Size-dependent clustering in phytoplankton communities in a three-dimensional environment

Coralie Picoche, William R. Young, Frédéric Barraquand

Institute of Mathematics of Bordeaux, University of Bordeaux and CNRS, Talence, France

Abstract

The coexistence of multiple phytoplankton species despite their reliance on similar resources is often explained with mean-field models assuming mixed populations. In reality, observations of phytoplankton indicate aggregation at all scales, possibly hindering competitive exclusion if patches are mostly monospecific, which reduces the probability of competing with other species. Here, we present a three-dimension, individual-based model including stochastic birth and death processes as well as a simplified model of viscosity and turbulence at the microscale. Parameters are based on phytoplankters' ecological and physical characteristics, which depend on the size of the organisms. We extract the spatial moment equations from the model and show a good fit between theoretical values and simulations. The hydrodynamics of the environment is key in our model: when the environment is turbulent, species are more mixed than in a viscous-only environment. Defining a zone of possible interactions as the overlap between nutrient depletion volumes around a particle, we show that the composition of the interacting neighbourhood is size-dependent. Larger cells are surrounded by cells from other species, while smaller organisms remain in mostly monospecific clusters. Intraspecific interactions are therefore more likely than interspecific interactions for smaller organisms, which is a major explanation of diversity maintenance. Our model should be taken as a first step towards multi-specific spatial models for phytoplankton to identify mechanisms behind coexistence: with basic ecological and hydrodynamics processes, we show that organisms of the same species do aggregate, but still encounter other species within range of interaction.

Keywords: aggregation; coexistence; individual-based model; phytoplankton; spatial moment equations; spatial point process

*corresponding author: `coralie.picoche@u-bordeaux.fr`

Introduction

Phytoplankton communities are among the most important photosynthetic groups on Earth, being at the bottom of the marine food chain, and responsible for half the global primary production (Field *et al.*, 1998). Their contribution to ecosystem functions is only matched by their contribution to biodiversity. Indeed, phytoplankton communities are characterized by a surprisingly high biodiversity, up to dozens of species within multiple genera and even classes in volumes as small as 10 mL. This observation has led to the formulation of the so-called "paradox of the plankton" (Hutchinson, 1961), which refers to the conflict between the observed diversity of species competing for similar resources in a seemingly homogeneous environment, and theoretical expectations of a few species outcompeting the others. Phytoplankton models for coexistence are now almost as diverse as their model organisms (Record *et al.*, 2014) and benefit from the support of coexistence theory which proposes mechanisms to escape from competitive exclusion, including non-linear responses to a temporally fluctuating environment (Li & Chesson, 2016; Chesson, 2018). Most of these models consider mean-field dynamics in a non-spatial setting, i.e. keep the premise of a homogeneous habitat. Such assumption needs to be challenged, as we know that spatialized models including a heterogeneous environment and/or non-mixed populations can lead to highly different outcomes from their nonspatial counterparts (Law *et al.*, 2003), or even from mean-field spatial models which do not take into account the limitation in interaction and movement ranges (Plank & Law, 2015). This is even more relevant as field observations themselves have higlighted phytoplankton patchiness for more than a century (Bainbridge, 1957), from the macro- to the micro-scale (Leonard *et al.*, 2001; Doubell *et al.*, 2006; Font-Muñoz *et al.*, 2017).

Phytoplankton patchiness can at least be partly explained by the hydrodynamics of their environment: their size being mostly below the size of the smallest eddy (i.e., the Kolmogorov scale) in a typical environment such as the ocean, they are embedded in viscous micro-structures at the individual level (Peters & Marrasé, 2000) while being displaced by a common flow at the larger scale, due to turbulence (Martin, 2003). Viscosity decreases diffusivity, which can in turn generate a heterogeneous landscape by limiting movements and mixing. At the other end of the hydrodynamics spectrum, studies on the fluid mechanics of phytoplankton show that a turbulent flow, sometimes associated with other phenomena such as motility can encourage the formation of clusters (Reigada *et al.*, 2003; Durham *et al.*, 2013; Arrieta *et al.*, 2015; Breier *et al.*, 2018). Turbulence also modifies phytoplankton interactions with their environment (Guasto *et al.*, 2012), by increasing nutrient consumption (Peters *et al.*, 2006) and encounter rates with either potential partners or predators (Kiørboe, 2018), therefore affecting organism reproduction.

This leads us to consider demography in the context of environmental variation created by hydrodynamics processes. As hydrodynamics processes generate heterogeneity in both the distributions and the environment at the scale of the individuals, individual-based, or agent-based, models can provide a better depiction of the behaviour of these organisms at the microscale (Hellweger & Bucci, 2009). In this framework, population growth is a result of all individual births, deaths and survivals, which can in turn create a patchy environment as individuals are born close to their parents which can lead to clusters of 'kin' (Young *et al.*, 2001). When hydrodynamics processes

are considered, provided that the separation between parents and offsprings induced by the flow is not too fast to offset family aggregation, i.e. if turbulence and/or diffusion are not too high, this mechanism maintains and enables individuals to form aggregates (Young *et al.*, 2001; Bouderbala *et al.*, 2018). Observed patchiness can therefore be reproduced in models taking into account basic mechanisms governing the life of phytoplankton organisms. This spatial heterogeneity can affect the dynamics at the community-level: even when competitors are equivalent, the combination of local dispersal and competition stabilizes populations when mortality rates are density-dependent, as individuals tend to survive longer when they are rarer (Detto & Muller-Landau, 2016). Intraspecific clustering increases the likelihood of a particle interacting with an organism of its own species (a conspecific) instead of a different species (a heterospecific). Knowing that a high intra-to-interspecific interaction strength ratio is associated with coexistence in models (Levine & HilleRisLambers, 2009; Barabás *et al.*, 2017) and often observed in the field (Adler *et al.*, 2018; Picoche & Barraquand, 2020), species-specific clustering might be key to diversity maintenance.

Many spatial models of individuals at the microscale already show the clustering of organisms within a single species (Young *et al.*, 2001; Birch *et al.*, 2007; Bouderbala *et al.*, 2018; Breier *et al.*, 2018). Models which do include multiple species often compare relative clustering of only two sets of organisms with opposite characteristics, e.g. increase or decrease in density with turbulence (Borgnino *et al.*, 2019; Arrieta *et al.*, 2020) or the ability to move by itself or not (Durham *et al.*, 2013). In order to understand the maintenance of highly diverse communities, we need to consider models with more species than two opposite archetypes. To our knowledge, Benczik *et al.* (2006) have modeled the only individual-based model including turbulence, but not diffusion, for up to 5 species, in which they show that physical characteristics of the cell (the inertia related to the size of the organism) allow for aggregation and coexistence. While taking diversity of organisms into account is the first step towards studying coexistence in spatial models, the paucity of studies with multiple species can be explained by the computational cost of modeling many, possibly interacting, individuals, especially because very large numbers are required for phytoplankton organisms which can reach concentrations up to 10^6 C/L. Numerical costs could however be partially offset by favoring mathematical analyses. Indeed, the random distributions of events emerging from a series of stochastic processes that constitute the type of model we describe here are known as spatial point processes, and can be studied by spatial statistics, via the evaluation of aggregated quantities such as correlations between pairs of points. (Illian *et al.*, 2008).

Here, we develop the Brownian Bug Model (BBM) from its previous two-dimension, monospecific version (Young *et al.*, 2001) to three-dimensions and multiple species. This individual-based model includes turbulence and diffusion, as well as births and deaths with parameter values adapted to the physical properties of the environment and the organisms we study. We consider two types of communities, which are either made of nanophytoplankton (small cells) or microphytoplankton (large cells), and populate them with 3 to 10 different species. As every realization of our model can be seen as the superposition of different point processes (one per species), we are able to provide both numerical simulations and analytical formula to quantify the degree of intra- and interspecific clustering

of organisms, via correlations between positions of organisms. To do so, we derive derive the spatial moments of the model (Plank & Law, 2015), which can be seen as the extension of moments of random variables. This enables us to characterize the composition of the neighbourhood of a particle at different distances from it, and use this information to make informed guesses about mechanisms supporting diversity within different communities of phytoplankton.

Models and spatial statistics

Brownian bug model

The Brownian Bug Model (BBM) describes the dynamics of particles going through demographic processes in a turbulent and viscous environment, in continuous space and time. It has been developed in its two-dimension, monospecific version in Young *et al.* (2001), which we now extend to three dimensions and to S species.

In this model, we consider a community of particles, each individual being characterized by its species i and its position $\mathbf{x}^T = (x\, y\, z)$. Within a given community, all species are equivalent and share the same parameters. The population dynamics are modeled by a linear birth-death process with birth rate λ and death rate μ. Each particle independently follows a Brownian motion with diffusivity D, and is advected by a common stochastic flow modeling the turbulence with stretching parameter γ, meaning that the separation $s(t)$ between two points previously on top of each other follows $s(t) \propto e^{3\gamma t}$. We focus here on ecologically relevant quantities which can be extracted from this model, both analytically and numerically.

For numerical simulations, this model needs to be discretized. During each time step of duration τ, events unroll as follow:

1. demography: each particle can either reproduce with probability $p = \lambda\tau$ (forming a new particle of the sames species i at the same position $\mathbf{x}$), die with probability $q = \mu\tau$, or remain unchanged with probability $1-p-q$.

2. diffusion: each particle moves to a new position $\mathbf{x}(t+t') = \mathbf{x}(t) + \delta\mathbf{x}(t)$ where each element of $\delta\mathbf{x}(t)$ follows a Gaussian distribution $\mathcal{N}(0, \Delta)$ with $D = \frac{\Delta^2}{2\tau}$ the diffusivity.

3. turbulence: each particle is displaced by a turbulent flow, following the Pierrrehumbert map (Pierrehumbert, 1994), adapted in its three-dimension version (Ngan & Vanneste, 2011).

$$\begin{cases} x(t+\tau) &= x(t+t') + U\tau/3 \cos\left(ky(t+t') + \phi(t)\right) \\ y(t+\tau) &= y(t+t') + U\tau/3 \cos\left(kz(t+t') + \theta(t)\right) \\ z(t+\tau) &= z(t+t') + U\tau/3 \cos\left(kx(t+\tau) + \psi(t)\right) \end{cases}$$

where U is the maximum velocity of the particle, $k = 2\pi/L_s$ is the wavenumber for the flow at the length scale L_s (see below) and $\phi(t)$, $\theta(t)$, $\psi(t)$ are random phases uniformly distributed between 0 and 2π.

Particles are distributed in a cube of side L, with periodic boundary conditions.

Characterization of the spatial distribution

Let W be the observation window we are working on (here, the whole space, as we do not subsample the space hereafter). The state of the system at time t can be described as a collection of S populations, where the population of species i is made of k_i particles randomly distributed in W, with positions $\boldsymbol{X}_i = [\boldsymbol{x}_{1,i}, \boldsymbol{x}_{2,i}, ...\boldsymbol{x}_{k_i,i}]$. It is a realization of the spatial point process deriving from a stochastic, spatialized individual-based model such as the Brownian Bug Model. One of the most common methods to describe a spatial point process is through its moments, that can be theoretically derived and allow us to check our simulations. However, the spatial moments of a process are merely mathematical quantities which then need to be related to ecological processes. This is the role of the dominance index, which we detail afterwards.

Spatial moments

The first moment is the intensity of the process, or concentration of particles $C_i(t) = \frac{N_i(W)}{V(W)}$ where $N_i(W)$ is the number of particles of species i in the cube and $V(W) = L^3$ is the volume of the cube; it does not give any information regarding the spatial distribution, and possibly spatial correlation between particles.

The second moment, hereafter referred to as the pair correlation function $g(r,t)$, or pcf, can be thought of as the mean number of points at distance r of a particle. If we define $C_i(t)$ and $C_j(t)$ the concentrations of species i and j respectively, and $P_{ij}(r,t)$ the probability of finding a particle of species i in the sphere dV_1 and a particle of species j in the sphere dV_2 where the centers of dV_1 and dV_2 are separated by distance r, we can write:

$$P_{ij}(r,t) = C_i(t)C_j(t)dV_1dV_2g_{ij}(r,t) \tag{1}$$

The pcf is commonly used to describe spatial point patterns (see formula for other standard processes in the Supplementary). In the Brownian Bug Model, when the birth rate λ is the same as the mortality rate μ, the pair density $G(r)$ (where $G(r) = C^2g(r)$) is a solution of eq. 2 (see (Picoche *et al.*, XXX) for detailed explanations on this derivation).

$$\frac{\partial G}{\partial t} = \frac{2D}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial G}{\partial r}\right) + \frac{\gamma}{r^2}\frac{\partial}{\partial r}\left(r^4\frac{\partial G}{\partial r}\right) + 2\lambda C\delta(\boldsymbol{r}) \tag{2}$$

By integration of eq. 2, the intraspecific pcf g_{ii} follows eq. 3 (see the Appendices). The system stabilizes in the presence of advection, but depends on time in its absence.

$$g_{ii}(r,t) = \begin{cases} 1 + \frac{\lambda}{4D\pi rC_i}\left\{1 - \mathrm{erf}\left(\frac{r}{\sqrt{8Dt}}\right)\right\} & \text{for } U = 0 \\ 1 + \frac{\lambda}{2\pi C_i}\left(\frac{\sqrt{\gamma}\arctan\left(\frac{\sqrt{\gamma}r}{\sqrt{2D}}\right)}{2^{3/2}D^{3/2}} + \frac{1}{2Dr} - \frac{\pi\sqrt{\gamma}}{2^{5/2}D^{3/2}}\right) & \forall U > 0 \end{cases} \tag{3}$$

As populations of different species do not directly interact, they constitute a random superposition of stationary point processes (Ilian 2008) and $g_{ij}(r,t) = 1 \; \forall i \neq j, U$.

Related to the pair correlation function is the Ripley's K function, through equation eq. 4.

$$g(r) = \frac{K'(r)}{4\pi r^2} \tag{4}$$

The Ripley's K-function $K(r)$ is the average number of points surrounding a particle within a sphere of radius r (Illian *et al.*, 2008).

$$\forall r \geq 0, \; K(r) = \frac{1}{C}\mathbb{E}\left(N\left(b(o,r)\backslash\{o\}\right)\right) \tag{5}$$

where C is the concentration of particles and $N\left(b(o,r)\backslash\{o\}\right)$ is the number of points of the process N in the sphere of radius r centred on o, not counting o itself.

This definition can be extended to multivariate processes. $C_j K_{ij}(r)$ is the mean number of points of species j in a sphere of radius r centred on a point of species i.

$$\forall r \geq 0, \; K_{ij}(r) = \frac{1}{C_j}\mathbb{E}_i\left(N_j\left(b(o,r)\right)\right) \tag{6}$$

where $\mathbb{E}_i$is the mean with respect to points of species i.

Combining eq. 3 and eq. 4, we can show that:

$$K_{ii}(r,t) = \begin{cases} \frac{4}{3}\pi r^3 + \frac{\lambda}{C_i D}\left(\frac{r^2}{2} - \frac{1}{2}erf(\frac{r}{\sqrt{8Dt}})(r^2 - 4Dt) - \frac{\sqrt{2Dt}r}{\sqrt{\pi}}e^{-r^2/8Dt}\right) & \text{for } U = 0 \\ \frac{4}{3}\pi r^3 + \frac{2\lambda}{C_i}\left(\frac{r^2}{6D} + \frac{\sqrt{\gamma}r^3 \arctan(\sqrt{\frac{\gamma}{2D}}r)}{6\sqrt{2}D^{3/2}} + \frac{\log\left(\gamma\frac{r^2}{2D}+1\right)}{6\gamma} - \frac{\sqrt{\gamma}\pi r^3}{12\sqrt{2}D\sqrt{D}}\right) & \forall U > 0 \end{cases} \tag{7}$$

For random superposition of stationary point processes, $K_{ij}(r,t) = \frac{4}{3}\pi r^3$.

Dominance index

The dominance index (defined in Table S1 in the Supporting Information of Wiegand *et al.*, 2007) is the ratio between conspecifics and all individuals surrounding a given particle.

Let $M_{i\cdot}(r)$ be the average number of individuals within a circle of radius r around an individual of species i, which can also be written with Ripley's K function as $M_{i\cdot}(r) = C_{\cdot} K_i(r)$. $M_{ii}(r)$ corresponds to the conspecific neighbourhood and $M_{io}(r)$ corresponds to individuals of all other species. We can then define D_i with eq. 8.

$$\begin{aligned} D_i(r) &= \frac{M_{ii}(r)}{M_{ii}(r)+M_{io}(r)} \\ &= \frac{C_i K_{ii}(r)}{\sum_{j=1}^{S} C_j K_{ij}(r)} \end{aligned} \tag{8}$$

When individuals of the same species i tends to cluster, $D_i(r)$ tends to 1 while it tends to the proportion of individuals of species i in the whole community when the distribution is uniform.

Using eq. 7 and 8, we can find a theoretical formula for the dominance index:

$$D_i(r,t) = \begin{cases} \dfrac{C_i\left[\frac{4}{3}\pi r^3 + \frac{\lambda}{C_i D}\left(\frac{r^2}{2} - \frac{1}{2}\operatorname{erf}(\frac{r}{\sqrt{8Dt}})(r^2-4Dt) - \frac{\sqrt{2Dt}r}{\sqrt{\pi}}e^{-r^2/8Dt}\right)\right]}{\sum_{j=1}^{S} C_j \frac{4}{3}\pi r^3 + \frac{\lambda}{D}\left(\frac{r^2}{2} - \frac{1}{2}\operatorname{erf}(\frac{r}{\sqrt{8Dt}})(r^2-4Dt) - \frac{\sqrt{2Dt}r}{\sqrt{\pi}}e^{-r^2/8Dt}\right)} & \text{for } U = 0 \\ \dfrac{C_i\left[\frac{4}{3}\pi r^3 + \frac{\lambda}{3C_i D}\left(r^2 + \frac{\sqrt{\gamma}r^3 \arctan(\sqrt{\frac{\gamma}{2D}}r)}{\sqrt{2D}} + \frac{D\log\left(\gamma\frac{r^2}{2D}+1\right)}{\gamma} - \frac{\sqrt{\gamma}\pi r^3}{2\sqrt{2D}}\right)\right]}{\sum_{j=1}^{S} C_j \frac{4}{3}\pi r^3 + \frac{\lambda}{3D}\left(r^2 + \frac{\sqrt{\gamma}r^3 \arctan(\sqrt{\frac{\gamma}{2D}}r)}{\sqrt{2D}} + \frac{D\log\left(\gamma\frac{r^2}{2D}+1\right)}{\gamma} - \frac{\sqrt{\gamma}\pi r^3}{2\sqrt{2D}}\right)} & \forall U > 0 \end{cases} \tag{9}$$

Parameters

We model two types of organisms: microphytoplankton (defined by a diameter between 20 and 200 µm, here 50 µm) and nanophytoplankton (defined by a diameter between 2 and 20 µm, here 3 µm). These two groups are characterized respectively by a low diffusivity, slow growth and lower concentration vs. high diffusivity, fast growth and higher concentration. Particles are displaced by a turbulent fluid whose velocity defines the time scale of the discretized model: we give here the reasoning behind parameter values, keeping in mind that our model can only be semi-quantitative. Main parameter definitions and values are given in Table 1.

Advection

We first consider the advection process, due to the turbulence of the environment. We only consider the Batchelor-Kolmogorov regime, i.e the space size is below the size of the smallest eddy, but above the smallest scale of fluctuations in nutrient concentrations. The defining scale of the environment therefore corresponds to a Reynolds number $Re \approx 1$.

$$Re = \frac{V}{k\nu} \approx 1$$

where $\nu = 10^{-6}$ m^2/s is the kinematic viscosity for water. The smallest wavenumber k corresponds to the largest length scale L_s (Kolmogorov scale), i.e. $k = 2\pi/L_s$, with $L_s \approx 1$ cm in the ocean (Barton *et al.*, 2014).

$$\begin{aligned} 1 &\approx \frac{V L_s}{2\pi\nu} \\ U &\approx \frac{2\pi\nu}{L_s} \end{aligned}$$

This means that U=6.3×10^{-4} m.s^{-1}=5.4×10^3 cm.d^{-1}. Using $U\tau/3 = 0.5$ cm as in Young *et al.* (2001), we have $\tau = 2.8 \times 10^{-4}$ d= 24 s. When $U\tau/2 = 0.0$, i.e. the environment is only diffusive, we keep the same value for τ.

The advection parameter in continuous time, γ, is computed through simulations, based on $s(t) \propto e^{3\gamma t} \rightarrow 1/3 \ln(s(t)) = \gamma t$ with $s(t)$ the separation between pairs of particles before stabilization. γ is estimated as the slope of $1/3 \langle \ln(s(t)) \rangle = f(t)$ with $\langle \ln(s(t)) \rangle$ being the average separation between 800 pairs of particles. For $U\tau/3 = 0.5$ cm, $\gamma = 1231$ s^{-1}.

Diffusion

If we use the Stokes-Einstein equations (Einstein, 1905, cited from Dusenbery, 2009), diffusivity can be computed according to the formula:

$$D = \frac{RT}{N_A} \frac{1}{6\pi\eta a} \tag{10}$$

where $R = 8.314$J.K^{-1}.mol^{-1} is the molar gas constant, $T = 293K$ is the temperature, $N_A = 6.0225 \times 10^{23}$ is Avogadro's number, $\eta = 10^{-3}$m^{-1}.kg.s^{-1} is the dynamic viscosity of water and a is the radius of the organism considered.

Using $D = \frac{\Delta^2}{2\tau}$,

$$\begin{aligned} \Delta &= \sqrt{2\tau D} \\ \Leftrightarrow \Delta &= \sqrt{\frac{RT}{N_A} \frac{\tau}{3\pi\eta a}} \end{aligned}$$

We consider $a_n = 1.5$ µm for nanophytoplankton particles and $a_d = 25$ µm for microphytoplankton particles, which allows us to compute Δ_n and Δ_d (see Table 1).

Ecological processes

We study the community at equilibrium, with the birth rate equal to the death rate, i.e. $p = q$. We use a diatom doubling rate of 1d^{-1} (Bissinger *et al.*, 2008) and consider the fastest-growing nanophytoplankton particles, corresponding to a diameter of 3 μm (Bec *et al.*, 2008), for which the doubling rate is between 2 and 3 d^{-1} (set to 2.5 d^{-1} here).

Parameter	Definition	Value
p_d, q_d	Probability of reproducing/dying for diatom particles	2.8×10^{-4}
p_n, q_n	Probability of reproducing/dying for nanophytoplankton particles	6.9×10^{-4}
$U\tau/3$	Stretching parameter proxy	$\{0, 0.5\}$ cm
Δ_d	Diffusion parameter for diatoms	6.4×10^{-5} cm
Δ_n	Diffusion parameter for nanophytoplankton particles	2.6×10^{-4} cm

Table 1: Definitions and values of the main parameters used in the three-dimension Brownian Bug Model, assuming the duration of a time step τ is 24 seconds.

Range of interaction

As we examine particle aggregation or segregation and their potential effects on interactions between species, we have to ascertain the volume around which a particle can be affected by the presence of other particles, or affect other particles. We only consider here interactions due to competition for nutrients, and therefore need to define a nutrient depletion volume. We approximate this volume as the sphere of radius r where $C(r) \leq 90\% C_\infty$ with C_∞ the background concentration of the nutrient (hereafter, the concentration boundary layer). The radius is maximized when the particle is in stagnant water (diffusion is the only hydrodynamics process). In this case, it corresponds to 10 times the radius of the particle (Jumars *et al.*, 1993; Karp-Boss *et al.*, 1996). An overlap between depletion zones can therefore happen at distances as far as 10 times the diameter of a particle. We consider this maximum value as our baseline, keeping in mind that turbulence reduces the size of the concentration boundary layer and cause an increasing nutrient flux to the cell (Arnott *et al.*, 2021), but determination of the exact shape of the nutrient depletion volume in the presence of turbulence is too complex to be adressed here (Karp-Boss *et al.*, 1996).

We consider a volume of 1000 cm^3 for diatoms and 10 cm^3 for nanophytoplankton (volumes are adapted to balance realistic concentrations and computation time) with periodic boundary conditions. Particles are uniformly distributed in the cube at the beginning of the simulation. We run an idealized simulation with 3 species with an even abundance distribution of about 10^4 C/L for diatoms (Picoche & Barraquand, 2020) and 10^6 C/L for nanophytoplankton particles (Edwards, 2019). We then model a more realistic community with 10 species having a skewed abundance distribution (between 55 000 and 400 particles C/L for microphytoplankton, according to observations of field abundance distributions in Picoche & Barraquand, 2020, and multiplied by 10^2 for nanophytoplankton). All simulations are run for 1000 time steps of duration τ.

Results

We show an example of nanophytoplankton spatial distributions with and without advection at the end of a simulation in Fig. 1: clustering is not visible to the naked eye, even when zooming in the observation volume, in the presence of advection, but removing turbulence helps visualising small aggregates of conspecifics. Microphytoplankton distributions are not so easy to analyse as no clusters can be detected from basic observations, whether advection is included or not (Fig. SX). Statistics are therefore needed to go further in detecting patterns of aggregation.

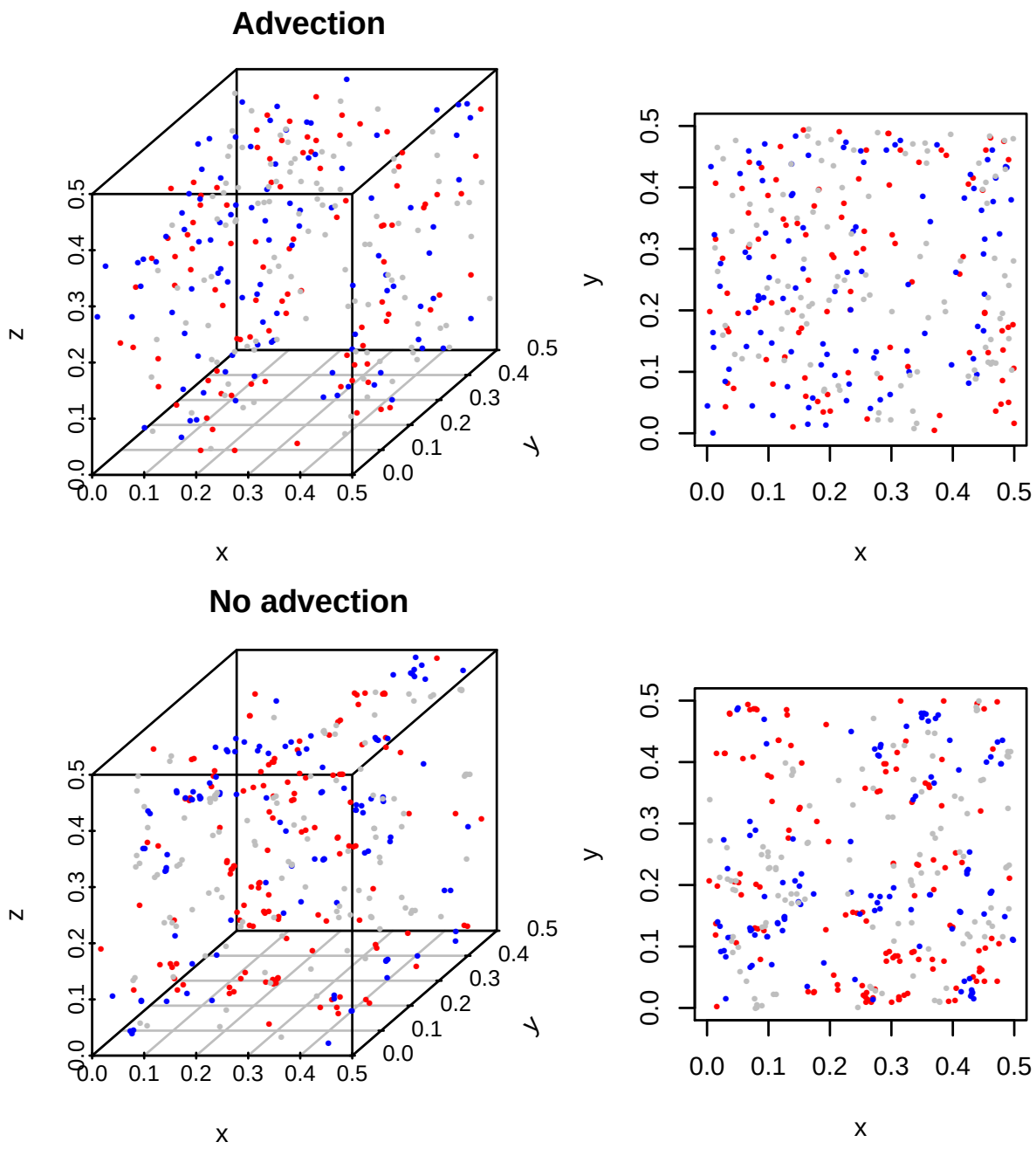

Figure 1: Spatial distributions of a 3-species community of nanophytoplankton with and without advection with density $\lambda = 10^3$ cm^{-3} after 1000 time steps. Each color corresponds to a different species. On the left-hand side, only a zoom on a 0.5×0.5 cm^2 cube is shown, and its projection on the x-y plane is shown on the right hand-side.

Ripley's functions extracted from numerical simulations match theoretical formula (Fig. 2) for both types of particles, which also indicates that dominance indices extracted from the simulation match theoretical expectation.

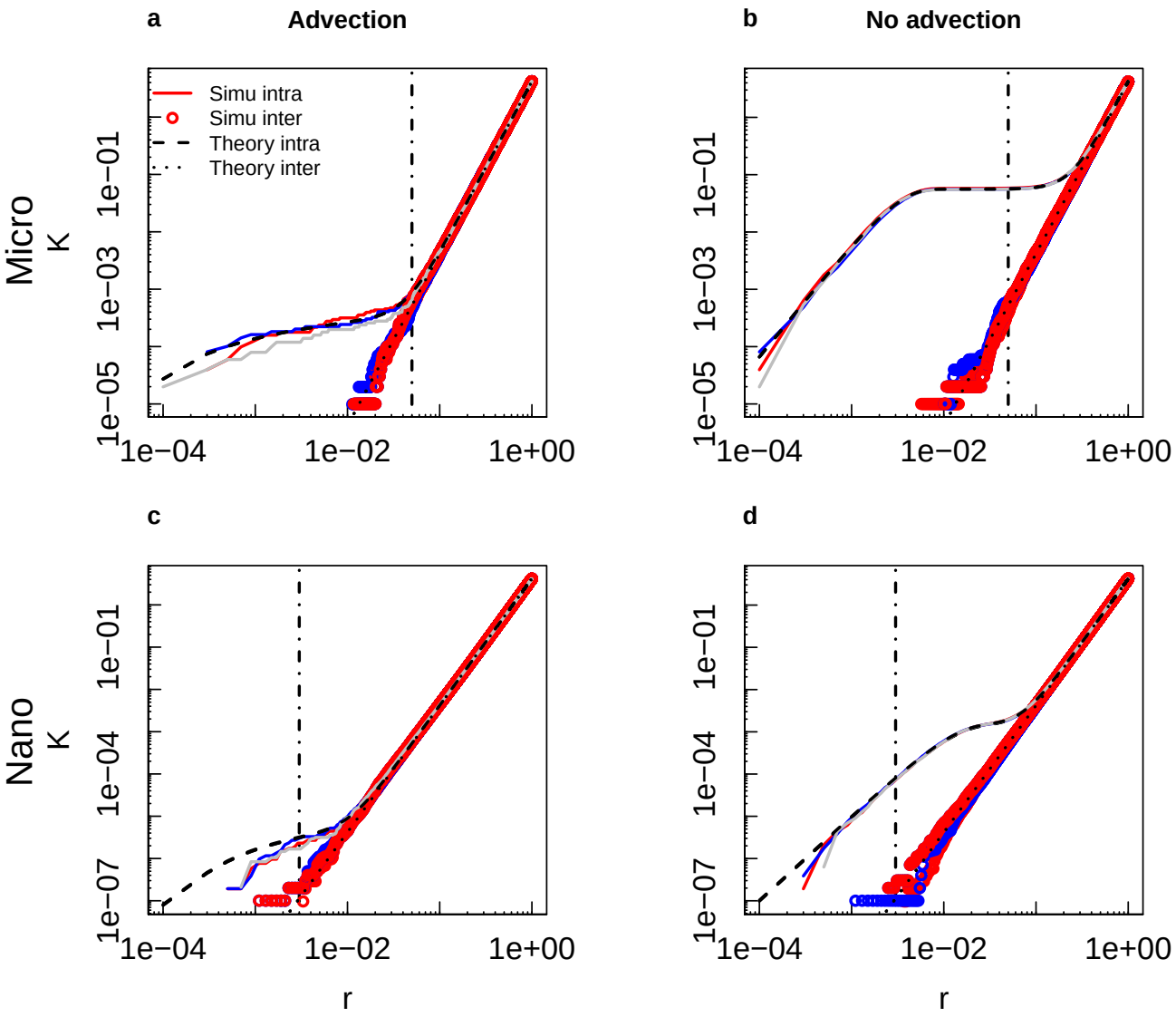

Figure 2: Ripley's K function for microphytoplankton (a-b) and nanophytoplankton (b-c) in a 3-species community with even distributions after 1000 timesteps, with (a, c) and without (b, d) advection. Each color represents a different species. The black dashed line corresponds to the 10-diameter threshold considered as the maximum range for nutrient-based competition.

Dominances indices all follow a similar pattern (Fig. 3 and 4). The dominance index is close 1 for small radii: there is always a scale at which a particle is surrounded almost only by conspecifics. The index then decreases sharply to converge at large radii (close to 1cm) to the proportion of the focus species in the whole community, as it would for a uniform distribution. Patterns differ at intermediate ranges of distances between particles.

In the presence of advection, the dominance index starts decreasing for a radius between 5 and 10 times lower than when advection is absent, i.e. organisms are closer to heterospecifics when their environment is turbulent. A quasi-uniform distribution is also reached for smaller radii with advection than without. Microphytoplankton species start mixing for distances larger than for nanophytoplankton species whatever the hydrodynamics around

them.

In a 3-species community with the same initial abundances, microphytoplankton dominance indices are between 0.37 and 0.47 at a distance equal to 10 diameters of a particle, while it is between 0.80 and 0.94 for nanophytoplankton species when advection is present. In its absence, dominance indices are all above 0.98 when the 10-diameter threshold is reached (Fig. 3). Microphytoplankton organisms are as likely to share their depletion volume with conspecifics as they are with heterospecifics, whereas nanophytoplankton organisms have almost only conspecifics around them in this space.

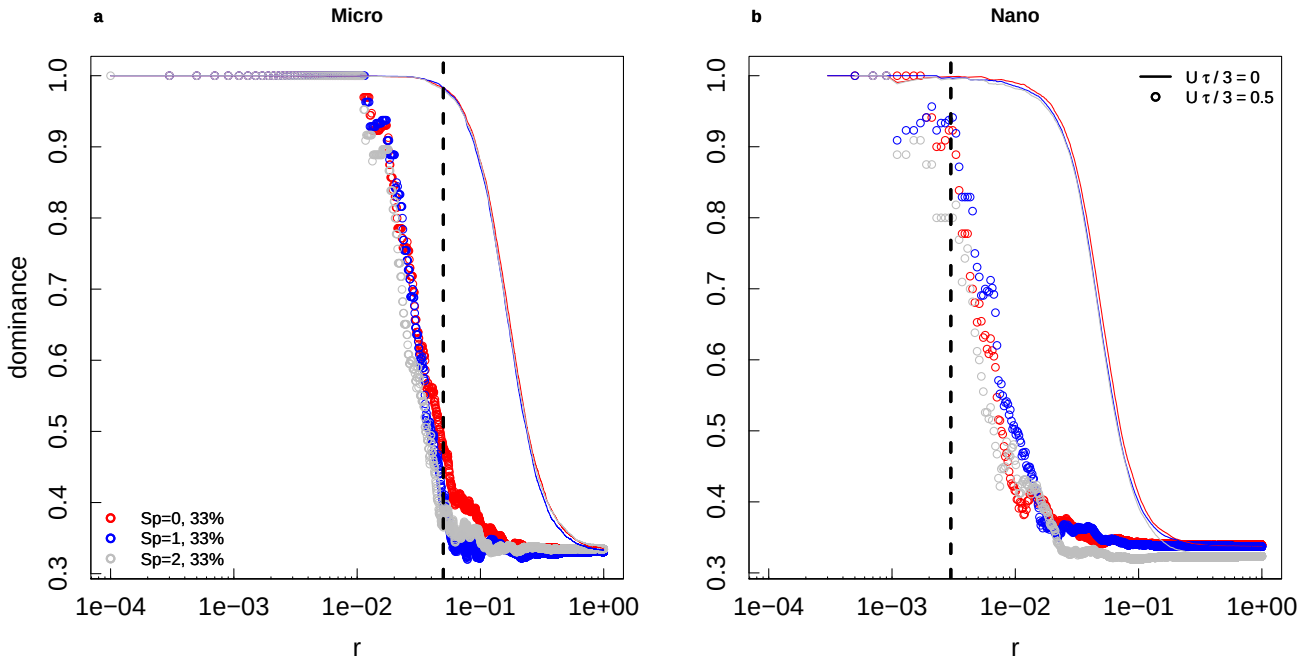

Figure 3: Dominance indices for microphytoplankton (a) and nanophytoplankton (b) in a 3-species community with even distributions after 1000 timesteps, with (circles) and without (lines) advection. Each color represents a different species. The black dashed line corresponds to the 10-diameter threshold considered as the maximum range for nutrient-based competition.

These differences between microphytoplankton and nanophytoplankton, and the role of advection, are even more pronounced when considering a 10 species-community with a skewed abundance distribution. (Fig. 4). In the presence of advection, microphytoplankton dominance indices at the 10-diameter threshold are between 0.34 (for the most abundant species) and 0.033 (for one of the least abundant species), while they are between 0.90 and 0.85 when advection is not taken into account. Nanophytoplankton species, too, are more mixed: dominance indices vary between 0.54 and 0.2 when the depletion-zone limit is reached (with an exception of 0 for one particular species which had no conspecific for radii below 10^{-2} cm) when particles are displaced by turbulence, while the same quantity is between 1 and 0.97 when they are only subject to diffusion.

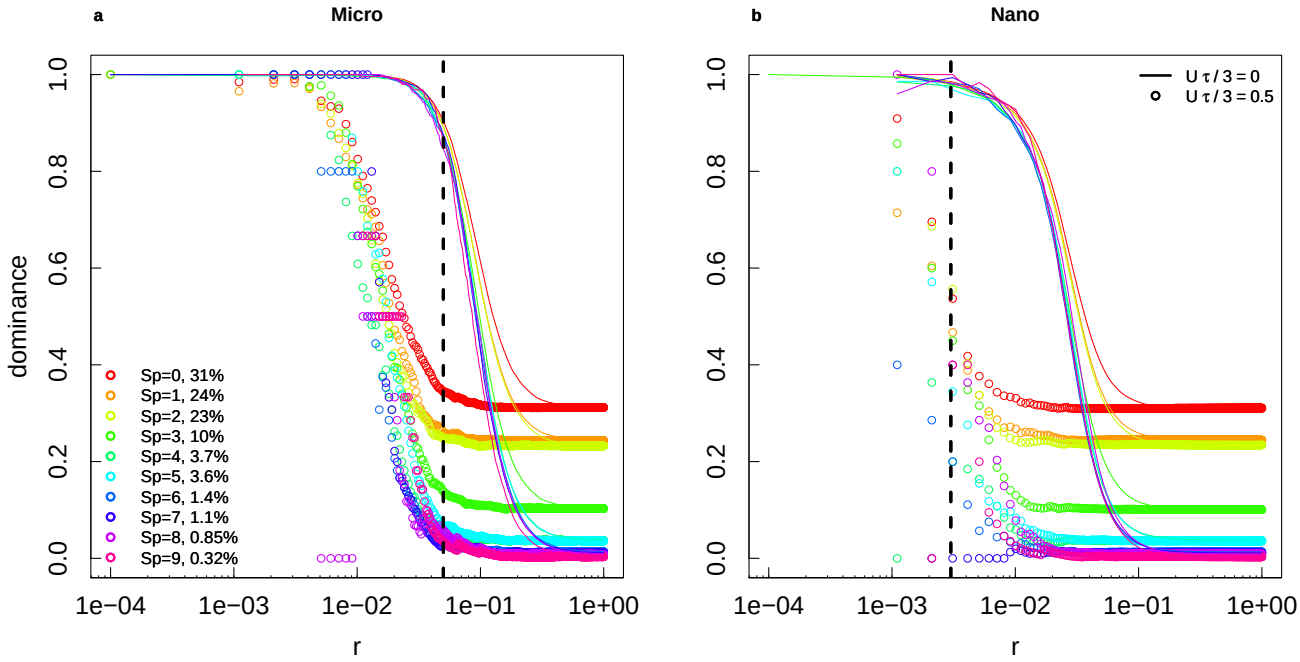

Figure 4: Dominance indices for microphytoplankton (a) and nanophytoplankton (b) in a 10-species community with a skewed abundance distribution (final proportion in the community is indicated in the legend) after 1000 timesteps, with (circles) and without (lines) advection. Each color represents a different species. The black dashed line corresponds to the 10-diameter threshold considered as the maximum range for nutrient-based competition.

Differences in distributions are not only due to particle sizes, but also to their abundances. When turbulence is taken into account, the dominance index decreases at lower radii around a particle when abundance is higher (Fig. 5 a-b): species which are more abundant tend to be omnipresent when they are mixed in the environment, which also makes them more likely to be close to a heterospecific, but they still also have more conspecifics close to them than the less abundant species. When turbulence is absent, the effect of abundance on the dominance index decrease is reversed: a 5% decrease is reached for a higher radius when species are more abundant for microphytoplankton, which may indicate that particle growth offsets diffusion in this case. The relationship is more complex for nanophytoplankton: the minimum radius for a 5% decrease in dominance index is reached for intermediate abundances; there is possibly a trade-off between occupation of the space due to mixing by diffusion and growth of particles around their conspecifics. The effect of average abundances on the composition of a particle neighbourhood within its depletion zone is less significant for nanophytoplankton than for microphytoplankton (Fig. 5 c-d), as a more abundant species of microphytoplankton is more present in its depletion zone than a less abundant species, while nanophytoplankton species of different average abundances tend to keep similar dominance indices.

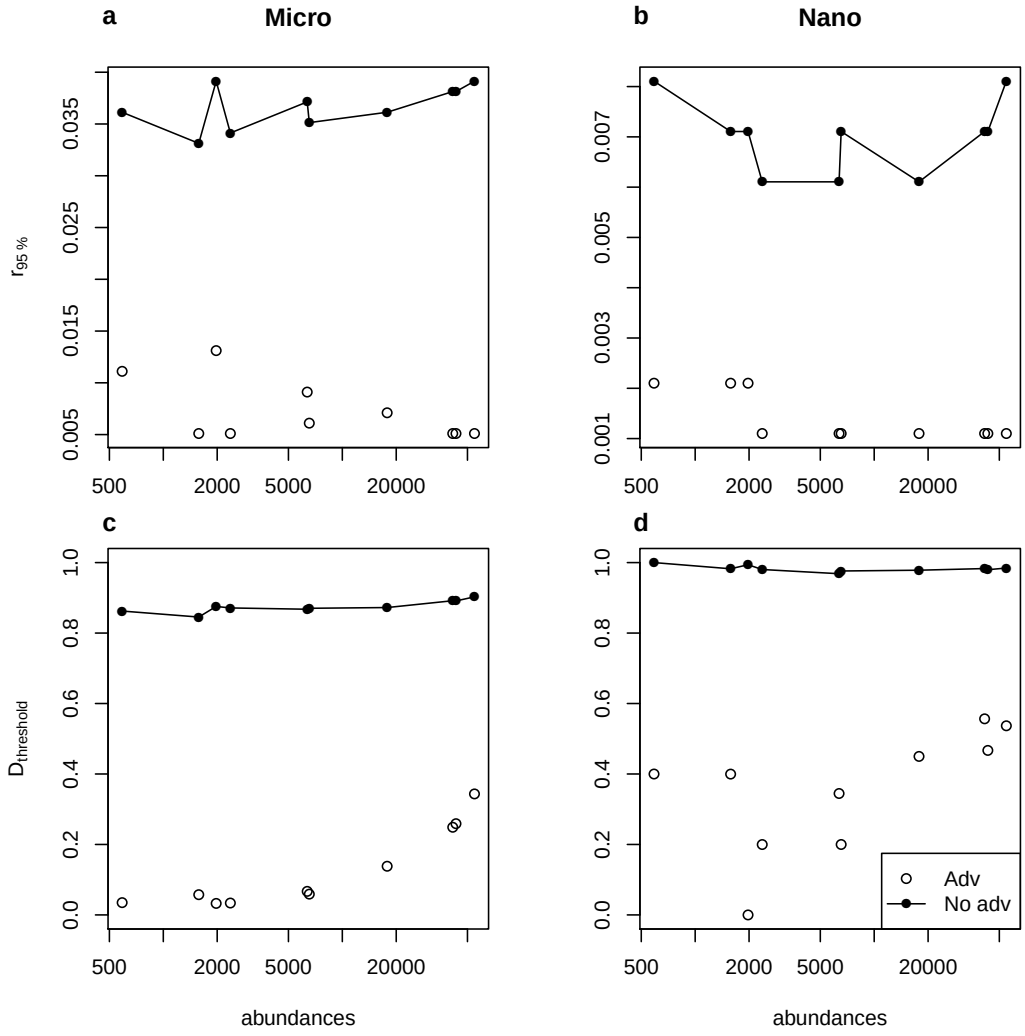

Figure 5: Minimum distances between points for dominance to drop below 95% (a and b) and dominance at a distance corresponding to the 10-diameter threshold (c and d) as a function of abundances (note the logarithmic scale on the x-axis) for microphytoplankton and nanophytoplankton with and without advection in a 10-species community with a skewed abundance distribution.

Discussion

We designed a stochastic, three-dimension, individual-based model of multiple species of particles in a viscous and turbulent flow, and conducted both the mathematical analyses and numerical simulations to quantify spatial correlations in particle distributions. We focused on the pair correlation function and Ripley's K-function, for which numerical and theoretical analyses showed a good agreement, and computed the corresponding dominance index to quantify the degree of intraspecific clustering and interspecific mixing. Comparing the distributions of organisms of different sizes and demography, we showed that the composition of the environment of a particle depends on its traits. While turbulence always leads to more mixing, microphytoplankters, larger cells with a lower diffusivity, were further away from heterospecifics than nanophytoplankters, but the zone in which competition for nutrient

was more likely also had a more diverse composition than for smaller cells.

Intraspecific clustering was a constant in our model: there was always a volume where a particle was only surrounded by conspecifics, and its radius was much lower than would be expected in a uniform distribution (Section SX). This is due to the prevalence of demographics processes at small scales, because a particle acts as a source point for other organisms of the same species, and hydrodynamics processes do not separate conspecifics fast enough to prevent aggregation. As a phytoplankter creates a gradient of depletion around itself, where nutrient concentrations are lower as we get closer to the cell (Karp-Boss *et al.*, 1996), competition for nutrient is likely stronger between closer neighbours, i.e. between conspecifics. Nutrients are not the only components that can be traded between particles: for example, sharing the phycosphere, a micro-environment where communities of bacteria exchange with the particle (Seymour *et al.*, 2017), can also impact organisms, both positively and negatively. The ratio between conspecifics and heterospecifics at different distances is therefore key in the relative proportions of interactions. Provided all species are equivalent with regards to their nutrient requirement and consumption, the possible interaction structure emerges from the spatial structure, and intraspecific interactions are more likely to be stronger than interspecific interactions, which is itself a crucial parameter in coexistence theory (Barabás *et al.*, 2017). While a high intra-to-interspecific interaction strength ratio is often associated with niche differences related to finite resources (nutrient, light, vitamins, etc.), we propose here that it can also emerge from the balance between demographics and hydrodynamics processes leading to clustering and mixing respectively.

The type of environment and organism considered affects the composition of a particle neighbourhood, which is especially relevant at distances in which depletion zones can overlap. It should be noted that interactions via zone of depletion overlap have sometimes been negated, but usely in the context of a uniform distribution (Siegel, 1998) and/or with diffusion only (Hulburt, 1970), which does not correspond to a realistic environment. Even though our model of turbulence is far from the complexity of 'real' turbulence, previous studies (e.g., Breier *et al.*, 2018) show that Navier-Stokes equation and simplified synthetic turbulence can give similar results in terms of particle aggregation. Here, in the absence of advection by turbulence, the neighbourhood is almost only made of conspecifics, for both microphytoplankton and nanophytoplankton, i.e. interactions are likely only intraspecific. When turbulence increases, mixing increases (Section SXX) and the neighbourhood composition is size-dependent: nanophytoplankton organisms mainly share this space with conspecifics (the dominance index remains close to 1, even near the overlap threshold) while microphytoplankton organisms can affect both conspecifics and heterospecifics (the dominance index is often below 50% at the overlap threshold, i.e. a particle's depletion zone is more likely to overlap more with a heterospecific's than a conspecific's). The size-dependence is the same if we consider the phycosphere previously mentioned: for a particle of microphytoplankton with a diameter of 50 µm, the phycosphere radius reaches 1200 µm, while it is nearly non-existent for particles below 3 µm (Jackson, 1987; Seymour *et al.*, 2017). This tends to confirm that microphytoplankters are more likely to interact with heterospecifics than nanophytoplankters, via exchanges of compounds necessary for their survival. The dominance of heterospecifics in a microphytoplankter's

neighbourhood may however be debated. Indeed, when computing the volumes of zones relevant for interactions (either for nutrients or for other components linked to the phycosphere), we use maximum volumes, corresponding to a diffusive-only flow. When turbulence increases, uptake increases and the size of the depletion zone decreases (Karp-Boss *et al.*, 1996). The proportion of change increases with the size of organism, meaning that larger cells such as microphytoplankters are much more sensitive to turbulence increases than smaller cells (Karp-Boss *et al.*, 1996; Barton *et al.*, 2014). If we adjusted depletion zone sizes with turbulence in mind, conspecifics might be as predominant for microphytoplankton as they are for nanophytoplankton *within* relevant zones. While there are certainly different mechanisms behind diversity maintenance for microphytoplankton and nanophytoplankton, we should therefore be cautious in saying that interspecific interactions could dominate for microphytoplankters.

Our model focuses on size as the defining trait of the communities we consider, with microphytoplankton being larger than nanophytoplankton by an order of magnitude. Indeed, size is considered as a major functional trait (Marañón, 2015), correlated with other capacities such as growth rate or nutrient affinities (Edwards *et al.*, 2012), or even average abundances (Agusti *et al.*, 1987). In this model, smaller sizes lead to a higher diffusivity, i.e. less clustering, and a higher growth rate, i.e. more clustering, but they also lead to a smaller zone of depletion. In reality, phytoplankton size classes cover a much wider range than showed here: nanophytoplankton sizes vary between 2 and 20 µm and microphytoplankton sizes, between 20 and 200 µm (Reynolds, 2006). Even within a single species, organism diameters may vary greatly (Harrison *et al.*, 2015). Taking into account this variability creates new mechanisms sustaining diversity: for example, size-based inertia in a chaotic flow creates size-specific aggregates, implying more conspecifics clusters (Benczik *et al.*, 2006). These results complement ours as they highlight the relevance of turbulence at microscale to produce new spatial niches for particles of different sizes. We can assume that, in our model, introducing a size gradient for species within a single phytoplankton community would create more complex spatial distributions. However, patterns of the dominance index were consistent between organisms: more than one order of magnitude between nanophytoplankton and microphytoplankton sizes did not change the emergence of clustering at small scales, relevant for interactions.

The model presented here is one of the first multispecific, spatial phytoplankton model focusing on the microscale, modeling only basic hydrodynamics and demographics processes, from which we can only guesstimate the differences in interactions due to distances between organisms. To our knowledge, no such model has been built with explicit interactions. In a monospecific context, density-dependence (Bouderbala *et al.*, 2018), predation (Bouderbala *et al.*, 2019), attraction between motile mechanisms (Breier *et al.*, 2018) have been modeled, based on the definition of a kernel of interactions, and these mechanisms also lead to clustering, especially when length scales associated to interactions and turbulence are close (Breier *et al.*, 2018). The mathematical analysis of such processes leads to much more complex moment equations (Bolker & Pacala, 1999), but remains feasible, and should be encouraged alongside with numerical simulations. Finally, clustering would be even stronger if taking into account a very common trait in microphytoplantkon, coloniality (Kiørboe *et al.*, 1990). With independent and equivalent individuals, our model

should therefore be taken as a first null model to which more complex, maybe more mechanistic models should be compared.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contribution statement

Acknowledgments

FB and CP were supported by the grant ANR-20-CE45-0004. CP was supported by a PhD grant from the French Ministry of Research.

Appendices

Derivation of the spatial characteristics of the Brownian Bug Model

We show here how to compute the monospecific pair correlation function and Ripley's K function of the Brownian Bug Model. Formula for standard processes are given in the Supplementary Information, for readers who want to familiarize with simpler models.

Proof of eq. 3

In 3 dimensions, when the birth rate λ is the same as the mortality rate μ, the pair density $G(r)$ (where $G(r) = C^2 g(r)$) is a solution of eq. 11 (see Young et al. and Picoche et al. for a detailed explanation).

$$\frac{\partial G}{\partial t} = \frac{2D}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial G}{\partial r}\right) + \frac{\gamma}{r^2}\frac{\partial}{\partial r}\left(r^4\frac{\partial G}{\partial r}\right) + 2\lambda C\delta(\boldsymbol{r}) \tag{11}$$

With advection In the presence of advection ($\gamma \neq 0$), a steady-state solution can be found.

$$\begin{aligned}
0 &= \frac{2D}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial G}{\partial r}\right) + \frac{\gamma}{r^2}\frac{\partial}{\partial r}\left(r^4\frac{\partial G}{\partial r}\right) + 2\lambda C\delta(\boldsymbol{r}) \\
\Leftrightarrow 0 &= 4\pi r^2\left(\frac{2D}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial G}{\partial r}\right) + \frac{\gamma}{r^2}\frac{\partial}{\partial r}\left(r^4\frac{\partial G}{\partial r}\right) + 2\lambda C\delta(\boldsymbol{r})\right) \\
\Leftrightarrow 0 &= 4\pi\left(2D\frac{\partial}{\partial r}\left(r^2\frac{\partial G}{\partial r}\right) + \gamma\frac{\partial}{\partial r}\left(r^4\frac{\partial G}{\partial r}\right)\right) + 4\pi r^2 2\lambda C\delta(\boldsymbol{r})
\end{aligned} \tag{12}$$

We can then integrate Eq. (11) over a small sphere centered on a particle, with radius ρ. Let us first note that

$$\begin{aligned} \int_{\mathbb{R}^3} \delta(\boldsymbol{r}) d^3\boldsymbol{r} &= 1 \\ \Leftrightarrow \int_0^{2\pi} \int_0^{\pi} \int_0^{\rho} \delta(\boldsymbol{r}') r'^2 \sin(\phi) \quad dr' \quad d\phi d\theta &= 1 \\ \Leftrightarrow 4\pi \int_0^{\rho} \delta(\boldsymbol{r}') r'^2 dr' &= 1 \end{aligned} \tag{13}$$

Using Eq. (12) and (13),

$$\begin{aligned} 0 &= 4\pi \left(2Dr^2 \frac{\partial G}{\partial r} + \gamma r^4 \frac{\partial G}{\partial r} \right) + 2\lambda C \\ \Leftrightarrow \frac{\partial G}{\partial r} &= -\frac{1}{4\pi} \frac{2\lambda C}{2Dr^2 + \gamma r^4} \end{aligned} \tag{14}$$

We can integrate between ρ and ∞, knowing that $G(\infty) = C^2$.

$$C^2 - G(\rho) = -\frac{2\lambda C}{4\pi} \int_{\rho}^{\infty} \frac{1}{2Dr^2 + \gamma r^4} dr \tag{15}$$

We first compute the primitive $A = \int \frac{1}{2Dr^2+\gamma r^4} dr$.

$$\begin{aligned} A &= \int \frac{1}{r^2(2D+\gamma r^2)} dr \\ &= \int \frac{1}{2Dr^2} - \frac{\gamma}{2D(2D+\gamma r^2)} dr \\ &= \frac{1}{2D} \int \frac{1}{r^2} dr - \frac{\gamma}{2D} \int \frac{1}{2D+\gamma r^2} dr \\ &= -\frac{1}{2Dr} - \frac{\gamma}{2D} \int \frac{1}{2D\left(1+\left(\sqrt{\frac{\gamma}{2D}} r\right)^2\right)} \end{aligned} \tag{16}$$

With a change of variable $u = \sqrt{\frac{\gamma}{2D}} r$, using $\int \frac{1}{1+u^2} = \arctan(u)$, we have:

$$A = -\frac{1}{2Dr} - \frac{\sqrt{\gamma} \arctan\left(\frac{\sqrt{\gamma} r}{\sqrt{2D}}\right)}{2\sqrt{2} D \sqrt{D}} + K \tag{17}$$

where K is a constant.

We can know compute $B = [A]_{\rho}^{\infty}$.

$$B = -\frac{\sqrt{\gamma}\pi}{4\sqrt{2}D\sqrt{D}} + \frac{1}{2D\rho} + \frac{\sqrt{\gamma} \arctan\left(\frac{\sqrt{\gamma}\rho}{\sqrt{2D}}\right)}{2\sqrt{2}D\sqrt{D}} \tag{18}$$

This leads to:

$$\begin{aligned} G(\rho) &= C^2 + \frac{2\lambda C}{4\pi} B \\ &= C^2 + \frac{\lambda C}{2\pi} \left[\frac{1}{2D\rho} + \frac{\sqrt{\gamma} \arctan\left(\frac{\sqrt{\gamma}\rho}{\sqrt{2D}}\right)}{2\sqrt{2}D\sqrt{D}} - \frac{\sqrt{\gamma}\pi}{4\sqrt{2}D\sqrt{D}} \right] \end{aligned} \tag{19}$$

Finally, the pair correlation function $g = G/C^2$ is defined as

$$g(\rho) = \frac{\lambda}{4\pi C D} \left(\frac{\sqrt{\gamma} \arctan\left(\frac{\sqrt{\gamma}\rho}{\sqrt{2D}}\right)}{\sqrt{2D}} + \frac{1}{\rho} - \frac{\pi\sqrt{\gamma}}{2\sqrt{2D}} \right) + 1 \tag{20}$$

Without advection When $U = 0$, $\gamma = 0$ and there is no steady solution. We can get back to Eq. (11).

$$\frac{\partial G}{\partial t} = \frac{2D}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial G}{\partial r}\right) + 2\lambda C\delta(\boldsymbol{r}) \tag{21}$$

Assuming an isotropic environment, this means

$$\frac{\partial G}{\partial t} - 2D\Delta G = 2\lambda C\delta(\boldsymbol{r}) \tag{22}$$

where $\Delta = \nabla^2$ is the Laplacian operator.

We therefore have

$$\mathcal{L}G(\boldsymbol{r}, t) = 2\lambda C\delta(\boldsymbol{r}) \tag{23}$$

where $\mathcal{L}$ is the linear differential operator $\partial_t - 2D\Delta$.

Using the Green's function theory, we know that $G(y) = \int H(y, s)2\lambda C\delta(s)ds$ where $H(y, s) = H(y - s)$ is the Green kernel (heat kernel).

$$\begin{aligned} G(\boldsymbol{r}, t) \quad & = 2\lambda C \int_{\mathbb{R}^3}\int_0^t H(\boldsymbol{r} - \boldsymbol{r'}, t')\delta(\boldsymbol{r'})dr'dt' \\ \Leftrightarrow \quad & = 2\lambda C \int_0^t H(\boldsymbol{r}, t')dt' \end{aligned}$$

A solution for the Green's function using $\mathcal{L} = \partial_t - 2D\Delta$ in 3 dimensions is $H(r, t) = \left(\frac{1}{4\pi 2Dt}\right)^{3/2}\exp(\frac{-r^2}{4\times 2Dt})$.

$G(r, t)$ can then be computed:

.

$$G(r, t) = 2\lambda C\left(\frac{\operatorname{erf}\left(\frac{r}{\sqrt{8tD}}\right)}{8\pi Dr} + K\right) \tag{24}$$

where erf is the error function. Using $G(r, 0) = C^2$ and $\lim_{x\to+\infty}\operatorname{erf}(x) = 1$ in Eq. (24),

$$\begin{aligned} C^2 \quad & = \quad 2\lambda C\left(\tfrac{1}{8\pi Dr} + K\right) \\ \Leftrightarrow \tfrac{C}{2\lambda} + \quad \tfrac{1}{8\pi Dr} & = \quad K \end{aligned} \tag{25}$$

We can finally compute G(r,t):

$$\begin{aligned} G(r, t) \quad & = \quad 2\lambda C\left(\frac{\operatorname{erf}\left(\frac{r}{\sqrt{8tD}}\right)}{8\pi Dr} + \frac{C}{2\lambda} + \frac{1}{8D\pi r}\right) \\ & = \quad \frac{\lambda C}{4\pi Dr}\left\{1 - \operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right)\right\} + C^2 \\ \Leftrightarrow g(r, t) \quad & = \quad \frac{\lambda}{4D\pi rC}\left\{1 - \operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right)\right\} + 1 \end{aligned} \tag{26}$$

Proof of eq. 7

We can integrate the pcf formula to compute Ripley's K-function, as $g(r) = \frac{K'(r)}{4\pi r^2}$.

With advection From eq. 20,

$$K(\rho) = 4\pi \int_0^\rho r^2 + \frac{\lambda}{2\pi C}\left[\frac{r}{2D} + \frac{\sqrt{\gamma}r^2 \arctan\left(\frac{\sqrt{\gamma}r}{\sqrt{2D}}\right)}{2\sqrt{2}D\sqrt{D}} - \frac{\sqrt{\gamma}\pi r^2}{4\sqrt{2}D\sqrt{D}}\right]dr \tag{27}$$

We define $A = \int_0^\rho r^2 dr$, $B = \int_0^\rho \frac{r}{2D}dr$, $C = \int_0^\rho r^2 \arctan\left(\frac{\sqrt{\gamma}r}{\sqrt{2D}}\right)dr$ and $E = \int_0^\rho \frac{\sqrt{\gamma}\pi r^2}{4\sqrt{2}D\sqrt{D}}dr$.

$$\begin{aligned} A &= \frac{1}{3}\rho^3 \\ B &= \frac{\rho^2}{4D} \\ E &= \frac{\sqrt{\gamma}\pi\rho^3}{12\sqrt{2}D\sqrt{D}} \end{aligned} \tag{28}$$

We can also compute $C = \int_0^\rho r^2 \arctan\left(\frac{\sqrt{\gamma}r}{\sqrt{2D}}\right)dr$. We can first change variable, with $u = \frac{r}{\sqrt{2D}}$,$dr = \sqrt{2D}du$.

$$\begin{aligned} C &= \int_0^{\rho/\sqrt{2D}} (\sqrt{2D}u)^2 \arctan(\sqrt{\gamma}u)\sqrt{2D}du \\ &= (2D)^{3/2}\int_0^{\rho/\sqrt{2D}} u^2 \arctan(\sqrt{\gamma}u)du \end{aligned} \tag{29}$$

We can integrate by parts, with $f = \arctan(\sqrt{\gamma}u)$ and $g' = u^2$.

$$\begin{aligned} C &= (2D)^{3/2}\left(\frac{\rho^3}{3(2D)^{3/2}}\arctan(\sqrt{\frac{\gamma}{2D}}\rho) - \int_0^{\rho/\sqrt{2D}} \frac{\sqrt{\gamma}u^3}{3(\gamma u^2+1)}du\right) \\ &= (2D)^{3/2}\left(\frac{\rho^3}{3(2D)^{3/2}}\arctan(\sqrt{\frac{\gamma}{2D}}\rho) - \frac{\sqrt{\gamma}}{3}\int_0^{\rho/\sqrt{2D}} \frac{u^3}{(\gamma u^2+1)}du\right) \end{aligned} \tag{30}$$

We can substitute $v = \gamma u^2 + 1$,$du = \frac{1}{2\gamma u}dv$.

$$\begin{aligned} \int_0^{\rho/\sqrt{2D}} \frac{u^3}{(\gamma u^2+1)}du &= \frac{1}{2\gamma^2}\int_1^{\gamma\rho^2/2D+1} \frac{v-1}{v}dv \\ &= \frac{1}{2\gamma^2}\int_1^{\gamma\rho^2/2D+1} 1 - \frac{1}{v}dv \\ &= \frac{1}{2\gamma^2}(\gamma\frac{\rho^2}{2D} - \log(\gamma\frac{\rho^2}{2D}+1)) \end{aligned} \tag{31}$$

Going back to C, we obtain:

$$\begin{aligned} C &= \frac{\rho^3 \arctan(\sqrt{\frac{\gamma}{2D}}\rho)}{3} - (2D)^{3/2}\frac{\sqrt{\gamma}}{3}\frac{1}{2\gamma^2}\left(\frac{\gamma}{2D}\rho^2 - \log(\gamma\frac{\rho^2}{2D}+1)\right) \\ &= \frac{\rho^3 \arctan(\sqrt{\frac{\gamma}{2D}}\rho)}{3} - \frac{\sqrt{2D}}{6\sqrt{\gamma}}\rho^2 + \frac{\sqrt{2}D^{3/2}}{3\gamma^{3/2}}\log(\gamma\frac{\rho^2}{2D}+1) \end{aligned} \tag{32}$$

Combining all equations:

$$\begin{aligned} K(\rho) &= \frac{4}{3}\pi\rho^3 + \frac{2\lambda}{C}\left(\frac{\rho^2}{4D} + \frac{\sqrt{\gamma}\rho^3 \arctan(\sqrt{\frac{\gamma}{2D}}\rho)}{6\sqrt{2}D^{3/2}} - \frac{\rho^2}{12D} + \frac{\log\left(\gamma\frac{\rho^2}{2D}+1\right)}{6\gamma} - \frac{\sqrt{\gamma}\pi\rho^3}{12\sqrt{2}D\sqrt{D}}\right) \\ &= \frac{4}{3}\pi\rho^3 + \frac{2\lambda}{C}\left(\frac{\rho^2}{6D} + \frac{\sqrt{\gamma}\rho^3 \arctan(\sqrt{\frac{\gamma}{2D}}\rho)}{6\sqrt{2}D^{3/2}} + \frac{\log\left(\gamma\frac{\rho^2}{2D}+1\right)}{6\gamma} - \frac{\sqrt{\gamma}\pi\rho^3}{12\sqrt{2}D\sqrt{D}}\right) \end{aligned} \tag{33}$$

Without advection From eq. 26,

$$\begin{aligned} K(\rho) &= \frac{\lambda}{DC}\int_0^\rho r\left\{1-\operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right)\right\}+4\pi r^2 dr \\ &= \frac{\lambda}{CD}\left(\frac{\rho^2}{2}-\int_0^\rho r\times\operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right)dr\right)+\frac{4}{3}\pi\rho^3 \end{aligned} \tag{34}$$

We first compute the primitive for $\int_0^\rho r\times\operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right)dr$. We define $u=\frac{r}{\sqrt{8Dt}}$, $dr=\sqrt{8Dt}du$.

$$\int_0^\rho r\times\operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right)dr = 8Dt\int_0^{\rho/\sqrt{8Dt}} u\times\operatorname{erf}(u)\,du \tag{35}$$

We can integrate by parts, with $f=\operatorname{erf}(u)$ and $g'=u$.

$$8Dt\int_0^{\rho/\sqrt{8Dt}} u\times\operatorname{erf}(u)\,du = 8Dt\left(\frac{\rho^2}{2}\frac{1}{8Dt}\operatorname{erf}(\frac{\rho}{\sqrt{8Dt}})-\frac{1}{\sqrt{\pi}}\int_0^{\rho/\sqrt{8Dt}}u^2e^{-u^2}du\right) \tag{36}$$

We integrate by parts again, this time with $f=u$ and $g'=ue^{-u^2}$, which leads to

$$\int u^2e^{-u^2}du=-\frac{ue^{-u^2}}{2}+\frac{1}{2}\int e^{-u^2}du=-\frac{ue^{-u^2}}{2}+\frac{\sqrt{\pi}\operatorname{erf}(u)}{4} \tag{37}$$

If we use eq. 37 in eq. 36:

$$\begin{aligned} 8Dt\int_0^{\rho/\sqrt{8Dt}} u\times\operatorname{erf}(u)\,du &= 8Dt\left(\frac{\rho^2}{2}\frac{1}{8Dt}\operatorname{erf}(\frac{\rho}{\sqrt{8Dt}})-\frac{\operatorname{erf}(\frac{\rho}{\sqrt{8Dt}})}{4}+\frac{1}{2\sqrt{\pi}}\frac{\rho}{\sqrt{8Dt}}e^{-\rho^2/8Dt}\right) \\ \Leftrightarrow \int_0^\rho r\times\operatorname{erf}\left(\frac{r}{\sqrt{8Dt}}\right)dr &= \frac{1}{2}\operatorname{erf}(\frac{\rho}{\sqrt{8Dt}})(\rho^2-4Dt)+\frac{\sqrt{2Dt}}{\sqrt{\pi}}\rho e^{-\rho^2/8Dt} \end{aligned} \tag{38}$$

We can now compute $K(\rho)$.

$$K(\rho) = \frac{\lambda}{CD}\left(\frac{\rho^2}{2}-\frac{1}{2}\operatorname{erf}(\frac{\rho}{\sqrt{8Dt}})(\rho^2-4Dt)-\frac{\sqrt{2Dt}\rho}{\sqrt{\pi}}e^{-\rho^2/8Dt}\right)+\frac{4}{3}\pi\rho^3 \tag{39}$$

References

Adler, P.B., Smull, D., Beard, K.H., Choi, R.T., Furniss, T., Kulmatiski, A., Meiners, J.M., Tredennick, A.T. & Veblen, K.E. (2018). Competition and coexistence in plant communities: intraspecific competition is stronger than interspecific competition. *Ecology Letters*, 21, 1319–1329.

Agusti, S., Duarte, C.M. & Kalff, J. (1987). Algal cell size and the maximum density and biomass of phytoplankton1. *Limnology and Oceanography*, 32, 983–986.

Arnott, R.N., Cherif, M., Bryant, L.D. & Wain, D.J. (2021). Artificially generated turbulence: a review of phycological nanocosm, microcosm, and mesocosm experiments. *Hydrobiologia*, 848, 961–991.

Arrieta, J., Barreira, A. & Tuval, I. (2015). Microscale patches of nonmotile phytoplankton. *Phys. Rev. Lett.*, 114, 128102.

Arrieta, J., Jeanneret, R., Roig, P. & Tuval, I. (2020). On the fate of sinking diatoms: the transport of active buoyancy-regulating cells in the ocean. *Phil. Trans. R. Soc. A.*, 378, 20190529.

Bainbridge, R. (1957). The size, shape and density of marine phytoplankton concentrations. *Biological Reviews*, 32, 91–115.

Barabás, G., Michalska-Smith, M.J. & Allesina, S. (2017). Self-regulation and the stability of large ecological networks. *Nature Ecology & Evolution*, 1, 1870–1875.

Barton, A.D., Ward, B.A., Williams, R.G. & Follows, M.J. (2014). The impact of fine-scale turbulence on phytoplankton community structure. *Limnology and Oceanography: Fluids and Environments*, 4, 34–49.

Bec, B., Collos, Y., Vaquer, A., Mouillot, D. & Souchu, P. (2008). Growth rate peaks at intermediate cell size in marine photosynthetic picoeukaryotes. *Limnology and Oceanography*, 53, 863–867.

Benczik, I.J., Károlyi, G., Scheuring, I. & Tél, T. (2006). Coexistence of inertial competitors in chaotic flows. *Chaos*, 16, 043110.

Birch, D.A., Tsang, Y.K. & Young, W.R. (2007). Bounding biomass in the fisher equation. *Phys. Rev. E*, 75, 066304.

Bissinger, J.E., Montagnes, D.J.S., Harples, J. & Atkinson, D. (2008). Predicting marine phytoplankton maximum growth rates from temperature: Improving on the Eppley curve using quantile regression. *Limnology and Oceanography*, 53, 487–493.

Bolker, B. & Pacala, S. (1999). Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. *The American Naturalist*, 153, 575–602.

Borgnino, M., Arrieta, J., Boffetta, G., De Lillo, F. & Tuval, I. (2019). Turbulence induces clustering and segregation of non-motile, buoyancy-regulating phytoplankton. *J. R. Soc. Interface*, 16, 20190324.

Bouderbala, I., El Saadi, N., Bah, A. & Auger, P. (2018). A 3d individual-based model to study effects of chemotaxis, competition and diffusion on the motile-phytoplankton aggregation. *Acta Biotheor*, 66, 257–278.

Bouderbala, I., El Saadi, N., Bah, A. & Auger, P. (2019). A simulation study on how the resource competition and anti-predator cooperation impact the motile-phytoplankton groups' formation under predation stress. *Ecological Modelling*, 391, 16–28.

Breier, R.E., Lalescu, C.C., Waas, D., Wilczek, M. & Mazza, M.G. (2018). Emergence of phytoplankton patchiness at small scales in mild turbulence. *Proc Natl Acad Sci USA*, 115, 12112–12117.

Chesson, P. (2018). Updates on mechanisms of maintenance of species diversity. *Journal of Ecology*, 106, 1773–1794.

Detto, M. & Muller-Landau, H.C. (2016). Stabilization of species coexistence in spatial models through the aggregation-segregation effect generated by local dispersal and nonspecific local interactions. *Theoretical Population Biology*, 112, 97–108.

Doubell, M.J., Seuront, L., Seymour, J.R., Patten, N.L. & Mitchell, J.G. (2006). High-resolution fluorometer for mapping microscale phytoplankton distributions. *Appl Environ Microbiol*, 72, 4475–4478.

Durham, W.M., Climent, E., Barry, M., De Lillo, F., Boffetta, G., Cencini, M. & Stocker, R. (2013). Turbulence drives microscale patches of motile phytoplankton. *Nat Commun*, 4, 2148.

Dusenbery, D. (2009). *Living at the microscale*. Harvard University Press.

Edwards, K.F. (2019). Mixotrophy in nanoflagellates across environmental gradients in the ocean. *Proceedings of the National Academy of Sciences*, p. 201814860.

Edwards, K.F., Thomas, M.K., Klausmeier, C.A. & Litchman, E. (2012). Allometric scaling and taxonomic variation in nutrient utilization traits and maximum growth rate of phytoplankton. *Limnology and Oceanography*, 57, 554–566.

Einstein, A. (1905). Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 4.

Field, C.B., Behrenfeld, M.J., Randerson, J.T. & Falkowski, P. (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science*, 281, 237–240.

Font-Muñoz, J.S., Jordi, A., Tuval, I., Arrieta, J., Anglès, S. & Basterretxea, G. (2017). Advection by ocean currents modifies phytoplankton size structure. *J. R. Soc. Interface*, 14, 20170046.

Guasto, J.S., Rusconi, R. & Stocker, R. (2012). Fluid mechanics of planktonic microorganisms. *Annual Review of Fluid Mechanics*, 44, 373–400.

Harrison, P.J., Zingone, A., Mickelson, M.J., Lehtinen, S., Ramaiah, N., Kraberg, A.C., Sun, J., McQuatters-Gollop, A. & Jakobsen, H.H. (2015). Cell volumes of marine phytoplankton from globally distributed coastal data sets. *Estuarine, Coastal and Shelf Science*, 162, 130–142.

Hellweger, F.L. & Bucci, V. (2009). A bunch of tiny individuals – individual-based modeling for microbes. *Ecological Modelling*, 220, 8–22.

Hulburt, E.M. (1970). Competition for nutrients by marine phytoplankton in oceanic, castal, and estuarine regions. *Ecology*, 51, 475–484.

Hutchinson, G.E. (1961). The paradox of the plankton. *The American Naturalist*, 95, 137–145.

Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. (2008). *Statistical analysis and modelling of spatial point patterns*. vol. 70. John Wiley & Sons.

Jackson, G.A. (1987). Simulating chemosensory responses of marine microorganisms1. *Limnology and Oceanography*, 32, 1253–1266.

Jumars, P.A., Deming, J., Hill, P., Karp-Boss, L., Yager, P. & Dade, W. (1993). Physical constraints on marine osmotrophy in an optimal foraging context. *Marine Microbial Food Webs*, 7, 121–159.

Karp-Boss, L., Boss, E. & Jumars, P.A. (1996). Nutrient fluxes to planktonic osmotrophs in the presence of fluid motion. *Oceanography and Marine Biology: An Annual Review*, 34, 71–107.

Kiørboe, T. (2018). *A mechanistic approach to plankton ecology*. Princeton University Press.

Kiørboe, T., Andersen, K.P. & Dam, H.G. (1990). Coagulation efficiency and aggregate formation in marine phytoplankton. *Marine Biology*, 107, 235–245.

Law, R., Murrell, D.J. & Dieckmann, U. (2003). Population growth in space and time: Spatial logistic equations. *Ecology*, 84, 252–262.

Leonard, C.L., Bidigare, R.R., Seki, M.P. & Polovina, J.J. (2001). Interannual mesoscale physical and biological variability in the north pacific central gyre. *Progress in Oceanography*, 49, 227–244.

Levine, J.M. & HilleRisLambers, J. (2009). The importance of niches for the maintenance of species diversity. *Nature*, 461, 254–257.

Li, L. & Chesson, P. (2016). The effects of dynamical rates on species coexistence in a variable environment: The paradox of the plankton revisited. *The American Naturalist*, 188, E46–E58.

Marañón, E. (2015). Cell size as a key determinant of phytoplankton metabolism and community structure. *Annual Review of Marine Science*, 7, 241–264.

Martin, A.P. (2003). Phytoplankton patchiness: the role of lateral stirring and mixing. *Progress in Oceanography*, 57, 125–174.

Ngan, K. & Vanneste, J. (2011). Scalar decay in a three-dimensional chaotic flow. *Phys. Rev. E*, 83, 056306.

Peters, F., Arin, L., Marrasé, C., Berdalet, E. & Sala, M. (2006). Effects of small-scale turbulence on the growth of two diatoms of different size in a phosphorus-limited medium. *Journal of Marine Systems*, 61, 134–148.

Peters, F. & Marrasé, C. (2000). Effects of turbulence on plankton: an overview of experimental evidence and some theoretical considerations. *Marine Ecology Progress Series*, 205, 291–306.

Picoche, C. & Barraquand, F. (2020). Strong self-regulation and widespread facilitative interactions in phytoplankton communities. *Journal of Ecology*, 108, 2232–2242.

Picoche, C., Young, W. & Barraquand, F. (XXX). [re] reproductive pair correlations and the clustering of organisms. *Journal of Ecology*, n/A, n/a.

Pierrehumbert, R. (1994). Tracer microstructure in the large-eddy dominated regime. *Chaos, Solitons & Fractals*, 4, 1091–1110.

Plank, M.J. & Law, R. (2015). Spatial point processes and moment dynamics in the life sciences: A parsimonious derivation and some extensions. *Bull Math Biol*, 77, 586–613.

Record, N.R., Pershing, A.J. & Maps, F. (2014). The paradox of the "paradox of the plankton". *ICES Journal of Marine Science*, 71, 236–240.

Reigada, R., Hillary, R.M., Bees, M.A., Sancho, J.M. & Sagués, F. (2003). Plankton blooms induced by turbulent flows. *Proc. R. Soc. Lond. B*, 270, 875–880.

Reynolds, C.S. (2006). *The ecology of phytoplankton*. Cambridge University Press.

Seymour, J.R., Amin, S.A., Raina, J.B. & Stocker, R. (2017). Zooming in on the phycosphere: the ecological interface for phytoplankton–bacteria relationships. *Nat Microbiol*, 2, 17065.

Siegel, D.A. (1998). Resource competition in a discrete environment: Why are plankton distributions paradoxical? *Limnology and Oceanography*, 43, 1133–1146.

Wiegand, T., Gunatilleke, C.V.S., Gunatilleke, I.A.U.N. & Huth, A. (2007). How individual species structure diversity in tropical forests. *Proceedings of the National Academy of Sciences*, 104, 19029–19033.

Young, W.R., Roberts, A.J. & Stuhne, G. (2001). Reproductive pair correlations and the clustering of organisms. *Nature*, 412, 328–331.