Accounting for soil architecture and microbial dynamics in microscale models: Current practices in soil science and the path ahead

V. Pot¹*, X. Portell², W. Otten², P. Garnier¹, O. Monga³, P.C. Baveye¹

1 Université Paris-Saclay, INRAE, AgroParisTech, UMR ECOSYS, Thiverval-Grignon, France
2 School of Water, Energy and Environment, Cranfield University, Cranfield, Bedfordshire, UK
3 IRD, UMI UMMISCO, Bondy, France

* corresponding author: Valérie Pot – valerie.pot@inrae.fr

Abstract

Macroscopic models of soil organic matter (SOM) turnover have faced difficulties in reproducing SOM dynamics or in predicting the spatial distribution of carbon stocks. These models are based on a largely inadequate linear response of soil microorganisms to bulk concentrations of nutrients and it is clear that a new approach to SOM modeling is required. Introducing explicit microbial activity and OM reactivity in macroscopic models represents a challenge because of the fine spatial scales at which the processes occur. To get a better grasp on interactions that take place at the micro-scale, a new generation of SOM models have been developed at the spatial scale of the soil micro-environments.
where microorganisms evolve. They are well adapted to challenge traditional hypotheses about the influence of soil architecture on soil microbial activity. Soil architecture provides the scene for a dynamic spatial accessibility of resources to microbes and the emergence of interactions between the actors of SOM decomposition. In this context, we review microscale models of microbial activity that have been designed for soils and soil analogs. To understand how these models account for spatial accessibility, we have looked in detail at how soil micro-environments are described in the different approaches and how microbial colonies are spatialized in these micro-environments. We present the advantages and disadvantages of the developed strategies and we discuss their limits.

**Keywords**: bacteria models, fungi models, microbial dynamics, microscale models, soil architecture, soil micro-environments, soil organic matter decomposition

**Highlights**:
- We review the state of the art in the development of microscale models of soil microbial processes
- Microscale models have integrated the tremendous progress in knowledge of OM cycling in soils
- We categorize modeling approaches in how they deal with complexity of soil architecture
- Modeling benchmarks are required to properly balance complexity and data uncertainty

1. Introduction
Over the last few years, a significant body of research has been devoted to improving the macroscopic models of the turnover of SOM that are integrated in large-scale C-cycling or climatic models of the Earth. To try to account for sub-macroscopic heterogeneity, these macroscopic SOM turnover models make use of empirical coefficients that are introduced to constrain turnover constants as a function of soil water content and, in rare cases, of physical separation between OM and microbes. In general, these tuning coefficients appear to provide sufficient degree of freedom to reach good estimates of turnover dynamics under stable pedoclimatic situations. However, these models tend to fail in transient situations, e.g., in response to increases in nitrogen fertilization or soil temperature (Sierra et al., 2012). A common overall assessment is that these models are not capable of predicting SOM dynamics accurately (Davidson et al., 2012; Sulman et al., 2018).

As a result, the assumptions embodied in these models have become increasingly challenged in recent years. Various authors have argued that the idea of splitting SOM into distinct pools based only on chemical composition and degree of recalcitrance no longer corresponds to the latest picture available of the nature of SOM (Abramoff et al., 2018; Basile-Doelsch et al., 2020; Blankinship et al., 2018; Dungait et al., 2012; Sainte-Marie et al., 2021; Yang et al., 2021). Among the pieces of evidence that led to this perspective, Kleber et al. (2011) showed that the oldest SOM contains a greater proportion of easily decomposable organic molecules, while the younger one contains organic structures that are supposed to be more stable, suggesting that “C age is not necessarily related to molecular structure or thermodynamic stability”. Other factors may explain the failure of macroscopic models to predict SOM turnover. Among them, are the physical protection of SOM through associations with poorly crystalline minerals (Kleber et al., 2005; Basile-Doelsch et al., 2020) and carbon saturation controlled by organo-mineral interactions (Kleber et al., 2007). Another avenue that has been explored to improve macroscopic models consists of accounting for microbial growth.
explicitly (Wieder et al., 2015) in order to better predict the transient behavior of carbon fluxes during environmental changes (Hadas et al., 1998; Lawrence et al., 2009; Sainte-Marie et al., 2021; Woolf and Lehmann, 2019).

Another possible explanation for the failure of simple macroscopic models to describe the transient dynamics of carbon and nitrogen fluxes in soils is that these models, from the start, are missing crucial aspects of the functioning of soils. This perspective is in line with the conclusion, reached by microbiologists sixty years ago, that microbial processes in soils are strongly influenced by the physical and chemical properties of microenvironments in which microorganisms reside. Alexander (1965), in particular, opined that without somehow taking into account the microscale heterogeneity of soils and the characteristics of these microenvironments, now also referred to as hotspots (Kuzyakov and Blagodatskaya, 2015), it would be impossible to describe accurately the multivarious microbial processes occurring in soils. This opinion has been confirmed by later experiments (e.g., Parkin, 1987), and essentially the same message has been reiterated by Young and Crawford (2004).

In the last fifteen years, it has become increasingly acknowledged that the spatially heterogeneous architecture of soil, the uneven distribution of water and air in the soil pores, and the heterogeneous spatial distribution of trophic resources and microorganisms contribute to the emergence of microbial processes in soils that cannot be accounted for solely with the macroscopic measurements that are currently carried out (Baveye et al., 2018; Heitkötter & Marschner, 2018; Joergensen & Wichern, 2018). During the same time frame, it has also become increasingly clear that the static picture of microorganisms confined to isolation in discrete microenvironments is inaccurate because microorganisms are able to move in order to find their sustenance (e.g., Berg & Brown, 1972; Juyal et al., 2018; Lacal et al., 2011; Kohlmeier et al., 2005; van Loosdrecht et al., 1990).
To our knowledge, Kuka et al. (2007) were the first to have attempted the development of a model of SOM dynamics that accounts for the influence of pore space distribution. However, their model did not account for positional relationships between the different classes of pores. Although model predictions match remarkably a set of observations in an independent field trial, the parameterization of fluxes necessarily lumps together the interactions between microorganisms and trophic resources, their spatial disconnection, and the role of diffusion and/or advection between the different classes of pores. To avoid this lumping together a new generation of SOM decomposition models has been developed subsequently, accounting for the three-dimensional architecture of soil and the heterogeneity of the distributions of microorganisms, water, air, SOM and reactive surfaces at the level of microbial habitats (Baveye et al., 2018). In two recent mini-reviews, König et al. (2020) reported on microscale models with a special focus on spatiotemporal dynamics of bacteria and bacterial functions in soil microhabitats, and Golparvar et al. (2021) reported on microscale models of reactive transport modeling in porous media.

In this general context, the key objective of the present review article is to assess the current practices for considering soil architecture and microbial dynamics in microscale models of soil microbial processes. We encompass models aimed at predicting SOM decomposition and greenhouse gases emission, but also models whose aim is to better understand ecological interactions between microbes as well as the emergence of the spatial organization of microbial colonies (bacteria and fungi) and biodiversity. Indeed, these studies participate to the general understanding of SOM turnover since for instance production of enzymes that are key factor in the depolymerization step of large polymers is dependent on ecological and spatial interactions and biodiversity. Although not related to SOM turnover, we also report some studies dealing with biofilm growth in artificial porous media for their advanced understanding of the interactions between bacteria cycle of life and water flow. We focus on
i) the main modeling strategies to describe soil architecture according to a complexity gradient in capturing the spatial heterogeneity at the scale of microhabitats (Section 2), and on ii) the main strategies adopted to describe soil microbial dynamics, including that of soil fungi (Section 3). Then, we discuss the advantages and disadvantages of these various strategies in Section 4. We address two major limitations of efforts in this area to date, and discuss critically several options that could be explored to fill knowledge gaps in the modeling of soil microbial functions (Section 5).

2. Main modeling strategies to describe soil architecture at the microbial habitat scale

To review how the soil micro-environments are considered in modeling efforts, it is convenient to classify the microscale models according to two criteria (Figure 1): i) the nature of the attributes of the grid nodes, and ii) the degree of complexity of the description of soil architecture. From the point of view adopted in this review, microscale models are classified as *implicit* when the node attributes consist of bulk soil averaged variables: porosity, water content, and effective diffusion coefficients. By contrast, microscale models are classified as *explicit*, also called “pore-scale” models, when the attributes are Boolean values associated with only one phase, i.e., water, air, or solid, and, for the former two phases, the diffusion coefficients are set equal to the corresponding molecular diffusion coefficients in water or air. Implicit models only allow a simplified description of soil architecture (Section 2.1), whereas the explicit strategy makes it feasible to reproduce actual, detailed soil architectures as well as generate idealized soil architectures (Section 2.2). A mixed strategy considering both implicit and explicit attributes is also sometimes adopted to describe detailed soil architectures (Section 2.3). For each of these distinct approaches, we report how models account for the abiotic...
processes that distribute air and water in the pore space, or lead to the transport of solutes (dissolved organic carbon (DOC), NO$_3^-$, O$_2$, etc) and gaseous products (CO$_2$, CH$_4$, N$_2$O), as well as the spread of reactive solutes on soil solid surfaces. Finally, we report how solid and dissolved SOM and microorganisms are spatially located in the soil architecture (Section 2.4).

Some of the studies cited in this review sometimes use the same model, either identically the same, or a slightly modified (simpler or more complex) version of it. The technical description of the models, their equations, and the parameters involved, are not always detailed enough to determine precisely where the differences reside. We therefore decided to present this analysis considering each literature source as an independent study.

### 2.1 Implicit microscale models

This strategy maintains the classical macroscopic approach, but it is applied at the level of microhabitats. The micro-environments are spatially discretized in a regular grid and the size of the grid nodes ranges from about 1 μm$^3$, i.e., roughly the average size of bacterial cells in soils, to about a few mm$^3$ (Table S1). This strategy is adopted in 25 out of the 78 studies reported on this review.

Ignoring totally the underlying soil architecture and diffusion processes, Masse et al. (2007) and Nunan et al. (2020) only distributed spherical spots of OM and of microorganisms in a 3D volume. Then, based on population ecology, the models rely on physical contacts between microbial spots and OM spots to achieve consumption and growth. To account for the spatial heterogeneity of the soil and the variable accessibility of the resource for the microbial decomposers, the size of the microbial spots can be varied (Garnier et al., 2008; Nunan et al., 2020).
Using a different approach, Folse & Allison (2012) model in a pseudo-2D grid the cooperation, competition, and formation of coalitions in a community of enzyme-producing microbes. Nutrients (carbon, nitrogen, and phosphorus) are unlimited and do not diffuse, whereas enzymes produced by microorganisms diffuse. Each node of the grid lacks information on its nature (pore or solid), or on its water and air content. All this information is encapsulated in a single effective diffusion coefficient of the soluble substrates. Other studies on the emergence of a spatial structure of bacterial colonies repeat this same approach (Kaiser et al., 2014, 2015; König et al., 2017, 2018, 2019; Wang & Allison, 2019).

In two other similar studies (Chakrawal et al., 2020; Pagel et al. 2020), the effective diffusion coefficients are calculated via the macroscopic relationship proposed by Millington & Quirk (1961), which considers hydration. However, spatial information about the water content at each grid node is not provided.

This particular aspect is addressed by Evans et al. (2016), who simulate variably saturated conditions to reproduce the Birch effect. However, soil water content is kept constant across all nodes of the grid. Knudsen et al. (2006) develop a 3D model of hyphal growth under similar conditions. Prescribed values for water content, nutrient availability, and temperature are attributed to each grid node. However, like Evans et al., (2016), Knudsen et al. (2006) do not allow these attributes to vary spatially. Thus, although these approaches go a step further to include spatial heterogeneity of hydration conditions in the simulated domain, neither explores this potential. This further step is taken in models studying the coexistence of competitive species under wet-dry cycles (Long & Or, 2005, 2009, Šťovíček et al., 2017). Long & Or (2005, 2009) vary the porosity and water content attributes spatially over their grid node according to a log-normal distribution of the variances of the parameters of the water retention curve proposed by van Genuchten (1980). Later, to describe more complex ecological interactions (competition and mutualism), Kim & Or (2016) calculate the water content of
each node with the help of a fractal dimension to describe the roughness of the porous surface and assume that pores have a conical geometry.

In the majority of implicit microscale models, physico-chemical processes are not considered at all. In the few articles that are exceptions (Babey et al., 2017; Banitz et al., 2015; Gras et al., 2010, 2011; Pagel et al., 2020), the sorption of dissolved OM components, organic pollutant (pesticide), and/or NH$_4^+$ are described as mass transfer processes.

2.2 Explicit microscale models

2.2.1 Case of detailed actual soil architecture

The description of soil architecture generally relies on cutting-edge imaging techniques (e.g., Baveye et al., 2018; Lucas et al., 2020), such as through the use of X-ray computed tomography (CT), either at synchrotron facilities or using table-top laboratory equipment. The quality and resolution of the images determines the extent to which they may provide 3D spatial information about, e.g., the geometry of the pore space, or of water and air interfaces, and about the location of solid particulate organic matter (POM). Segmentation of CT images (e.g., Hapca et al., 2013; Houston et al., 2013; Schlüter et al., 2014) is usually carried out in order to assign to each voxel one of the phases of the soil: void or solid, with voids filled with water or air, solid being OM, minerals or matrix.

Of the 78 references reported in this review, only 9 studies encompass the actual architecture of soils, using CT soil images as direct input data (Table S1). Attempts to model fungal growth have considered the architecture of both repacked soil samples (Falconer et al., 2007, 2015; Pajor et al., 2010) and undisturbed soil samples (Cazelles et al., 2013; Falconer et al., 2012; Kravchenko et al.,
 whereas two studies on SOM decomposition by bacteria have used CT information obtained in undisturbed soil samples (Portell et al., 2018; Vogel et al., 2015, 2018).

Measurements of air-water interfaces in soil using Synchrotron CT are available under very specific conditions (e.g., Brusseau et al., 2007; Pot et al., 2015). However, models based on an actual soil architecture so far have described unsaturated conditions via a mechanistic approach using a two-phase lattice-Boltzmann model (LBM) that considers Van der Waals forces (Shan & Chen, 1993, 1994) and capillary forces (Raiskinmäki et al., 2000). Phase separation simulations under static conditions (Appert et al., 1991) can be performed for different water saturation conditions following the approach proposed by Genty & Pot (2013). A 3D map of a continuous LBM fluid density is obtained, the highest and the lowest densities of which can be associated to the water and air phases respectively (Portell et al., 2018; Vogel et al., 2015). Then the soluble carbon substrate diffuses only in the nodes filled with water according to the molecular diffusion coefficient of the substrate in water. The coefficient according to which a fungal colony is assumed to diffuse through soil pores is maximal in the case of lowest density (Falconer et al., 2012).

2.2.2. Case of simplified actual soil architecture

The models described in the previous subsection rely on a huge amount of information contained in soil images, which is constantly increasing due to steady improvements of imaging techniques. Today, 3D CT images that are routinely acquired require several tens of gigabytes of storage for a size of up to about $2000^3$ voxels (Helliwell et al 2019; Koebernick et al., 2017; Quigley et al., 2018; San Jose Martinez et al., 2018). The numerical solvers of the pore-scale models that directly include CT images as input data, such as LBM or finite volume resolutions, require large computational and memory capabilities. Thus, a more efficient strategy in terms of computer usage consists of reducing the
information amount by simplifying soil architecture. Two directions have been adopted to that end (Figure 1). Either models retain the salient connectivity and size features of the pore space morphology using morphological or irregular pore network models, or they generate grids or pore networks that reproduce the statistical properties of actual soil samples as given by soil bulk measurements like particle size distribution or percolation experiments. The studies adopting this strategy, 18 out of the 78 references, are reported in increasing order of simplification of soil architecture.

Morphological models are based on paving the pore space with an ensemble of simplified geometrical primitives (usually balls, e.g., Vogel et al., 2005). This is done after having detected the median axes of the pore space. Only the minimal set of maximal balls that make a piecewise approximation of the pore space is retained (e.g., Al-Kharusi & Blunt 2007; Monga et al., 2007; Silin & Patzek, 2006). This set constitutes the nodes of an adjacency graph considering the neighborhood information of each ball or node where the decomposition of OM is calculated (Monga et al., 2008, 2009, 2014; Ngom et al., 2011). This ensemble is much more compact in terms of numerical data than all the voxels of the CT image. Other 3D geometric primitives such as ellipsoids and generalized cylinders (Ngom et al., 2012; Kemgue et al., 2019) have also been proposed, however they have not yet been used to simulate SOM decomposition. In this strategy, unsaturated conditions are modeled using the Young-Laplace law. The geometric primitives whose size does not exceed the size that would be expected to drain at a prescribed water potential are filled with water leading to an explicit 3D distribution of water and air at the nodes of the graph (Pot et al., 2015). The soluble carbon substrate diffuses only in the water-filled nodes according to the molecular diffusion coefficient of the substrate in water (Monga et al., 2014).

Irregular pore-network (PN) models are another valuable alternative to reduce the amount of information contained in the soil images to a more manageable size. These models consist of an
ensemble of nodes connected to each other by links whose length, radius and number can be varied (Blunt 2001; Xiong et al., 2016). The generated networks also keep most of the characteristics of the pore space morphology: the pore size distribution, a mean degree of pore branching or connectivity (e.g., Vogel & Roth, 2001), and the spatial heterogeneity of the clustering of pores and connectivity (e.g., Blunt et al., 2002). While these irregular PN models are often used nowadays in computer fluid mechanics (e.g., Martinez-Mendoza et al., 2019), the model of Perez-Reche et al. (2012) is the only one that takes advantage of the potential of this approach to simulate microbial invasion in soils.

Simpler PN models based on a regular arrangement of the nodes are more often used, e.g., face-centered cubic grids that retain only the pore size distribution and the mean degree of pore branching of a soil. To simulate growth in cases involving a single bacterial cluster to several aerobic and anaerobic communities, Ebrahimi & Or (2014, 2015, 2016, 2017, 2018) generate 3D regular angular PNs where the coordination number of the nodes, the size distribution of the links and the porosity are defined according to natural soil systems values or fitted in order to recover the van Genuchten (1980) water retention curve parameters of a given soil (Table S1). The width and spanning angle of triangular links can also be calibrated to soil water retention curves (Rosenzweig et al., 2013, 2014). This strategy is much less representative of soil architecture than the approaches using irregular PN models or morphological models, as noted by Blunt et al. (2002) and Ebrahimi et al. (2013). However, since regular PNs do not describe the spatial arrangement of pores, some links can be randomly removed to create spatial heterogeneity (Ebrahimi & Or, 2014). The length of the links can be changed to account for tortuosity using helical bonds (Ebrahimi & Or, 2016). As with morphological models, the Young-Laplace law is used to explicitly distribute water and air in the PN. The choice of a triangular cross-section of the links makes it possible to more realistically simulate variable water contents than would be possible if the links were spherical or cylindrical. In the latter cases, the application of the Young-
Laplace law leads to a binary response (empty or water-filled) and can thus promote greater fragmentation of the liquid phase.

The flexibility of the latter approach permits also to generate multi-scale porous systems. Taking advantage of a commercially available and widely used PN simulator (Xpore), and its predecessor (Pore-cor), Laudone et al. (2011) develop a dual-porosity PN model. The positional relationships of the macroporosity and microporosity are generated in order to fit several fundamental properties of soil, namely water retention, aggregate size distribution, porosity of the microporous and macroporous zones, and percolation characteristics (Laudone et al., 2013; Matthews et al., 2006). We observe that it is also one of the few models that consider diffusion of gases (CO₂, N₂O and N₂) emitted within the pore space.

Finally soil architecture can be described by using a regular grid whose node attributes are associated to pore or solid phases on the basis of statistical estimates of the pore sizes derived from a set of published CT images of soil aggregates (Resat et al., 2012). However, in this approach, the actual arrangement of pores is lost as pores are randomly placed on the grid with an additional constraint that pores be continuously connected. Initial random assignment of nodes into pore and solid phases are also considered in two pore-scale models that simulate temporal modifications of soil architecture under microbial activity. Over time, solid nodes with a weak structure are spatially rearranged to maximize the overall stability of solid node clusters (Crawford et al., 2012; Ray et al., 2017). This stability is defined according to the local rate of production of stabilizing agents, whose concentration is a function of the local respiration rate.

2.2.3. Case of idealized soil architecture
Columns filled with glass beads have occasionally been used to physically mimic soil architectures. To describe these artificial packings, imaging (Graf von der Schulenburg et al., 2009; Peszynska et al., 2016; Pintelon et al., 2012) or 2D mathematical algorithms of grain arrangement (Kapellos et al., 2007) can be used as input data. With such algorithms it is also easy to introduce spatial heterogeneity, such as the occurrence of fractures (Tartakovsky et al., 2009). A regular PN has also been generated based on pore size distribution of a column packed with glass beads (Gharasoo et al., 2012, 2014). These geometries are mostly used to study the permeability reduction due to the accumulation of biomass and its metabolic by-products in coarse-textured porous media under water-saturated conditions. Water flow, advection, and reaction-diffusion of soluble substrate are thus key processes in addition to biological processes. To consider the higher viscosity of the biomass and associated exopolymers compared to that of water, the coefficient of diffusion of the solutes through it can be reduced compared to that in water (e.g., Tang et al., 2013).

To create synthetic rough, porous surfaces, 2D regular PN models with triangular channels are generated based on typical pore sizes and connectivity. Variable water-saturated conditions can be calculated with the Young-Laplace law. This approach has been adopted to model the motility of bacteria and expansion of bacterial colonies (Wang & Or 2010), as well as the coexistence (Borer et al., 2018; Long & Or 2007; Wang & Or 2012, 2013) and self-organization (Wang & Or, 2014) of bacteria. Wang & Or (2012, 2013), simplify the time-consuming explicit calculation of solute diffusion-reaction processes in the links of the PN by directly calculating an effective diffusion coefficient at each node using the Millington & Quirk (1961) expression using the node water status as input. This simplification brings this contribution closer to an implicit description of soil architecture (see Section 2.1).
Random assignment of pore and solid nodes is also commonly used to generate soil analogs. Falconer et al. (2008) create porous media of porosity varying between 0 and 1 and simulate fungal invasion. Other authors have considered instead a random assignment of collections of nodes made impervious to mycelium and nutrient diffusion (Boswell et al., 2007; Boswell, 2008; Boswell & Hopkins, 2009; Vidal-Diez de Ulzurrun et al., 2017). In a simpler form, the fungus models of Boswell et al. (2003) and Boswell & Hopkins (2009) do not consider impermeable materials but a spatially heterogeneous initial distribution of the carbon substrate in the form of different tessellations. In these cases, diffusion of the substrate external to the fungus is not permitted.

Different sizes and lengths of geometrically highly simplified, linear pores in 2D are used to assess the role of salient soil pore geometry, from shallow pores to deep crevices with small opening, on the growth and survival of bacteria (Resat et al., 2012). Using the COMSOL Multiphysics software applied to a pore of uniform geometry, Schmidt et al. (2018) assess the influence of the spatial location of bacterial colonies along the pore on the biodegradation of an organic substrate that moves through it.

Finally, experimental micro-models of diverse geometrical forms like a succession of rectangular water-saturated microhabitats (Keymer et al., 2006), packing of cylinders, and etched-glass networks (Borer et al., 2019) have been reproduced in a number of computational models (Borer et al., 2019; Centler et al., 2011; Knutson et al., 2005; Tang et al., 2013). Due to simplification of the reproduced micro-models into a 1D grid, Centler et al. (2011) use effective diffusion coefficients for the substrate and bacteria, together with effective chemotactic sensitivity constants towards the substrate and the bacteria.

2.3 Mixed explicit/implicit microscale models
A final strategy mixing implicit and explicit attributes of the grid nodes to describe actual soil architecture from soil imaging is adopted in two studies (Yan et al., 2016, 2018). In this approach, a dual threshold segmentation of the CT images is used to obtain ternary images where the gray voxels whose value lies between the two thresholds that define voids and solids are associated with a mix of pore and solid phases. These latter nodes have an implicit porosity attribute whose value depends on the gray level scale (Figure 1). The diffusion coefficients of the dissolved OM, O\textsubscript{2} and CO\textsubscript{2} are equal to their molecular diffusion coefficient in liquid and air phase for the explicit water and air voxels, respectively. Whereas for the implicit unsaturated voxels, the diffusion coefficients are calculated according to empirical laws (Archie, 1942; Hunt, 2004). Mixed explicit and implicit diffusion coefficient attributes of the grid nodes are thus used in this approach. Additionally, the sorption processes of OM are described by mass transfer rates using effective surface reactivity parameters.

2.4 Spatial description of the initial distribution of SOM and microorganisms in the reviewed models

It is critical to consider how OM is distributed within the soil architecture. Although combined staining and/or chemical and imaging techniques are promising tools (e.g., Maenhout et al., 2021; Peth et al., 2014; Roose et al., 2016) they are far yet from routinely providing data. Therefore, the placement of SOM has been mainly considered through modeling scenarios. When SOM consists of dissolved nutrients, it is most of the time homogeneously applied at the nodes of the grid (e.g., König et al., 2018) and/or applied at specific locations in the simulated domain to produce gradients or counter gradients between carbon and O\textsubscript{2} (e.g., Ebrahimi & Or, 2015). Metabolites or enzymes produced by microbial activity are located at the nodes of the producers (e.g., Folse & Allison, 2012), from where they diffuse.
When SOM consists of solid pieces, different distribution scenarios have been used ranging from aggregated to dispersed distributions (e.g., Falconer et al., 2015). Likewise, microorganisms are either homogeneously located in the simulated domain (e.g., Gharasoo et al., 2014), randomly distributed (e.g., Portell et al., 2018), or locally placed (e.g., Ebrahimi and Or, 2015, Falconer et al., 2015; Kravchenko et al., 2011), and the effect of different distributions can be explored. They can also be placed according to the size of pores (e.g., Monga et al., 2009). Models using explicit attributes and describing actual soil architectures tend to locate bacteria in water-filled nodes where they are assumed to live (e.g., Vogel et al., 2015) while fungi preferentially invade air-filled nodes (e.g., Falconer et al., 2012). Only one model (Pagel et al., 2020) uses different probabilistic distributions based on Raynaud & Nunan (2014) who integrate data in the visualization of microorganisms in soils (Nunan et al., 2003).

3. Main modeling strategies to describe the microbial dynamics in microscale models

In this section, we focus first on the nature of SOM pools, the type of microorganisms and the strategies used to distribute them spatially in their microhabitats. Then we focus on biological processes, encompassing substrate decomposition, microbial substrate uptake, microbial growth, microbial metabolic processes and mobility. It would make sense to report the microbial dynamics modeling approaches in terms of the soil spatial representation since the latter is at the heart of the reported microscale models. However, the number and the degree of details of the biological processes are not linked with the degree of representativeness of actual soil architecture or the choice of implicit
or explicit grid node attributes. Rather, it is the research question to which all these microscale models try to answer that has dictated the strategy. This is evidenced in Table 1 which lists the references from the most to the least accurate soil spatial representation. Table 1 provides details of each of the biological processes, the nature of SOM pools, microorganisms and modeling strategy to describe microorganisms colonies.

3.1 Nature of substrates and microorganisms

Most of the studies (51 out of the 78 references) consider a single simple soluble substrate (monomer) that is directly taken up by microbes (Table S1, e.g., Long & Or, 2007). Other studies (24 references) consider a preliminary step of decomposition (depolymerization) of complex solid substrate (polymers) into soluble organic compounds that are taken up by microbes (e.g., Pagel et al., 2020). Of the 78 references, two studies do not specify the nature of microorganisms (Perez-Reche et al., 2012; Crawford et al., 2012), and only 13 relate to filamentous fungi, while the others relate to bacteria, including 9 references that deal with accumulation of bacterial biomass and associated exopolymers. Bacteria models generally simulate a single specific species or strain (in 47 references, e.g., Knutson et al., 2005), or up to eight different genotypes (Folse & Allison, 2012) or 10 different taxa (Nunan et al., 2020). Fungi models simulate a single species except for Falconer et al. (2008) that are the only ones to study more than one fungal species, comparing compatible and/or antagonist species. Most bacterial models describe i) generic functional groups such as anaerobic bacteria, aerobes, methanogens (e.g. Ebrahimi & Or, 2018), ii) generic competitive species such as competitive, generalists, or cheaters for the same resource (e.g. Kaiser et al., 2014; Monga et al., 2008; Pagel et al., 2020) or from a range of specialists of a single resource to generalists of up to ten different resources (Nunan et al., 2020). Three
studies consider commensal or mutualistic relations in microbial consortia. In their simplest form, the models consider a chain of up to three substrates and microbe species, with one of the species dependent on one substrate produced by other species (Wang & Or, 2014; Kim & Or, 2016). In its most complex form, the model of Borer et al. (2019) which is based on the COMETS model of Harcombe et al. (2014), encompasses a reduced metabolic network involving more than a hundred metabolites and reactions.

3.2 Modeling strategies to describe microbial colonies

3.2.1 Case of bacteria

About a third of the references (21 out of 62 studies dealing with bacteria) model the life cycle processes at the scale of bacterial colonies that are spatially distributed in the soil pore space. This population-based strategy calculates the growth of a single (lumped) mass of bacteria representative of the microorganisms contained in a given microhabitat. Cell division is not considered and mortality processes (starvation or predation) are assumed to be continuous processes (e.g., Vogel et al., 2015). The updating scheme is synchronous, i.e., uptake, growth, mortality occur at the same time for the whole microcolony.

Another strategy, adopted by 37 references, models the life cycle processes at the scale of individuals, which are the bacteria cells (e.g., Graf von der Schulenburg et al., 2009). In this individual-based, or agent-based modeling approach (Hellweger & Bucci, 2009), cell division processes are discrete events, a feature that facilitates the discretization of the spatial expansion of the population. Starvation mortality or predation processes are discrete events (the whole cell dies, not a fraction of it). In addition, interspecific and intraspecific variability in the individual properties are considered (e.g.,
size or mass threshold to start cell division). Synchronous updating (e.g., Kreft et al., 1998) or asynchronous updating (e.g., Ginovart et al. 2005) schemes can be used. Despite the fact that, in the presence of recycling processes, differences can appear between both schemes (Caron-Lormier et al., 2008), the type of updating of individual-based models is not always stated, which complicates reimplementaion and, ultimately, reproducibility of the research.

Meta-agents, or “superindividuals”, i.e., ensembles of single bacterial cells, are used to mitigate the large computation costs associated with the description of processes at the single-cell level (Banitz et al., 2015; Borer et al., 2019; Masse et al., 2007). This coarse-grained description is also used by Long & Or (2005) who adopt the random walker model developed by Ben-Jacob et al. (1994). Using this approach, the question remains how to identify the size of the meta-agents that makes it possible to retain the influence of the individual behavior of bacteria.

3.2.2 Case of fungi

The fungal “individuality” is much more difficult to define as compared to bacteria. Fungi are multicellular organisms that spread in the pore space. Fungi grow as an interconnected network of hyphae, named mycelium or fungal colony. The shape of the whole fungal colony thus needs to be considered. None of the reviewed fungal models considers single cells. Continuous models and individual-based models of fungal growth have been distinguished (Boswell & Hopkins, 2009, Table S1). In the former approach, the fungal mycelium is considered as a continuous biomass (e.g., Boswell et al., 2003; Falconer et al., 2007), whereas the individual-based approach considers individual hyphae explicitly and describes the discrete (random) events involved when they grow and branch (e.g., Knudsen et al., 2006; Boswell et al., 2007).
3.3 Description of biological processes

3.3.1 SOM hydrolysis

In some models (e.g., Long & Or, 2005), the depolymerization step of SOM is assumed to be independent of microbial growth and be described by a fixed hydrolysis rate. Often, models assume that depolymerization depends on enzyme production (e.g., Allison, 2005) or microbial growth (e.g., Falconer et al., 2015). Complex biogeochemical reactions including mixed kinetics and thermodynamic equilibrium reactions, such as oxic degradation of SOM (Aguilera et al., 2005), are calculated for the soluble substrate, using the Biogeochemical Reaction Network Simulator in one study (Gharasoo et al., 2012).

The mathematical description of SOM hydrolysis follows the three conceptual classes described by Wutzler & Reichstein (2008), which represent respectively no-, linear- and non-linear dependencies on decomposer biomass. In the first class, constant rates or more often Michaelis-Menten kinetics (e.g., Kaiser et al., 2014) are used. The latter permits an indirect dependence on biomass uptake and growth through microorganisms that produce enzymes. Chakrawal et al. (2020) use the three conceptual classes in a comparative study.

3.3.2 Microbial substrate uptake and growth

In its simpler form, and depending on the research question, growth of biomass is neglected (in six references, e.g., Laudone et al., 2011; Crawford et al., 2012) and a constant biomass value is used. A large number of studies (36 references) use the Monod equation to describe the non-linear growth of decomposer biomass (e.g., Banitz et al., 2011) whereas linear kinetics is adopted in eight references (e.g., Boswell, 2008; Gharasoo et al., 2014). The logistic growth is also described in two studies
(Centler et al., 2011; Gharasoo et al., 2014). Fixed rates are used to calculate growth in the six references that adopt the offer-demand approach (e.g., Masse et al., 2007; Monga et al., 2008). In five studies, growth is calculated from metabolic pathway reactions (e.g., Gras et al., 2011; Borer et al., 2019). Finally, in one case, growth is implicitly modeled by increasing the size of the area where the cell consumes the resource assuming the increase to be an inverse function of the catabolic breadth of the cell taxon, i.e., the range of organic substrates that can be used (Nunan et al., 2020).

Microbial substrate uptake and growth can also be limited by the availability of other nutrients (N, P) and O₂. Two approaches are adopted. Either microbe stoichiometry constraint limits uptake (e.g., Allison, 2005) or the concentration of the limiting factor is introduced in the Monod formalism. In the latter case, the dual-Monod equation is generally used (e.g., Yan et al., 2016) or the minimum of the Monod equations established for each substrate (Borer et al., 2018; Kim & Or, 2016). In one reference, a multi-substrate Monod kinetics is used to describe the uptake of small polymers and monomers (Pagel et al., 2020).

### 3.3.3 Microbial metabolism

Almost two third of the references consider microbial maintenance (e.g., Boswell et al., 2003; Resat et al., 2012). Growth respiration is generally implicitly considered by using the assimilation yield, whereas more rarely it is integrated in a lumped respiration term that does not distinguish maintenance and growth respiration processes (e.g., Vogel et al., 2018). The production of metabolites by microorganisms can be directly calculated in the form of enzymes (e.g., Evans et al., 2016), by-products (Kim & Or, 2016), generic inhibitor (Falconer et al., 2008; Kravchenko et al., 2011) or sticky agents (Ray et al., 2017), which can undergo transformation (decay) or sorption processes.
The production of greenhouse gases is not always modeled (27 references), and when it is, the majority of studies deal only with CO₂ (e.g., Nguyen-Ngoc et al., 2013), whereas a minority also calculate denitrification and emission of N₂O (Ebrahimi & Or, 2015, 2016, 2018; Laudone et al., 2011) or CH₄ (Ebrahimi & Or, 2017, 2018).

3.3.4 Microbial death, recycling, and dormancy

Microbial decay processes are mostly described by starving death (48 references, e.g., Allison, 2005) and additionally by predation (e.g., Portell et al., 2018). However, only one reference explicitly simulates predator growth (Pagel et al., 2020). The recycling of dead biomass in solid- (e.g., Gras et al., 2010), soluble- (e.g., Wang & Allison, 2019), and in both solid and soluble forms (e.g., Folse & Allison, 2012; Monga et al., 2014) appears in 17 references. Dormancy is only accounted for in 12 references (e.g., Resat et al., 2012).

3.3.5 Microbial mobility

Colonization is a key behaviour of fungi and it is driven by hyphae invasion. Advective and diffusive movements including apical growth and branching (e.g., Boswell, 2008) or solely diffusive movements (e.g., Falconer et al., 2007) are described. The hyphae segments can be constrained to grow and branch on a regular lattice (Boswell, 2008; Boswell & Hopkins, 2009; Boswell et al. 2007) or the geometrical constraint on orientation of hyphae growth can be relaxed when using a lattice-free approach (Knudsen et al., 2006; Vidal-Diez de Ulzurrun et al., 2017). Fungal networks generated with these lattice-free models are more realistic morphologically than the ones generated with the lattice-based approach but they require much more computational resources (Boswell & Hopkins, 2009).
A large majority of the bacterial models account for movement. Motility can be considered as a simple random motion from one node containing bacterial cells to a neighbor node (e.g., Ginovart et al., 2005). The choice of the neighbor can be further constrained by hydration status (Long & Or, 2005, 2009) of the targeted node, its chemical nature (cellulose adhesion in Resat et al., 2012) or the size of pore (Perez-Reche et al., 2012). Cell motion is described only during the cell division process (e.g., Kaiser et al., 2014) or in situations of excess biomass (e.g., Knutson et al., 2005; Graf von der Schulenburg, 2009). Several studies have accounted for cell propulsion due to flagellated motility where hydration status can reduce the cell velocity (e.g., Wang & Or, 2010), chemotaxis process toward substrate (e.g., Ebrahimi & Or, 2015) or towards bacteria (e.g., König et al., 2017). Two studies (Banitz et al. 2011, 2016) consider bacterial diffusion along analogs of fungal network hyphae, whereas others account for the effect of shear stress which can reduce the movement of bacteria cells (e.g., Tartakovsky et al., 2009).

3.3.6 Carbon translocation by fungi and anastomosis

Carbon translocation along fungal mycelium is driven by the concentration gradients of the internal substrate along the hyphae (e.g., Falconer et al., 2007) while an active translocation can additionally be described (e.g., Boswell, 2008). Anastomosis is explicitly considered by removing hyphal tips when they move to a node that already contains a hypha (e.g., Boswell, 2008). Hyphal interactions that can result in the fusion of two individual colonies or maintain mycelial individuality of somatically incompatible genotypes are described in Falconer et al. (2008).

4. Discussion about the main modeling strategies
4.1 Soil architecture

In all the reported studies, the inclusion of a spatial component is at the heart of the latest generation of microscale models to make ecological interactions emerge and explain soil microbial activity. The value of studies that use implicit approaches to summarize the environment is to emphasize the role of space and dispersal in ecological interactions, which can lead to patterns and hotspots that would not be observed in well-mixed system. The advantage of this approach is that these models rely on a few, easily measured bulk soil variables (Table 1). However, they cannot establish new mechanistic relationships between soil architecture descriptors and soil microbial response since they are dependent on empirical macroscopic laws like those relating the effective diffusion coefficients to hydration.

At the opposite end of the spectrum, methods that use explicit attributes can obviously capture most of the spatial heterogeneity of the clustering and connectivity of pores. This approach uses new knowledge of the microscale intricacies of soil architecture, which until recently had remained a black box for a long time. It offers the possibility to identify relevant 3D soil architecture descriptors that explain soil microbial activity and in fine will help upscale microbe activity from the micro-scale to the core-scale. The level of accuracy of the description of spatial heterogeneity needed in this approach has opened the brought-together scientists of various and complementary fields (soil science, mathematics, computer vision, ecologists, physicists). However, this strategy needs a huge amount of input data and, as a consequence, simulations require tremendous computing calculations and memory operations, which limits the size of the investigated domain that can be considered (Table 1). In the case of idealized geometry, like those of experimental micromodels or packings of glass beads, explicit attributes permits to exactly reproduce the geometry and helps validate the models (Table 1). The disadvantage is that actual soil architecture cannot be investigated.
The value of the approaches that describe simplified actual soil architecture or idealized porous media (which encompass regular grid and regular PN models) is to explore simplified scenarios of the complexity of pore spatial arrangement to investigate its potential role in soil microbial response (Table 1). The great advantage of methods based on morphological or PN models is that once the simplified networks have been defined, processes can be simulated much more efficiently allowing for larger samples at greater speed to be tested. A disadvantage is that these methods tend to focus on single aspects of the pore space such as connectivity, volume, pore-size distribution, but may not capture well the pore-solid interface area, which is critical for many processes in soil. It is also not clear at this stage whether regular PN are sufficient to catch the complexity of pore arrangement as compared to the more realistic irregular PN. Indeed, it is far from certain that predicted key factors in SOM microbial decomposition like accessibility between OM and decomposers is correctly captured in regular PN. Furthermore, only a limited number of soil types has been sufficiently described to provide reliable statistics on the pore space morphology. They are necessary data to generate regular PN, and this renders quantification of the soil microbial activity by this approach very questionable.

We observe that the more realistic the description of the 3D soil architecture is, the more mechanistic is the description of the distribution of water and air in the pore space, calculated from physical laws that consider interfacial and capillary forces, giving rise to realistic air-water interfaces (Table 1). It is surprising that none of the PN models reported in this review calculate mechanistically the invasion of the air phase in pores as a function of the evaporation rate, as for example Yiotis et al. (2001) did. They all rely on the empirical Young-Laplace relation, which is based on a cylindrical representation of pore geometry. Finally, knowledge on the spatial location of microbes in soils (e.g., Eickhorst & Tippkötter, 2008; Nunan et al., 2003) has not been taken into account so far in microscale models. At best, scenarios based on statistical models of spatial distribution of bacteria (Raynaud &
Nunan, 2014) have been used. Likewise, data on the spatial location of tiny SOM patches using staining techniques (e.g., Peth et al., 2014) or multimodal data acquisition on spatial location of chemical elements (Hapca et al., 2011) and bacteria (Schlüter et al., 2019) have not been used although they can be easily implemented in approaches considering explicit description of soil architecture.

4.2 Microbial dynamics

Given the large variety of processes described no consensus seems to emerge on the optimum choice of a spatial modeling strategy. Almost one third of the reported studies used population-based approaches, and the choice of the approach was not related to the size of the grid node volumes that could reach down to a few µm³ (Table S1). At this scale, a limited number of individuals can be contained in a node and this should in principle call for an individual-based approach rather than a population-based one, meant to represent large numbers of microorganisms. Indeed, population-based and individual-based model results will diverge when there is a low number of individuals, because in this case, the randomness of discrete events can result in a behavior that is different from the average (Boswell & Hopkins, 2009). This is particularly true when the microenvironments of soil microhabitats are depleted in nutrients which would prevent the development of a large number of individuals. Randomness of events, like motility used as a strategy to alleviate starving conditions, could make emerge a different picture than in population-based models. The same criticism could be levied to a large extent toward individual-based microscale models, which still adopt the Monod equation to describe bacterial growth, even though this usage of the equation for small groups of cells is not warranted at all, and other formalisms exist (Bermúdez et al. 1989). Nevertheless, irrespective of the modeling formalism used, encapsulating microbial growth at the single cell level allows to account for
interspecific and/or intraspecific variability as well as the influence of random effects, which can have a high influence at low microbial numbers.

In this review we voluntarily omitted to report on numerical schemes used to solve the partial differential equations (PDEs) describing biological as well as physical processes, adopting in this respect the point of view that the methods used have been intensively verified and assessed by numerical analysts (e.g., Ginzburg et al., 2008; Hecht, 2013) and it is not the purpose of this review to evaluate them, in part because benchmarks have to be developed for microscale models simulating microbial activity.

An encouraging finding is that microscale models have integrated the tremendous progress in knowledge of OM cycling in soils. Even though the vast majority of models simulate the uptake of simple monomers, when they take into account the solid form of SOM they all include a depolymerization step to produce dissolved OM that is taken up by microorganisms. The description of this process is generally less complex when actual soil architecture is considered, perhaps to keep the overall complexity of the models within computationally manageable bounds, whereas more complicated reactions of depolymerization are found when soil architecture is simplified or idealized.

Similarly, microscale models have attempted to integrate the knowledge gained on ecological interactions between species. Even though the vast majority of models simulate a single species, when they consider different species, they include ecological interactions between species from the simpler case of competition for food to more complex interactions (mutualism or commensalism). Once again, the simpler representations of ecological interactions tend to be adopted when actual soil architecture is described. We also note that none of the microscale models have so far coupled fungi and bacteria except the models of Kaiser et al. (2014) or Pagel et al. (2020), which do it through different parameterizations of growth and uptake and not through different expansion rules. As fungi and
bacteria or archea operate at different spatial scales (a fungal colony can translocate nutrients over cm’s), the combination of both in a single model is currently challenging and hampered by computational limitations.

Most of the microscale models have also integrated some of the increasing knowledge about microbial mobility in soil (e.g., Juyal et al., 2020). However, bacterial models that describe in detail the complexity of soil architecture have tended to disregard this process. No consensus has been reached in general on the type of motility that needs to be included in models, as random motion and/or chemotaxis can be found in experiments, mostly depending on the research question being addressed. Advanced descriptions of spatial distribution of soil microorganisms derived from ecological knowledge have also started to be considered in the latest generation of microscale models (Pagel et al., 2020).

Although tremendous progress has been made toward understanding the role of O₂ on microbial activity (e.g., Keiluweit et al., 2016), few studies have integrated that knowledge. Most models consider decomposition of SOM under optimal O₂ diffusion and do not describe O₂ diffusion in the liquid phase. Similarly, although the role of air-filled pores in the emission of N₂O has been demonstrated experimentally (e.g., Rabot et al., 2016), of the five microscale models that describe denitrification processes, only one includes diffusion of the emitted gases (N₂ and N₂O) in the liquid and gas phases within the soil. Whereas diffusion of the produced CH₄ is ignored at the pore-scale in virtually all models, it is considered by one 1D-upscaled model at the scale of the soil profile (Ebrahimi & Or, 2017).

Perhaps the most striking feature of existing models is the fact that the majority of them so far have not accounted for the presence of roots, despite the rhizosphere being the most active region in soil. Models simulating root surface colonization by bacteria have focused on the description of the root
itself and its growth while the rhizosphere is either considered as inactive grid nodes, in the sense that bacteria are set inactive until the root invades the soil node (Muci et al., 2012) or it is considered as a boundary influx of soil bacteria (Dupuy & Silk, 2016). Contrastedly, models addressing the transport of nutrients have had a stronger focus on the rhizosphere by using as input data 3D CT images of root hairs in soil pores and they have demonstrated the benefits of such image-based modeling approaches in understanding the uptake of phosphorus (e.g., Keyes et al., 2013; Roose et al., 2016). The development of this type of model that includes the growth and metabolism of, e.g., beneficial bacteria, pathogens or mycorrhiza is therefore urgently required to advance our understanding of the rhizosphere and open up opportunities to control it to benefit crop production.

5. Limitations of microscale models and future prospects

We identified two major limitations that constitute severe challenges to microscale modeling efforts. On one hand, available input data on soil architecture need to be sufficiently representative of reality and to capture its essential features. On the other hand, the parameterization of biotic components of these models needs to be sufficiently robust. In both respect, we shall identify research avenues that could help fill current knowledge gaps and choose the most appropriate modeling strategies.

5.1 Available input data on soil architecture

The approaches that uses imaging of soil architecture are attractive because they preserve as much as is practically feasible the real spatial arrangement of soil pores and provide realistic statistics on soil physical properties. However, different image treatments (noise reduction, conversion into low number of bits, segmentation algorithms) are applied to obtain binarized or ternary CT images on which most
of the image analyses are performed. These treatments can produce modifications of the bulk measures of pore morphology such as porosity, specific surface area, or Euler connectivity (Houston et al., 2013). Furthermore, small local changes of the size of key throats governing the water flux can dramatically affect simulated hydrodynamic properties (Gackiewicz et al., 2019; Pot et al., 2020). Imaging data are also limited by the pixel resolution attained by the different imaging tools, which can mask a significant portion of soil porosity. At best, existing microscale models simulate microbial activity at a scale of a few microns and while they can only capture a small part of the tremendous complexity of soil architecture they miss out a lot of the sub-resolution pore space in which many bacteria, archaea, and viruses reside. Baveye et al. (2017) report different strategies that can be used to consider the unresolved pores from CT data and to describe their role in the carbon transformation. Among them is the use of all the information contained in the gray-level pixels instead of using binarized images, or to mix the traditional, binary approach with gray-scale modeling, as in the work of Yan et al. (2016, 2018). However, an implicit approach is necessary to assign a value for instance to the diffusion coefficients in these pixels as a function of hydration or gas status. Multi-fractal approaches were mentioned by Baveye et al. (2017) as a possible way to simulate the statistical distribution of unresolved pores, but further analysis by Perrier et al. (2020) has recently demonstrated that their potential was limited in this respect. Although we are still far from having handle on the sub-resolution porosity, the approaches that have been explored so far at least confirm the importance of ‘hidden’ smaller pores in understanding soil processes and thereby encourage future research efforts. Recent advances in imaging techniques should bring better knowledge of pore arrangement at the infra-micron scale (Felde et al., 2020; Schlüter et al., 2019) and address their potential role in shaping soil microbial activity.
An area of research that has recently emerged is the utilization of neural networks to predict processes in rocks, using training data sets made of 3D binary CT images (Santos et al., 2020). To improve training, geometrical features such as the Euclidean or transform distance map, the maximum inscribed sphere flow, and tortuosity maps of the global paths along the flow direction are included in the training data set. Rock architecture is far more homogeneous than is soil architecture, however the convolutional neural network was able to deal to a reasonable extent with the higher heterogeneity observed in rocks (dead-end pores, disconnected pores) compared to the initial training data set of bead packs.

Another limitation comes from the fact that the spatial arrangement of soil constituents is highly dynamic, continuously changing under the effect of biotic activity (including roots, macro-, meso-, and micro-fauna) and abiotic stresses (e.g., drying-rewetting cycles). Time-resolved methods to image the dynamics of soil architecture (e.g., Bottinelli et al., 2016) provide experimental data that can potentially be used as direct model inputs. Recent attempts to simulate the evolution of soil architecture as a result of microbial activity (Ray et al., 2017) or physico-chemical processes (Rupp et al., 2019), together with research focusing on the role of roots in modifying the geometrical arrangement of soil particles (Aravena et al., 2014; Kolb et al., 2017) will certainly open new avenues to bridge knowledge gaps.

### 5.2 Parameterization of the biotic component

A major concern that has not been discussed in any of the reviewed references deals with the parametrization of the biological component. When models simulate the activity of a well-defined microorganism, parameters describing its physiology are usually extracted from experimental data measured in pure culture system. In doing so, models can only tackle a tiny part of the soil biodiversity.
Representative species of known functional groups are however modeled. A major difficulty in extracting parameters from experiments is that the underlying mathematical formulations for the growth kinetics used to describe the experimental growth curves are unevenly informative and parameters extracted from these experimental studies can thus be in conflict with the formulation adopted in a given microscale model (Vogel et al., 2018). We thus advocate to report them in experimental studies. But even so, it is far from certain that the physiology of bacteria remains the same whether the cells live freely or are attached to soil surfaces (e.g., Fletcher, 1979; Kirchman and Mitchell, 1982; Lerch et al., 2017; Ulířová & Šantrůčková, 2003). Additionally, other parameters associated to mortality and dormancy are generally obtained by calibration of macroscopic models (e.g., Ingwersen et al., 2008) to experimental data from soil incubations with composite microbiota and carbon sources (Vogel et al., 2018). These limitations render the use of these experimental data to parameterize the biotic component very questionable.

The mobility of bacteria or fungi is also very difficult to parameterize. For bacteria, at best, data on motility exist in artificial media (e.g., agar surfaces, micro-models), which may not be directly transferable to motility in soil pores. Direct measurements in soils have demonstrated the impact of soil architecture on the ability of bacteria to move (Juyal et al., 2018). These data, which describe the movement of a population of bacteria, are directly transferable to models that do not adopt an explicit description of soil architecture. For models accounting explicitly for soil architecture, this information has to be used to calibrate/validate biotic parameters accounting for microbial movement, irrespective of the formalism, either population- or individual-based.

In the case of fungi dispersal rates, hyphal expansion, and branching angles have been observed on 2D surfaces and incorporated into fungal models (Stacey et al., 2001). In theory these types of parameterization can be expanded to 3D, however it has been shown that branching angles in a 3D pore
structure can significantly deviate from those reported in literature, suggesting that expansion of fungal colonies has more complex feedback mechanisms with the environment than is currently considered in models (Soufan et al., 2018).

6 Conclusions

Given the advantages and disadvantages of the different strategies used to describe soil micro-environments, we recommend modeling benchmarks. Indeed, an explicit approach will not be able to cope with the tremendous amount of data needed, and simplified approaches conserving the key spatial arrangement of pores appears as a necessary alternative. Like Vogel et al. (2005) did in comparing a pore-network model, a lattice-Boltzmann model and a morphological model to simulate the capillary pressure of a soil sample whose topology was obtained by CT, and Pot et al. (2015) did in comparing a lattice-Boltzmann model and a morphological model to simulate experimental spatial distribution of air-water menisci as observed in synchrotron CT images, similar benchmark analyses should be carried out with respect to soil microbial activity. Pure diffusion conditions (as those existing in standardized soil incubation experiments) but also addition of intermittent water fluxes should be tested in modeling scenarios. The latter should benefit from recent progress made in simplifying modelling of water flow at pore scale for unsaturated conditions (Zhang et al., 2016).

In light of the complexity and uncertainty of soil processes as well as of imperfections of data available, we also advocate for the use of a multimodal description of soil microhabitats along with a modeling strategy involving data at different levels of description such as the so-called pattern-oriented modeling (Grimm et al. 2005). The non-destructive imaging of soil properties and processes is rapidly
evolving and we need to find new ways of combining and including data emanating from these techniques in the modeling process.

In this review we addressed the current practices used in soil science to account for soil architecture and microbial dynamics at the scale of microbial habitats and address their advantages and disadvantages. We will address in a forthcoming paper the insights gained by these models on soil functions and discuss to what extent it is important to include the explicit description of soil architecture in microscale models of soil microbial activity.

7 Fundings

This work was supported by a grant from the French ANR to project Soil μ-3D (ANR-15-CE01-0006). WO and XP received support from the National Environment Research Council (NE/P014208/1 and NE/S004920/1)

8 Data Sharing and Data Accessibility statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

9 References


Falconer, R.E., Otten, W. & White, N. (2015). Toward modeling the resistance and resilience of "Below-ground" fungal communities: A mechanistic and trait-based approach. *Advances in Applied Microbiology* 93, 1-44. DOI: 10.1016/bs.aambs.2015.08.001


microbes determines ecosystem interactions and spatial dynamics. *Cell Reports* 7, 1104-1115. DOI: 10.1016/j.celrep.2014.03.070


**Table 1:** Advantages and disadvantages of the different strategies used to describe soil micro-environments in microscale models of soil microbial dynamics. PN and OM are pore network and organic matter, respectively.
<table>
<thead>
<tr>
<th>Modeling strategy for the description of soil microenvironments</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| Explicit node attributes / actual soil architecture | 3D realistic soil architecture  
3D realistic water-air distribution  
3D realistic diffusion pathways  
3D realistic location of OM and microorganisms  
Can deliver relationships between 3D soil architecture descriptors and soil microbial response | High volume of input data  
Computations intensive  
Small-scale simulation |
| Explicit node attributes / simplified soil architecture | 3D realistic diffusion pathways  
3D realistic location of OM and microorganisms  
Larger-scale simulation  
Can deliver relationships between 3D soil architecture descriptors and soil microbial response | High volume of input data  
Approximate pore-solid interface  
Approximate air-water interface |
| Case of morphological and irregular PN models | Low volume of input data (bulk soil parameters)  
Larger-scale simulation | Approximate 3D soil architecture  
Approximate 3D diffusion pathways |
| Case of regular PN models | Low volume of input data (bulk soil parameters)  
Deliver simplified prototypes of models | Ignore actual 3D soil architecture  
Ignore actual 3D water-air distribution  
Ignore actual 3D diffusion pathways |
| Case of regular grid | | |
| Explicit node attributes / idealized soil architecture | Exact reproduction of experimental designs  
Model assessment | Approximate actual 3D soil architecture |
| Mixed attributes / actual soil architecture | 3D realistic soil architecture  
3D realistic water-air distribution  
3D realistic location of OM and microorganisms  
Take into account unresolved pores | High volume of input data  
Computations intensive  
Small-scale simulation |
| Implicit node attributes / simplified soil architecture | Low volume of input data (bulk soil parameters)  
Larger-scale simulation  
Include space and dispersal in ecological interactions | Ignore actual 3D soil architecture  
Ignore actual 3D water and air distribution  
Ignore actual 3D diffusion pathways |
Classification of pore-scale models of soil processes, according to the degree of representativeness of soil architecture (from idealized porous media to simplified architectures eventually to actual soils) and to the nature of the grid node attributes (implicit, explicit and mixed, i.e., both explicit and implicit). Implicit attributes are bulk porosity, water content and effective diffusion coefficients which values are represented by different levels of gray. Explicit attributes are local characteristics of the soil phases: solid, pores filled with air or pores filled with water, and molecular diffusion coefficients in air and water, which are represented by different colors. PN and CT are pore network and computed tomography, respectively. We added two potential examples of what would be a microscale model made of mixed attributes describing an idealized and a simplified soil architecture. The first one would be a packing of porous spheres described by explicit nodes (pore) and implicit nodes of constant local porosity and explicit nodes (the pore phase). The second one would be a random assignment of explicit nodes (pore and solid) and implicit nodes of varying local porosity that would fit the porosity of a soil sample. Microscale models using implicit attributes to describe either an idealized or a detailed soil architecture would be unrealistic cases and they are therefore indicated by the mathematical empty set symbol $\emptyset$. 
Nature of the grid node attributes

<table>
<thead>
<tr>
<th>Explicit</th>
<th>Mixed</th>
<th>Implicit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-defined grid assignment</td>
<td>Random grid assignment</td>
<td>Regular PN assignment</td>
</tr>
<tr>
<td>Packing of spheres</td>
<td>Micromodel</td>
<td>Irregular PN assignment</td>
</tr>
<tr>
<td>Genty &amp; Pia (2016)</td>
<td>adapted from Bai et al. (2012)</td>
<td>adapted from Agmon et al. (2012)</td>
</tr>
<tr>
<td>Potential example with packing of spheres</td>
<td>Potential example with a random grid assignment</td>
<td>Direct CT assignment</td>
</tr>
<tr>
<td>adapted from Van et al. (2012)</td>
<td></td>
<td>adapted from Aalto et al. (2011)</td>
</tr>
</tbody>
</table>

Degree of representativeness of soil architecture

Idealized | Simplified | Detailed

Ø | Ø