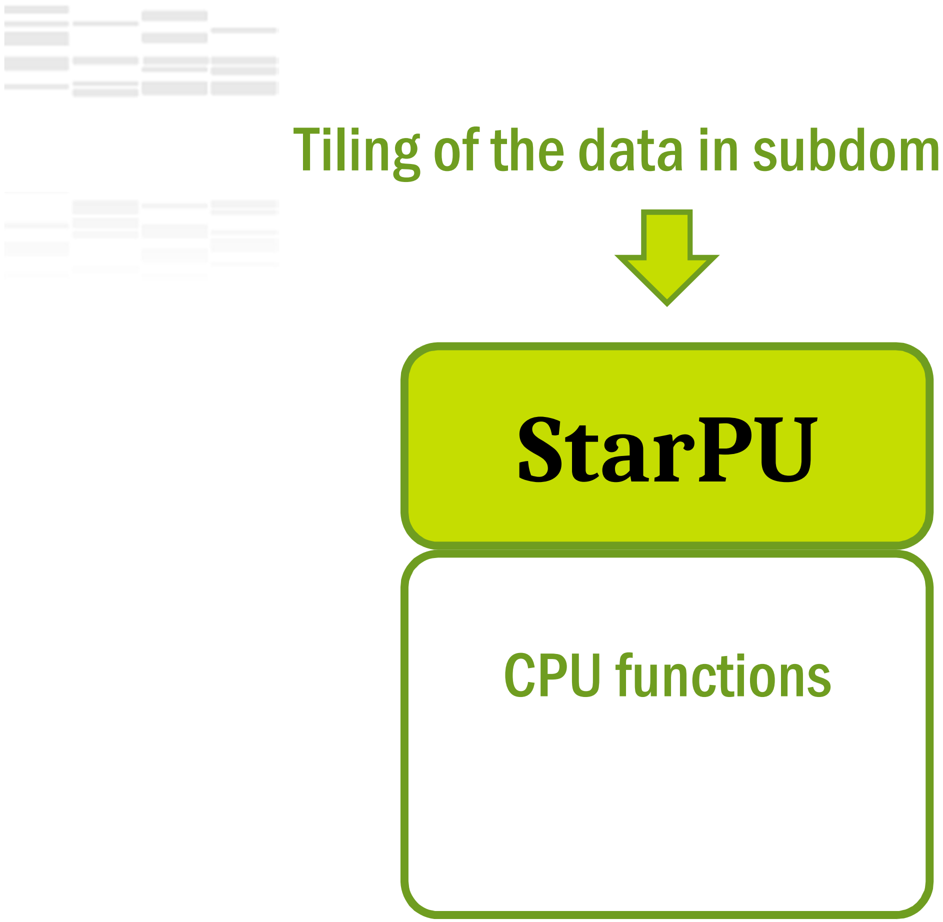
- 3D CT image of 200^3 nodes.
- 10 % of porosity.
- 1 000 000 bacterial nodes
- 5 to 25 POM nodes.
- O2, DOC, NH4, NO2/ NO3

125-150 hours of processing time.

In any case, a non practical amount of time to perform simulation experiments

Parallelisation strategy of HPC-IlBioS

Parallelisation strategy of HPC-IlBioS


Tiling of the data in subdomains
StarPU
CPU functions

StarPU

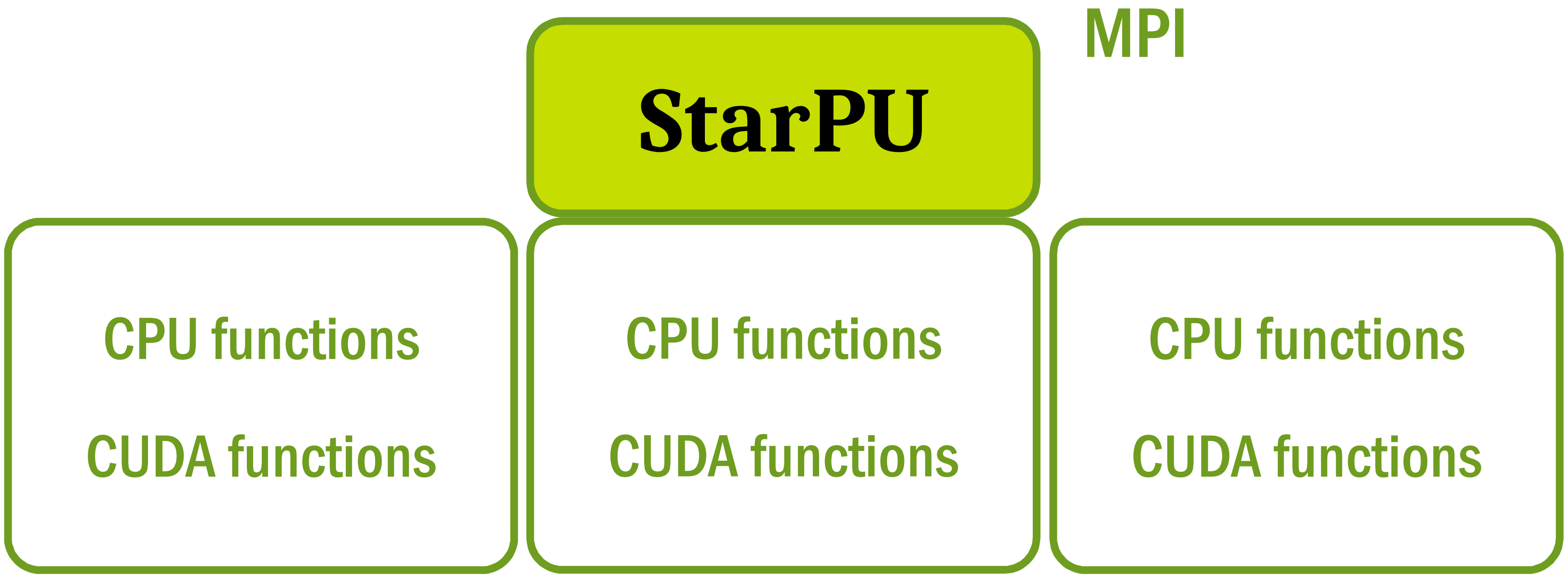
CPU functions

CUDA functions

100 x

Speed up for lattice-Boltzmann models using GPU

(Banari et al., 2014)



Future need of increasing the resolution of the 3D CT images

Thank you very much for your attention.

Any question?

Anyhow, a lot to be done yet …

IlBioS for an in depth study of the soil!!!