https://doi.org/10.5194/egusphere-2026-1841
Preprint. Discussion started: 13 April 2026

Modelling the effects of microbial spatial heterogeneity on soil metabolic networks

Julie-Maï Paris[1], Naoise Nunan[2], and Xavier Raynaud[2]

[1]ENS PSL, 24 rue Lhomond, 75005 Paris, France
[2]Sorbonne Université, Université Paris Cité, Univ Paris Est Créteil, CNRS, IRD, INRAE, Institut d'écologie et des sciences de l'environnement de Paris, IEES, F-75005 Paris, France

Correspondence: Xavier Raynaud (xavier.raynaud@sorbonne-universite.fr)

Abstract. Bacteria and archaea with a wide variety of metabolisms are found in huge numbers in soil and are engaged in both intra- and inter-specific interactions. The exchange and transformation of resources within the metabolic networks formed by microbes are affected by the heterogeneity of soil, as well as by the diversity and abundance of microorganisms. Therefore, depending on environmental conditions and the relative abundance of species and resources present, not all theoretically possible interactions are likely to be realized in practice. Most studies that aim to reconstruct metabolic networks in soil do not account for the potential spatial separation between organisms, caused by the heterogeneous physical structure of the soil environment, and thus only represent some kind of "average network" that may not reflect reality. Here, we further developed a simple geometric model to study how bacterial spatial distributions can alter the functioning of metabolic networks and the emerging behaviours that can arise from them. We show that the spatial distribution of bacteria impacts the transformation of resources depending on the distance to which cells can alter their environment, and that it can lead to behaviours that blur the line between cooperation and competition between bacterial species when both interact with a third partner species. We found that when bacterial density is low, have short interaction ranges and are distributed independently of one another, the use of available resources is less efficient and more variable. On the other hand, when population densities are high and interactions occur over longer distances, more complex interactions emerge. For example, apparent competitors can enter into mutualistic interactions merely due to their spatial configuration around a common partner. Overall, we show that the spatial distribution of organisms may be an important regulatory factor of the functioning of microbial communities and may determine the rates at which resources are transformed in soils.

1 Introduction

A single gram of soil can host more than 10^9 bacterial and archaeal cells[1] (Raynaud and Nunan, 2014). This very large number of microorganisms is known to give rise to a huge diversity of genomes (Torsvik et al., 1990; Ma et al., 2023; Torsvik and Øvreås, 2002), species (Delgado-Baquerizo et al., 2018)) and metabolisms (Nunan et al., 2020). In various environments, such as soils, the ocean or the gut of animals, bacteria interact and these interactions can take the form of exchanges of metabolites,

---

[1] Through this article, the use of the word bacteria or microbe can designate any unicellular species, bacteria, archaea or eukaryotes.

competition for resources or the release of signalling molecules, amongst others (Hense et al., 2007; D'Souza et al., 2018; Baveye et al., 2018; Kundu et al., 2019). These interactions result in a vast diversity of microbial communities, i.e. multi-species assemblages in which microbes live in close enough proximity to interact with each other (Konopka, 2009). They also shape the composition of communities through competition among species for resources, but also through cooperative interactions: it has been demonstrated that reciprocal cooperative behaviour among bacteria is both common and important for community assembly (Schink, 2002; Mee et al., 2014; D'Souza et al., 2014; Zelezniak et al., 2015). Metabolic interactions within communities allow cells to maintain smaller genome sizes, as they have fewer metabolic genes, and the communities are likely more resource efficient (D'Souza et al., 2018; Nunan et al., 2020; Machado et al., 2021). Such metabolic interactions may be prevalent in soils: a metabolic modelling analysis of genomes suggested that communities based on both competition and cooperation are found in soil (Machado et al., 2021). Furthermore, it has been suggested that the difficulty in culturing soil bacteria could stem from species having to acquire essential metabolites, particularly amino acids, through cross-feeding (Pande and Kost, 2017).

However, it is far from easy to determine the species and metabolites involved in metabolic networks, and results often reveal very complex networks: Kundu et al. (2019) for example, showed that the cross-feeding network of a termite (*Nasutitermes corniger*) microbiota uses 205 microbes and 265 metabolites. Furthermore, soil is a highly spatially structured environment, which is generally not taken into account in studies on interactions within bacterial communities. However, the spatial organisation of species within a synthetic community has been show to affect the intensity of the interactions among species and, thus, the stability of the community Kim et al. (2008).

In soils, pores define a complex 3D architecture containing more or less air and water depending on pore neck diameters and conditions (Kremen et al., 2005). Resources are heterogeneously distributed within or at the surface of pores. Bacteria and archaea are also heterogeneously distributed on pore surfaces (O'Donnell et al., 2007) and can only interact with each other within a short spatial range that depends, for example, on the extent of local diffusion or bacterial movement (Dal Co et al., 2020). Therefore, depending on the species and resources present locally, not all parts of potential metabolic networks may actually be expressed (Kremen et al., 2005). It is difficult to implement empirical studies of spatial effects on microbial interactions due to the opaque nature of soil and because most observations tend to destroy the integrity of the soil's spatial structure (Baveye et al., 2018).

Modelling approaches are useful for identifying and understanding the mechanisms and interactions that underlie soil processes (see for example (Allison, 2005)). Networks are convenient for modelling. A network is a set of entities potentially interacting as a system and, in network theory, are represented as nodes that are connected by edges (Biggs et al., 2015). This representation is ideal for describing metabolic networks. For instance, nitrification can be described as a linear network in which ammonium oxidizing archaea (AOA) and ammonium oxidizing bacteria (AOB) convert NH_3 to NO_2^- and nitrite oxidizing bacteria (NOB) NO_2^- to NO_3^- (Gee et al., 1990).

Here, we propose to implement a spatially explicit metabolic network in order to determine how the spatial organisation of bacterial communities affects metabolic flows. We used the model SHISHAMO (Spatial Heterogeneity In Soils, Hotspots And Micro Organisms), a spatially explicit, individual-based model, which was first introduced in Nunan et al. (2020), in order to study the impact of soil heterogeneity on a simple linear metabolic network. We supposed that increased spatial heterogeneity leads to clusters of microbes where metabolism is increased as well. We also modelled a more complex network with loops to try to reproduce purely-spatial interactions in our model.

2 Methods

2.1 Model general description

SHISHAMO gives an explicit spatial representation of soil microbial metabolic networks for which the user can set up multiple parameters to simulate a wide variety of metabolic networks and different spatial configurations of bacteria. SHISHAMO is written with R (R Core Team, 2023) and uses different libraries including SPATSTAT (Baddeley and Turner, 2005) and MATRIX (Bates et al., 2024).

2.1.1 Spatial representation

SHISHAMO represents microbes and resources in a finite 2D space. Bacterial cells and their effect on their environment are represented by discs where the disc centre is the cell location and the radius – or range – is the distance within which a cell interacts with its environment (e.g., resource acquisition or release). In the model, diffusion to and from the cell is simplified as an homogeneous spread over the disc. Units in the model are arbitrary. However, in this paper 1 pixel is equivalent to 1 μm.

2.1.2 Metabolic representation

Bacterial species are characterized by the types of metabolite they consume and release in the environment. Each species, in turn (the order of which is set by the user), consumes resources from within and releases resources into all their discs. The consumption and release depend on the species' metabolism. The products of cell metabolism are spread homogeneously across the disc. This approach allows us to represent diffusion in a simple and rapid fashion in which the complexities of diffusion and the dependency on time are removed.

SHISHAMO calculates resource transformation species by species except in the case where two species compete for the same resource. In this case, the resource is shared beforehand between the species and distributed among the different cells equally when cells belong to the same species or depending their respective metabolisms when the species are different. We set our simulations so that every cell always gets the same amount of resource from a pixel over which their discs overlap,

https://doi.org/10.5194/egusphere-2026-1841
Preprint. Discussion started: 13 April 2026

whatever the species. Steps are summarized in Figure 1.

Both linear metabolic networks (Fig 1) and cyclic networks - i.e. networks with at least one cycle - can be studied with SHISHAMO. In the case of cyclical networks, after each species has taken and transformed the available resources once, cells can then transform the product of other species if they can consume it. For example, suppose two species, S_1 and S_2 and two resources R_1 and R_2. S_1 takes up R_1 and produces R_2 and S_2 takes up R_2 and produces R_1. Once the first round of transformation, it is possible that some cells from S_2 produced R_2 that S_1 could now transform, only this time, the amount of resource available is lower because of diffusion to places outside the bacteria range (an exception could be in the case of a simulation where the whole modelled area is covered by discs). Equilibrium is always attained in linear networks, i.e. networks with no loops. For example in Figure 2, R_1 cannot be regenerated, so if the species act in the order of arrows (S_1 then S_2), no repetition of the aforementioned steps is needed. When a cycle exists, there can always be some of the initial resource regenerated by the cycle, which means that running steps once for each species may give back a situation where microbes could still transform resources. Therefore, steps have to be repeated a certain amount of times until an equilibrium is reached; diffusion ensures that less resource is available each set of repetitions. The number of step repetitions to reach equilibrium is set by the user. An example of this effect is shown in Figure S2 in the supplementary materials.

SHISHAMO does not represent time and processes such as death, growth or movement. The model does not yet represent reactions that require multiples resources, such as peptide production with two amino acids.

2.2 Model parametrisation

2.2.1 Area

All simulations were carried out in a square measuring 1000×1000 pixels. In all simulations we assumed that one resource (R_1) was homogeneously spread at start with a quantity of 1 per pixel, and none of the other resources were present.

2.2.2 Cell distribution

Three point patterns are implemented to model bacterial distributions :

- **CSR (complete spatial-randomness)**: the bacteria of each species are randomly distributed across space according to a continuous uniform distribution (Illian, 2008).
- **LGCP (Log-Gaussian Cox Processes)**: LGCP are inhomogeneous Poisson processes. A Gaussian probability field is generated according to which the points are distributed (Illian, 2008). LGCP can be used to model aggregated spatial point patterns caused by a stochastic environmental heterogeneity (Møller et al., 1998) as is found in soils (O'Brien et al., 2016; Raynaud and Nunan, 2014). In this case, an independent Gaussian field is generated for each species so that all species are independently distributed.

https://doi.org/10.5194/egusphere-2026-1841
Preprint. Discussion started: 13 April 2026

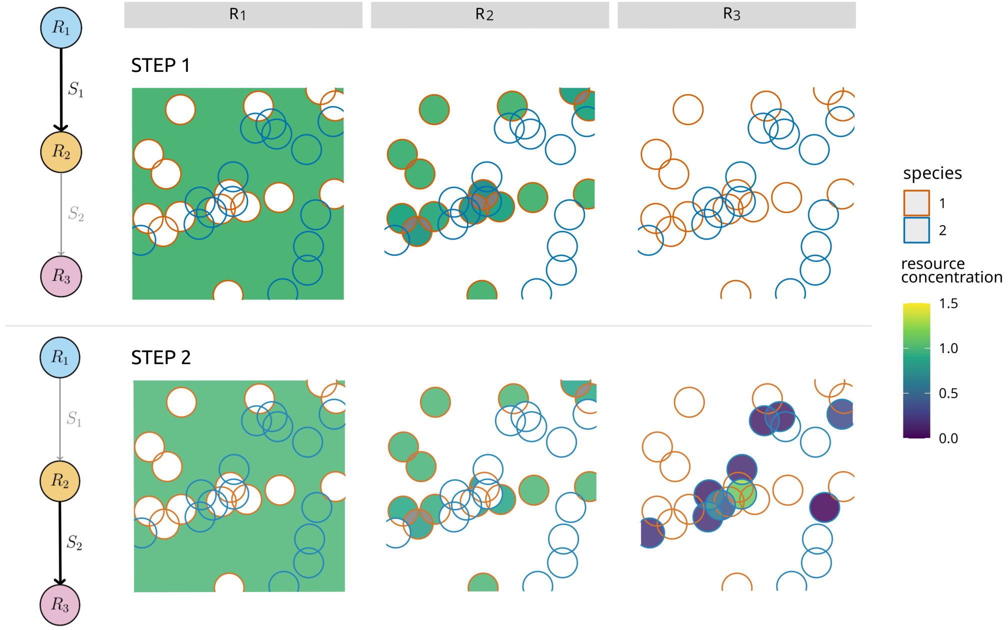

Figure 1. Diagram illustrating how SHISHAMO works in the case of a linear network with 2 species and 3 resources (R_1, R_2 and R_3) and no periodic edge-correction, with corresponding graph on the left. Resource distribution and amounts (colour) are represented from left to right. At start, R_1 is spread everywhere with a concentration of 1 per pixel. Step 1 (top): S_1 takes up all R_1 at its disposal and produces R_2. Step 2 (bottom): S_2 takes up all R_2 at disposal and produces R_3

- **Aggregated LGCP** : in this case, all species are distributed according to the same Gaussian field, so that their distributions are not independent. This is equivalent to distributing all cells with an LGCP then randomly assigning the individuals to species. In soils where bacteria tend to be aggregated (Nunan et al., 2003), this distribution can be especially relevant (Raynaud and Nunan, 2014).

The mathematical details of the model can be found in the supplementary materials. In general, aggregated LGCP generates less scattered cells than LGCP which is also more aggregated than CSR. To limit border effects, the model uses a periodic edge-correction: after cells are distributed, if a disc overlaps with the border of the area, the overlapping section is carried over to the other side, effectively turning the square into a torus. This correction is not applied directly to the field of probabilities that generates the spatial distributions of points.

2.2.3 Model runs

All simulations were launched on Sorbonne Université supercalculator MeSU (https://sacado.sorbonne-universite.fr). Because of the spatial discretization (pixels) of the discs and computer rounding errors, we assumed that, when comparing model simulations to theoretical relationships, two values where considered equal when the difference between them was less than 1

https://doi.org/10.5194/egusphere-2026-1841
Preprint. Discussion started: 13 April 2026

%.

Within each simulation, the number of repetitions to reach equilibrium was set to 10 times the number of species for the analysed networks. Convergence was always reached for all the metabolic networks presented here.

2.3 Analysis of example metabolic networks

2.3.1 Linear network

We studied the effect of the spatial distribution of cells on the behaviour of a linear network with 2 species and 3 resources (Figure 2). The three distributions of cells described above were considered. All discs had the same radius which was set to 10, 20 or 30 pixels. The number of cells per species was identical and was varied from the same 20 to 300 to study the effect of cell density. Each combination of parameters (point pattern, disc radius and number of cell per species) was run 10 times, which amounts to a total of 900 simulations.

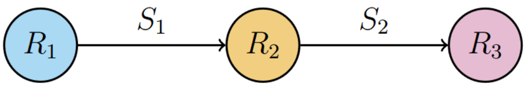

Figure 2. Graph representation of a linear network with 2 species and 3 resources

We compared the model outputs with theoretical formulas that are expected for such a network. Let A_T be the total surface of the modelled area, R_1 a resource spread homogeneously on this surface and c_s the expected area coverage of each species S_s. The amount of resource is written with a bar on top of the name of the resource (e.g. the amount of R_1 is written $\bar{R}_1$).

As the model considers that an individual cell is able to use all the resource within its disc, the theoretical amount of resources for a simple 3 resource linear network with cells distributed under CSR can be written as follows:

$$\begin{cases} \bar{R}_1 = A_T(1-c_1) \\ \bar{R}_2 = A_T c_1 (1-c_2) \\ \bar{R}_3 = A_T c_1 c_2 \end{cases} \tag{1}$$

These are the theoretical values that are expected under CSR. These equations were used to calculate the percentage difference in resource concentration between CSR and the other cell distributions. Details and demonstration about the way we obtained this formulas can be found in the supplementary material.

2.3.2 Network of two interconnected loops

Mutualism is a common behaviour in microbial networks (Machado et al., 2021) . The "eclipse dilemma" explores the case of two species competing for a same resource but that can help each other indirectly when they share a mutualistic partner. Even though there is competition for the resource produced by the partner, because the competitors supply the common mutualistic

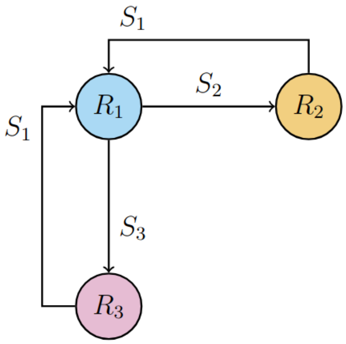

Figure 3. Graphic representations of the eclipse dilemma the way it was implemented in SHISHAMO

partner, some spatial configurations can make competitors help each others, hence the name "mutualistic competitors" (Harcombe et al., 2014).

The eclipse dilemma represents an example of a simple network that can lead to relatively complex behaviour. We represented the eclipse dilemma in SHISHAMO as a network with 2 interconnected loops: S_2 consumes R_1 and produces R_2, S_3 consumes R_1 and produces R_3, and S_1 can consume both R_2 and R_3 to regenerate R_1 (Figure 3). Such metabolic interaction network has been described for interactions between populations of *E. Coli* and *S. enterica* (Harcombe et al., 2014).

To explore how spatial heterogeneity affects the eclipse dilemma, we ran SHISHAMO with 3 species, 5 different disc radii, 10, 20, 30, 40 and 50 pixels and all three distribution patterns, CSR, LGCP and aggregated LGCP. All simulations were run twice with the same exact parameters and the same cell distributions. In a first series of simulations, we only considered the interactions between S_1 and S_2. S_3 cells were present in the model but inactivated (i.e. no production of R_3; "inactive S_3": $\circ$). In a second simulations, S_3 was active and took up R_1 and produces R_3 ("active S_3": $\bullet$). The two simulation sets allowed for the comparison of the production of R_2 with and without eclipse, while avoiding a bias due to spatial distribution randomness. The loop was run for 20 iterations and each species had 20, 47, 73 or 100 cells in different simulation runs. Each parameter combination was run 10 times, which amounts to a total of 1200 simulations.

The metabolism of S_2 was computed by calculating the sum of R_2 produced by all S_2 microbes over all time steps ($\sum_t \bar{R}_2$). For each cell, the difference $D_{\circ,\bullet}$ (%) between the situation with inactive and active S_3 was expressed as:

$$D_{\circ,\bullet} = \frac{\sum_t \bar{R}_{2,\bullet} - \sum_t \bar{R}_{2,\circ}}{\sum_t \bar{R}_{2,\circ}} \times 100 \tag{2}$$

This difference is positive if the species S_2 produces more R_2 in the presence of its mutualistic competitor S_3, negative if it produces less, and zero if no impact is observed. When an S_2 disc and an S_3 disc intersect, the two bacteria have to share resources with each other. Therefore, the theoretical negative impact of two competitors is linked to the distance between

https://doi.org/10.5194/egusphere-2026-1841
Preprint. Discussion started: 13 April 2026

them and their range, and can be estimated based on geometric considerations. The maximal loss a bacterium can suffer from the intersection with another single bacterium is -50%, i.e. they share resource between the two of them on each pixels. The theoretical % loss L can be calculated from the area of the intersection between the two discs. As we assumed that all cell discs had the same radius d, L can be calculated as a function of distance between cells as in Schnell et al. (2021):

$$L(d,r) = -50 \times \left(2r^2 \arccos\left(\frac{d}{2r}\right) - \frac{1}{2}d \cdot \sqrt{4r^2 - d^2}\right) \qquad \text{if } 0 \leq d < 2r \qquad (3a)$$

$$= 0 \qquad \text{if } d \geq 2r \qquad (3b)$$

We carried out similar simulations to compare the metabolism of all S_1 cells with active and inactive S_3. The amounts were normalized by the $\bar{R}_1$ maximum value in the simulation.

3 Results

3.1 Impact of bacterial spatial distribution on resource transformations along a linear network

Figure 4 shows the quantity of resources at the end of simulations in the case of a simple linear network with 2 species and 3 resources. When the spatial distribution of species follows complete spatial randomness, the higher the number of cells per species and the bigger the disc radius, the less R_1 is left (Figure 4a). It should be noted that doubling the radius does not multiply the consumption by 4, which can be explained by the increased probability of disc overlap among bacteria from S_1. In the case of R_2 and R_3, the more cells there are and the bigger their disc radii, the higher R_2 and R_3 concentrations are at the end. However, the quantity of R_2 reaches a maximum at about 200 cells when the range is 30 pixels. With lower ranges the shape of the curve suggests that a plateau would be reached at cell values greater than 300. This is due to the balance between the production of R_2 by species S_1 and consumption of R_2 by species S_2. When these parameters are high, S_1 covers a large area and produces a lot of R_2 but then S_1 and S_2 are more likely to intersect, increasing R_3 at the expense of R_2. This is expected because if all species were present everywhere then only R_3 would remain at the end. It can also be seen on Figure 4a. that there is little deviation between the theoretical model and the simulation.

Figures 4b. and 4c. show the difference in concentration between the theoretical CSR model and LGCP, and aggregated LGCP, respectively. In both case, the quantity of R_1 left increases by less than a percent compared with CSR. The deviations from CSR are similar for the two distributions. The concentration of R_2 decreases by a few tenths of percent relative to CSR in both LGCP distributions. The deviations vary slightly depending on the disc radii. Though the variations are small and overlap, an influence of the cell number and radius is apparent, with slight differences between LCP and aggregated LGCP. With LGCP, species are independently distributed but when an intersection occurs, because cells from a same species are aggregated, areas where they intersect are more likely to happen. Aggregated LGCP favours intersections by distributing cells according to the same probability field. Still, the variations of both compared to CSR remain under -1 %. Therefore, less R_2 will remain when

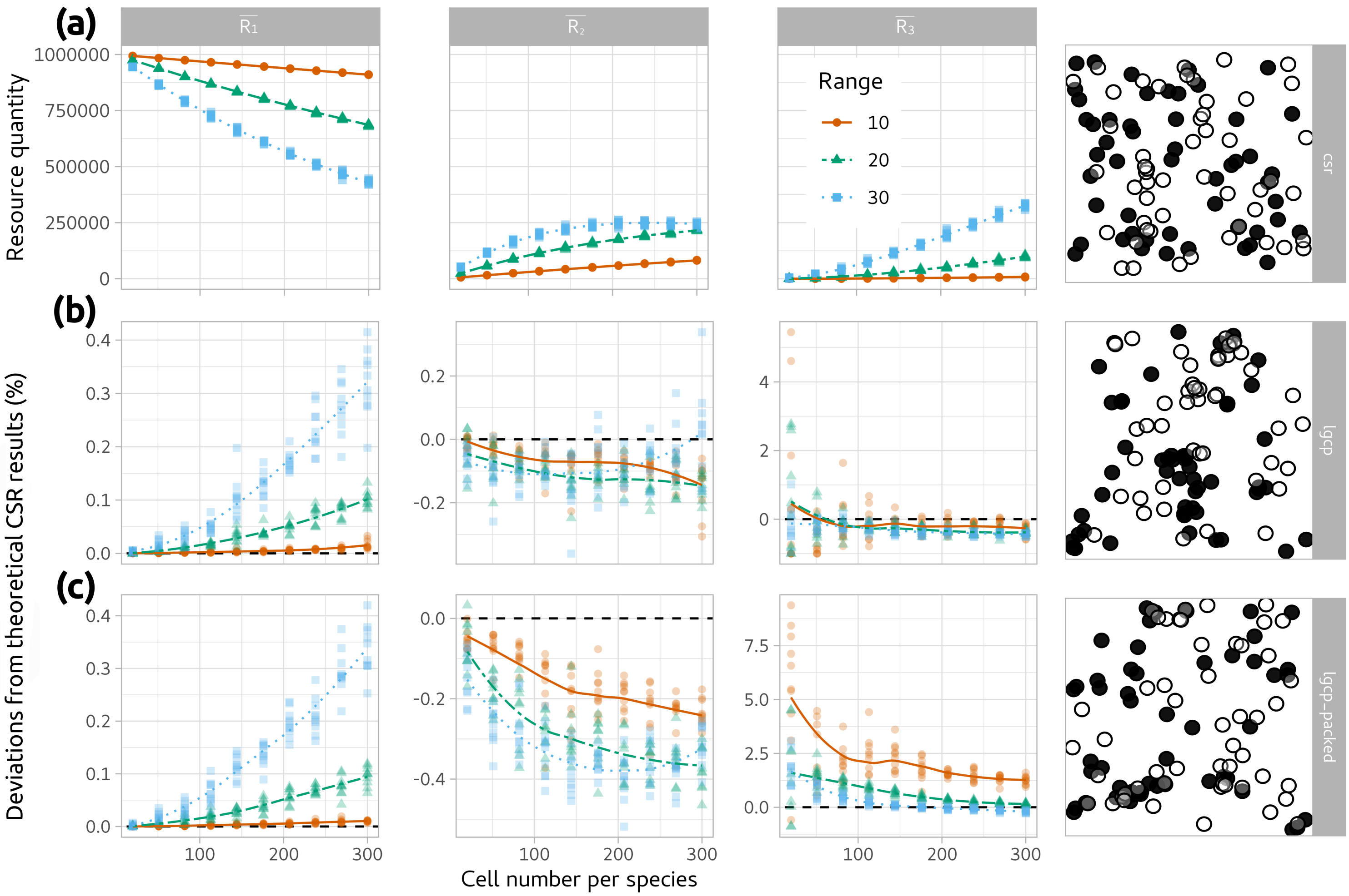

Figure 4. Results for a simple linear network (Figure 2)

From left to right: results for R_1, R_2 and R_3

From top to bottom: (a) resource quantity at the end of the simulation in a CSR distribution with lines representing the theoretical model, (b) deviations from theoretical CSR results (Equation 1) with lines being LOESS (LOcally Estimated Scatterplot Smoothing) to help visualization for LGCP and (c) aggregated LGCP.

- - ● - - : range = 10 pixels, – –▲– – : range = 20 pixels, · · ■ · · : range = 30 pixels

there are more discs with larger radii in more aggregated distributions.

The behaviours of LGCP and aggregated LGCP are more distinct for the end resource R_3. In the case of LGCP, the variation remained very close to 0 and within one percent most of the time. However, the deviations are much larger for aggregated LGCP, especially when there are fewer cells and the disc radii are smaller. This situation leads to more R_3 being produced compared to CSR. This can be explained by the fact that in aggregated LGCP, all bacteria from the two species tend to be aggregated, which is not the case for the other distributions where species are independently distributed (Section 2.3.1). Therefore, when the area covered by each species is small, this distribution favours interactions between species, but this effect is not as strong when there are a lot of bacteria as the probability of disc overlap is high, even under CSR. This also explains why the variability is higher for low ranges and cell densities, as this intersection is less probable (between -1% and +8% for a 10 pixel range with 20 bacteria per species versus around 0% for a 30 pixel range with 300 bacteria per species).

Overall, these results indicate that resource dynamics are affected by three distinct properties: the number of cells, their spatial distributions, and the range over which they interact with their environment (cell radii).

3.2 Network of interconnected loops and emergent behaviours

We present here one combination of parameters (47 cells per species, aggregated LGCP) among all the combinations we ran. The other sets of parameters gave similar results. We analyse the impacts that the S_3 activity has on both the mutualistic competitor S_2 and on their common partner S_1 for this same combination of parameters.

3.2.1 Impact on mutualistic competitor S_2

Figure 5 shows the difference in % between the situation with and without S_3, $D_{\circ,\bullet}$, against the distance between each S_2 cell and the closest S_3 bacterium ($d_{2,3}$). These values are compared to the theoretical loss of resource L (Eqs 3) to assess the impact of the eclipse. If the observed difference is below the theoretical loss curve, this means the presence of S_3 has a more negative impact on the metabolism of S_2 than expected (i.e. increased competition), and if is it above, this means that presence of S_3 has a beneficial impact on the metabolism of S_2 (i.e. mutualism).

Different areas of the graph are associated with the different behaviours. On the left of the vertical dotted line (which represents $x = 2r$), the closest S_3 disc intersect with the S_2 disc. Gray points close to the gray line means $D_{\circ,\bullet}$ follows the theoretical loss L (Equation 3). The most likely explanation is a single intersection between a pair of S_2 and S_3 cells (diagram 2. in Fig 5). If the point is above the L line, then $D_{\circ,\bullet}$ is less negative, if not positive. This means that the two bacteria have to share resources (i.e. compete), but if there is another partner, for example a S_1 bacteria which partnership with the S_3 bacterium is also beneficial for S_2 (diagram 1. in Fig 5) In this case, the benefit reduces the loss but it can also actually increase the S_2 bacterium metabolism. Below the curve, the S_2 cell has to share more resources than a single S_3 cell could take. It probably points to another S_3 cell whose centre is further from the first one's (diagram 3. in Fig 5). On the right of the dotted line, the closest S_3 disc does not intersect with S_2 cells, therefore $L = 0$ (Equation 3). If $D_{\circ,\bullet} \neq 0$, more than two cells are involved in a configuration which benefits S_2 (diagram 4. in Fig 5) or not (diagram 6. in Fig 5). Which case between 4. and 6. happens depends on how many cells are involved, and how big are the intersections. The examples drawn here are not exhaustive and there is a high complexity in interaction patterns.

https://doi.org/10.5194/egusphere-2026-1841
Preprint. Discussion started: 13 April 2026

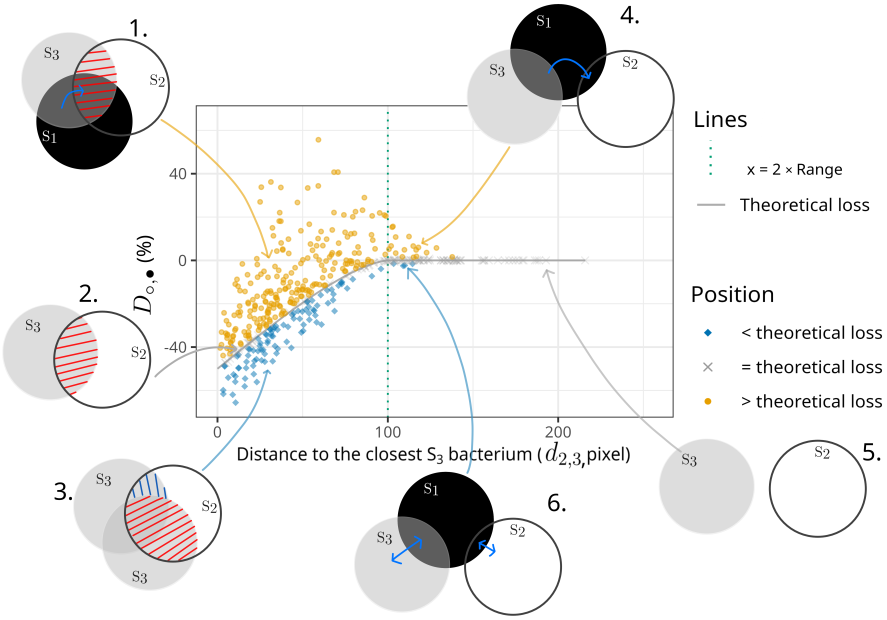

Figure 5. Impact of active S_3 on S_2 versus the distance to the closest mutualistic competitor S_3 (aggregated LGCP, $r = 50$ pixels, 47 cells per species, compilation of 10 simulations)

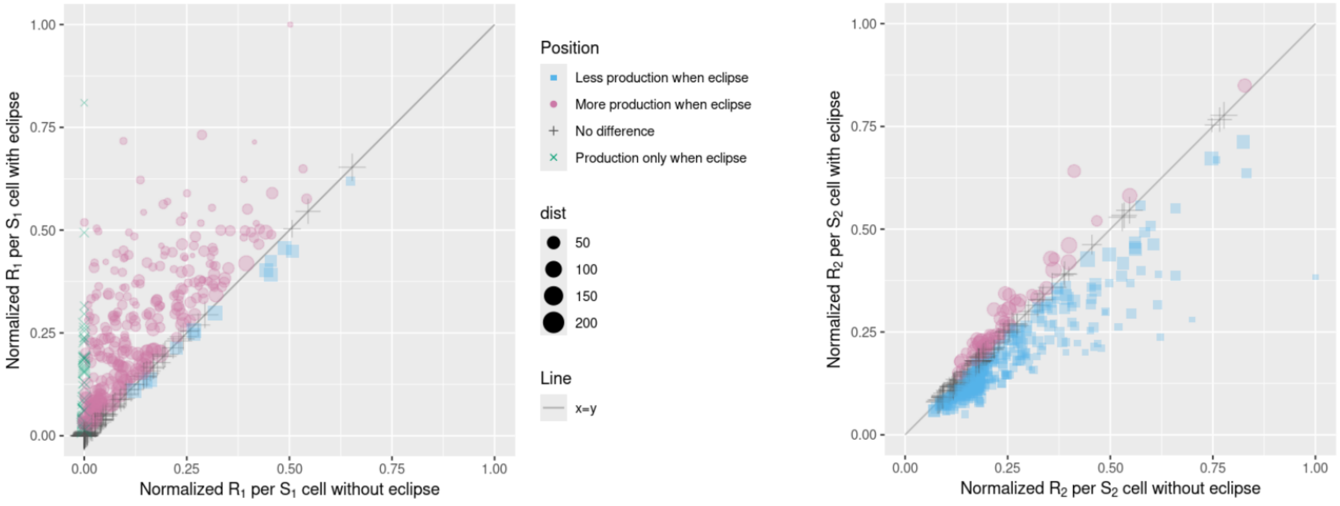

Figure 6. Impact of active S_3 on S_1 (left) and S_2 (right) (aggregated LGCP, $r = 50$ pixels, 47 cells per species, compilation of 10 simulations)

3.2.2 Impact on common partner S_1

Figure 6 presents the impact of the eclipse for the common partner S_1. Values are normalized by the highest R_1 production value over the 10 simulations for these parameters. Points on the 1:1 line correspond to cells are not impacted by the eclipse (26.8 %; 126 points). Points above the 1:1 line (60.0% of the cells; 282 points) indicate that eclipse has a positive effect, i.e. the metabolism of S_1 increased thanks to the presence of S_3. Moreover, points on the vertical axis $x = 0$ are cells for which the eclipse phenomenon was necessary for these cells to have access to resources (10.4% of the cells; 49 points). In those case, the impact tends to be stronger when the distance increases but it then occurs less often. Finally, only 2.8% (13 points) show slightly reduced metabolism (i.e. competition).

Therefore, as we expected for S_1 cells, having more cross-feeding partners have a strong positive impact on their access to resources, but in some cases, the spatial heterogeneity will cause a slight decrease, increasing the metabolism variability across space.

Figure 6 also shows another visualization of S_2 results for comparison with S_1. It uses the same data as Figure 5 but method for Figure 5 cannot be applied so S_1 because of points with a 0 value before eclipse. For S_2, values are normalized by the highest R_2 production value over the 10 simulations for these parameters. The distance to the closest S_3 cell tend to decrease the closer the point are to that lines, which shows that the eclipse effect is reduced when the distance increases, as Figure 5 shows. More than half cells produce less resources because of the presence of S_3 (56.8 %; 267 points). About a third do not show a difference above 1% (31.5 %; 148 points) but more 11.7 % (55 points) of S_2 cells has an increase in their metabolism thanks to the eclipse.

4 Discussion

4.1 Spatial heterogeneity representation in SHISHAMO

As seen in the Results, microbial spatial distributions have a strong influence on resource consumption and production, even more so when bacteria are scarce or their interaction range is low. When interaction ranges are low, bacterial cells can only acquire resources from a small region of the environment. The presence of other species within the interaction range, as would be the case in aggregated distributions, can have a large impact, either through competition or through metabolite exchanges, which can reduce or stimulate activity, respectively. Spatial heterogeneity not only modulates the types and strength of interactions that can occur (Hense et al., 2007; Kim et al., 2008), but it also underlines emergent behaviours, such as the eclipse dilemma (Harcombe et al., 2014), which SHISHAMO managed to reproduce. The review from Warrier et al. (2026) shows a good overview of the multiple impacts that spatial structure, biotic or abiotic, can have on microbe communities in all sorts of environments.

https://doi.org/10.5194/egusphere-2026-1841
Preprint. Discussion started: 13 April 2026

For our simulations, we have used Log Gaussian Cox Processes to obtain aggregated cells. Such processes have been used to adequately describes cell distribution in soil thin sections, although without addressing which species could be observed where in those sections (Raynaud and Nunan, 2014). However, soil thin sections contains both habitable (pore) and non habitable (soil matrix) environments while SHISHAMO only consider the habitable part of the environment. Other bacteria distributions processes that consider how cell reproduce and that take species into account could be tested to explore the role of spatial distributions in the model outputs.

4.2 Linking SHISHAMO to real metabolic networks

To date, the literature contains very few examples of explicit metabolic networks, even less in soils (but see Xu et al., 2019). This is due, in part, because metabolic networks tend to be very intricate and therefore hard to reconstruct (see for example the rebuilt network in Kundu et al., 2019); so it was not possible to compare the performance of SHISHAMO with empirical data. For the time being, most synthetic communities (SynComm) also remain quite complex and it is rarely possible to find their details in literature and therefore to try implementing them in SHISHAMO. More work is needed to bridge the gap between model and real systems, and it has to be carried out while taking soil heterogeneity into account to truly understand interactions between bacteria: species co-occurrence alone is insufficient to trace back the actual interactions (Goberna and Verdú, 2022).

Despite the overall complexity of real metabolic networks, some interpretation of our results can still be done with very simple networks such as nitrification, which is a linear network like the one we used in Section 3.1. At microscale, bacteria tend to be aggregated together (Raynaud and Nunan, 2014). For nitrification, this could mean that AOA/AOB tend to gather, so that less NH_3 is consumed, but because AOA/AOB and NOB could also tend to be closer to each other, the NO_2^- consumption would be more efficient. In the end, it amounts to a slight increase in NO_3^- production with less NO_2^- remaining. This means that spatial constraint by pore can make NO_3^- production more efficient. Simulations by Grundmann et al. (2001) based on soil micro-samples also suggests this intensification of the nitrification process when bacteria are clustered. More generally, for non cyclic network, the intermediate resources are found in lower quantity where bacteria are more aggregated.

4.3 Further perspectives for model results

With the help of experimental data and more precisely statistical parameters of microbial distribution in a soil sample, it may be possible to integrate results from microscopic-scale models like SHISHAMO to macroscopic scale. Statistical results from SHISHAMO or similar models to quantify how metabolic networks are affected by distributions could be then used to make the soil compartment reflects more accurately the metabolism of microbial communities (Baveye et al., 2018). Some studies already explore the links between soil structures, microbial distribution (Raynaud and Nunan, 2014) and metabolism (Kremen et al., 2005; Li et al., 2024). Maybe in the long term, this could lead to a better understanding of soils overall with applications in land surface models.

While this paper explored questions of metabolic networks and heterogeneity in a 2D configuration, SHISHAMO does not consider 3D. 3D structure of soil and pore diameters influence not only the air and water content (Kremen et al., 2005) but also microbial diversity and metabolic pathways (Li et al., 2024). This represents an extra consideration that researchers have to keep in mind when reconstructing networks, building SynComm or studying the role of heterogeneity in soil metabolic networks, as Grundmann et al. (2001) did for their own study of nitrification. Integrating pore variability and microbe heterogeneity to larger models could thus not only help improve more global models, such as soil compartment in climate models, but also help finding solutions to mitigate the impact of soil destructive agricultural practices like tilling (Li et al., 2024; Wardak et al., 2022).

Apart from spatial point distribution questions, the networks themselves are a complex topic. SHISHAMO has been written to make it easy for the user to implement the metabolic networks they work with. Metabolic networks are represented by directed graphs with nodes representing resources, and the adjacency matrixes of those graphs can be used to easily represent whole metabolic network graphs (Pavlopoulos et al., 2011). Each species is characterized by a matrix that represents its resource transformations, and can be associated to one or more arcs of the graph. It could be possible to use mathematical theorems from graph theory - more specifically network theory - and network algorithms to study new metabolic networks, especially networks with a high number of species and resources where it is complicated to detect cycles and patterns.

Finding theoretical equations could become almost inextricable for more complex cases: complex networks, resource sharing between species or multiple different reaction rates between species and metabolites. Even if theoretical formulae were found to handle such cases, applying them to an unknown network would be arduous. Given an adjacency matrix, it is memory and time consuming to define the graph structure, especially its cycles. This is possible with algorithms like Depth-first search (Lacroix et al., 2008) but takes a lot of time as the size of networks increases which is often the case in real metabolic systems. Simplification of the system by definition of stable sub-networks of bacteria could be a way to alleviate complexity (Pascual-García et al., 2020), but even then, redistribution and resource sharing between species may be a hurdle, therefore the study of complex systems step-by-step like in SHISHAMO remains relevant.

Code availability. SHISHAMO model code is available at https://github.com/xraynaud/SHISHAMO

.

Appendix A: Abbreviations and notations

A list of all abbreviations used in this article can be found in Table A1.

Table A1. List of used abbreviations and notations

Short name	Full name
AOA	Ammonium Oxidizing Archaea
AOB	Ammonium Oxidizing Bacteria
A_T	Total arena area
c_s	Expectancy of coverage proportion by species S_s
CSR	Complete Spatial Randomness
$d_{1,3}$	Distance between an S_1 bacterium and its closes neighbor from S_3
$d_{2,3}$	Distance between an S_2 bacterium and its closes neighbor from S_3
$D_{\circ,\bullet}$	Percentage difference of R_2 produced by one cell over a simulation without and with eclipse
$L(d,r)$ or L	Theoretical loss of resources 2 discs of radius r and whose centres are separated by a distance d
LGCP	Log-Gaussian Cox Processes
LOESS	LOcally Estimated Scatterplot Smoothing
NOB	Nitrite Oxidizing Bacteria
$n_{S,i,j}$	Number of bacteria from species S on the pixel (i,j)
N_s	Number of bacteria from species s
nSp	Number of species
SHISHAMO	Spatial Heterogeneity In Soils, Hotspots And Micro Organisms
S_s	Species number s
SynCom	Synthetic community
$\sum_t \bar{R}_r$	Sum of R_r produced over all time steps by a cell
r	Range or radius
R_r	Resource number r
$\circ$	No eclipse
$\bullet$	Eclipse

Author contributions. NN and XR conceived the ideas, XR and JMP developed the model, JMP analysed the model and wrote the manuscript. All authors edited the manuscript.

Competing interests. We declare that no competing interests are present.

Acknowledgements. This work benefited from the PEPR Faircarbon CANETE project (ANR-22-PEXF-0008) as well as the SoilPacMan ANR-DFG joint project (ANR project no. ANR-22-CE92-0039-01 and DFG project no. 505667051).

https://doi.org/10.5194/egusphere-2026-1841
Preprint. Discussion started: 13 April 2026

References

Allison, S. D.: Cheaters, diffusion and nutrients constrain decomposition by microbial enzymes in spatially structured environments, Ecology Letters, 8, https://doi.org/10.1111/j.1461-0248.2005.00756.x, 2005.

Baddeley, A. and Turner, R.: spatstat: An R Package for Analyzing Spatial Point Patterns, Journal of Statistical Software, 12, 1–42, https://doi.org/10.18637/jss.v012.i06, 2005.

Bates, D., Maechler, M., and Jagan, M.: Matrix: Sparse and Dense Matrix Classes and Methods, https://CRAN.R-project.org/package=Matrix, r package version 1.6-5, 2024.

Baveye, P. C., Otten, W., Kravchenko, A., Balseiro-Romero, M., Beckers, E., Chalhoub, M., Darnault, C., Eickhorst, T., Garnier, P., Hapca, S., Kiranyaz, S., Monga, O., Mueller, C. W., Nunan, N., Pot, V., Schlüter, S., Schmidt, H., and Vogel, H.-J.: Emergent Properties of Microbial Activity in Heterogeneous Soil Microenvironments: Different Research Approaches Are Slowly Converging, Yet Major Challenges Remain, Frontiers in Microbiology, 9, 1929, https://doi.org/10.3389/fmicb.2018.01929, 2018.

Biggs, M. B., Medlock, G. L., Kolling, G. L., and Papin, J. A.: Metabolic network modeling of microbial communities, WIREs Systems Biology and Medicine, 7, 317–334, https://doi.org/10.1002/wsbm.1308, 2015.

Dal Co, A., van Vliet, S., Kiviet, D. J., Schlegel, S., and Ackermann, M.: Short-range interactions govern the dynamics and functions of microbial communities, Nature Ecology & Evolution, 4, 366–375, https://doi.org/10.1038/s41559-019-1080-2, publisher: Nature Publishing Group, 2020.

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., Maestre, F. T., Singh, B. K., and Fierer, N.: A global atlas of the dominant bacteria found in soil, Science, 359, 320–325, https://doi.org/10.1126/science.aap9516, 2018.

D'Souza, G., Waschina, S., Pande, S., Bohl, K., Kaleta, C., and Kost, C.: Less is more: selective advantages can explain the prevalent loss of biosynthetic genes in bacteria, Evolution; International Journal of Organic Evolution, 68, 2559–2570, https://doi.org/10.1111/evo.12468, 2014.

D'Souza, G., Shitut, S., Preussger, D., Yousif, G., Waschina, S., and Kost, C.: Ecology and evolution of metabolic cross-feeding interactions in bacteria, Natural Product Reports, 35, 455–488, https://doi.org/10.1039/C8NP00009C, 2018.

Gee, C. S., Pfeffer, J. T., and Suidan, M. T.: *Nitrosomonas and Nitrobacter* Interactions in Biological Nitrification, Journal of Environmental Engineering, 116, 4–17, https://doi.org/10.1061/(ASCE)0733-9372(1990)116:1(4), publisher: American Society of Civil Engineers, 1990.

Goberna, M. and Verdú, M.: Cautionary notes on the use of co-occurrence networks in soil ecology, Soil Biology and Biochemistry, 166, 108 534, https://doi.org/10.1016/j.soilbio.2021.108534, 2022.

Grundmann, G. L., Dechesne, A., Bartoli, F., Flandrois, J. P., Chassé, J. L., and Kizungu, R.: Spatial Modeling of Nitrifier Microhabitats in Soil, Soil Science Society of America Journal, 65, 1709–1716, https://doi.org/10.2136/sssaj2001.1709, _eprint: https://acsess.onlinelibrary.wiley.com/doi/pdf/10.2136/sssaj2001.1709, 2001.

Harcombe, W. R., Riehl, W. J., Dukovski, I., Granger, B. R., Betts, A., Lang, A. H., Bonilla, G., Kar, A., Leiby, N., Mehta, P., Marx, C. J., and Segrè, D.: Metabolic Resource Allocation in Individual Microbes Determines Ecosystem Interactions and Spatial Dynamics, Cell Reports, 7, 1104–1115, https://doi.org/10.1016/j.celrep.2014.03.070, publisher: Elsevier, 2014.

Hense, B. A., Kuttler, C., Müller, J., Rothballer, M., Hartmann, A., and Kreft, J.-U.: Does efficiency sensing unify diffusion and quorum sensing?, Nature Reviews Microbiology, 5, 230–239, https://doi.org/10.1038/nrmicro1600, 2007.

Illian, J., ed.: Statistical analysis and modelling of spatial point patterns, Statistics in practice, John Wiley, Chichester, England ; Hoboken, NJ, ISBN 978-0-470-01491-2, oCLC: ocn180852093, 2008.

Kim, H. J., Boedicker, J. Q., Choi, J. W., and Ismagilov, R. F.: Defined spatial structure stabilizes a synthetic multispecies bacterial community, Proceedings of the National Academy of Sciences, 105, 18 188–18 193, https://doi.org/10.1073/pnas.0807935105, 2008.

Konopka, A.: What is microbial community ecology?, The ISME Journal, 3, 1223–1230, https://doi.org/10.1038/ismej.2009.88, publisher: Nature Publishing Group, 2009.

Kremen, A., Bear, J., Shavit, U., and Shaviv, A.: Model Demonstrating the Potential for Coupled Nitrification Denitrification in Soil Aggregates, Environmental Science & Technology, 39, 4180–4188, https://doi.org/10.1021/es048304z, 2005.

Kundu, P., Manna, B., Majumder, S., and Ghosh, A.: Species-wide Metabolic Interaction Network for Understanding Natural Lignocellulose Digestion in Termite Gut Microbiota, Scientific Reports, 9, 16 329, https://doi.org/10.1038/s41598-019-52843-w, 2019.

Lacroix, V., Cottret, L., Thébault, P., and Sagot, M.-F.: An Introduction to Metabolic Networks and Their Structural Analysis, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 5, 594–617, https://doi.org/10.1109/TCBB.2008.79, conference Name: IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2008.

Li, Z., Kravchenko, A. N., Cupples, A., Guber, A. K., Kuzyakov, Y., Philip Robertson, G., and Blagodatskaya, E.: Composition and metabolism of microbial communities in soil pores, Nature Communications, 15, 3578, https://doi.org/10.1038/s41467-024-47755-x, 2024.

Ma, B., Lu, C., Wang, Y., Yu, J., Zhao, K., Xue, R., Ren, H., Lv, X., Pan, R., Zhang, J., Zhu, Y., and Xu, J.: A genomic catalogue of soil microbiomes boosts mining of biodiversity and genetic resources, Nature Communications, 14, 7318, https://doi.org/10.1038/s41467-023-43000-z, 2023.

Machado, D., Maistrenko, O. M., Andrejev, S., Kim, Y., Bork, P., Patil, K. R., and Patil, K. R.: Polarization of microbial communities between competitive and cooperative metabolism, Nature Ecology & Evolution, 5, 195–203, https://doi.org/10.1038/s41559-020-01353-4, 2021.

Mee, M. T., Collins, J. J., Church, G. M., and Wang, H. H.: Syntrophic exchange in synthetic microbial communities, Proceedings of the National Academy of Sciences, 111, E2149–E2156, https://doi.org/10.1073/pnas.1405641111, 2014.

Møller, J., Syversveen, A. R., and Waagepetersen, R. P.: Log Gaussian Cox Processes, Scandinavian Journal of Statistics, 25, 451–482, https://doi.org/10.1111/1467-9469.00115, 1998.

Nunan, N., Wu, K., Young, I. M., Crawford, J. W., and Ritz, K.: Spatial distribution of bacterial communities and their relationships with the micro-architecture of soil, FEMS Microbiology Ecology, 44, 203–215, https://doi.org/10.1016/S0168-6496(03)00027-8, 2003.

Nunan, N., Schmidt, H., and Raynaud, X.: The ecology of heterogeneity: soil bacterial communities and C dynamics, Philosophical Transactions of the Royal Society B: Biological Sciences, 375, 20190 249, https://doi.org/10.1098/rstb.2019.0249, 2020.

O'Brien, S. L., Gibbons, S. M., Owens, S. M., Hampton-Marcell, J., Johnston, E. R., Jastrow, J. D., Gilbert, J. A., Meyer, F., and Antonopoulos, D. A.: Spatial scale drives patterns in soil bacterial diversity, Environmental microbiology, 18, 2039–2051, https://doi.org/10.1111/1462-2920.13231, 2016.

O'Donnell, A. G., Young, I. M., Rushton, S. P., Shirley, M. D., and Crawford, J. W.: Visualization, modelling and prediction in soil microbiology, Nature Reviews Microbiology, 5, 689–699, https://doi.org/10.1038/nrmicro1714, 2007.

Pande, S. and Kost, C.: Bacterial Unculturability and the Formation of Intercellular Metabolic Networks, Trends in Microbiology, 25, 349–361, https://doi.org/10.1016/j.tim.2017.02.015, 2017.

Pascual-García, A., Bonhoeffer, S., and Bell, T.: Metabolically cohesive microbial consortia and ecosystem functioning, Philosophical Transactions of the Royal Society B: Biological Sciences, 375, 20190 245, https://doi.org/10.1098/rstb.2019.0245, 2020.

Pavlopoulos, G. A., Secrier, M., Moschopoulos, C. N., Soldatos, T. G., Kossida, S., Aerts, J., Schneider, R., and Bagos, P. G.: Using graph theory to analyze biological networks, BioData Mining, 4, 10, https://doi.org/10.1186/1756-0381-4-10, 2011.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2023.

Raynaud, X. and Nunan, N.: Spatial Ecology of Bacteria at the Microscale in Soil, PLoS ONE, 9, e87 217, https://doi.org/10.1371/journal.pone.0087217, 2014.

Schink, B.: Synergistic interactions in the microbial world, Antonie van Leeuwenhoek, 81, 257–261, https://doi.org/10.1023/A:1020579004534, 2002.

Schnell, R., Klingwort, J., and Farrow, J.: Locational privacy-preserving distance computations with intersecting sets of randomly labeled grid points, International Journal of Health Geographics, 20, https://doi.org/10.1186/s12942-021-00268-y, 2021.

Torsvik, V., Goksøyr, J., and Daae, F. L.: High diversity in DNA of soil bacteria, Applied and Environmental Microbiology, 56, 782–787, https://doi.org/10.1128/aem.56.3.782-787.1990, 1990.

Torsvik, V. and Øvreås, L.: Microbial diversity and function in soil: from genes to ecosystems, Current Opinion in Microbiology, 5, 240–245, https://doi.org/10.1016/S1369-5274(02)00324-7, 2002.

Wardak, D. L. R., Padia, F. N., De Heer, M. I., Sturrock, C. J., and Mooney, S. J.: Zero tillage has important consequences for soil pore architecture and hydraulic transport: A review, Geoderma, 422, 115 927, https://doi.org/10.1016/j.geoderma.2022.115927, 2022.

Warrier, V., Chen, Y., Rappaport, E., Haruta, S., Youk, H., and Momeni, B.: Interplay of Spatial Structure and Interactions in Microbial Communities, Environmental Microbiology, 28, e70 262, https://doi.org/10.1111/1462-2920.70262, 2026.

Xu, X., Zarecki, R., Medina, S., Ofaim, S., Liu, X., Chen, C., Hu, S., Brom, D., Gat, D., Porob, S., Eizenberg, H., Ronen, Z., Jiang, J., and Freilich, S.: Modeling microbial communities from atrazine contaminated soils promotes the development of biostimulation solutions, The ISME Journal, 13, 494–508, https://doi.org/10.1038/s41396-018-0288-5, 2019.

Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D. R., Bork, P., and Patil, K. R.: Metabolic dependencies drive species co-occurrence in diverse microbial communities, Proceedings of the National Academy of Sciences, 112, 6449–6454, https://doi.org/10.1073/pnas.1421834112, 2015.