

Supplementary information for Modelling the effects of microbial spatial heterogeneity on soil metabolic networks

Julie-Maï Paris, Naoise Nunan, and Xavier Raynaud

1 Spatial distribution definitions

1.1 Poisson point process

A Poisson point process consists of points randomly located across an n -dimension space where all points are independently generated. A Poisson point process depends on its intensity. The intensity represents the number of points per unit of area or
5 volume. When the intensity is constant over space, the Poisson point process is qualified of homogeneous (Chiu et al., 2013) and inhomogeneous otherwise. The intensity can then be a function of space for instance (Daley and Vere-Jones, 2005; Illian, 2008).

1.2 LGCP

LGCP are inhomogeneous Poisson point processes of which the intensity is a Gaussian random measure. LGCP may take
10 different forms but here, the exponential one was chosen. LGCP depends on 3 parameters: mean (μ), variance (σ), and scale (β) of the Gaussian random measure (for multiple species, these parameters can be set independently but were always similar for all species in this study). The intensity of the point processes and the aggregation scale are determined by these three parameters. The average intensity (λ) of an LGCP is given by: $\lambda = e^{\mu + \sigma^2/2}$ (Illian, 2008).

The function *rLGCP* from R library SPATSTAT is used.

15 1.3 CSR

CSR is a special case of LGCP in which λ is constant. Thus CSR are homogeneous Poisson point processes (Illian, 2008). The function *runifpoint* from R library SPATSTAT is used (Baddeley and Turner, 2005).

1.4 Aggregated LGCP

Aggregated LGCP distributions use the intensity λ of a previously generated LGCP. The function *rpoispp* from R library
20 SPATSTAT is used (Baddeley and Turner, 2005).

2 Theoretical formulas for linear networks

Let $r \in \mathbb{R}^*$. We consider a linear network with r resources and $r - 1$ species (Figure S1). We use an arena of surface A_T . Quantity of resources are noted with a bar on top. We use this writing to distinguish generic names (no bar) from resource quantities (with a bar).

25

We consider that all cells from all species are independently distributed following a CSR law and $\forall n < r - 1, S_n$ transforms R_n into R_{n+1} .

Each species S_n occupies a proportion of space c_n between 0 and 1. Proportions have no units.

30

We initially set the quantity of R_1 at $\bar{R}_1 = \bar{A}_T$ (uniform spread of concentration 1 over the surface A_T). The total quantities for each resource over the arena at start is :

$$\begin{cases} \bar{R}_1 = \bar{A}_T \\ \bar{R}_n = 0, \text{ if } n \neq 1 \end{cases} \quad (1)$$

We can show that for a linear network, the final total quantities for each resource are:

$$35 \quad \begin{cases} n = 1, \bar{R}_1 = \bar{A}_T(1 - c_1) \\ \forall n \in [2, r - 2], \bar{R}_n = \bar{A}_T(1 - c_n) \prod_{m=1}^{n-1} c_m \\ n = r - 1, \bar{R}_{r-1} = \bar{A}_T \prod_{m=1}^{n-1} c_m \end{cases} \quad (2)$$

We detail the proof here. At start, we have:

$$\begin{cases} \bar{R}_1 = \bar{A}_T \\ \bar{R}_n = 0, \text{ ofr } n > 1 \end{cases} \quad (3)$$

S_1 transforms R_1 into R_2 . The quantity transform is proportional to the space occupied by S_1 :

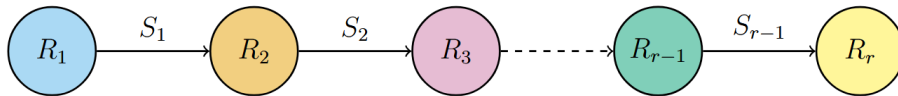


Figure S1. Graph representation of a linear network with $r - 1$ species and r resources

$$\begin{cases} \bar{R}_1 &= \bar{A}_T - \bar{A}_T c_1 = \bar{A}_T(1 - c_1) \\ \bar{R}_2 &= \bar{A}_T c_1 \\ \bar{R}_n &= 0, \text{ for } n > 2 \end{cases} \quad (4)$$

40 S_2 transforms R_2 into R_3 . The quantity \bar{R}_2 is equal to $\bar{A}_T c_1$. On the whole arena of surface \bar{A}_T , the quantity of R_2 is equal to $\bar{A}_T(1 - c_1)$, therefore, the mean surface concentration of R_2 over the arena is equal $\frac{\bar{A}_T c_1}{\bar{A}_T}$. The area covered by S_2 is equal to $A_T c_2$. Resource R_2 can only be transformed where S_2 is present. Cells are independently distributed, therefore, the mean quantity of resource transformed by S_2 is the product of the area covered by S_2 , c_2 , times the mean surface concentration in R_2 :

$$45 \quad \frac{\bar{A}_T c_1}{A_T} \times A_T c_2 = \bar{A}_T c_1 c_2 \quad (5)$$

This gives us:

$$\begin{cases} \bar{R}_1 &= \bar{A}_T(1 - c_1) \\ \bar{R}_2 &= \bar{A}_T c_1 - \bar{A}_T c_1 c_2 = \bar{A}_T c_1(1 - c_2) \\ \bar{R}_3 &= \bar{A}_T c_1 c_2 \\ \bar{R}_n &= 0, \text{ for } n > 3 \end{cases} \quad (6)$$

The previous thinking can be carried out for all next steps by recurrence.

$$\begin{cases} \bar{R}_1 &= \bar{A}_T(1 - c_1) \\ \bar{R}_2 &= \bar{A}_T c_1(1 - c_2) \\ \bar{R}_3 &= \bar{A}_T c_1 c_2(1 - c_3) \\ \dots & \\ \bar{R}_{r-1} &= \bar{A}_T c_1 c_2 \dots c_{r-2}(1 - c_{r-1}) \\ \bar{R}_r &= \bar{A}_T c_1 c_2 \dots c_{r-1} \end{cases} \quad (7)$$

50 2.1 Model convergence in circular networks

Figure S2 shows a example of convergence test for SHISHAMO.

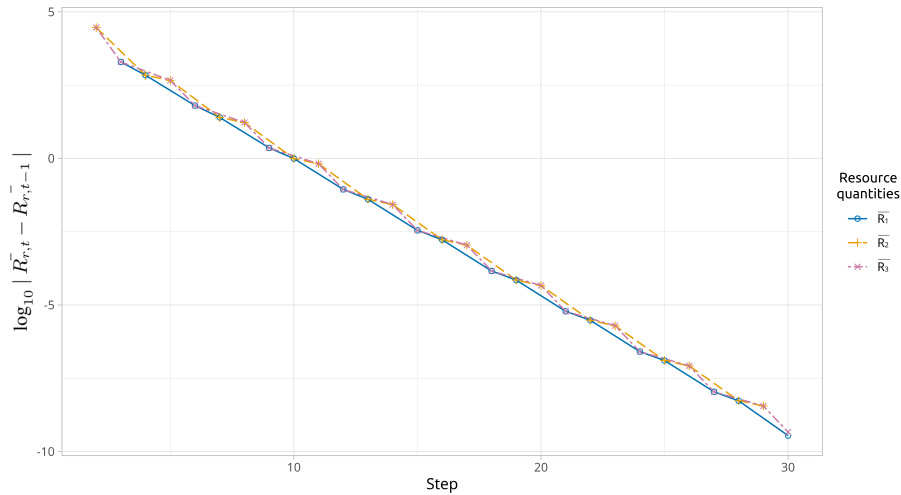


Figure S2. Test of resource convergence: \log_{10} of the absolute difference of the total amount of each resource between the current and the previous steps versus $\log_{10} |\bar{R}_{r,t} - \bar{R}_{r,t-1}|$. Simple circular network with 3 species S_1, S_2, S_3 , respectively consuming R_1, R_2 and R_3 and producing R_2, R_3 and R_1 , 30 bacteria per species, $r = 50$, CSR, 1 simulation.

Across steps, the difference of the total amount of each resource decreases exponentially. With a number of $10 \times nSp$ steps, we considered the difference small enough to end the simulation.

References

- 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.
- 55 Chiu, S. N., Stoyan, D., Kendall, W. S., and Mecke, J.: Stochastic Geometry and Its Applications, Wiley, Hoboken, 3rd ed edn., ISBN 978-1-118-65825-3, oCLC: 852758220, 2013.
- Daley, D. J. and Vere-Jones, D.: An introduction to the theory of point processes. Vol. 1: Elementary theory and methods, vol. 1 of *Probability and its applications*, Springer, New York, NY, 2. ed., corr. print edn., ISBN 978-1-4757-8109-0 978-0-387-95541-4, 2005.
- 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.
- 60