

CEMRACS 2016: project HPC-ILBioS

**Understanding the best HPC strategy to couple a lattice-Boltzmann fluid dynamics model and an individual-based model of soil bacteria**

**Project leader:**

Dr Xavier Portell-Canal ([xavier.portell-canal@grignon.inra.fr](mailto:xavier.portell-canal@grignon.inra.fr))

**Participants involved in the project:**

Dr Xavier Portell-Canal

Dr Valérie Pot

Dr Philippe Baveye

Dr Patricia Garnier

**Table of contents:**

<i>Abstract:</i> .....	1
1. Introduction – State of the Art.....	1
2. Description of the project.....	3
3. Objectives of the project .....	3
4. Expected participants and presence at CIRM .....	4
5. Status of the funding.....	4
References.....	4

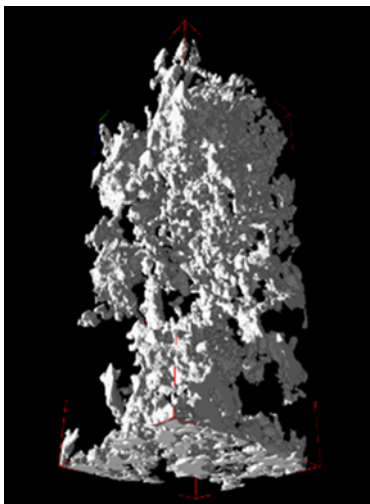
### Abstract:

The model ILBioS combines a lattice-Boltzmann model and an individual-based model of the soil bacteria, and it holds the potential to allow us to gain a more mechanistic knowledge of the processes taking place in the soil. Due to the particularities of these two modeling approaches, HPC techniques involving both CPU and GPU programming are required to allow simulations in a reasonable time.

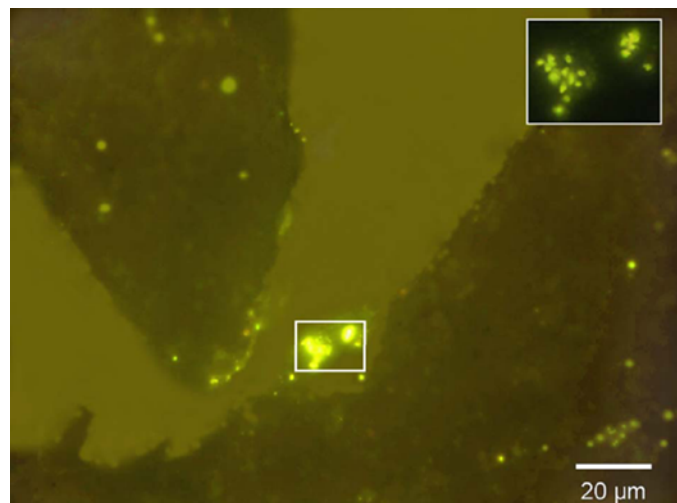
## 1. Introduction – State of the Art

Greenhouse gases emissions, a global concern. There is a significant concern related to the release by agriculture fields to the atmosphere of important amounts of GreenHouse Gases (GHG) like CO<sub>2</sub> and N<sub>2</sub>O. CO<sub>2</sub> is released largely from microbial mineralization of soil organic matter [1] and N<sub>2</sub>O is generated by the microbial transformation of nitrogen in soils, especially under wet conditions [2]. Many models describe the emission of GHG, including CO<sub>2</sub> and N<sub>2</sub>O. These models are poorly predictive because they are ignoring important *microscopic* aspects of soils, in particular their high level of heterogeneity at the microbial habitat and pore scale, caused by soil structure [3] which can lead to a spatial disconnect between soil carbon and nitrogen, energy sources, oxygen, and the organisms that are involved in carbon and nitrogen transformations.

The soil biological and physical heterogeneity. Soils are more biodiverse than any other ecosystem on earth. In the soil matrix, microbial communities exist and are active in a complex 3D physical framework (**Fig. 1**) that can cause a variety of niches, or microhabitats, that are more or less suitable for microbial growth, activity and survival [e.g., 4]. The structure of solid and pore spaces results in a complex distribution of oxygen, water and gradients of solutes spanning distances as small as a few micrometres, and therefore in radically different *local conditions* at very fine scales [e.g., 5]. In this physical framework, microorganisms (**Fig. 2**) tend to be found in microcolonies and occupy only a very small fraction of the total pore volume (a few ‰), and they have a highly heterogeneous spatial distribution and that has been investigated only recently [e.g., 6].



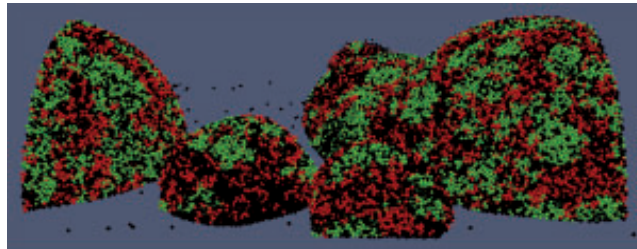
**Figure 1.** X-ray 3D Computed Tomography image of a small fraction of soil. The soil pore space is shown in white.



**Figure 2.** Bacteria in the top soil of a Podsol. Bacteria tend to appear in microcolonies, therefore, due to the soil heterogeneity, even bacteria near in the space may experience very different conditions, and impacting, finally, differently the GHG emissions at higher scales. From Eickhorst & Tippkötter [7].

Lattice-Boltzmann modeling together with the biological agent-based modeling technique: a modeling approach able to deal with the biological and physico-chemical soil complexity.

Innovative modeling tools to study the consequences of the microscopic soil complexity on the GHG emissions have to be developed. These models have to be able to take into account the substrates *mass transport processes* taking place in the soil, their high *microbiological diversity* and the *heterogeneity of the porous matrix*. Fluid dynamics can be accurately simulated using a lattice-Boltzmann model (LBM), which is appropriate to deal with the complex soil pore space geometry. Microorganisms in soils can be modeled in a natural way using the individual-based modeling approach [8], also known as agent-based modeling in other disciplines. An Individual-based Model (IbM, see also **Fig. 3**) simulates individual cell *properties* (mass, individual lag time, age, etc.) and the rules defining the *behavior* of the individual (substrates uptake, growth, reproduction, mobility, birth, death, etc.). Although global phenomena can also be modeled, it is common practice that the behavior of the individual depends –and affects– the media locally. Therefore, the approach is well suited to deal with the soil micro-heterogeneity.



**Figure 3.** Individual based simulation showing cooperative relationship between two bacterial species. In black, death cells. From Kang *et al.* [9].

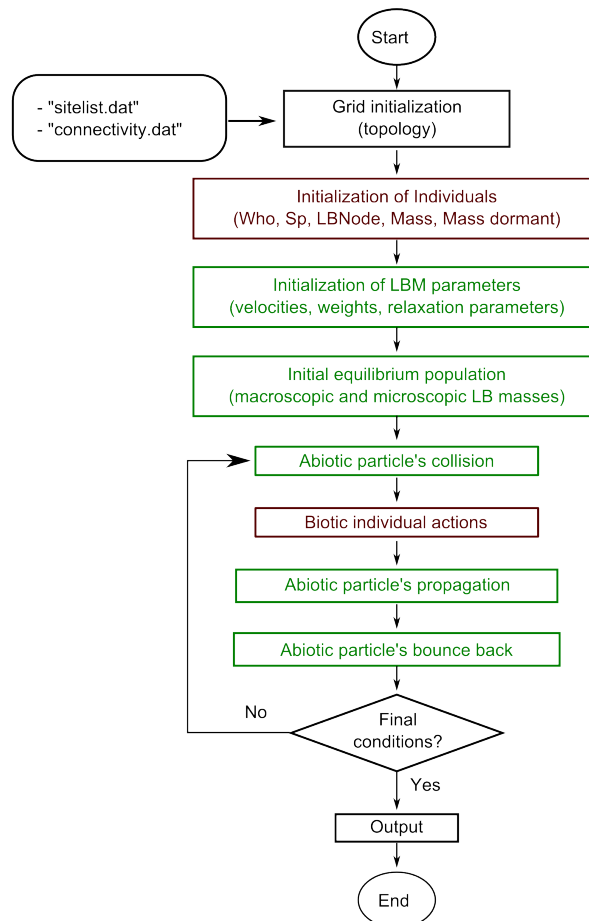
Main computing challenges faced by a LBM-IbM modeling approach for the soil. Although a joined LBM-IbM approach is an invaluable tool to gain a mechanistic understanding of the processes taking place into the systems, their application to study natural soils suffers from some challenges. For the project in hands, the most relevant of them are:

1. *Size of the simulated media.* In order for the simulated outputs to be usable to upscale to larger scales, the model has to simulate the biogeochemical functioning of larger core samples (several cm<sup>3</sup>). To obtain a realistic description of the soil matrix usable by the LBM, 3D computed tomography images (see **Fig. 1**) should be used. Assuming a modest scan resolution of 10 μm<sup>3</sup> per voxel –still largely insufficient– the size of the modelled media becomes intractable using traditional computing methods.
2. *Biological diversity and number of simulated individuals.* The amount of computation required by an IbM is proportional to the number of individuals simulated, typically around 10<sup>9</sup> bacteria per gram in soils. In addition to the number of individuals, many functional groups of microorganisms has to be taken into account. Each functional group may have different and alternative metabolisms producing and consuming a variety of chemical substrates. These substrates have to be taken into account into the LBM, increasing the need of computation power.
3. *Best coupling strategy.* It is well known that the LBM is very well adapted to be run using GPUs since it reduces to a data parallel algorithm. On the contrary, algorithms required for the IbM approach tend to have certain parts not easily parallelizable. Therefore, the benefits of using a pure GPU programming

approach to run a combined LBM-IbM approach are not clear and needs to be investigated.

## 2. Description of the project

We have designed and implemented in C language a LBM-IbM approach taking into account the model LBioS of Vogel *et al.* [10] as a core model. The working name of ILBioS has been adopted for the approach. The initial ILBioS code designed to be run into the CPU will be ported to GPU. The general workflow of the modeling approach is shown in the **Fig. 4** below.



**Figure 4.** General workflow of ILBioS, the modeling approach coupling a lattice-Boltzmann model and an Individual-based model.

## 3. Objectives of the project

The main aim of this CEMRACS project is to achieve reasonable simulations of microbial degradation in 3D real soils images of  $1000^3$  voxels using the model ILBioS. A thorough investigation of the best high performance computing approach is expected in this project. One major deliverable is an increase of speed-up by more than x100 of the current serial code.

The main aim of the project will be faced establishing three specific objectives:

1. To analyze and, when possible, to improve the individual-based model parts of ILBioS having the highest impact in the computing time and model scalability.
2. To find out the best LBM-IbM coupling strategy in function of the computing

resources available. This analysis will be dependent on the computing facilities available at the time of realization of the project but would preferably involve a small/medium cluster of computers and a supercomputer architecture.

3. Design of the test scenarios required to demonstrate the best strategy to the scientific community. If possible, the tests will be run during the CEMRACS. If this is not possible, the tests will be run later and the results synthesized into an article to be published into an indexed journal.

#### 4. Expected participants and presence at CIRM

Dr Xavier Portell-Canal, a post-doc scientist at UMR ECOSYS of INRA ([xavier.portell-canal@grignon.inra.fr](mailto:xavier.portell-canal@grignon.inra.fr)), will be present during the 6 weeks of CEMRACS. He is an expert in individual-based modelling and has background in computer sciences. He will be in charge of the project management with the student.

Dr Valérie Pot, scientist at UMR ECOSYS of INRA ([valerie.pot@grignon.inra.fr](mailto:valerie.pot@grignon.inra.fr)), will not be physically present, but will provide support on the lattice-Boltzmann approach and on the description of diffusion and soil respiration processes in the 3D soil architecture at the microhabitat scale.

Dr Philippe Baveye, full professor at AgroParisTech an scientist at UMR ECOSYS of INRA ([philippe.baveye@agroparistech.fr](mailto:philippe.baveye@agroparistech.fr)), will not be physically present, but will provide general scientific support, especially on soil physics.

Dr Patricia Garnier, scientist at UMR ECOSYS of INRA ([patricia.garnier@grignon.inra.fr](mailto:patricia.garnier@grignon.inra.fr)) and general coordinator of the ANR project Soil $\mu$ 3D (see section 5). She will not be physically present, but will coordinate the funding requirements and provide occasional scientific support.

#### 5. Status of the funding

The ANR project Soil $\mu$ 3D 'Emergent properties of soil microbial functions: Upscaling from 3D modeling and spatial descriptors of pore scale heterogeneity' (2015-2019) will provide the funding required for the student and Dr Portell during the summer school.

#### References

- [1] Janzen, H.H. Carbon cycling in earth systems – a soil science perspective. *Agriculture, Ecosystems and Environment* 104, pages 399- 417, 2004.
- [2] Smith, K.A. & F. Conen. Impacts of land management on fluxes of trace greenhouse gases. *Soil Use and Management* 20, pages 255-263, 2004.
- [3] Dungait, J.A.J., D.W. Hopkins, A.S. Gregory, & A.P Whitmore. Soil organic matter turnover is governed by accessibility not recalcitrance. *Global Change Biology*, 18, 1781-1796, 2012.
- [4] Chenu, C., J. Hassink. & J. Bloem. Short term changes in the spatial distribution of microorganisms in soil aggregates as affected by glucose addition. *Biology and Fertility of Soils* 34, 349-356, 2001.
- [5] Carminati, A., A. Kaestner, P. Lehmann & H. Flühler. Unsaturated water flow across soil aggregate contacts. *Advances in water resources* 31, pages 1221-1232, 2008.
- [6] Raynaud, X. & N. Nunan. Spatial ecology of bacteria at the microscale in soil. *Plos One*, 9, 2014.
- [7] Eickhorst, T. & R. Tippkötter. Detection of microorganisms in undisturbed soil by combining fluorescence *in situ* hybridization (FISH) and micropedological methods. *Soil Biology and Biochemistry* 40, pages 1284-1293, 2008.
- [8] O'Donnell, A.G., I.M. Young, S.P. Rushton, M.D. Shirley & J.W. Crawford. Visualization, modelling and prediction in soil microbiology. *Nature Reviews Microbiology* 5, pages 689-699; 2007.
- [9] Kang, S., S. Kahan, J. McDermott, N. Flann & I. Shmulevich. Biocellion: accelerating computer simulation of multicellular biological system models. *Bioinformatics* 30, pages 3101–3108, 2014.

- [10] Vogel L., D. Makowski, P. Garnier, *et al.* Modeling the effect of soil meso- and macropores topology on the biodegradation of a soluble carbon substrate. *Advances in Water Resources* 83, pages 123-136, 2015.