Computational library for the Nutrient-Unicellular-Multicellular plankton modeling framework v. 1.0

A. Papapostolou[1], A.V. Almgren[1], T.F. Hansen[1], A. Kandylas[1], C. Serra-Pompei[1], A.W. Visser[1], and K.H. Andersen[1]

[1]Center for Ocean Life, Natl. Institute for aquatic resources, Technical University of Denmark

Correspondence to: Amalia Papapostolou (gnamalia@hotmail.com)

Abstract. The Nutrient-Unicellular-Multicellular (NUM) model is a trait-based model of unicellular and multicellular plankton that uses body size as the main structuring variable for community composition. The central feature is that body size is used for structuring predator-prey interactions and to scale parameters. For unicellular plankton, trophic strategies across the full range from osmotrophy, phototrophy, phagotrophy and mixotrophy, are an emergent outcome of the model. Another distinguishing feature is that the multicellular component, represented by copepods, includes ontogeny, which is crucial in shaping population dynamics. In addition, the framework encompasses a nutrient pool consisting of nitrogen, silica and dissolved organic carbon (DOC) which interacts dynamically with the plankton community in three model setups: a chemostat simulating the photic zone, a water-column, and a global setup. Here we present a user-friendly Fortran-Matlab library which makes the NUM model accessible as a practical tool for marine ecologists or biogeochemical modellers. The model output is validated with *in situ* and satellite data and demonstrates its applicability in chemostat, water-column, and global setups.

1 Introduction

Marine planktonic food webs play an important role in the ocean's biogeochemical cycles, influence the Earth's climate regulation, and determine potential fisheries yields (Stock et al., 2014). Plankton participate in the global carbon cycle by accounting for roughly half the global carbon fixation through photosynthesis, and by sequestering carbon via sinking dead detrital matter (Basu and Mackey, 2018; Boyd et al., 2019). The production of new biomass from photosynthesis also determines the amount of carbon available to higher trophic levels of fish production, which globally support about 17% of human protein consumption (Waite et al., 2014).

The central problem that ecosystem models face is how to represent the immense diversity of species in a tractable yet mechanistically meaningful way. The dominant approach has been to assemble similar kinds of species into functional groups. In the simplest formulation, the planktonic community is divided into phytoplankton and zooplankton in "NPZ" models (Franks, 2002). A finer representation of diversity is introduced by subdividing plankton into functional types (Quere et al., 2005) e.g. flagellates, dinoflagellates, ciliates, and diatoms (Butenschön et al., 2016), or into size groups, e.g. small- and large phytoplankton (Neumann, 2000) or sizes of zooplankton (Stock et al., 2014). The addition of each functional group fills another piece in the puzzle towards a complete representation of plankton diversity, however, it also brings a new set of parameters, which must be specified. The expanded set of parameters lends flexibility because the parameters for each functional group can be calibrated to reflect the dominant species in a particular region (Turner et al., 2014). While such calibrations make it possible to achieve good fits with observations, for global applications or for applications of environmental change, where the dominant species may change, the calibration may be driven outside its calibration envelope. Nevertheless, the functional group approach represents a robust solution to the representation of diversity that also retains a direct connection to the taxonomic representation of life.

An alternative approach has been to represent plankton diversity solely through differences in their size (cell size or body size) (e.g. Ward and Follows, 2016; Chakraborty et al., 2020; Andersen and Visser, 2023; Hansen et al., 2024). Using only size to represent diversity eschews any representation of taxonomic diversity, even down to not distinguishing between phyto- and zooplankton. Practically, body size is discretized in a number of size groups with an ordinary differential equation to represent each size group. Although size-based models may well have the same number of state variables as functional-type models, they have only one set of parameters that applies across all groups through scaling relations (Andersen et al., 2016). However, while size is indeed recognized as the "master" trait, pure size-based representations ignore key functional groups such as diatoms (Cadier et al., 2020), bacteria, or obligate phytoplankton and zooplankton. In practice plankton models often adopt a mixture of size-based and functional-group type of models, e.g., using size for diversity inside functional groups of phyto- and/or zooplankton (Banas, 2011; Stock et al., 2014; Ward et al., 2014, 2018), or diatoms (Terseleer et al., 2014). This method enables functional-group type of models to reduce their parameter set through scaling relations with size.

The approach followed here is to represent diversity through functional traits (Kiørboe et al., 2018). In the idealized case, a trait-based approach ignores all taxonomical differences between organisms and only describes differences between organisms through a small set of functional traits (Westoby and Wright, 2006). A trait-based representation of diversity requires a decision

about which trait-axes to model, i.e., which traits are most important for the functional diversity of the community. While there is wide agreement about cell or body size as the master trait, various other trait axes have been proposed for the secondary axes, such as investments into resource uptakes: phototrophy and nutrient harvesting (Bruggeman and Bolding, 2014) or including also phagotrophy (Berge et al., 2017; Chakraborty et al., 2017, 2020), degree of gelatinousness (Lemoine et al., 2025), feeding mode (active vs. passive) (Prowe et al., 2019; Serra-Pompei et al., 2020), or vacuolation (Cadier et al., 2020). When traits are discrete, as for active/passive feeding or for unicellular plankton with or without a silicate shell, the trait-based representation becomes essentially identical to the functional type approach. In that respect the differences between the two approaches are mainly conceptual – whether the representation of diversity is oriented towards taxonomical differences or functional traits.

Here we present an implementation of the trait-based Nutrient-Unicellular-Multicellular (NUM) plankton model framework (Serra-Pompei et al., 2020). The NUM framework attempts to make a pure trait-based plankton model. It uses cell size as the dominant trait for unicellular plankton and silicate shell as a discrete trait (essentially a diatom functional group). For multicellular plankton it uses adult size as the dominant trait and active/passive feeding mode as a secondary discrete trait. The NUM framework focuses on the ecology of the plankton and puts less focus on the chemistry. It therefore only includes a very basic representation of nutrient chemistry.

A distinguishing feature of the NUM model is a representation of multicellular plankton which explicitly resolves the developmental stages of copepods. Such representations remain rare (but see Record et al., 2013) for two reasons: First, multicellular zooplankton can increase by several orders of magnitude in body mass from eggs to adults, with a growth rate which depends on the availability of food, among other factors. Representing multiple stages requires several state variables to represent just one species. Second, copepod diversity varies strikingly in different regions (Rombouts et al., 2009) and representing all dominant species in global models is not feasible. The Nutrient-Unicellular-Multicellular (NUM) model solves the first problem by using an efficient representation of life history as a physiologically structured model (De Roos et al., 2008) and the latter problem by describing the diversity of copepods by two traits: their adult body size and their feeding mode (passive or active) (Serra-Pompei et al., 2020).

A central ambition of the NUM model is to base all parameters on "first principles", thereby avoiding free parameters that require calibration. The first principles can be related directly to physical laws (diffusion, fluid mechanics, etc.), energetics of chemical reactions, geometry, or mass conservation. This ambition is realized for many of the parameters for the unicellular compartment (Andersen and Visser, 2023). The challenges increase for the multicellular compartment, and here most parameters are determined by cross-species analyses. The description of the parameter values has been conducted elsewhere (Serra-Pompei et al., 2020; Cadier et al., 2020; Hansen and Visser, 2019; Andersen and Visser, 2023) and the focus here is on the implementation of the NUM framework as a computational platform for plankton ecology on a global scale.

A technical difficulty concerns the reduction of computational time for executing global simulations of full plankton ecosystem models. To provide marine ecologists without expertise in global biogeochemical modelling the ability to perform global simulations we based the global simulations on the transport matrix method (Khatiwala, 2007). The NUM model library version 1.0 is based on the NUM model as described in Serra-Pompei et al. (2022) with a number of key additions: 1) the inclusion of an explicit diatom functional group and silicate dynamics; 2) inclusion of labile dissolved organic carbon and (emergent)

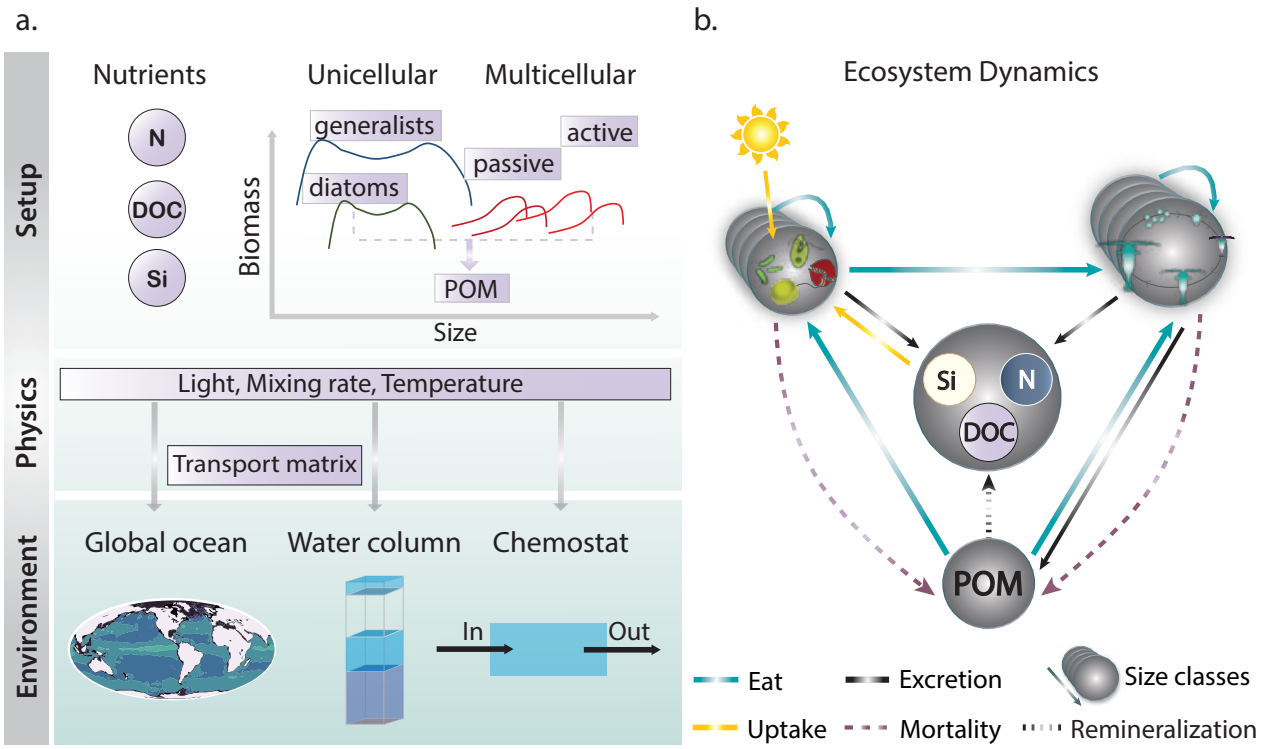

Figure 1. Conceptual sketch of the NUM framework. Panel (a) illustrates the ecological setup and the coupling to physical environments. Panel (b) describes the ecosystem dynamics. Several size-classes are modeled in each plankton group. The number of size classes and plankton groups can be defined at the beginning of each model run.

heterotrophic bacteria; 3) emergent respiration of unicellular plankton due to uptake regulation; 4) a fast Fortran implementation of the core functions; 5) a matlab front-end interface with three model setups: chemostat, water column, and global; 6) a flexible configuration of plankton communities, ranging from single-celled group (e.g. unicellular generalists, including osmotrophs, phototrophs, phagotrophs, and mixotrophs) to a full setup with unicellular generalists, diatoms, sinking particular organic matter, and a complete representation of the multicellular copepod community.

In the following, we first provide an overview of the NUM model structure and implementation, followed by a description of the key equations in the model. Next, we compile global calibration and validation data and compare it with global model simulations. Finally, we present examples of the model output, from global patterns to cell-level metabolism, and discuss open issues and potential applications of the NUM library.

2 The Nutrient-Unicellular-Multicellular library version 1.0

The core NUM model consists of a set of coupled ordinary differential equations for the state variables. These equations are all coded in a Fortran library for computational efficiency, while the code is executed from a matlab interface. The equations

require a representation of the physical environment, which could be a global or regional circulation model or a simple chemostat (Andersen and Visser, 2023). The matlab front-end to the Fortran library, interfaces with three environments: a chemostat, a water-column, and the entire globe. All the setups are derived from a global transport matrix that represents a discretized version of the underlying geophysical partial differential transport equations (Fig. 1a). The combination of computationally efficient compiled code with an interpreted language allows to run global simulations on a standard laptop.

The NUM state variables belong to one of three groups: biogeochemical tracers (referred to as "nutrients", including dissolved organic carbon DOC, nitrogen, and silicate), unicellular plankton (generalists or diatoms), multicellular zooplankton groups (copepods), and particulate organic matter (POM) (Fig. 1b). The three nutrient variables are the minimal set needed to represent growth limitation for phytoplankton (incl. diatoms which are limited by silicate) and the carbon source for osmoheterotrophs (bacteria). The unicellular and multicellular groups are represented by "size spectra" encompassing a number of state variables that each models the dynamics of one size group (Fig. 1a). By implementing several sizes of the unicellular and multicellular groups, the size spectrum of the entire plankton community emerges. Size – defined as cell mass of unicellular organisms and body mass of multicellular organisms – is the key trait that describes resource encounter and uptake, as well as the physiology of individual plankton organisms (metabolism, growth, and reproduction) (Hillebrand et al., 2022). Further, size determines the central process in the model, which is predation by larger organisms on smaller ones (Wirtz, 2012). By using size as the core structuring variable, each organism group is described by one set of parameters defined through size scalings. This approach reduces the overall number of parameters and permits a flexible number of state variables.

The four size spectrum groups described below are generalists, diatoms, copepods, and dead particulate organic matter (POM).

2.1 Unicellular generalists

The "generalist" group represents all unicellular plankton including bacteria, phytoplankton, phagotrophs, and mixotrophs (organisms which combine trophic strategies, such as photosynthesis, diffusive uptake of nutrients and predation on smaller cells), but excluding diatoms. The generalist group represents all unicellular plankton (heterotrophic and photosynthetic bacteria, flagellates, dinoflagellates, ciliates, etc.) except diatoms, which we simulate as a separate group (explained in the following section). The generalists are all modeled as potential mixotrophs that acquire carbon and nutrients through a combination of osmotrophy (taking up dissolve organic carbon), autotrophy (photosynthesis), and/or phagotrophy. Whether a given size group represents bacteria, phytoplankton, etc., depends on the principal mode of carbon acquisition , which is determined by a combination of cell size and the environment. The trophic strategies of each size class is therefore an emergent property and varies dynamically over space and time. The generalist spectrum is thus a trait-based model with cell size being the trait.

2.2 Diatoms

Large diatoms, that thrive in eutrophic environments, have historically been correlated with food webs that lead to major fisheries (Ryther, 1969). Thus, the inclusion of diatoms in the model is critical for the estimation of fish production potential. Diatoms are simulated as a separate group because of their importance in marine systems (they account for about 40% of marine

primary production (Tréguer et al., 2018)), and their distinct characteristics that affect their physiology and predation mortality. Diatoms enclose a vacuole that increases their volume, allowing them to have larger physical size than other cells of the same biomass. In the model, this trait allows them to increase their volume per carbon without additional nutrient demands allowing them to have a higher per-carbon diffusive uptake than non-vacuolated cells. Further, diatoms are engulfed in a silicate shell, which makes silicate another limiting nutrient. The hard silica shell also reduces predation pressure on the diatoms (Hamm et al., 2003; Pančić et al., 2019). The rigidity of the hard silica shell deprives the cell of the plasticity needed to engulf prey and excludes diatoms from eating prey. Diatoms therefore cannot perform phagotrophy and only obtain nutrients through diffusive uptake. The silicate shell also induces growth limitation of diatoms by dissolved Si, besides light and nitrogen.

2.3 Copepods

Multicellular zooplankton are the key link between lower and higher trophic levels, as they ingest unicellular plankton and are themselves prey for fish. Among the different zooplankton taxa in the ocean, copepods are the most abundant metazoans (Humes, 1994). Copepods undergo size changes throughout their life stages and thereby occupy different ecological niches. These ontogenetic niche shifts substantially impact population dynamics (Verity and Smetacek, 1996). For instance, copepod growth rates are strongly correlated with spring bloom intensity (Durbin and Durbin, 1992). To account for these size changes, each population of copepods is represented by a size spectrum that spans from the offspring size to the adult size.

Each population of copepods is characterized by two traits: the adult size and their feeding strategy (active or passive). In this way, the model represents the large differences in adult size observed in the oceans (Brun et al., 2016). Modelling the diversity of copepods with these two main traits makes it possible to represent the entire copepod community in the global ocean with just a few populations that span the trait space of adult size and feeding modes.

2.4 Particulate organic matter

Dead organic matter (detritus) is represented by a size spectrum of particulate organic matter (POM) that includes dead organisms, cell fragments from lysis, and fecal pellets produced by multicellular plankton. POM may be represented by a size spectrum to enable the calculation of sinking velocities for different POM sizes. However, in the simulations presented here we simulated just a single size class of POM.

3 Model description

Here we describe the main processes in the NUM model. We first describe the size-based predation among modeled organisms and by larger "higher trophic level" (HTL)organisms that are not represented by the model. Then we describe the cell model (generalists and diatoms), the multicellular model, the POM model, and the biogeochemical model (Fig. 1). The full equations for each model component are listed in the supplementary material (App. B).

3.1 Predation

Predation is the central process that connects the plankton size classes in NUM. Each size class i is characterized by its mass m_i (either cell mass or body mass in μg carbon) and the organisms in the class prey on smaller organisms m_{prey} with a log-normal preference function:

$$\phi_{kl}(m, m_{\mathrm{prey}}) = p_{kl} \exp\left[-\frac{1}{2\sigma_k^2}(\ln(\frac{m}{\beta_k m_{\mathrm{prey}}}))^2\right], \tag{1}$$

where β is the preferred predator:prey mass ratio, σ is the width, and p_{kl} is the preference of the predator group k towards the prey group l (all parameters are specified in Tables B2, B4, and B5 in the supplementary). As each size class represents a finite range of sizes (with m representing the geometric mean; see Fig. C1), the actual preference is the average of the encounter over all prey and predator masses, so the average encounter θ_{ij} between two size classes i (predator) and j (prey) becomes a more complex function (App. D).

The available food for size class i is the sum over all other size classes weighted by the preference (θ_{ij}): $F_i = \sum_i \theta_{ij} B_j$ where B_j is the biomass concentration of prey (μg C L^{-1}). The actual consumption is limited by a maximum mass-specific maximum consumption rate: j_{Fmax} (day^{-1}):

$$j_{\mathrm{F}.i} = \epsilon_{\mathrm{F}} f_i j_{\mathrm{Fmax}.i} \quad \text{with} \quad f_i = \frac{a_{\mathrm{F}.i} F_i}{a_{\mathrm{F}.i} F_i + j_{\mathrm{Fmax}.i}}, \tag{2}$$

where $a_{\mathrm{F}.i}$ is the mass-specific clearance rate of the predator (L(μg C day)$^{-1}$) and ϵ_{F} is the assimilation efficiency. The predation results in a predation mortality on the prey size class j:

$$\mu_{\mathrm{pr}.j} = \sum_i \theta_{ij} \frac{j_{\mathrm{F}.i} B_i}{\epsilon_{\mathrm{F}} F_i}. \tag{3}$$

In addition to the predation mortality inflicted by larger plankton, the largest size classes are also exposed to predation by higher trophic levels that are not explicitly represented in the model, such as fish. The exposed size classes are those larger than m_{HTL} (μg C), described by the selectivity function p_{HTL} :

$$p_{\mathrm{HTL},i} = \frac{1}{1 + (m_i/m_{\mathrm{HTL}})^{-2}}. \tag{4}$$

Here we use a value of the $m_{\mathrm{HTL}} = 1\,\mu$g C ($\approx 500\mu$m) as the size of prey of forage fish in the size range 1-10 g wet weight (Andersen, 2019, Chap. 2). Further, in the standard setup presented here, we constrain the higher trophic level (HTL) mortality to act only on copepods. If a setup is run without copepods the HTL mortality would represent the predation by the copepods on the unicellular organisms. The higher trophic level mortality can either be a constant or proportional to the biomass concentration in the size class:

$$\mu_{\mathrm{HTL},i} = \begin{cases} \mu_{\mathrm{HTL},0}\, p_{\mathrm{HTL},i} & \text{constant} \\ \mu_{\mathrm{HTL},0}\, p_{\mathrm{HTL},i} \frac{B_i}{\ln(1/z_i)} & \text{“quadratic”} \end{cases} \tag{5}$$

The density dependent mortality is termed “quadratic” to follow common nomenclature, though the mortality is in reality linear and it is only the loss term (mortality multiplied by biomass) that is quadratic. An increasing mortality with biomass is

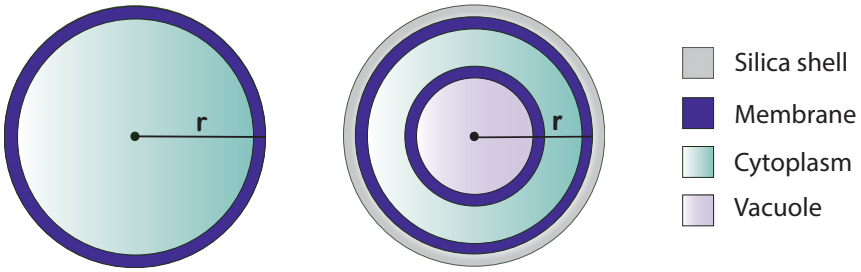

Figure 2. Geometry of generalist and diatom cells

commonly argued to be the result of the varying abundance of predators — a high biomass leads to more predators and *vice versa* (Heath et al., 2014). However, this argument is incorrect because it fails to recognize that marine ecosystems generally follow a Sheldon spectrum (Sheldon et al., 1977), where the ratio of biomass between predators and prey is a constant. As the mortality is also a function of the predator:prey biomass ratio, it follows that the mortality is independent of biomass (see also Andersen and Beyer (2006)). Nevertheless, the quadratic formulation of mortality has the practical advantage that it helps stabilize the dynamics in the simulation – running without a quadratic loss term results in competitive exclusion among copepod groups and limited co-existence. We therefore use the quadratic mortality formation here in this capacity. The term $1/\ln(1/z_i)$ is a correction for the number of size classes used. Size classes are evenly distributed on a logarithmic size grid and in that case the number of size classes in a fixed size range is proportional to this correction term, where z_i is the ratio between the lower and the upper size range in each size class. It should be noted that higher trophic level mortality does not respect the increased prey vulnerability due to motility that predation mortality follows. This assumption follows from higher trophic level organisms being visual predators and the hydrodynamic signals from prey motility do not affect the detection of prey (Howe et al., 2018; Lange Jensen et al., 2023).

3.2 Cell model

The unicellular spectra can be either generalists or diatoms. The only differences between the two spectra are that diatoms have a vacuole, that they cannot perform phagotrophy, and that they have a lowered predation risk ($p < 1$) (as explained above). The present description of diatoms with a fixed vacuole size is a simplified version of the description with dynamic vacuoles in Hansen and Visser (2019); Cadier et al. (2020). The full set of equations and parameters for the unicellular model are given in Tables B1 and B2.

The cell model describes the encounter and uptake of resources and metabolism of an individual plankton cell as modulated by the cell's size (mass) m (μg C). The size of the cell defines traits, and traits in combination with the environment define its trophic strategy, whether it is an osmo-heterotroph (a bacterium feeding mainly on DOC), a light- or nutrient-limited phototroph, or, for the generalists, a mixotroph (combining osmo-, photo-, and phagotrophy) or a heterotroph (living mainly on phagotrophy) (Andersen et al., 2016).

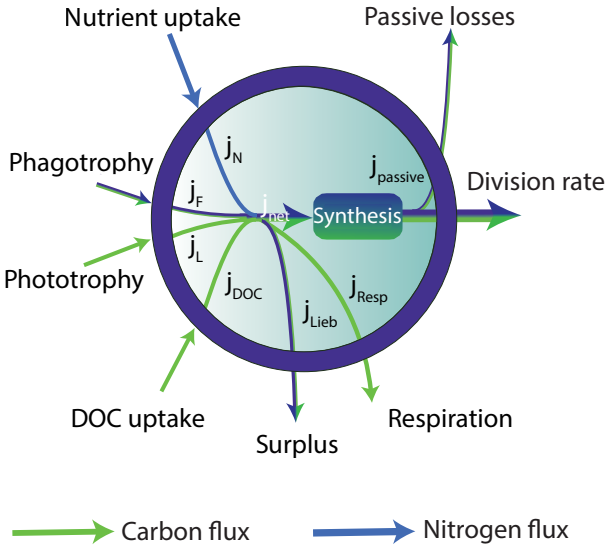

Figure 3. Illustration of the cell metabolism of a generalist with uptakes of resources as carbon (blue arrows) and nutrients (green arrows), losses due to respiration or passive losses, biosynthesis leading to cell division rate g, resulting in a net division rate. The diatom cell is similar, but with an additional uptake of silicate j_{Si} and no phagotrophic uptake.

Regarding their morphology, we assume that cells are spherical and consist of a cytoplasm surrounded by a membrane of thickness δ (Fig. 2). The diatom cells additionally have a vacuole that occupies a fraction v_{vac} of the cell volume and is surrounded by an additional inner membrane. The radius of a cell with mass m is approximately (assuming $\delta \ll r$) (eq. B5):

$$r = \left(\frac{3}{4\pi} \frac{m}{\rho} \frac{1}{1 - v_{\mathrm{vac}}} \right)^{1/3}, \tag{6}$$

where ρ is the density of the cytoplasm (μg C μm^{-3}). As part of the cell's mass is used for the membrane(s), the fraction of the cell mass available for active metabolism for diatoms is calculated in (eq. B6) and for generalists is (Andersen and Visser, 2023):

$$\nu = 1 - 3\frac{\delta}{r}. \tag{7}$$

The cell model describes three processes: encounter with resources, uptake of resources and metabolism, and mortality:

3.2.1 Encounter

The cell takes up carbon via photosynthesis (light), while both carbon and nitrogen are taken up via phagotrophy (i.e. eating prey; only generalists) and diffusion, in the form of dissolved organic carbon (DOC) and dissolved inorganic nutrients respectively (Fig. 3). Diatoms also take up silicate to build their silica shell. The encounter with the resources (light L, DOC, N, prey

F, and silicate Si) is described as specific mass fluxes j_{X} (day^{-1}):

$$j_{\mathrm{enc.X}} = a_{\mathrm{X}}(m) X \rho_{\mathrm{C:X}}, \text{ with } X \in \{\mathrm{DOC}, \mathrm{N}, \mathrm{L}, \mathrm{F}, \mathrm{Si}\} \tag{8}$$

where $\rho_{\mathrm{C:X}}$ is the ratio between carbon and the element of the resource X and a_{X} is the mass specific affinity.

The affinity for each resource uptake is determined by the geometric properties of the cell and is expressed as a function of mass (m) and the radius r. The dependency of affinities on size is analysed comprehensively in Andersen and Visser (2023) and given in Table B1. Generally, the diffusive uptake is limited by the cell radius, light uptake by the cell cross-sectional area, and feeding by the cell volume.

3.2.2 Metabolism

The cell division rate depends on light and available resources. Uptake of resource X has a metabolic carbon cost $\beta_{\mathrm{X}} j_{\mathrm{X}}$. As the cell has different sources of carbon (from photosynthesis, DOC, and food) and nutrients (from dissolved N and food), each with different metabolic costs of uptake, it regulates the uptakes j_{X} up to the encountered uptakes $j_{\mathrm{enc.X}}$ down to the effective uptakes j_{net} to maximize division rate under the constraint of a fixed C:N:Si stoichiometry (Table B1 "resource uptakes"). The regulation of light absorption acts as a simple representation of light-adaptation, while the regulation of the nutrient uptakes represent Liebig's law. The emergent respiration is:

$$j_{\mathrm{resp}} = j_{\mathrm{R}} + \sum_{\mathrm{X}} \beta_{\mathrm{X}} j_{\mathrm{X}} + \beta_{\mathrm{g}} g, \tag{9}$$

where j_{R} is the basal metabolism and $\beta_g g$ is the growth metabolism as a fraction of the division rate g. The cell division rate is limited by its capacity for biosynthesis $j_{\max}$ (Eq. B1.15):

$$g = j_{\max} \frac{j_{\mathrm{net}}}{j_{\mathrm{net}} + j_{\max}} - j_{\mathrm{passive}}, \tag{10}$$

where $j_{\max}$ is proportional to the active cell mass v (7), and j_{passive} are passive losses across the cell surface.

3.2.3 Population growth

Besides predation losses (eqs. 3 and 5) the cell is exposed to viral lysis with a mortality μ_2 proportional to the cell biomass:

$$\mu_{2.i} = \mu_2 B_i / \ln(\Delta m_i) \tag{11}$$

where Δm_i is the width of the size class (Fig. C1). The correction with the width of the size class is needed because the biomass B_i describes the biomass in a size class i, and the biomass will vary depending on the width of the size class (increasing the size range (width) of a size class by a factor increases the biomass in the bin by the same factor).

The final growth equation for the unicellular plankton is:

$$\frac{\mathrm{d}B_i}{\mathrm{d}t} = \left(g_i - \mu_{\mathrm{pr},i} - \mu_{\mathrm{HTL},i} - \mu_{2,i}\right) B_i, \quad i \in \mathrm{uni} \tag{12}$$

where $\mathrm{uni} = \{\text{all unicellular groups}\}$.

3.3 Multicellular model

Copepods represent multicellular zooplankton and the copepod community consists of different populations. Each population is characterized by the adult body-mass and the feeding mode (active or passive). The spectrum for the population resolves the life stages of copepods, from nauplii to adults, following the description in Serra-Pompei et al. (2020):

$$\frac{\mathrm{d}B_i}{\mathrm{d}t} = J_{\mathrm{in},i} + (j_{\mathrm{out}.i} + g_i - \mu_{\mathrm{pr}.i} - \mu_0 - \mu_{\mathrm{HTL},i})B_i, \quad i \in \mathrm{multi} \tag{13}$$

where $\mathrm{multi} = \{\text{all multicellular groups}\}$. Here J_{in} (μg C L^{-1} day^{-1}) is the mass flux into the size class, which equals the mass flux out of the previous class, except for the first class where it is the reproductive flux. The other fluxes are "pure" rates in units of 1/day: $j_{\mathrm{out}.i}$ is the flux of out a size class, g_i is the biomass accumulation rate, and μ_0 is background mortality. The fluxes and the mortalities are calculated from the consumption by predation (eq. 2) and the predation mortality (eq. 3), and are detailed in Table B3.

3.4 Particulate organic matter (POM)

Particulate organic matter is produced by losses from plankton (Fig. 1). The POM has the same C:N ratio as the plankton, meaning that POM does not transport silicate. Instead, all silicate that should have gone to POM is considered lost to the deep sea and therefore also lost from the model. POM is generated from several sources (Fig. 4): 1) a fraction $1 - \gamma_2$ of cells dying from lysis; 2) Multicellular plankton produce POM from sloppy feeding, fecal pellets (incomplete assimilation) and from the losses due to inefficient reproduction (egg mortality); 3) A fraction $\gamma_{\mathrm{POM}} = 1 - \gamma_{\mathrm{HTL}}$ of the losses from higher trophic level mortality is assumed to be fecal pellets routed to POM. Particulate organic matter is lost from remineralization with a rate r_{POM} and from grazing by copepods:

$$\begin{aligned}\frac{\mathrm{d}B_i}{\mathrm{d}t} = &\sum_j \left[(1-\gamma_2)\mu_{2.j} + (1-\gamma_{\mathrm{F}})\frac{1-\epsilon_{\mathrm{F}}}{\epsilon_{\mathrm{F}}} j_{\mathrm{F}.j} + (1-\gamma_{\mathrm{HTL}})\mu_{\mathrm{HTL},j}\right] B_i \\ &+ \sum_{S \,\in\, \text{multi adults}} (1-\epsilon_{\mathrm{R}}) g_S B_S - (r_{\mathrm{POM}} + \mu_{\mathrm{pr}.i})\, B_i, \quad i \in \mathrm{POM},\end{aligned} \tag{14}$$

where index S indicates the last (adult) size class in each copepod population. Parameters are given in Table B5.

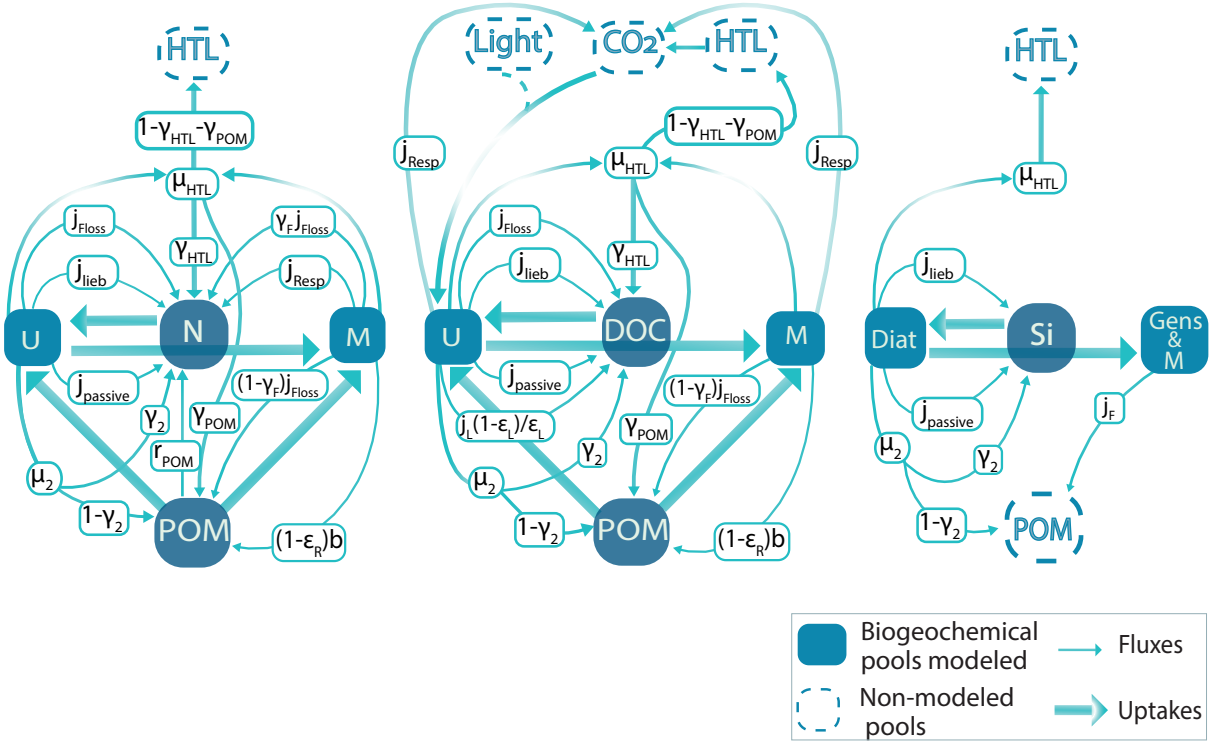

Figure 4. Representation of nitrogen, carbon and silica fluxes in unicellular (generalists and diatoms) (U) and multicellular (M) plankton. All rates are in units day^{-1}. In the chemostat setup POM sinks out to the deep (Eq. 20). The j's represent direct losses, the μ's represent mortalities, γ's represent fractioning of a flux into different pools (e.g. between direct remineralization and to POM), $b = J_{\mathrm{in.1}}/m_{\mathrm{egg}}$ is the copepod birth rate, and ε_{F} is feeding efficiency. $\gamma_{\mathrm{POM}} = 1 - \gamma_{\mathrm{HTL}}$, and j_{lieb} refers to the losses of N and Si when growth is negative, and surplus carbon from photosynthesis to DOC. Note that POM does not contain silicate, instead all silicate fluxes that should have gone into POM are considered lost to the deep sea, therefore also absent in the model. Similarly, we assume that none of the silica in diatoms is assimilated by their predators and it is all remineralized.

3.5 Bio-geochemical dynamics

The dissolved state variables are taken up through diffusive encounter and replenish with the losses from the plankton groups (Fig. 4). The equations for the dissolved phases of nitrogen, DOC, and silicate are:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \left[\sum_{i\in\mathrm{uni}} \left(-j_{\mathrm{N},i} + j_{\mathrm{Nloss},i} + j_{\mathrm{passive},i} + \gamma_{\mathrm{F}} j_{\mathrm{Floss},i} + \gamma_2 \mu_{2,i}\right) B_i + \gamma_{\mathrm{HTL}}\mu_{\mathrm{HTL}} \sum_{i\in\mathrm{multi}} B_i + r_{\mathrm{POM}} \sum_{i\in\mathrm{POM}} B_i\right] \frac{1}{\rho_{\mathrm{C:N}}} \tag{15}$$

$$\frac{\mathrm{d}C}{\mathrm{d}t} = \sum_{i\in\mathrm{uni}} \left(-j_{\mathrm{DOC},i} + j_{\mathrm{C.loss},i} + j_{\mathrm{passive},i} + j_{\mathrm{Lloss},i} + \gamma_{\mathrm{F}} j_{\mathrm{Floss},i} + \gamma_2 \mu_{2,i}\right) B_i \tag{16}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \left[\sum_{i\in\mathrm{diatoms}} \left(-j_{\mathrm{Si},i} + j_{\mathrm{SIloss}.i} + j_{\mathrm{passive},i} + \gamma_2 \mu_{2,i}\right) B_i\right] \frac{1}{\rho_{\mathrm{C:Si}}}. \tag{17}$$

We have checked that all nutrients are conserved within an error close to numerical noise in Eqs. 12-17, while accounting for losses of silicate to the deep sea.

3.6 Temperature

Temperature T (°C) influences the physical and the metabolic processes of all groups in the model. The influence of temperature is described via Q_{10} corrections to the parameters from a reference temperature of 10° as:

$$Q(T) = Q_{10}^{(T-10^\circ)/10^\circ}. \tag{18}$$

The only physical process influenced by temperature is the diffusion rate of dissolved matter (DOC, Si, and nitrogen) towards the cell, which has a $Q_{10} = 1.5$ (Serra-Pompei et al., 2019). This means that the parameter α_{N} is corrected by temperature (Table B2). The other temperature correction is on metabolic rates by $Q_{10} = 2$. The affected metabolic rates are: the basal metabolic rate (j_{R} for unicellulars and k_{R} for multicellulars), the maximum synthesis rates ($\alpha_{\max}$ and h respectively), and the remineralization rate of POM r_{POM}. Applying temperature corrections on the individual processes means that the temperature response of division rate and growth are emergent and typically do not follow a Q_{10} relation (Serra-Pompei et al., 2019; Andersen and Visser, 2023).

3.7 Embedding within ocean physics

The core NUM model library solves the state-variable equations (12-17) and can be embedded into a full circulation model. The library provides three simulation environments: chemostat (steady state and seasonal), water-column, and global. All environments, except the steady state chemostat, are implemented based on the transport matrix method (TMM) (Khatiwala et al., 2005; Khatiwala, 2007). The transport matrix method is a computationally cost-efficient offline numerical scheme to simulate ocean biogeochemical tracers, where a sparse matrix represents the advective-diffusive transport of a passive tracer. The transport matrix approach is a reliable alternative to conventional ocean general circulation models that reporduces reasonably well both the mean spatial and seasonal patterns of biogeochemical tracers in the ocean (Kvale et al., 2017). The NUM library employs the MITgcm transport matrices in a coarse resolution of $2.8^\circ \times 2.8^\circ$ and 15 vertical layers (Dutkiewicz et al., 2005), and in higher resolution of $1^\circ \times 1^\circ$ with 21 vertical layers (ECCO) (Forget et al., 2015; Transport Matrix Configurations). Both setups provide monthly transport matrices with a transport time step of 0.5 day. Sinking of POM is resolved with an implicit first-order upwind scheme (Press et al., 2007, Chap. 19). For the global and water-column simulations, the state-variable equations are solved with a forward Euler scheme with a simple predictor-corrector step to avoid negative values of the dissolved tracers, and a time step of 0.1 day. The bottom boundary conditions are closed for all living biological state variables and for DOC. POM is allowed to sink into the bottom. Nutrients (nitrogen and silicate) are nudged towards a fixed concentration at the bottom:

$$u_{t+1} = u_t + \kappa \frac{\Delta t}{\Delta z}(u_{\mathrm{bottom}} - u_t), \tag{19}$$

where u is the state variable (N or S), t is the time step, Δt is the transport time step (0.5 day), Δz is the thickness of the bottom layer, $\kappa = 1/365$ m day^{-1} is the relaxation rate, and u_{bottom} is the bottom concentration taken from World Ocean Atlas climatologies (Reagan et al., 2023).

The chemostat environment models the upper mixed layer which mixes with an unresolved deep layer, with a mixing rate d (day^{-1}) (Evans and Parslow, 1985). The deep layer concentrations of nutrients are fixed, N_{deep} and S_{deep}, and the concentration of DOC is zero. Unicellular organisms are mixed out of the upper layer, while the multicellular organisms are assumed to stay in the upper layer. POM is mixed out of the surface layer but also sinks out with a rate v/M, where v (m day^{-1}) is the sinking velocity and M (m) the thickness of the mixed layer. The governing equations are solved with the internal matlab Rosenbrock 2nd order stiff solver:

$$\frac{\mathrm{d}B_i}{\mathrm{d}t} = f_i(L,T,\mathbf{B}) + \underbrace{d(B_{\mathrm{deep},i} - B_i)}_{\text{for nutrients and uni.}} - \underbrace{\frac{v}{M}B_i}_{\text{for POM}}, \tag{20}$$

where the first term $f_i(L,T,\mathbf{B})$ is the right-hand-side of the state-variable equations (12-14), which depends on light L, temperature T and the set of all the state variables $\mathbf{B}$. Finally, the chemostat can be run in a seasonal environment where light, temperature, and mixing rate d vary with time and are extracted from the transport matrix. In high latitude environments it is necessary to disable the mixing of unicellular plankton to the deep layer to allow them to survive the winter period.

3.8 Model setup and numerical solution

The model framework allows for a flexible combination of the four groups. A model setup requires at least one unicellular group (generalists or diatoms), and may include POM and a number of copepod groups. The default NUM model setup used here involves all four groups. Each group contains a number of size classes and the model results depend, to some degree, on the number of size classes used in the unicellular groups (Fig. 5a,b) and in the copepod groups (panels c,d). Further, it depends on the number of active-copepod and passive-copepod groups (panels e,f). The number of size classes and copepod groups in the default NUM model setup is chosen such that the overall results, in terms of size spectra (Fig. 5a,c,d) and ecosystem function (panel b,d,f), do not change considerably upon the addition of more classes or groups. The setup includes 10 size classes of generalists, 10 size classes of diatoms, two passive copepod groups (adult sizes 0.2 and 5 μg C; prosome lengths 320 and 1040 μm), and three active copepod groups (1, 32, and 1000 μg C; prosome lengths from 584 and 7200 μm), where each group is discretized in 6 size-classes representing the growth in body mass. The number of state variables adds up to a total of 54.

Running the NUM model from the matlab interface proceeds in three steps: 1) setting up the simulation type (i.e., choosing which size spectra groups to simulate), 2) setting up the simulation environment (chemostat, water column, or global), 3) running the model, and 4) plotting the output:

```
p = setupNUMmodel( bParallel=true ); % Model setup
p = parametersGlobal( p );           % Model environment
sim = simulateGlobal( p );           % Run simulation
```

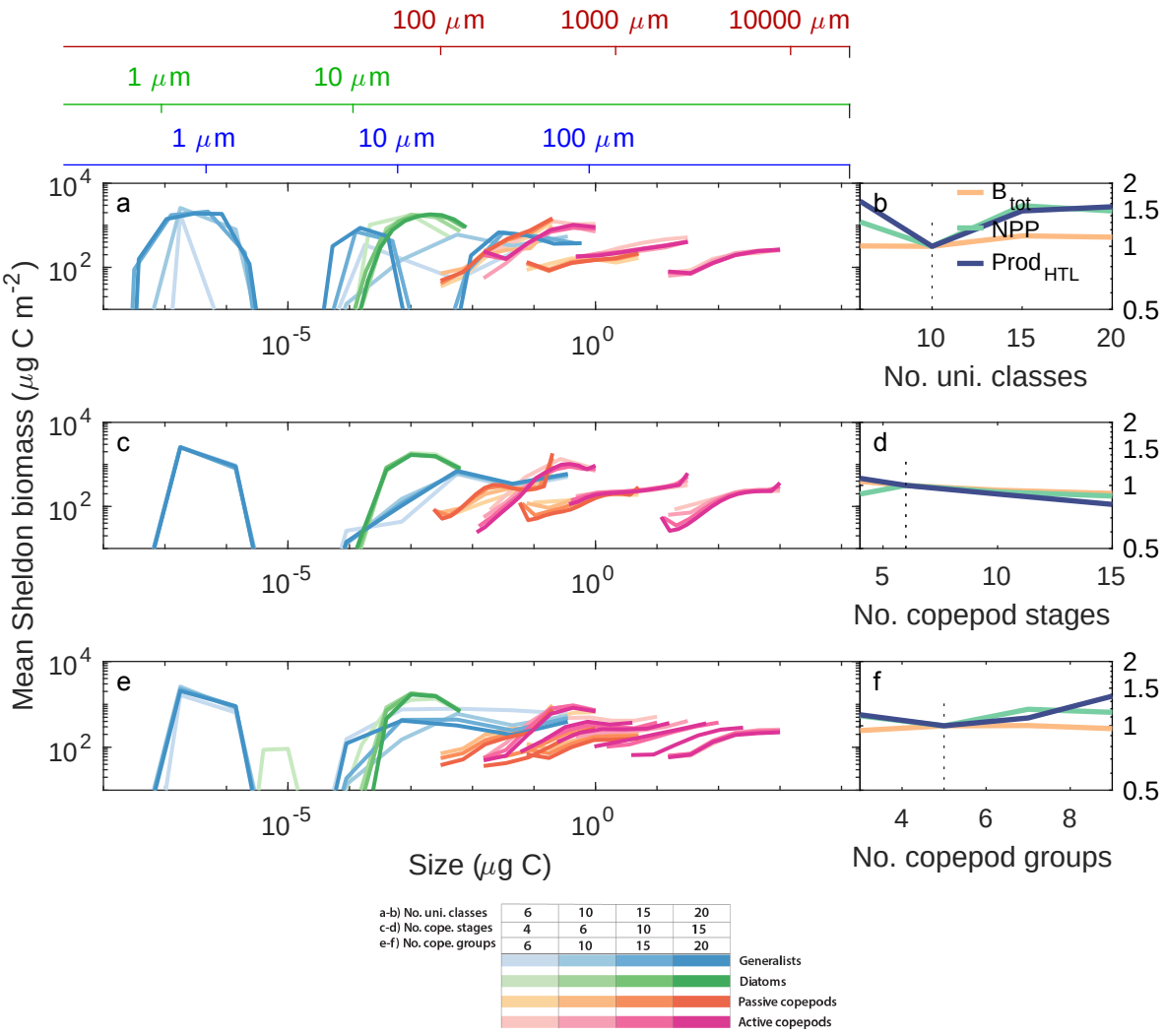

Figure 5. Sensitivity of the model to number of unicellular size classes (a-b), number of copepod stages (c-d), and number of copepod groups (e-f). The left column of panels show the Sheldon size distributions (biomass normalized by the width of each size bin; Andersen and Visser (2023, Box V)) and the right column shows total biomass (orange), NPP (green), and HTL production (dark blue) normalized by the value in the standard setup (vertical dashed line). All runs are from a simulation of the water column model at 60°N and 40°W. Color intensity indicates number of classes/groups. In the left panels, blue shades represent generalists, green diatoms, orange-red passive copepods and magenta active copepods. The colored axes on top show the lengths of the three groups: equivalent spherical diameter for generalists (blue) and diatoms (green), and prosome length for copepods (red).

```
plotSimulation( sim );
```

All parameters are encapsulated in the `p` structure, which is passed to the simulation (in this example a global simulation). The simulation returns the entire output in the `sim` structure which is passed on to other functions for analysis or plotting. These two structures are documented on the github site and all functions are documented in their `help` pages.

3.9 Calibration and evaluation data

All parameters of the plankton groups are based on first principles or cross-species comparisons (Andersen and Visser, 2023; Serra-Pompei et al., 2020). While these parameters are uncertain (explored by Hansen et al. (2024)), they are not expected to vary spatially. However, some of the extrinsic parameters either have no first-principle arguments or data, and/or are varying spatially. These extrinsic parameters together mold the overall function of the community: the global average light extinction coefficient by water k_w without feedback from plankton, the average POM sinking velocity v, and the higher trophic level mortality $\mu_{\mathrm{HTL.0}}$. These variables conjointly determine the total biomass and production of the global model: lower light extinction increases gross primary production, lower sinking velocity decreases the remineralisation depth and thereby increases the amount of primary production. Decreasing the mortality by higher trophic levels increases the copepod biomass, which influences the unicellular spectrum through a trophic cascade (Fig. F5). In nature, the values of these three variables vary on a global scale, but here they are kept constant. Consequently, we must calibrate to find reasonable global average values. The calibration algorithm minimizes the error between modeled and observed values, based on the following steps:

First we compute the mean error between insitu data Q_{obs} to modeled output Q_{model} for three metrics: picoplankton($i = 1$), POC($i = 2$), copepod biomass ($i = 3$) in varying space and time for each metric, such that the number of available measurements of each metrics i is n_i, and each distinct measurement is j. This calculation is represented in the first addend in equation 21. The datasets are described below. The datasets used in this part of the algorithm are: 1) Picophytoplankton biomass (organisms with a diameter less than 2 μm) between 0 and 10 m depth from water samples between 1997 and 2014, across all latitudes mainly in the Atlantic Ocean (fig. E3) (Martínez-Vicente et al., 2017); 2) Particulate Organic Carbon (POC) biomass of all plankton and detritus particles with a radius between $0.35 - 15$ μm from the GO-POPCORNv2 dataset (Tanioka et al., 2022) from cruises (2011-2020) across all major ocean basins, spanning from 70°S to 55°N (fig. E2); 3) Copepod biomasses across the Atlantic Meridional Transect cruises (López and Anadón, 2008).

$$\varepsilon = \frac{1}{6}\left[\sum_{i=1}^{3}\frac{1}{n_i}\sum_{j}^{n_i}\log\left(\frac{Q_{\mathrm{model}.ij}}{Q_{\mathrm{obs}.ij}}\right) + \sum_{k=1}^{3}\log\left(\frac{P_{\mathrm{NPPmodel}.k}}{P_{\mathrm{NPPobs}.k}}\right)\right]. \tag{21}$$

Second, we select three distinct locations in the ocean that correspond to an oligotrophic, a eutrophic and a seasonally varying environment. (specify) to compare the mean error between the annual averages of satellite derived NPP and modelled NPP (P_{NPPmodel}). This calculation is shown in the second addend of equation 21. P_{NPPobs} corresponds to annual averages (2002-2021) from satellite observations processed with the CAFE model, selected due to its better performance on a global average relative to other satellite-derived NPP models (Silsbe et al., 2016). The three sites chosen for comparison are: oligotrophic

location (22°N, 158°E), a eutrophic location (5°S, 5°E), and a seasonal location (60°N, 40°W). At these sites, the annual averaged NPP over the entire water column from CAFE is approximately 200, 1000, and 500 mg C yr^{-1}, respectively.

The net primary production is calculated in the model as the amount of carbon that is fixed by photosynthesis minus what is used for respiration:

$$P_{\mathrm{NPP}} = \sum_{i \in \mathrm{uni}} \max\{0, (j_{\mathrm{L}.i} - j_{\mathrm{resp}.i}) B_i\}. \tag{22}$$

This definition appears simple but there are two important comments. First, all mixotrophic metabolic costs are included in the respiration (j_{resp}) that is subtracted from the fixed carbon (j_{L}), including the costs of feeding and DOC uptake. This is different from the usual thinking about primary production, which implicitly assumes that primary production only occurs by obligate phototrophs. Second, it should be noted that the definition does not include all primary production. Some carbon is lost to DOC during the uptake process (Fig. 4b), which can fuel osmotrophic biomass production. The DOC part of primary production can contribute a substantial portion of measured primary production (Karl et al., 1998). Adding this potential carbon production into the definition can be achieved by replacing j_{L} with $j_{\mathrm{L}}\epsilon_{\mathrm{L}}^{-1}$. However, we refrain from including DOC NPP, which may result in underestimating measured NPP at oligotrophic sites.

After the initial calibration of the three parameters we evaluate the performance of the NUM model using a second set of observations. For the performance evaluation we use observations of nano- and microplankton from three Atlantic Meridional Transects cruises (Rodríguez-Ramos et al., 2015), a meta-analysis of global mesozooplankton biomass distribution (Moriarty et al., 2013), and annual averages of surface nutrients from the World Ocean Atlas (Reagan et al., 2023). To validate the global diatom distributions we use the Copernicus Global Ocean Colour satellite product (CMEMS, 2025). This product provides the surface chlorophyll divided between groups. The ratio between the total chlorophyll and the diatom chlorophyll is a good proxy for the ratio between modeled phytoplankton and diatom biomasses. However, the satellite observations only capture the first few meters of the surface, whereas the top layer in the global model is mean over the top 50 m. This evaluation data-set will therefore be inaccurate in the areas with a deep chlorophyll maximum.

4 Results

Simulations are initiated with concentrations of inorganic nitrogen and silicate from the World Ocean Atlas (Reagan et al., 2023). Biomass distributions equilibrate quickly and after 10 years are well converged (Fig. F6b, d). Deep nutrient concentrations are on a long slow transient, particularly away from seasonal environments (Fig. F6a, c). The value of the bottom boundary condition exhibits a very weak effect on the biomass and net primary production (Fig. F7). The simulations presented in the following are after 10 years of simulation with the standard NUM model setup (Section 3.8).

We first show the model calibration followed by the model global biomass distributions. We then assess performance with evaluation data, show how the three model environments (global, water column, and chemostat) compare, and finally show outputs related to the physiological dynamics of the organisms in the model.

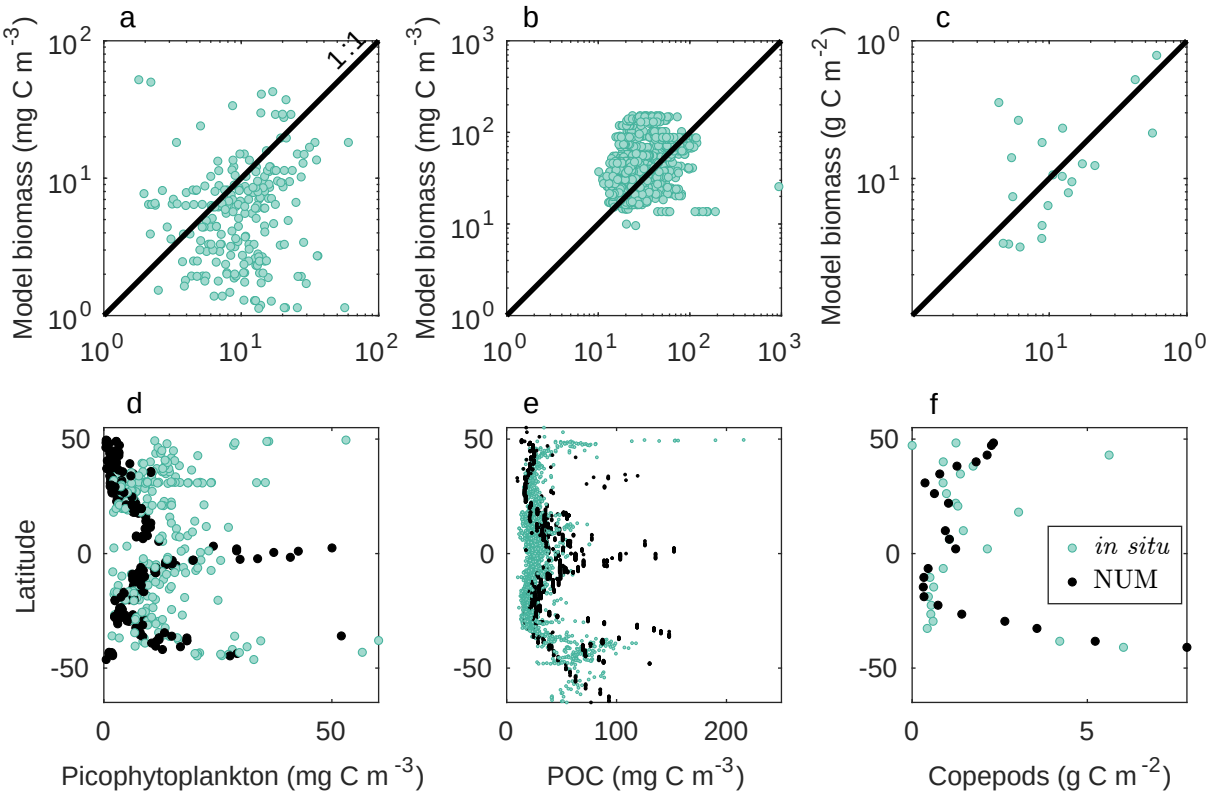

Figure 6. Comparison of model biomass output to *in situ* calibration data for biomasses of picophytoplankton (a, d), particulate organic carbon (b, e), and copepods (c, f). The top row shows modeled biomasses on the y-axis and observations on the x-axis. The bottom panels show the latitudinal variation of biomasses. Mean biases are -0.70 (a), 0.38, (b), and 0.016 (c), calculated from eq. 21.

4.1 Model calibration

The calibration gave an optimal light extinction coefficient of 0.06 m^{-1}, a sinking velocity of 19 m day^{-1}, and a coefficient of higher trophic level mortality of 0.017 L(μg C $day)^{-1}$. The response of the model to changes in each calibration parameter is shown in Fig. F5. The calibrated model produces biomasses of picoplankton, POC, and copepods which are in the correct order of magnitude of the observations (Fig. 6a-c). Though the three parameters are calibrated to be constant in time and space, the model does produce latitudinal patterns in accordance with the observed patterns (Fig. 6d-f: elevated biomass at temperate latitudes (-50° and 50°) and around Equatorial upwelling. The modeled copepod biomass approximates observations at temperate latitudes, however, it underestimates copepod biomass in the Equator.

4.2 Modeled biomass and productivity

The model captures general global-scale biomass patterns, with higher concentrations of unicellular and multicellular plankton in high-latitude seasonal environments and equatorial upwelling areas and the lowest biomasses in oligotrophic gyres (Fig. 7).

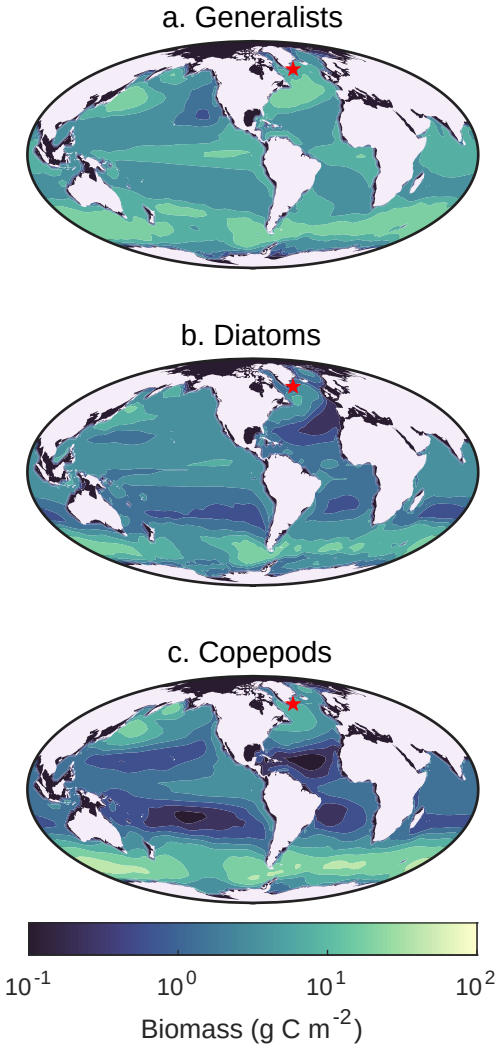

Figure 7. Global annual average biomass of (a) generalists, (b) diatoms and (c) copepods. The red star indicates the location (60°N,40°W) that is used for evaluate the seasonal dynamics in later figures.

These patterns are mirrored in satellite-derived NPP estimates (Fig. 8). Satellite production of NPP is subject to a systematic bias (Dierssen, 2010), and further the panels b-e highlight the big variation between satellite observation products. Nevertheless, one unrealistic pattern in the model that clearly stands out is the very high biomass and production in the Southern Ocean.

4.3 Model evaluation

Turning to the evaluation data from the Atlantic transects, we see that the model captures the right order of magnitude of nano- and micro-plankton biomasses summed together (Fig. 9). In these data, the equatorial upwelling is less pronounced which is also reproduced by the model. However, the model overestimates plankton biomass at higher latitudes compared to the AMT

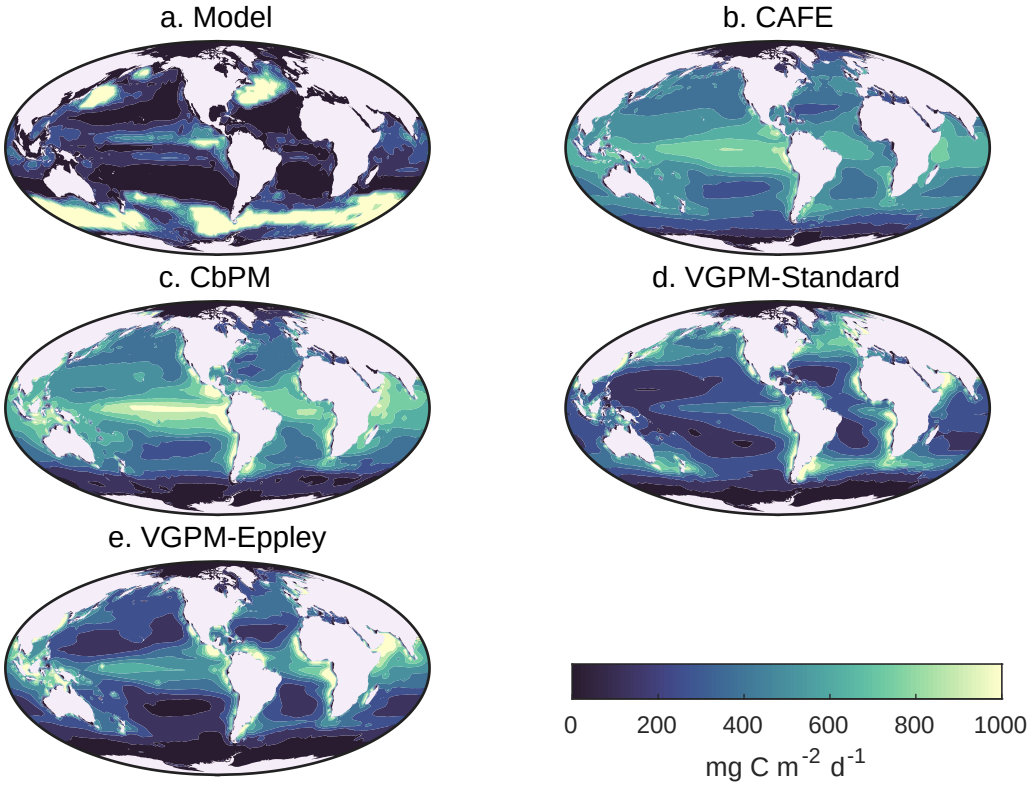

Figure 8. (a) NUM annual mean NPP,(b) CAFE estimated annual mean NPP (Silsbe et al., 2016), (c) CbPM estimated annual mean NPP (Westberry et al., 2008), (d) VGPM-Standard estimated annual mean NPP (Behrenfeld and Falkowski, 1997), and (e) VGPM-Eppley estimated annual mean NPP (Behrenfeld and Falkowski, 1997), all datasets from (Ocean Productivity). Units are mg C m^{-2} d^{-1}.

14 cruise (Fig. 9d), where samples were collected during May. At the global scale, the simulated average macrozooplankton concentration (copepods, larvae, and krill with a prosome length > 2 mm) is of the correct order of magnitude compared to observations, although the correlation is weak (Fig. 10).

The model resolves nitrogen in a simplified manner without accounting for its different forms (i.e. nitrate, nitrite, ammonia). We have chosen nitrate for the comparisons because it is considered to be the main nitrogenous compound utilized for primary production (Lenton and Watson, 2000). The model captures the spatial patterns in annual mean nitrogen and silicate concentrations at the surface (Fig. 11) with high values in the North Atlantic and Southern Ocean and lower at the oligotrophic gyres. In addition, it reproduces elevated concentration in equatorial upwelling areas. However, the modeled values are generally much lower than the observed values. This is likely due to very efficient uptakes of nutrients by unicellular plankton in the model, which leads to low limiting nutrient concentrations (low "R^*" *sensu* Tilman (1980)). This effect was also observed in the analysis of the generalist level (Andersen and Visser, 2023).

The model captures the high surface diatom:phytoplankton ratio in high latitudes but overestimates diatoms in mid latitudes (Fig. 12). This concurs with observations in the Pacific (Endo et al., 2018) and the Atlantic Oceans (Marañón et al., 2000). The

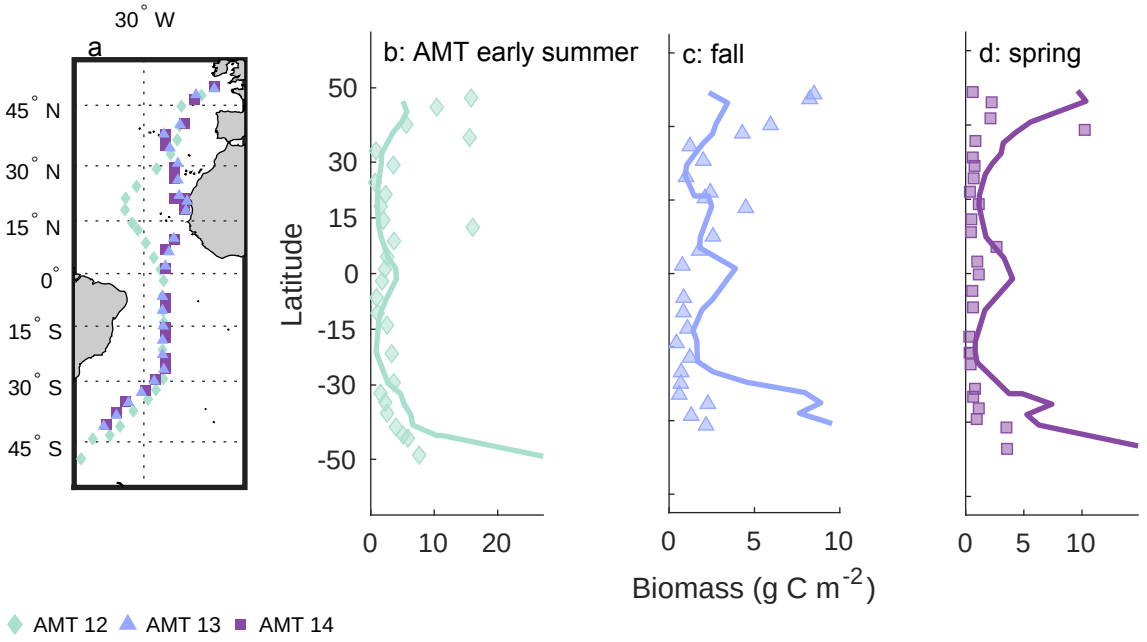

Figure 9. Biomass of nano- (2-20 μm radius) and micro-plankton (20-200 μm radius) of the model (line) compared to AMT data (markers). a: Transects of the different AMT cruises, b: AMT 12: May-June, c: AMT 13: September-October, d: AMT 14: May. The modeled biomass is the depth-integrated average of the respective months for each cruise, extracted from the final year of a 10-year global simulation.

overestimation in mid latitudes could be an artifact of the satellite measurements only seeing the very surface waters, whereas the top layer in the simulations performed here is the upper 50 meter. Further, data from the Continuous Plankton Recorder (CPR) survey (Reid et al., 2003) shows that diatoms constitute a fraction up to about 50% of the total phytoplankton biomass throughout the North Atlantic (Barton et al., 2013), in correspondence with the model simulations.

4.4 Model simulations

The unicellular NUM model's output differs from traditional biogeochemical plankton models: instead of having state variables of bacteria, phytoplankton, and zooplankton, the NUM model resolves generalists and diatoms. While the diatoms are obligate phytoplankton (including DOC uptake), the trophic strategies of the generalists are dynamic and an emerging property of the model. However, common terminology and most models explicitly resolve state variables of bacteria, phytoplankton, and zooplankton. To relate the model's output to the aforementioned groups, we can estimate bacteria, phytoplankton, and zooplankton biomass from the generalists and diatoms, by weighing each size class with the fraction of carbon assimilation

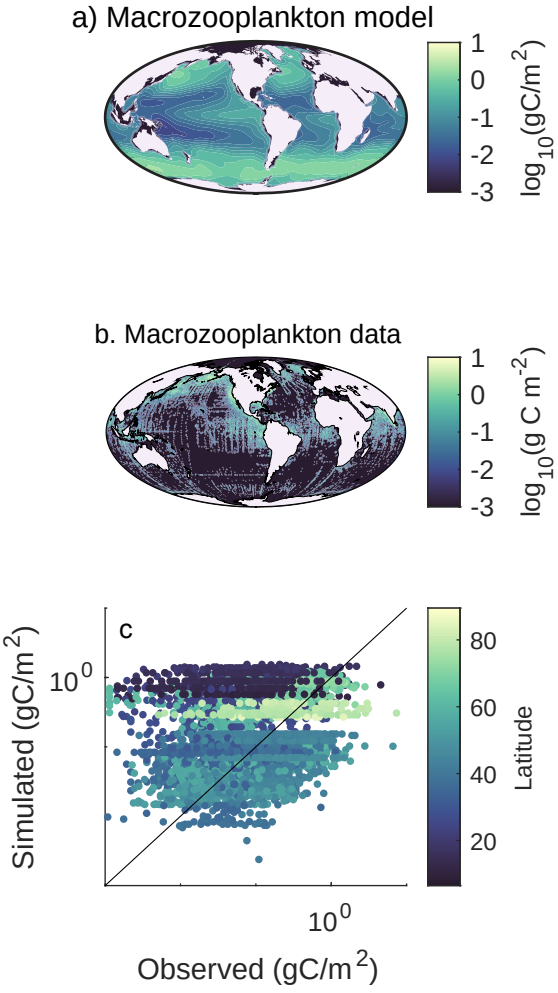

Figure 10. Annual mean macrozooplankton biomass, a) model output (integrated over the top 170 m) and b) observation data from (Moriarty et al., 2013) (integrated over the top 170 m), compared at different latitudes in panel c).

from DOC (bacteria-like), photosynthesis (phytoplankton-like), and phagotrophy (zooplankton-like):

$$B_{\mathrm{bact}} = \sum_i \frac{j_{\mathrm{DOC}}}{j_{\mathrm{C}}} B_i \tag{23}$$

$$B_{\mathrm{phyto}} = \sum_i \frac{j_{\mathrm{L}}}{j_{\mathrm{C}}} B_i \tag{24}$$

$$B_{\mathrm{zoo}} = \sum_i \frac{j_{\mathrm{F}}}{j_{\mathrm{C}}} B_i, \tag{25}$$

where $j_{\mathrm{C}} = j_{\mathrm{DOC}} + j_{\mathrm{L}} + j_{\mathrm{F}}$ (Fig. 13a-c).

The copepod community is dominated by active copepods in higher latitudes (Fig. 14) and by passive (ambushing) copepods in oligotrophic gyres. In equatorial regions the presence of active copepods varies in different oceans. In the Equatorial Atlantic

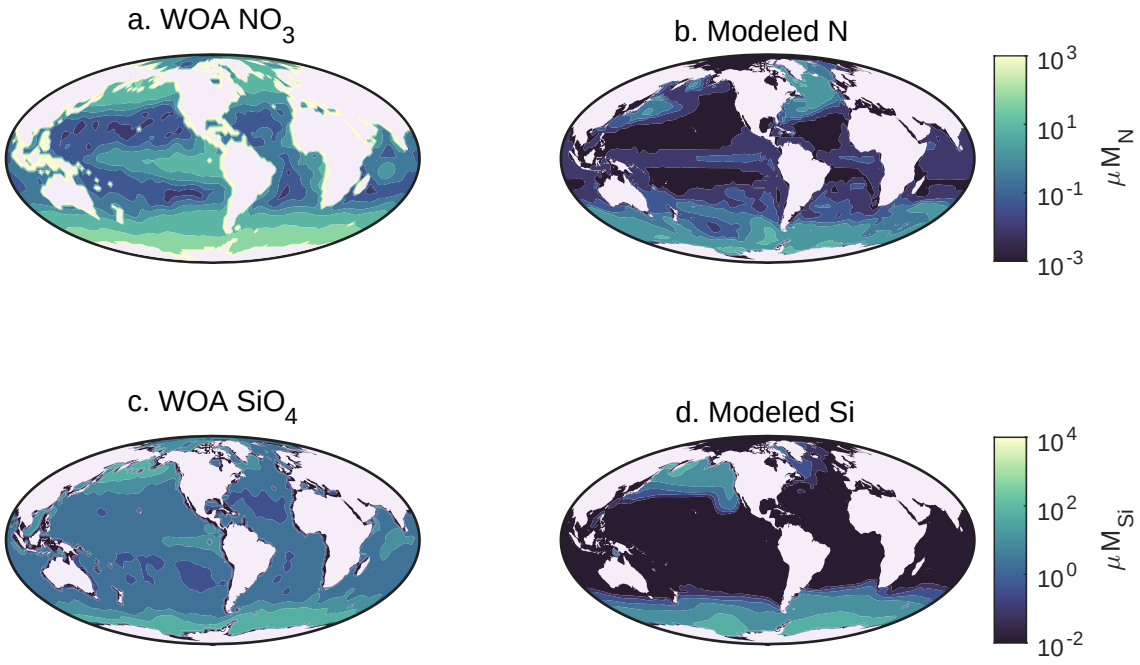

Figure 11. Annual mean concentrations of (a-b) nitrate and (c-d) silicate in the top 5 m. World Ocean Atlas (WOA) (Reagan et al., 2023) observations (left) compared to model output (right).

active copepods occupy between 50 and 60% of the multicellular plankton community. Lastly, active copepods occupy 60 to 80% of copepods along boundary currents in the Pacific, namely the California Current east and Kuroshio Current west.

4.5 Comparison of environment setups

We compared runs of the water-column and seasonal chemostat model setups to the output from a global simulation at a high latitude location (Fig. 15) that is seasonally stratified. We chose a seasonally stratified location because it exhibits the succession of oligotrophic and eutrophic conditions throughout the year. The seasonal dynamics of total biomass show similar dynamics in the global and water-column simulations (panels a and c): a bloom in spring throughout the water column, followed by a deeper production maximum in summer, and terminated by a smaller autumn bloom. The spring bloom of diatoms is terminated first by silicate limitation, and the generalist bloom is terminated a bit later by nitrogen limitation (Fig. F4 b,a). Contributing to the demise of the bloom is the grazing by copepods, which become established over summer (panels f-h). The deep maximum is less well represented in the global simulation, with only 13 vertical levels, than the water column simulation with 21 levels.

The size spectra show an overall similar community structure in the global ocean, water column and chemostat simulations (Fig. 15b,d,e). The plankton biomass is generally higher in the chemostat simulations than in the global and water-column simulations. The physical mixing in the global and water-column models constitutes a loss term by exporting planktonic organisms to the deep water, where they die unless they are mixed back into the photic zone. In the seasonal chemostat

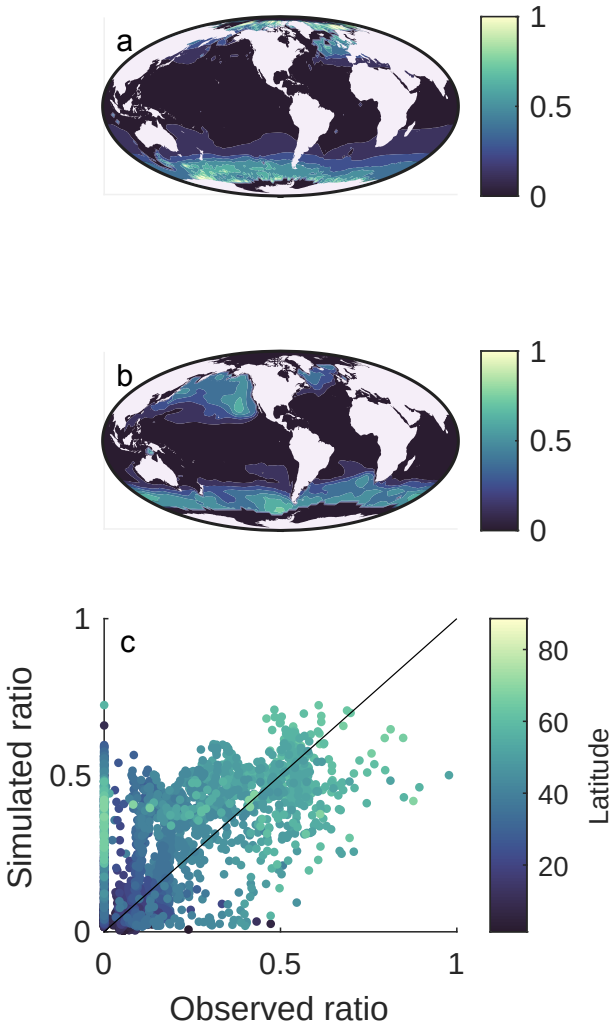

Figure 12. Observed (a) and simulated (b) global surface distribution of the diatom:phytoplankton biomass ratio (eq. 24). Panel c compares the observations and simulations with the color indicating latitude. The observation is a climatological mean from 1998 to 2024 of surface chlorophyll (see text for a description of the observation data set.).

this mixing does not occur. That healthy growing diatoms tend to have near-neutral buoyancy is well established (Gross and Zeuthen, 1948; Smayda, 1970). How they do it is also well established (Boyd and Gradmann, 2002). It is only when cells become stressed (e,g, low nutrients) and die that they start to succumb to gravity (Waite et al., 1992). The modelling assumption that diatoms while growing are neutrally buoyant is relatively robust. The difference in the modelling of mixing losses between chemostat simulations and the two other environments should be kept in mind when using the environments. Nevertheless, the output of the water column illustrates the same dynamics as the global at the same site, indicating that the water column model – which is naturally much faster than a global simulation – performs well. It should be noted, though, that the water column

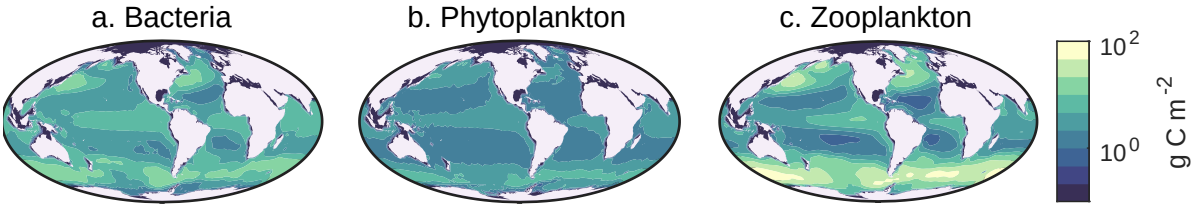

Figure 13. Annual depth-integrated mean biomass in the model: (a) bacteria biomass, (b) phytoplankton, (c) zooplankton, including unicellular plankton (mixotrophs and phagotrophs) and copepods, all calculated from Eqs. 23-25.

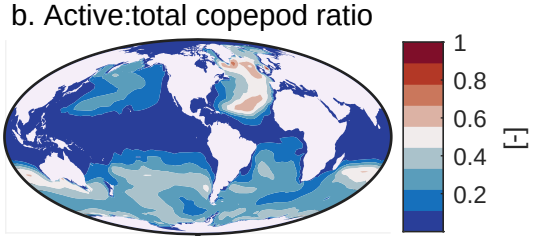

Figure 14. Ratio of active-feeding copepods to total copepod biomass.

simulations perform best in seasonal environments. In less seasonal environments, the absence of lateral advection appears to result in too low production (Figs. F1,F2).

4.6 Individual-level processes

The NUM library makes it possible to extract descriptions of the rates that drive the dynamics of each size class. Fig. 16 shows the rates of the generalists, diatoms, and copepods from the chemostat simulation in Fig. 15d, with uptakes of resources (or prey) illustrated as positive rates and losses including mortalities as negative rates. Overall, the rates of uptakes and losses are much higher in the spring bloom (left column) than in the summer (right column). The spring panels portray strong predation pressure on the unicellular community (dashed red lines in panels a and b), indicating that the bloom is decaying at this time. The high predation is also reflected in high feeding rates of the copepod community (red lines in panel c and d). Therefore, the spring copepod community is acquiring biomass at the expense of the unicellular community. In contrast, during summer, the division and loss rates in the unicellular community are roughly equal (grey lines in and panel e and f). This indicates a community in balance with little changes in biomass. Further, in the unicellular community, the figure shows how the trophic strategies change with cell size (Andersen et al., 2016): for the smallest cells, the acquisition of carbon is governed by DOC uptake (magenta lines), and dissolved nutrients (blue lines), in particular in the spring bloom (panels a and b). Larger sizes of generalists combine phototrophy (yellow lines) for carbon uptake for respiration with predation (red lines) for nitrogen and carbon for biosynthesis (and possibly respiration). Larger sizes of diatoms are purely phototrophic (negligible uptake of DOC) and are limited by nitrogen and silicate uptakes during summer.

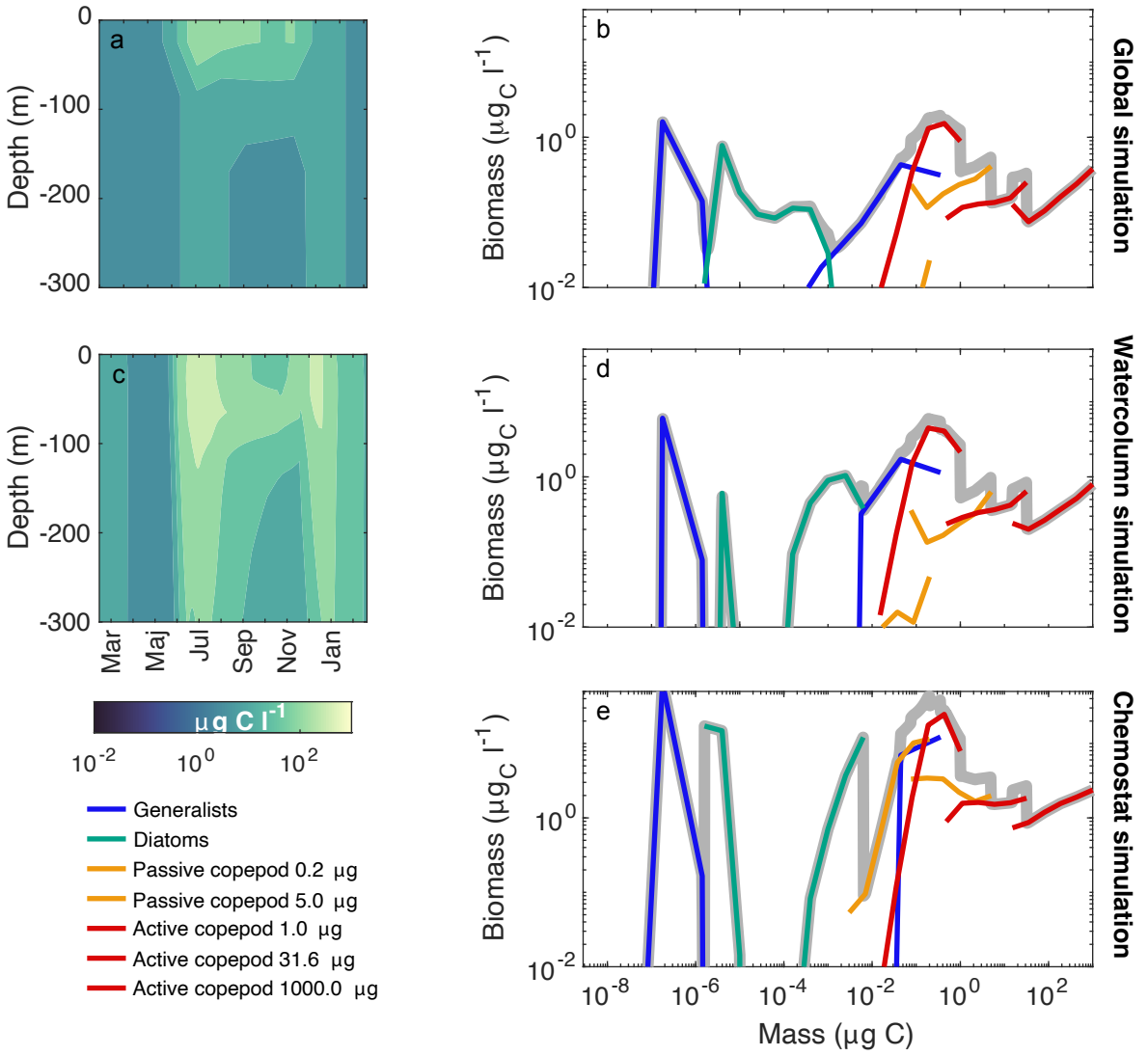

Figure 15. Model output for a 10-year simulation in the global, water column, and chemostat environments at a seasonally stratified location (60°N, 40°W) (red star in Fig. 7). (a) Total biomass in a water column extracted from the global simulation. (c) Total biomass in the same water column, but extracted from the water-column simulation. (b),(d),(e): Sheldon spectrum of the community from global, water-column, and chemostat simulations at 5 meters depth and averaged over the last year. See Fig. 5 for an indicative length axis.

The regulation of uptakes by unicellular plankton (section 3.2.2) entails that the respiration budget is resolved explicitly among the carbon used for synthesizing new biomass and different respiratory costs of basal metabolism, uptakes of dissolved nutrients and carbon, specific dynamic action of feeding (digestion), and costs of synthesizing new biomass (Fig. 17). In general, the synthesis of new biomass (the division rate) is one third of the total metabolic budget (Fenchel, 2013), while the respiratory costs are distributed among the different uptake costs, and a small portion is allocated to basal metabolism.

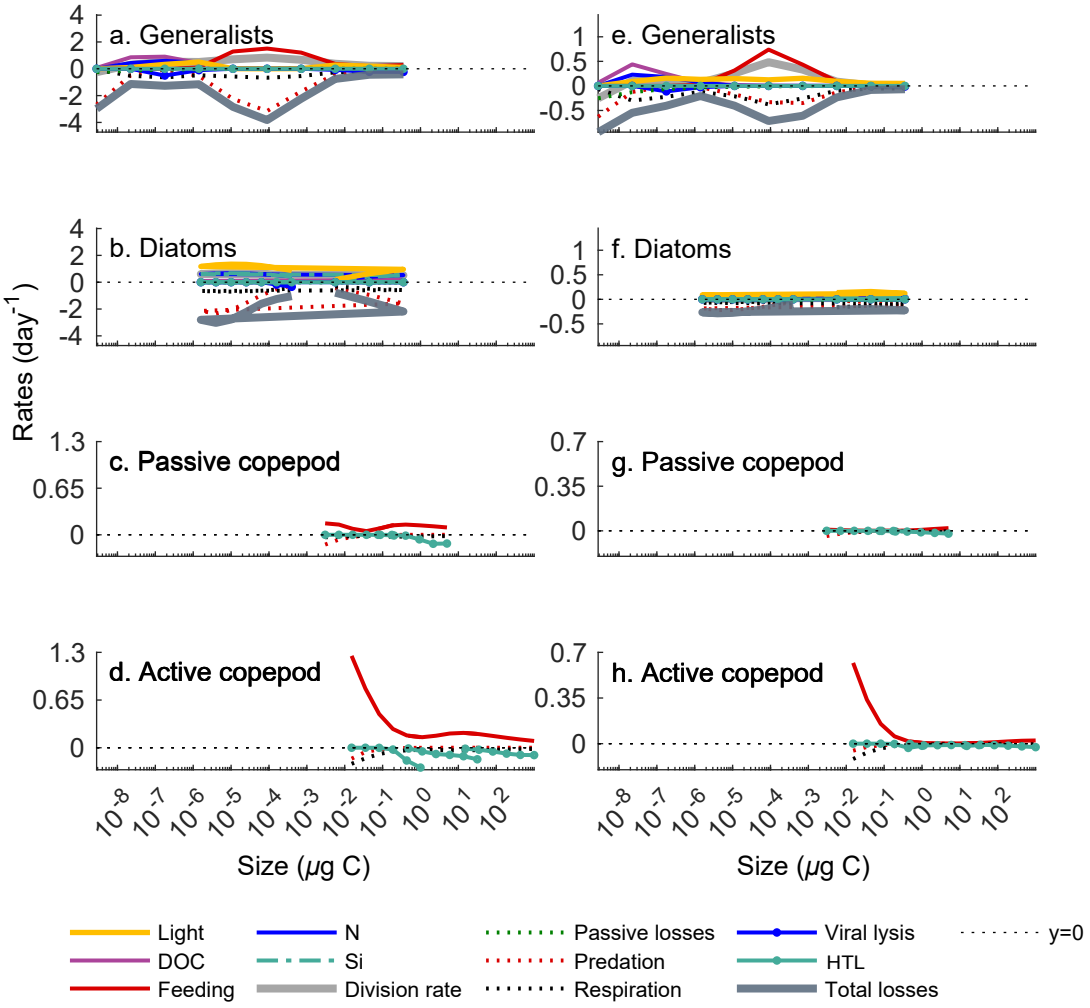

Figure 16. Gains and losses as a function of cell/body size in the spring bloom (April, left) and the summer (July, right). The lines above the horizontal dashed line shows gains either of carbon (from DOC uptake or photoharvesting; magenta and yellow), nutrients, silicate (blue and green), or a combination of carbon and nutrients from predation (red). The lines below the dashed lines shows losses of carbon from respiration (dashed black) or mortalities from predation (dashed red) or viral lysis (blue). The division rate is shown with thick grey. The output is from the chemostat simulation in Fig. 15d. Note the differences in the y-axes ranges between the left and right columns.

5 Discussion

We have presented a computational library for ecological modelling of plankton communities that resolves their utilization of resources, their trophic strategies, their size-structured population dynamics, and their production of dissolved and particulate organic carbon. The library includes a high-level interface to simulate the communities on global scale, in a water column,

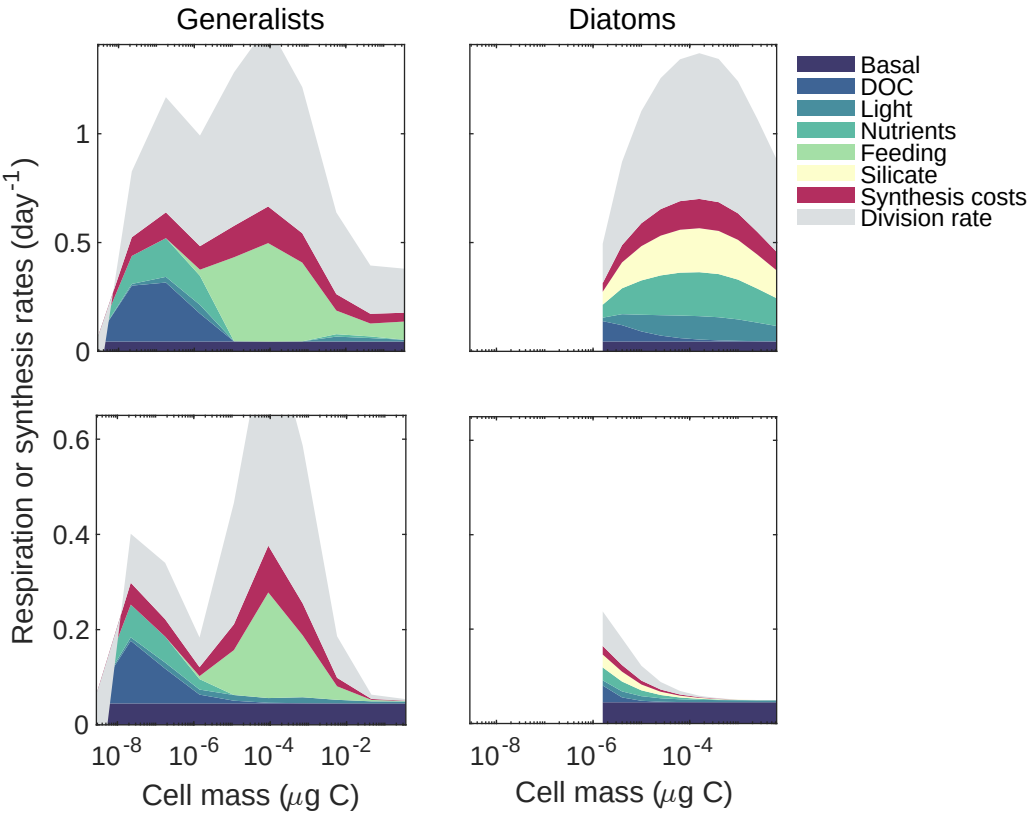

Figure 17. Metabolic budgets for generalists (left) and diatoms (right) in April (top panels) and July (bottom panels) extracted from Fig. 16. The division rate represents the uptakes that are used for synthesis of new biomass.

or in a chemostat. Because the model includes copepods, which are often weakly represented in biogeochemical models, the model lends itself to linking with fish production models, or to modelling the carbon pump arising from copepod fecal pellets and plankton carcasses. Overall the NUM model is an intermediate-complexity model laboratory for ecological simulations that is accessible to marine ecologists without biogeochemical or physical oceanography background. The focus here has been on the presentation of the general and flexible programming library and the calibration of the core NUM model setup.

The library has been designed with flexibility in mind. It can be run fully from a high-level programming language (mainly matlab). As the core library is written in Fortran, it is straight-forward to implement the model in a full circulation model (Hansen et al., 2024), for example using the FABM interface (Bruggeman and Bolding, 2014). Further, new size-spectrum groups can be added to the library. While POM described here is a single-state variable, the model does include a module to handle POM as an entire size spectrum. A more complex POM module is currently under development (Visser et al., 2024). Future extensions would be the inclusion of an iron nutrient tracer to limit the production in the Southern Ocean, a diazotroph spectrum to include nitrogen fixation, and a description of calcifiers.

Another flexible aspect of the NUM library is the possibility of performing simulations in three environmental settings: global, water column, and chemostat. Water-column or chemostat simulations are useful for making fast model development or trial simulations before moving to full global simulations. However, it should be noted that even though the three settings are based on the same transport matrix there are differences in the output between them. First, the seasonal variation of total biomass both in the global and water column simulations at a high-latitude location concur: a spring bloom throughout the water column, followed by a deeper plankton biomass maximum in summer (Yasunaka et al., 2021), and a moderate short autumn bloom. However, the global simulations reproduction of the deep maximum is coarser than in the water column, most probably due to a lower vertical resolution. While the structure of the water-column simulations corresponds relatively well with the global simulations in a strongly seasonal environment, they do not match well outside of high-latitude environments, probably due to the absence of advective processes in the water-column simulation. Further, we observe disparity in the size spectra of the community extracted from global, water column, and chemostat models. In particular, the dominance of large diatoms in chemostat simulations, emerges from the absence of mixing losses of plankton in the chemostat simulations, while mixing losses are present in water-column simulations. This finding points to the importance of a better understanding of plankton's ability to maintain a position in the photic zone despite vertical mixing.

Diatoms have been included here as a new size spectrum group. The module is based on an extension of the generalist module with the addition of a vacuole and a silicate shell. In this manner, the diatom module does not use any diatom-specific parameters, apart from the cost of silicate uptake, the size of the vacuole, the minimum and maximum size, and the vulnerability to predation, but shares all other parameters with the generalists. Further, silicate is not incorporated into the POM, it is therefore lost from the model. In light of this very simple description of diatoms, and the important omission of non-sinking silicate POM, the module performs quite well, as it captures the emergence of large diatoms during spring blooms in seasonal environments and their dominance in high latitudes (figs. 7b,14a). Nevertheless, while larger phototrophic diatoms appear in high latitude seasonal regions, they are missing in upwelling regions in the global and water-column simulations (though they do appear in chemostat simulations). The absence of phototrophic diatoms in upwelling regions simulated in the model is implausible, given their known prevalence in such nutrient-rich waters (Benoiston et al., 2017). We conjecture that the absence of large diatoms is due to mixing losses; in nature diatoms are able maintain their position in the water-column by active buoyancy regulation. However, in the global and water-column simulations, diatoms are mixed out. In contrast, the chemostat simulation ignores mixing losses of unicellular plankton. A future fix could be to include buoyancy control in the diatom module to avoid mixing diatoms out and thereby limiting their mortality. Finally, silicate should be included into the POM group. This would require another state variable to account for the variable stoichiometry of POM.

All processes in each size-spectrum group are parameterized from considerations from first principles or from cross-species analysis of laboratory observations. The only calibrated parameters are those that form the closure of the model with respect to light absorption (light absorption coefficient), nutrient turn-over rate (sinking speed of POM), and mortality on higher trophic levels. Each of these parameters are expected to vary significantly across the global ocean, but here they are just calibrated to average values. Despite this simplification, the model reproduces important global-scale patterns of plankton biomass, nutrient fields, and net primary production with some accuracy. There are however notable deviations, and these are important to keep

in mind when the results are interpreted: 1) The model produces unrealistically high production in the Southern Ocean. This may be due to the lack of iron as a nutrient, which is known to be limiting production in the Southern Ocean (Bazzani et al., 2023). This results in overestimation of the biomass of phytoplankton and concomitantly the predators thereof, 2) as mentioned above, the model does not represent well large diatoms outside seasonal environments.

The model simulates a lower net primary production in the oligotrophic gyres than observed by satellites. There are two reasons for this underestimation. First, the measure of NPP calculated by the model (22) is not entirely consistent with the empirical definition. The empirical measure is based on bottle incubations where the biomass growth of the entire plankton community is measured. The measure used here is closer to net community production, because the heterotrophic respiration is also subtracted. Further, the measure ignores the DOC that is created in the assimilation process (the $\epsilon_{\mathrm{L}}/(1-\epsilon_{\mathrm{L}})j_{\mathrm{L}}$ term in Fig. 4b). This DOC is readily taken up and used to fuel production. The NPP measure used here therefore underestimates NPP to some degree. However, it is not trivial to design a formulation of NPP that directly matches the experimental procedure of measuring NPP. Second, primary production in the oligotrophic gyres is limited by the diffusion of deep nutrient up into the photic zone. The rate of diffusion is determined by the depth at which sinking POM is remineralized – deeper remineralization depths lead to lower rates of diffusion and lower NPP. Here, the remineralization depth is determined by the sinking velocity and the POM remineralization rate. Oligotrophic gyres are characterized by a dominance of small cells and few copepods, leading to small POM particles which would be remineralized at a shallower depth than in highly productive regions. Using only a single globally averaged sinking velocity fails to resolve this variation and overemphasizes the difference in NPP between low and high productive regions. This can be solved by introducing several POM size groups with different sinking velocities. This capability is already present in the library, but has not been exploited in the calibration presented here.

The evaluation of copepod global trait-distribution is challenging, with only a handful of global studies on copepod trait-distribution conducted (Brun et al., 2016; Prowe et al., 2019; Benedetti et al., 2023). We compare our results (Fig. 14b) with the functional traits identified in the global ocean in Benedetti et al. (2023), adopting the classification proposed by Brun et al. (2016), which categorizes cruise-current-feeding, current-feeding and cruise-feeding copepod species as active feeding copepods and ambush-feeding species as passive feeding copepods. Our results show that the highest ratio of active copepods is found in the Southern Ocean, North Pacific, and Arctic Ocean. This aligns with observations reporting the highest proportions of current-cruise-feeding, current-feeding, and cruise-feeding species in these regions (Benedetti et al., 2023). Our model also reproduces a lower fraction of active-feeding copepods in oligotrophic areas, such as the Pacific Equatorial band and the Indian Ocean, which belong to the same regional categorization (Benedetti et al., 2023) and exhibit the highest proportions of ambush-feeding species. The pattern of active copepods in high latitudes and passive copepods in low latitudes is different that observed and modelled by Prowe et al. (2019). We believe that the pattern in passive/active copepods were due to a very limited set of observations, in particular at low latitudes. The biomass of large copepods is lower in oligotrophic regions and increases at higher latitudes, particularly in the Southern Ocean, North Atlantic and Northwest Pacific. In contrast, the Northeast Pacific is nitrogen-depleted in the model (Fig. 11b), with lower unicellular plankton biomass there (Fig. 7a,b), which limits the capacity of this region to sustain active copepods.

In the model design we have striven for conceptual simplicity while being mindful that simplifying assumptions sometimes lead to limitations. Apart from the limitations already mentioned, a prominent one is the unresolved motility behaviour that is known to have significant consequences on trophic and population dynamics in the plankton (Mariani et al., 2013; Kenitz et al., 2017; Pinti et al., 2022). While some aspects have been incorporated in the feeding mode parameterization of multicellular components, this has not been extended to the unicellular components. Neither has vertical migration - either diurnal or seasonal been included – a behavior that among other things, appears to play a significant role in the biological carbon pump (Hansen and Visser, 2016; Pinti et al., 2023a, b). In some cases we have purposefully avoided modelling aspects of plankton ecology that are phenomenologically well known but lack a mechanistic understanding. The underlying philosophy we pursue is to base all model descriptions on first principles as much as possible (Andersen and Visser, 2023). In some cases the underlying trade-offs remain elusive. What, for instance is the cost-benefit of calcifying plankton such as coccolithophres (Monteiro et al., 2016)? Rather than introducing extra parameters to fix these issues, we want to see how far a generalized model can come in explaining observed global patterns of ocean plankton ecology. If the broad contours of these patterns can be explained, then future marine ecosystems under climate change can be predicted with the same level of certainty.

Code and data availability. The code repository for this version of the framework is available via Zenodo at https://doi.org/10.5281/zenodo.14280889 (Papapostolou (2025)). This version corresponds to the official release version 1.0 and includes the code used to generate all figures in the manuscript https://github.com/amaliapap/NUM_v1_ComputationalLibrary (https://doi.org/10.5281/zenodo.15856680). The code for the general framework and an actively developed version of the code is also hosted on GitHub at https://github.com/Kenhasteandersen/NUMmodel, where full documentation is available through the wiki and help pages for all functions.

Author contributions. AP contributed the diatom model, collected and analyzed data for validation, conducted simulations and made most figures. KHA developed the general model framework and some figures. The manuscript was jointly written by AP and KHA with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This publication was funded by the European Union (GA 101059915 BIOcean5D and GA 869383 ECOTIP), by the VKR Center of Excellence Ocean Life, by the Simon's Foundation grant 931976, and by the NFR project 334996 "Pelagic". Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them. This study has been conducted using E.U. Copernicus Marine Service Information: doi 10.48670/moi-00278. Thanks to Stephanie Dutkiewicz for hosting

AP during a research stay at the Department of Earth, Atmospheric, and Planetary Sciences at the Massachusetts Institute of Technology, and for discussions on evaluation of global plankton models.

References

Andersen, K., Berge, T., Gonçalves, R., Hartvig, M., Heuschele, J., Hylander, S., Jacobsen, N., Lindemann, C., Martens, E., Neuheimer, A., Olsson, K., Palacz, A., Prowe, A., Sainmont, J., Traving, S., Visser, A., Wadhwa, N., and Kiørboe, T.: Characteristic Sizes of Life in the Oceans, from Bacteria to Whales, Annual Review of Marine Science, 8, 217–241, https://doi.org/10.1146/annurev-marine-122414-034144, 2016.

Andersen, K. H.: Fish ecology, evolution, and exploitation: A new theoretical synthesis, Princeton University Press, 2019.

Andersen, K. H. and Beyer, J. E.: Asymptotic Size Determines Species Abundance in the Marine Size Spectrum, The American Naturalist, 168, 54–61, https://doi.org/10.1086/521238, 2006.

Andersen, K. H. and Visser, A.: From cell size and first principles to structure and function of unicellular plankton communities, Progress in Oceanography, 213, 102 995, 2023.

Banas, N. S.: Adding complex trophic interactions to a size-spectral plankton model: Emergent diversity patterns and limits on predictability, Ecological Modelling, 222, 2663–2675, 2011.

Barton, A. D., Pershing, A. J., Litchman, E., Record, N. R., Edwards, K. F., Finkel, Z. V., Kiørboe, T., and Ward, B. A.: The biogeography of marine plankton traits, Ecology Letters, 16, 522–534, https://doi.org/10.1111/ele.12063, 2013.

Basu, S. and Mackey, K. R.: Phytoplankton as key mediators of the biological carbon pump: Their responses to a changing climate, Sustainability, 10, 869, 2018.

Bazzani, E., Lauritano, C., and Saggiomo, M.: Southern Ocean iron limitation of primary production between past knowledge and future projections, Journal of Marine Science and Engineering, 11, 272, 2023.

Behrenfeld, M. J. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, Limnology and oceanography, 42, 1–20, 1997.

Benedetti, F., Wydler, J., and Vogt, M.: Copepod functional traits and groups show divergent biogeographies in the global ocean, Journal of Biogeography, 50, 8–22, 2023.

Benoiston, A.-S., Ibarbalz, F. M., Bittner, L., Guidi, L., Jahn, O., Dutkiewicz, S., and Bowler, C.: The evolution of diatoms and their biogeochemical functions, Philosophical Transactions of the Royal Society B: Biological Sciences, 372, 20160 397, 2017.

Berge, T., Chakraborty, S., Hansen, P. J., and Andersen, K. H.: Modeling succession of key resource-harvesting traits of mixotrophic plankton, The ISME Journal, 11, 212–223, https://doi.org/10.1038/ismej.2016.92, number: 1 Publisher: Nature Publishing Group, 2017.

Boyd, C. and Gradmann, D.: Impact of osmolytes on buoyancy of marine phytoplankton, Marine Biology, 141, 605–618, 2002.

Boyd, P. W., Claustre, H., Levy, M., Siegel, D. A., and Weber, T.: Multi-faceted particle pumps drive carbon sequestration in the ocean, Nature, 568, 327–335, https://doi.org/10.1038/s41586-019-1098-2, 2019.

Bruggeman, J. and Bolding, K.: A general framework for aquatic biogeochemical models, Environmental Modelling & Software, 61, 249–265, https://doi.org/10.1016/j.envsoft.2014.04.002, 2014.

Brun, P., Payne, M. R., and Kiørboe, T.: Trait biogeography of marine copepods–an analysis across scales, Ecology Letters, 19, 1403–1413, 2016.

Brzezinski, M. A.: The Si: C: N ratio of marine diatoms: interspecific variability and the effect of some environmental variables 1, Journal of Phycology, 21, 347–357, 1985.

Butenschön, M., Clark, J., Aldridge, J. N., Allen, J. I., Artioli, Y., Blackford, J., Bruggeman, J., Cazenave, P., Ciavatta, S., Kay, S., et al.: ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels, Geoscientific Model Development, 9, 1293–1339, 2016.

Cadier, M., Hansen, A. N., Andersen, K. H., and Visser, A. W.: Competition between vacuolated and mixotrophic unicellular plankton, Journal of Plankton Research, 42, 425–439, https://doi.org/10.1093/plankt/fbaa025, 2020.

Chakraborty, S., Nielsen, L. T., and Andersen, K. H.: Trophic Strategies of Unicellular Plankton, The American Naturalist, 189, E77–E90, https://doi.org/10.1086/690764, 2017.

Chakraborty, S., Cadier, M., Visser, A. W., Bruggeman, J., and Andersen, K. H.: Latitudinal Variation in Plankton Traits and Ecosystem Function, Global Biogeochemical Cycles, 34, https://doi.org/10.1029/2020GB006564, 2020.

CMEMS: Global Ocean Colour (Copernicus-GlobColour), Bio-Geo-Chemical, L4 (monthly and interpolated) from Satellite Observations (1997-ongoing), Tech. rep., Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS), https://doi.org/10.48670/moi-00281, 2025.

De Roos, A. M., Schellekens, T., Van Kooten, T., Van De Wolfshaar, K., Claessen, D., and Persson, L.: Simplifying a physiologically structured population model to a stage-structured biomass model, Theoretical population biology, 73, 47–62, 2008.

Dierssen, H. M.: Perspectives on empirical approaches for ocean color remote sensing of chlorophyll in a changing climate, Proceedings of the National Academy of Sciences, 107, 17 073–17 078, 2010.

Durbin, A. G. and Durbin, E. G.: Seasonal changes in size frequency distribution and estimated age in the marine copepod Acartia hudsonica during a winter-spring diatom bloom in Narragansett Bay, Limnology and Oceanography, 37, 379–392, 1992.

Dutkiewicz, S., Follows, M. J., and Parekh, P.: Interactions of the iron and phosphorus cycles: A three-dimensional model study, Global Biogeochemical Cycles, 19, 2005.

Endo, H., Ogata, H., and Suzuki, K.: Contrasting biogeography and diversity patterns between diatoms and haptophytes in the central Pacific Ocean, Scientific Reports, 8, 10 916, 2018.

Evans, G. T. and Parslow, J. S.: A Model of Annual Plankton Cycles, Biological Oceanography, 3, 327–347, https://doi.org/10.1080/01965581.1985.10749478, 1985.

Fenchel, T.: Ecology of Protozoa: The biology of free-living phagotropic protists, Springer-Verlag, 2013.

Forget, G., Campin, J.-M., Heimbach, P., Hill, C., Ponte, R., and Wunsch, C.: ECCO version 4: An integrated framework for non-linear inverse modeling and global ocean state estimation, Geoscientific Model Development, 8, 3071–3104, 2015.

Franks, P. J.: NPZ models of plankton dynamics: their construction, coupling to physics, and application, Journal of Oceanography, 58, 379–387, 2002.

Gross, F. and Zeuthen, E.: The buoyancy of plankton diatoms: a problem of cell physiology, Proceedings of the Royal Society of London. Series B-Biological Sciences, 135, 382–389, 1948.

Hamm, C. E., Merkel, R., Springer, O., Jurkojc, P., Maier, C., Prechtel, K., and Smetacek, V.: Architecture and material properties of diatom shells provide effective mechanical protection, Nature, 421, 841–843, 2003.

Hansen, A. N. and Visser, A. W.: Carbon export by vertically migrating zooplankton: an optimal behavior model, Limnology and Oceanography, 61, 701–710, 2016.

Hansen, A. N. and Visser, A. W.: The seasonal succession of optimal diatom traits, Limnology and Oceanography, 64, 1442–1457, https://doi.org/10.1002/lno.11126, _eprint: https://aslopubs.onlinelibrary.wiley.com/doi/pdf/10.1002/lno.11126, 2019.

Hansen, T. F., Canfield, D. E., Andersen, K. H., and Bjerrum, C. J.: The unicellular NUM v. 0.91: A trait-based plankton model evaluated in two contrasting biogeographic provinces, Geoscientific Model Development Discussions, 2024, 1–39, 2024.

Heath, M. R., Speirs, D. C., and Steele, J. H.: Understanding patterns and processes in models of trophic cascades, Ecology letters, 17, 101–114, 2014.

Hillebrand, H., Acevedo-Trejos, E., Moorthi, S. D., Ryabov, A., Striebel, M., Thomas, P. K., and Schneider, M.-L.: Cell size as driver and sentinel of phytoplankton community structure and functioning, Functional Ecology, 36, 276–293, 2022.

Howe, H. B., McIntyre, P. B., and Wolman, M. A.: Adult zebrafish primarily use vision to guide piscivorous foraging behavior, Behavioural Processes, 157, 230–237, 2018.

Humes, A. G.: How many copepods?, Hydrobiologia, 292, 1–7, 1994.

Karl, D. M., Hebel, D. V., Björkman, K., and Letelier, R. M.: The role of dissolved organic matter release in the productivity of the oligotrophic North Pacific Ocean, Limnology and Oceanography, 43, 1270–1286, 1998.

Kenitz, K. M., Visser, A. W., Mariani, P., and Andersen, K. H.: Seasonal succession in zooplankton feeding traits reveals trophic trait coupling, Limnology and Oceanography, 62, 1184–1197, 2017.

Khatiwala, S.: A computational framework for simulation of biogeochemical tracers in the ocean, Global Biogeochemical Cycles, 21, 2007.

Khatiwala, S., Visbeck, M., and Cane, M. A.: Accelerated simulation of passive tracers in ocean circulation models, Ocean Modelling, 9, 51–69, 2005.

Kiørboe, T., Møhlenberg, F., and Hamburger, K.: Bioenergetics of the planktonic copepod Acartia tonsa: relation between feeding, egg production and respiration, and composition of specific dynamic action, Mar Ecol Prog Ser, 26, 85–97, 1985.

Kiørboe, T., Visser, A., and Andersen, K. H.: A trait-based approach to ocean ecology, ICES Journal of Marine Science, 75, 1849–1863, https://doi.org/10.1093/icesjms/fsy090, publisher: Oxford Academic, 2018.

Kvale, K. F., Khatiwala, S., Dietze, H., Kriest, I., and Oschlies, A.: Evaluation of the transport matrix method for simulation of ocean biogeochemical tracers, Geoscientific Model Development Discussions, 2017, 1–25, 2017.

Lange Jensen, L., Bjørn, T., Hein Korsgaard, A., Pertoldi, C., and Madsen, N.: Influence of Turbidity on Foraging Behaviour in Three-Spined Sticklebacks (Gasterosteus aculeatus), Fishes, 8, 609, 2023.

Lemoine, J., Ayata, S.-d., Jaspers, C., and Lombard, F.: Biomass-to-volume ratio as a central continuous functional trait for marine zooplankton, Limnology and Oceanography, 70, 2673–2687, 2025.

Lenton, T. M. and Watson, A. J.: Redfield revisited: 1. Regulation of nitrate, phosphate, and oxygen in the ocean, Global biogeochemical cycles, 14, 225–248, 2000.

López, E. and Anadón, R.: Copepod communities along an Atlantic Meridional Transect: Abundance, size structure, and grazing rates, Deep Sea Research Part I: Oceanographic Research Papers, 55, 1375–1391, https://doi.org/10.1016/j.dsr.2008.05.012, 2008.

Marañón, E., Holligan, P. M., Varela, M., Mouriño, B., and Bale, A. J.: Basin-scale variability of phytoplankton biomass, production and growth in the Atlantic Ocean, Deep Sea Research Part I: Oceanographic Research Papers, 47, 825–857, 2000.

Mariani, P., Andersen, K. H., Visser, A. W., Barton, A. D., and Kiørboe, T.: Control of plankton seasonal succession by adaptive grazing, Limnology and Oceanography, 58, 173–184, 2013.

Martínez-Vicente, V., Evers-King, H., Roy, S., Kostadinov, T. S., Tarran, G. A., Graff, J. R., Brewin, R. J., Dall'Olmo, G., Jackson, T., Hickman, A. E., et al.: Intercomparison of ocean color algorithms for picophytoplankton carbon in the ocean, Frontiers in Marine Science, 4, 378, 2017.

Monteiro, F. M., Bach, L. T., Brownlee, C., Bown, P., Rickaby, R. E., Poulton, A. J., Tyrrell, T., Beaufort, L., Dutkiewicz, S., Gibbs, S., et al.: Why marine phytoplankton calcify, Science Advances, 2, e1501 822, 2016.

Moriarty, R., Buitenhuis, E., Le Quéré, C., and Gosselin, M.-P.: Distribution of known macrozooplankton abundance and biomass in the global ocean, Earth System Science Data, 5, 241–257, 2013.

Neumann, T.: Towards a 3D-ecosystem model of the Baltic Sea, Journal of Marine Systems, 25, 405–419, 2000.

Ocean Productivity: Ocean Productivity Home Page, http://sites.science.oregonstate.edu/ocean.productivity/index.php, accessed: 2024-04-25, 2024.

Pančić, M., Torres, R. R., Almeda, R., and Kiørboe, T.: Silicified cell walls as a defensive trait in diatoms, Proceedings of the Royal Society B, 286, 20190 184, 2019.

Pančić, M., Torres, R. R., Almeda, R., and Kiørboe, T.: Silicified cell walls as a defensive trait in diatoms, Proceedings of the Royal Society B: Biological Sciences, 286, 20190 184, https://doi.org/10.1098/rspb.2019.0184, 2019.

Papapostolou, A.: Computational library for the Nutrient-Unicellular-Multicellular plankton modeling framework (code), https://doi.org/10.5281/zenodo.14280889, 2025.

Pinti, J., Visser, A. W., Serra-Pompei, C., Andersen, K. H., Ohman, M. D., and Kiørboe, T.: Fear and loathing in the pelagic: How the seascape of fear impacts the biological carbon pump, Limnology and Oceanography, 67, 1238–1256, 2022.

Pinti, J., DeVries, T., Norin, T., Serra-Pompei, C., Proud, R., Siegel, D. A., Kiørboe, T., Petrik, C. M., Andersen, K. H., Brierley, A. S., et al.: Model estimates of metazoans' contributions to the biological carbon pump, Biogeosciences, 20, 997–1009, 2023a.

Pinti, J., Jónasdóttir, S. H., Record, N. R., and Visser, A. W.: The global contribution of seasonally migrating copepods to the biological carbon pump, Limnology and Oceanography, 68, 1147–1160, 2023b.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: Numerical Recipes 3rd Edition: The Art of Scientific Computing, Cambridge University Press, USA, 3 edn., ISBN 0521880688, 2007.

Prowe, A. F., Visser, A. W., Andersen, K. H., Chiba, S., and Kiørboe, T.: Biogeography of zooplankton feeding strategy, Limnology and Oceanography, 64, 661–678, 2019.

Quere, C. L., Harrison, S. P., Colin Prentice, I., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Cotrim Da Cunha, L., Geider, R., Giraud, X., et al.: Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models, Global Change Biology, 11, 2016–2040, 2005.

Reagan, J. R., Boyer, T., García, H., Locarnini, R., Baranova, O., Bouchard, C., Cross, S., Mishonov, A., Paver, C., Seidov, D., and Wang, Zhankunand Dukhovskoy, D.: World Ocean Atlas 2023, NOAA National Centers for Environmental Information Dataset, https://www.ncei.noaa.gov/archive/accession/0270533, [Accessed 26-07-2024], 2023.

Record, N., Pershing, A., and Maps, F.: Emergent copepod communities in an adaptive trait-structured model, Ecological modelling, 260, 11–24, 2013.

Reid, P. C., Colebrook, J. M., Matthews, J. B. L., and Aiken, J.: The Continuous Plankton Recorder: concepts and history, from plankton indicator to undulating recorders, Progress in Oceanography, 58, 117–173, 2003.

Rodríguez-Ramos, T., Marañón, E., and Cermeño, P.: Marine nano- and microphytoplankton diversity: redrawing global patterns from sampling-standardized data, Global Ecology and Biogeography, 24, 527–538, https://doi.org/10.1111/geb.12274, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/geb.12274, 2015.

Rombouts, I., Beaugrand, G., Ibañez, F., Gasparini, S., Chiba, S., and Legendre, L.: Global latitudinal variations in marine copepod diversity and environmental factors, Proceedings of the Royal Society B: Biological Sciences, 276, 3053–3062, https://doi.org/10.1098/rspb.2009.0742, publisher: Royal Society, 2009.

Ryther, J. H.: Photosynthesis and Fish Production in the Sea, Science, 166, 72–76, https://doi.org/10.1126/science.166.3901.72, 1969.

Serra-Pompei, C., Hagstrom, G. I., Visser, A. W., and Andersen, K. H.: Resource limitation determines temperature response of unicellular plankton communities, Limnology and Oceanography, 64, 1627–1640, https://doi.org/10.1002/lno.11140, _eprint: https://aslopubs.onlinelibrary.wiley.com/doi/pdf/10.1002/lno.11140, 2019.

Serra-Pompei, C., Soudijn, F., Visser, A. W., Kiørboe, T., and Andersen, K. H.: A general size- and trait-based model of plankton communities, Progress in Oceanography, 189, 102 473, https://doi.org/10.1016/j.pocean.2020.102473, 2020.

Serra-Pompei, C., Ward, B. A., Pinti, J., Visser, A. W., Kiørboe, T., and Andersen, K. H.: Linking plankton size spectra and community composition to carbon export and its efficiency, Global Biogeochemical Cycles, https://doi.org/10.1029/2021GB007275, 2022.

Sheldon, R. W., Sutcliffe Jr., W. H., and Paranjape, M. A.: Structure of Pelagic Food Chain and Relationship Between Plankton and Fish Production, Journal of the Fisheries Research Board of Canada, 34, 2344–2353, https://doi.org/10.1139/f77-314, publisher: NRC Research Press, 1977.

Silsbe, G. M., Behrenfeld, M. J., Halsey, K. H., Milligan, A. J., and Westberry, T. K.: The CAFE model: A net production model for global ocean phytoplankton, Global Biogeochemical Cycles, 30, 1756–1777, 2016.

Smayda, T. J.: The suspension and sinking of phytoplankton in the sea, Oceanogr. Mar. Biol. Ann. Rev., 8, 353–414, 1970.

Stock, C. A., Dunne, J. P., and John, J. G.: Global-scale carbon and energy flows through the marine planktonic food web: An analysis with a coupled physical–biological model, Progress in Oceanography, 120, 1–28, https://doi.org/10.1016/j.pocean.2013.07.001, 2014.

Tanioka, T., Larkin, A. A., Moreno, A. R., Brock, M. L., Fagan, A. J., Garcia, C. A., Garcia, N. S., Gerace, S. D., Lee, J. A., Lomas, M. W., et al.: Global ocean particulate organic phosphorus, carbon, oxygen for respiration, and Nitrogen (GO-POPCORN), Scientific data, 9, 688, 2022.

Terseleer, N., Bruggeman, J., Lancelot, C., and Gypens, N.: Trait-based representation of diatom functional diversity in a plankton functional type model of the eutrophied southern North Sea, Limnology and Oceanography, 59, 1958–1972, 2014.

Tilman, D.: Resources: a graphical-mechanistic approach to competition and predation, The American Naturalist, 116, 362–393, 1980.

Transport Matrix Configurations: Transport Matrix Configurations, http://kelvin.earth.ox.ac.uk/spk/Research/TMM/TransportMatrixConfigs/, accessed: 2024-04-25, 2024.

Tréguer, P., Bowler, C., Moriceau, B., Dutkiewicz, S., Gehlen, M., Aumont, O., Bittner, L., Dugdale, R., Finkel, Z., Iudicone, D., Jahn, O., Guidi, L., Lasbleiz, M., Leblanc, K., Levy, M., and Pondaven, P.: Influence of diatom diversity on the ocean biological carbon pump, Nature Geoscience, 11, 27–37, https://doi.org/10.1038/s41561-017-0028-x, bandiera_abtest: a Cg_type: Nature Research Journals Number: 1 Primary_atype: Reviews Publisher: Nature Publishing Group Subject_term: Biogeochemistry;Carbon cycle;Marine biology;Ocean sciences Subject_term_id: biogeochemistry;carbon-cycle;marine-biology;ocean-sciences, 2018.

Turner, E. L., Bruesewitz, D. A., Mooney, R. F., Montagna, P. A., McClelland, J. W., Sadovski, A., and Buskey, E. J.: Comparing performance of five nutrient phytoplankton zooplankton (NPZ) models in coastal lagoons, Ecological Modelling, 277, 13–26, 2014.

Verity, P. G. and Smetacek, V.: Organism life cycles, predation, and the structure of marine pelagic ecosystems, Marine Ecology Progress Series, 130, 277–293, 1996.

Visser, A., Almgren, A., and Kandylas, A.: SISSOMA (v1): modelling marine aggregate dynamics from production to export, preprint, Modelling, 2024.

Waite, A. M., Thompson, P. A., and Harrison, P. J.: Does energy control the sinking rates of marine diatoms?, Limnology and Oceanography, 37, 468–477, 1992.

Waite, R., Beveridge, M., Brummett, R., Castine, S., Chaiyawannakarn, N., Kaushik, S., Mungkung, R., Nawapakpilai, S., and Phillips, M.: Improving productivity and environmental performance of aquaculture, WorldFish, 2014.

Ward, B. A. and Follows, M. J.: Marine mixotrophy increases trophic transfer efficiency, mean organism size, and vertical carbon flux, Proceedings of the National Academy of Sciences, 113, 2958–2963, 2016.

Ward, B. A., Dutkiewicz, S., and Follows, M. J.: Modelling spatial and temporal patterns in size-structured marine plankton communities: top–down and bottom–up controls, Journal of Plankton Research, 36, 31–47, https://doi.org/10.1093/plankt/fbt097, 2014.

Ward, B. A., Wilson, J. D., Death, R. M., Monteiro, F. M., Yool, A., and Ridgwell, A.: EcoGEnIE 1.0: plankton ecology in the cGEnIE Earth system model, Geoscientific Model Development, 11, 4241–4267, https://doi.org/10.5194/gmd-11-4241-2018, publisher: Copernicus GmbH, 2018.

Westberry, T., Behrenfeld, M. J., Siegel, D. A., and Boss, E.: Carbon-based primary productivity modeling with vertically resolved photoacclimation, Global Biogeochemical Cycles, 22, https://doi.org/10.1029/2007GB003078, 2008.

Westoby, M. and Wright, I. J.: Land-plant ecology on the basis of functional traits, Trends in ecology & evolution, 21, 261–268, 2006.

Wirtz, K. W.: Who is eating whom? Morphology and feeding type determine the size relation between planktonic predators and their ideal prey, Marine Ecology Progress Series, 445, 1–12, 2012.

Yasunaka, S., Ono, T., Sasaoka, K., and Sato, K.: Global distribution and variability of subsurface chlorophyll a concentration, Ocean Science Discussions, 2021, 1–22, 2021.

Appendix B: Model equations and parameters

Equations and parameters for unicellular plankton are in Tables B1 and B2, for copepods in Tables B3 and B4, and parameters for POM in Table B5.

B1 Diatom membrane fraction

Let r be the radius of the entire cell including the outer membrane, r_{vac} the vacuole radius, ν_{vac} the fraction of the cell volume occupied by the vacuole, ρ the cell density, and δ the thickness of each membrane (see fig. 2). The total cell volume is:

$$V = \frac{4}{3}\pi r^3 \Leftrightarrow r = \left(\frac{3}{4\pi}V\right)^{1/3} \tag{B1}$$

The vacuole volume is:

$$V_{\mathrm{vac}} = \frac{4}{3}\pi r_{\mathrm{vac}}^3 \Leftrightarrow \nu_{\mathrm{vac}}\frac{V_{\mathrm{vac}}}{V} = \frac{r_{\mathrm{vac}}^3}{r^3} \Leftrightarrow r_{\mathrm{vac}} = r\nu_{\mathrm{vac}}^{1/3} \tag{B2}$$

The volume of the membranes is:

$$V_{\mathrm{m}} = \underbrace{\frac{4}{3}\pi(r^3-(r-\delta)^3)}_{\text{outer membrane}} + \underbrace{\frac{4}{3}\pi((r_{\mathrm{vac}}+\delta)^3 - r_{\mathrm{vac}}^3)}_{\text{inner membrane}} \tag{B3}$$

The Taylor series expansion of $f(x) = r^3-(r-x)^3$ around $\delta = 0$ is $3r^2\delta$.
The approximation for the volume of membranes is then:

$$V_{\mathrm{m}} \approx 4\pi r^2\delta + 4\pi r_{\mathrm{vac}^2}\delta \overset{\text{(B2)}}{=} 4\pi\delta(1+\nu_{\mathrm{vac}}^{2/3})r^2 \tag{B4}$$

Assuming the vacuole density to be 0, the mass of the cell is:

$$m = \rho(V - V_{\mathrm{vac}}) \overset{\text{(B2)}}{=} \rho V(1-\nu_{\mathrm{vac}}) \Leftrightarrow r = \left(\frac{3}{4\pi}\frac{m}{\rho}\frac{1}{1-\nu_{\mathrm{vac}}}\right)^{1/3} \tag{B5}$$

The membrane fraction of the cell's mass is then:

$$\begin{aligned}\nu_{\mathrm{m}} &= \frac{\rho V_{\mathrm{m}}}{m} \overset{\text{(B4),(B5)}}{=} \frac{4\pi\delta(1+\nu_{\mathrm{vac}}^{2/3})r^2}{4/3\pi(1-\nu_{\mathrm{vac}}r^3)} \\ &= 3\frac{\delta}{r}\frac{1+\nu_{\mathrm{vac}}^{2/3}}{1-\nu_{\mathrm{vac}}} \overset{\text{(B5)}}{=} 3\delta\left(\frac{4\pi}{3}\right)^{1/3}\left(\frac{m}{\rho}\right)^{-1/3}\frac{1+\nu_{\mathrm{vac}}^{2/3}}{(1-\nu_{\mathrm{vac}})^{2/3}} \\ &= 6^{2/3}\pi^{1/3}\delta\left(\frac{m}{\rho}\right)^{-1/3}\frac{1+\nu_{\mathrm{vac}}^{2/3}}{(1-\nu_{\mathrm{vac}})^{2/3}}\end{aligned} \tag{B6}$$

and the active biomass fraction is: $\nu = 1-\nu_{\mathrm{m}}$.

Table B1: Equations for the unicellular models (generalists and diatoms)

Description	Equation	
Geometry		
Radius	$r = \left(\frac{3}{4\pi}\frac{m}{\rho}\frac{1}{1-v_{\mathrm{vac}}}\right)^{1/3}$	B1.1
Active biomass fraction	$\nu = 1 - \left[6^{2/3}\pi^{1/3}\delta\left(\frac{m}{\rho}\right)^{-1/3}\frac{1+\nu_{\mathrm{vac}}^{2/3}}{(1-\nu_{\mathrm{vac}})^{2/3}}\right]$	B1.2
Affinities		
Diffusive affinity	$a_{\mathrm{D}} = \frac{\alpha_{\mathrm{N}} r^{-2}}{1+(\frac{r}{r_{\mathrm{N}*}})^{-2}}\frac{1}{1-v_{\mathrm{vac}}}$	B1.3
Light affinity	$a_{\mathrm{L}} = \frac{\alpha_{\mathrm{L}}}{r}\left[1-\exp\left(-\frac{r}{r_{\mathrm{L}*}}\right)\right]\frac{1-\nu}{1-v_{\mathrm{vac}}}$	B1.4
Food affinity	$a_{\mathrm{F}} = \alpha_{\mathrm{F}}$	B1.5
Metabolism		
Max. synthesis rate	$j_{\max} = \alpha_{\max}(1-\nu)$	B1.6
Basal metabolism	$j_{\mathrm{R}} = \alpha_{\mathrm{R}}\alpha_{\max}$	B1.7
Passive losses	$j_{\mathrm{passive}} = c_{\mathrm{passive}}/r$	B1.8
Resource encounter		
Nutrient encounter rate	$j_{\mathrm{enc}.X} = a_{\mathrm{D}}\rho_{\mathrm{C}:X}N$ with $X \in \{\mathrm{N}, \mathrm{DOC}, \mathrm{Si}\}$	B1.9
Light encounter rate	$j_{\mathrm{enc.L}} = \epsilon_{\mathrm{L}} a_{\mathrm{L}} L$	B1.10
Food encounter rate	$j_{\mathrm{enc.F}} = \epsilon_{\mathrm{F}} j_{\mathrm{Fmax}}\frac{a_{\mathrm{F}}F}{a_F j_{\mathrm{Fmax}} + a_{\mathrm{F}}F}$ with $j_{\mathrm{Fmax}} = \frac{c_{\mathrm{F}}}{r}$	B1.11
Synthesis		
Available carbon	$j_{\mathrm{C}} = \sum_X j_{\mathrm{enc}.X}(1-\beta_X) - j_{\mathrm{R}}$	B1.12
Potential nutrient uptake	$j_{\mathrm{nut}} = \max\{0, \min\{j_{\mathrm{enc.N}}, j^{(a)}_{\mathrm{enc.Si}}, \frac{j_{\mathrm{C}} - j_{\mathrm{enc.F}}(1+\beta_{\mathrm{g}})}{1+\beta_{\mathrm{g}}+\beta_{\mathrm{N}}+\beta_{\mathrm{Si}}}\}\}$	B1.13
Net uptakes	$j_{\mathrm{net}} = \min\{j_{\mathrm{nut}} + j_{\mathrm{enc.F}}, [j_{\mathrm{C}} - (\beta_{\mathrm{N}}+\beta_{\mathrm{Si}})j_{\mathrm{nut}}]/(1+\beta_{\mathrm{g}})\}$	B1.14
Division rate	$g = \begin{cases} j_{\mathrm{C}} - j_{\mathrm{passive}} \text{ for } j_{\mathrm{C}} < 0 \\ j_{\mathrm{net}} - j_{\mathrm{passive}} \text{ for } j_{\mathrm{net}} < j_{\mathrm{passive}} \\ j_{\max}\frac{j_{\mathrm{net}}}{j_{\max}+j_{\mathrm{net}}} \text{ for } j_{\mathrm{net}} \geq j_{\mathrm{passive}} \end{cases}$	B1.15
Resource uptakes		
Nutrient uptake	$j_{\mathrm{N}} = j^{(a)}_{\mathrm{Si}} = \max\{0, j_{\mathrm{net}} - j_{\mathrm{enc.F}}\}$	B1.16
Feeding	$j_{\mathrm{F}} = \min\{j_{\mathrm{enc.F}}, \frac{j_{\mathrm{net}} + \beta_{\mathrm{g}}\max\{0, j_{\mathrm{net}}\} + j_{\mathrm{R}} + \beta_{\mathrm{N}} j_{\mathrm{N}} + \beta_{\mathrm{Si}} j_{\mathrm{Si}}}{1-\beta_{\mathrm{F}}}\}$	B1.17
DOC uptake	$j_{\mathrm{DOC}} = \frac{j_{\mathrm{net}} + \beta_{\mathrm{g}}\max\{0, j_{\mathrm{net}}\} + b_{\mathrm{N}} j_{\mathrm{N}} + b_{\mathrm{Si}} j_{\mathrm{Si}} - j_{\mathrm{F}}(1-\beta_{\mathrm{F}})}{(1-\beta_{\mathrm{DOC}})j_{\mathrm{enc.DOC}} + (1-\beta_{\mathrm{L}})j_{\mathrm{enc.L}}} j_{\mathrm{enc.DOC}}$	B1.18
Light uptake	$j_{\mathrm{L}} = \frac{j_{\mathrm{net}} + \beta_{\mathrm{g}}\max\{0, j_{\mathrm{net}}\} + b_{\mathrm{N}} j_{\mathrm{N}} + b_{\mathrm{Si}} j_{\mathrm{Si}} - j_{\mathrm{F}}(1-\beta_{\mathrm{F}})}{(1-\beta_{\mathrm{DOC}})j_{\mathrm{enc.DOC}} + (1-\beta_{\mathrm{L}})j_{\mathrm{enc.L}}} j_{\mathrm{enc.L}}$	B1.19
Losses		

Surplus N exudation	$j_{\mathrm{Nloss}} = j^{(a)}_{\mathrm{SIloss}} = \max\{0, j_{\mathrm{N}} + j_{\mathrm{F}} - j_{\mathrm{passive}} - g\}$	B1.20
Feeding losses (to POM, N, and DOC)	$j_{\mathrm{Floss}} = j_{\mathrm{F}}(1-\epsilon_{\mathrm{F}})/\epsilon_{\mathrm{F}}$	B1.21
Photosynthesis losses (to DOC)	$j_{\mathrm{Lloss}} = j_{\mathrm{L}}(1-\epsilon_{\mathrm{L}})/\epsilon_{\mathrm{L}}$	B1.22
Viral lysis	$\mu_2 = \mu_{2.0} B_i / \ln \Delta_i^{(b)}$	B1.23

$^{(a)}$Only for diatoms

$^{(b)}$Correction for the size-range of a size group where Δ_i is the ratio between the upper and lower sizes of a size group.

Table B2: Parameters for the unicellular model (generalists and diatoms). All parameters from Andersen and Visser (2023) unless noted. Values in parentheses are for diatoms. Temperature corrections are given with Q_{10} regulation (18).

Symbol	Description	Value & Unit
General parameters		
$\rho_{C:N}$	C:N mass ratio of cell	5.68
$\rho_{\mathrm{C:Si}}$	C:Si ratio of cell	$(3.4)^{(a)}$
ν_{vac}	Vacuole fraction	0 $(0.8)^{(b)}$
ρ	Carbon density	0.4×10^{-6} μg C μm^{-1}
δ	Cell wall thickness	0.05 μm
	Size range	$\begin{cases} 10^{-9}...1\ \mu\mathrm{g\ C} & \text{generalists} \\ 10^{-6}...0.01\ \mu\mathrm{g\ C} & \text{diatoms}^{(c)} \end{cases}$
Resource uptakes:		
α_{F}	Food affinity scaling factor	0.018 (0) Lμg C^{-1}d^{-1}
c_{F}	Max phagotrophy coef.	30 μm/d
α_{L}	Light affinity scaling factor	0.3
r^*_{L}	Light aff. cross-over	7.5 μm
α_{N}	Nutrient affinity scaling factor	$0.972\,Q_{1.5}(T)$ Ld^{-1}(μg C)$^{-1}$$\mu$m^2
r^*_{N}	Diffusive aff. cross-over	0.4 μm
Metabolism:		
$\alpha_{\max}$	Maximum synthesis rate	$1.5\,Q_2(T)$ d^{-1}
α_{R}	Basal resp. coef.	$0.03\,Q_2(T)$
ϵ_{L}	Light uptake efficiency	0.8
ϵ_{F}	Food assimilation efficiency	0.8
β	Preferred pred:prey mass ratio	500
σ	Feeding range	1.30
β_{N}	Uptake cost of N	0.3 $\mathrm{g_C/g_N}$
β_{DOC}	Uptake cost of DOC	0.3

β_{Si}	Uptake cost of Si	$(0.3\ \mathrm{g_C/g_{Si}})$
β_{L}	Uptake cost of CO_2	0.08
β_{F}	Uptake cost of feeding	0.3 (0)
β_{g}	Synthesis cost	0.2
Losses:		
μ_2	quadratic mortality coefficient	$0.004\ (\mu\mathrm{g\ C\ L^{-1}})^{-1}$
c_{passive}	passive leakage coef.	0.03
p	Vulnerability	$1\ (0.25)$ [(d)]

[(a)] If the thickness of silicate shell of diatoms were independent of the size of the diatom, the Si:C ratio would decline with size. However, the thickness of the shell actually increases with size, probably to maintain the ability to withstand the crushing force of copepod mandibles (Pančić et al., 2019). Data shows that the C:Si ratio is roughly independent of size with an average value around 3.4 (Fig. B1b).

[(b)] The vacule fraction increases slightly with size. Here we use a constant value of 0.8 which resonably well approximates date (Fig. B1a).

[(c)] The size range of diatoms are roughly from 5×10^{-6} to $0.01\ \mu\mathrm{g\ C}$ (Pančić et al., 2019; Brzezinski, 1985)

[(d)] The vulnerability parameter deteremines the diatom:phytoplankton ratio (Fig. F3). While we know that the vulnerability of diatoms to copepod predation varies with the thickness of the shell (Pančić et al., 2019) there is no laboratory information on the average relative vulnerability of diatoms relative to other relevant prey like dinoflagellates. Hence, the value of the vulnerability is rather arbitrarily set to 0.25.

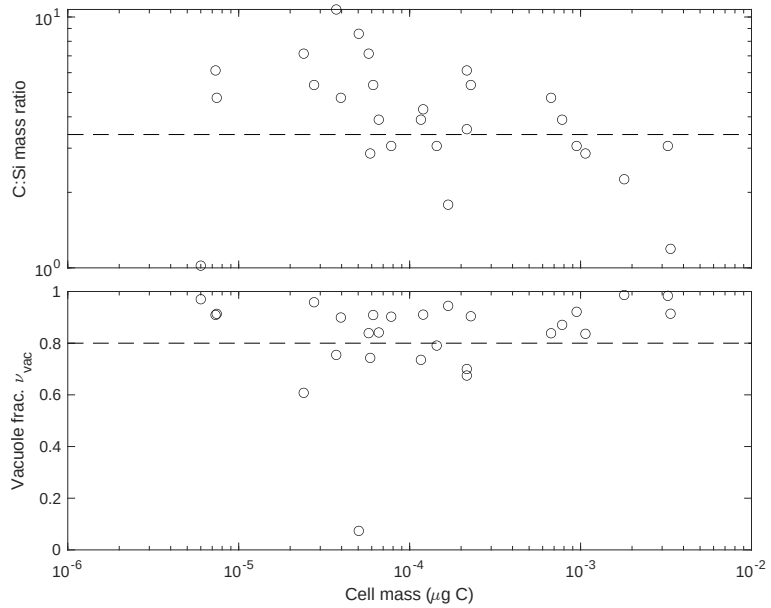

Figure B1. Analysis of diatom C:Si ratio (a) and vacuole fraction (b) as a function of carbon mass. The horisontal dashed lines show the values used in the diatom module. Data from Brzezinski (1985).

Table B3. Equations for the copepod growth and mortality model. Subscript s refers to the stage within the population and S is the adult stage. All variables are in dimensions of per time, except J_{in} which is biomass per time.

Description	Equation	
Maximum consumption	$j_{\mathrm{Fmax}.s} = hm_s^n$	B3.1
Available energy	$j_{\mathrm{avail}.s} = \epsilon j_{\mathrm{F}.s} - j_{\mathrm{resp}.s}$	B3.2
Respiration	$j_{\mathrm{resp}.s} = k_{\mathrm{R}}\epsilon j_{\mathrm{Fmax}.s} + k_{\mathrm{SDA}} j_{\mathrm{F}.s}$ [b]	B3.3
Biomass growth	$\max\{0, j_{\mathrm{avail}.s}\}$	B3.4
Growth and reproductive flux	$J_{\mathrm{in}.s} = \begin{cases} \epsilon_{\mathrm{r}} g_S B_S, & s = 1 \\ j_{\mathrm{out},s-1} B_{s-1}, & s > 1 \end{cases}$	B3.5
Somatic growth rate	$j_{\mathrm{out}.s} = \begin{cases} \frac{g_s - \mu_s}{1-(m_s^-/m_s^+)^{1-\mu_s/g_s}}$ [a], & $s < S$ \\ 0, & $s = S$ \end{cases}$	B3.6
Starvation mortality	$\mu_{\mathrm{st}.s} = \min\{0, j_{\mathrm{avail}.s}\}$	B3.7

[a] m_s^-/m_s^+ is the ratio between the lower and upper size of a size class; see Fig. C1.

[b] The formulation of respiration is slightly revised from the original formulation in Serra-Pompei et al. (2020). It consists of two terms: a fixed basal respiration that is proportional to the maximum respiration and a "specific dynamics action" that is proportional to consumption.

Table B4. Parameters for the copepod model. All parameters are from Serra-Pompei et al. (2022) except where noted.

Symbol	Description	Value		Unit
h	Max. consumption coef.	$0.29\,Q_2(T)$	$0.97\,Q_2(T)$	
ϵ_{F}	Assimilation eff.	0.67		
k_{R}	Basal metabolism coef.	$0.01\,Q_2(T)$ [(a)]		μ g C$^{1/4}$day^{-1}
k_{SDA}	Specific dyn. action coef.	0.16		
β	Preferred pred:prey mass ratio	100	10000	
σ	Size range pref.	1	1.5	
q	Clearance rate exponent	0.75		
α_{F}	Clearance rate coef.	0.29	0.97	l·μ g C$^{1/4}$day^{-1}
m_{egg}	Offspring mass	$0.01 m_S$		
ϵ_{r}	Reproductive efficiency	0.25		
p	Vulnerability	0.2	1	

[(a)] This value represents the starvation metabolism. Following Kiørboe et al. (1985) the starvation metabolism is approximately 18% of the standard metabolism, which is approx. 20% of max metabolism. This gives a very small value of $k_{\mathrm{R}} = 0.006$, which implies that copepods can survive a long time. We increase the value to $k_{\mathrm{R}} = 0.01$ such that large copepods can survive approx. 200 days which is sufficient to surve the winter at high latitudes.

Table B5. Parameters involved in POM and nutrient cycling

Symbol	Description	Value & unit
r_{POM}	Remineralization rate	$0.07\,Q_2(T)$ day^{-1}
γ_2	Frac. of virulysis to POM	0.5
γ_{HTL}	Frac. of HTL mortality to POM	0.5
γ_{F}	Frac. of feeding losses to POM	0.1
p	Palatability of POM	0.1

Appendix C: Computational grid for size groups

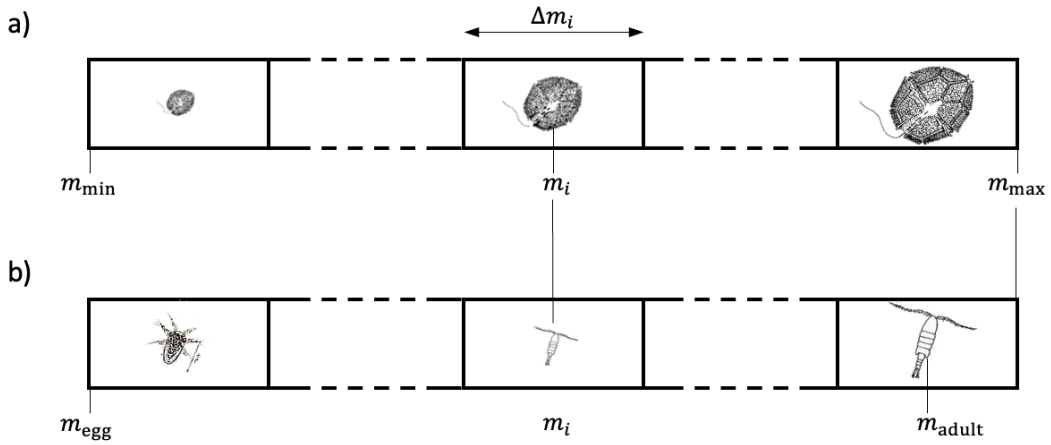

Figure C1. Specification of the computational grid used for unicellular organisms (a) and multicellular organisms (b). The grid for each unicellular plankton group (generalists and diatoms) contain n contiguous size groups in the range from $m_{\min}$ and $m_{\max}$. Each size group is defined by its central mass (geometric mean) m_i. For multicellular plankton the grid for each group is from m_{egg}. The last size group has the central mass m_{adult}. This means that copepods actually becomes adult at smaller mass than m_{adult}, and that the last size group spans a range of masses.

Appendix D: Predation kernel

The effective prey preference function θ_{ij} between size classes of predators i and prey j should deal with the different width's of size classes depending on the size grid spacing. Following Andersen and Visser (2023) this is calculated by integrating over the prey size preference (Eq. 1). The encountered prey in size class j by all predators in class i is:

$$E_{ij} = \int_{m_i^-}^{m_i^+} a_\mathrm{F} m \int_{m_j^-}^{m_j^+} \phi(m,w) B(w)\,\mathrm{d}w\, B(m)/m\,\mathrm{d}m, \tag{D1}$$

where m_i^- and m_i^+ represent the upper an lower bounds of size class i. $B(m)$ represents the normalized biomass spectrum. We assume a Sheldon distribution, i.e., $B(m) \propto m^{-1}$. With the discrete prey and predator groups we can write the encountered food as:

$$E_{ij} = a_\mathrm{F} m_i \theta_{ij} B_j N_i \tag{D2}$$

Where B_j is the total biomass in class j, $B_j = \int B(w)\,\mathrm{d}w$ and N_i is the total abundance of predators $N_i = \int B(m)/m\,\mathrm{d}m$. Equating the two terms and isolating θ_{ij} gives:

$$\begin{aligned}\theta_{ij} = {} & \frac{\sqrt{\Delta}}{(\Delta-1)\log(\Delta)} \\ & \left[\left(\frac{1}{2}s\left(e^{-\frac{\log^2\left(\frac{\Delta z}{\beta}\right)}{s}} + e^{-\frac{\log^2\left(\frac{\beta\Delta}{z}\right)}{s}} - 2e^{-\frac{\log^2\left(\frac{z}{\beta}\right)}{s}}\right) - \right.\right. \\ & \frac{1}{2}\sqrt{\pi}\sqrt{s}\left(\log\left(\frac{\Delta z}{\beta}\right)\mathrm{erf}\left(\frac{\log(\beta)-\log(\Delta z)}{\sqrt{s}}\right) + \log\left(\frac{\beta\Delta}{z}\right)\mathrm{erf}\left(\frac{\log(z)-\log(\beta\Delta)}{\sqrt{s}}\right) + \right. \\ & \left.\left.\left.\log\left(\frac{z}{\beta}\right)\mathrm{erf}\left(\frac{\log\left(\frac{z}{\beta}\right)}{\sqrt{s}}\right)\right)\right)\right]\end{aligned} \tag{D3}$$

where $s = 2\sigma^2$ and $z = m_i/m_j$ and $\Delta = m^+/m^-$. In this way θ_{ij} represent the total preference of size class i and j, including a compensation for the width of size classes and the preference p_{kl} between predator from group k on group l (from Eq. 1). Technically the preference should be expressed as θ_{klij} but the dependence on the group preferences is suppressed to improve readability.

Appendix E: Data sources and processing

E1 POC biomass: GOPOPCORN

GO-POPCORN v2 consists of samples from 12 cruises from 2011 to 2020. We used POCavg_uM that range between 0.7 and 30μm. We compare the POC data with all plankton and detritus particles with a radius between 0.35-15 μm at the same coordinates, month and depth.

E2 Picophytoplankton biomass

We export the insitu phytoplankton biomass (Prochlorococcus, Synechococcus and Picoeukaryotic) at different transects and chlorophyll-a (depth-averaged between 0-10m and ignoring the deeper samples to match satellite measurements), that where measured at different months (ref. "Intercomparison of Ocean Color Algorithms from Picophytoplankton Carbon in the Ocean"). The cell counts were from water samples from up to 200m depth, between 1997 and 2014. Insitu data are compared to the models output at the top layer.

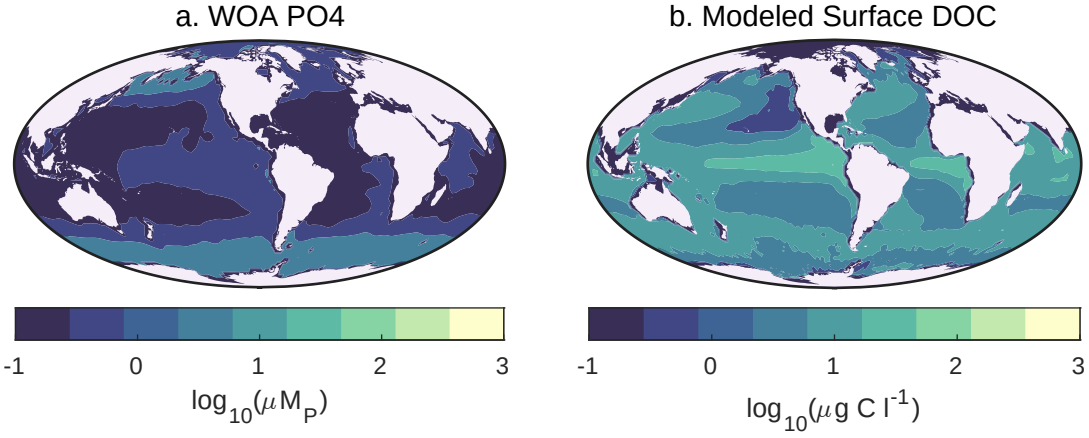

Figure E1. Annual mean concentrations of (a) phosphate (μg P l^{-1}) from World Ocean Atlas (Reagan et al., 2023) and (b) dissolved organic carbon (μg Si l^{-1}) at top 5m.

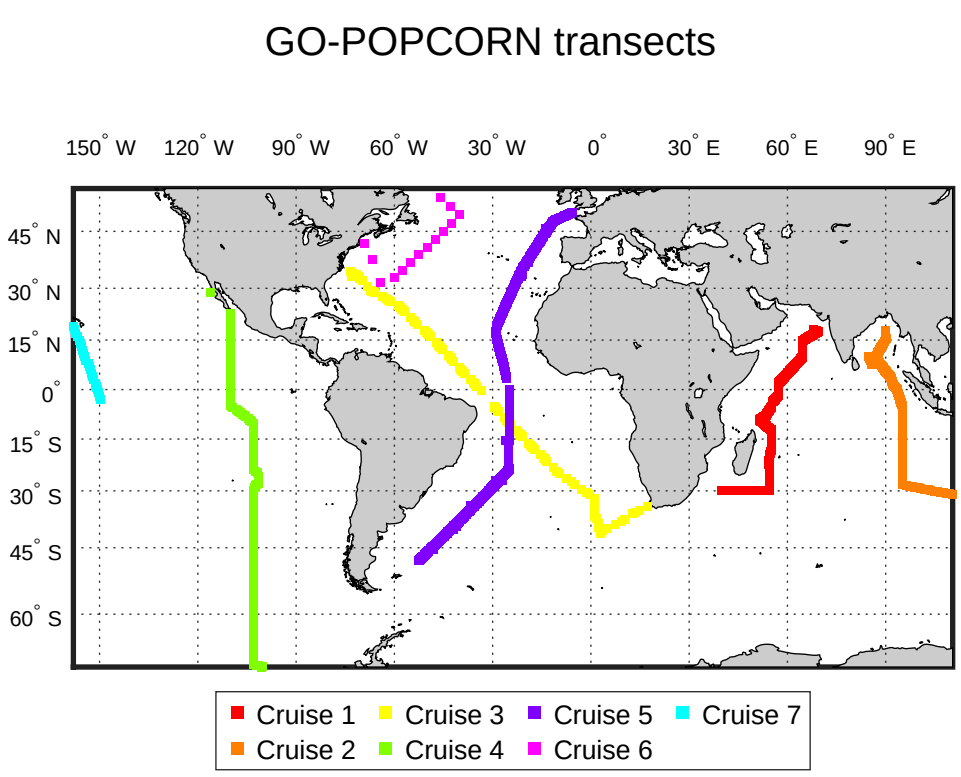

Figure E2. GOPOPCORN Cruises transects

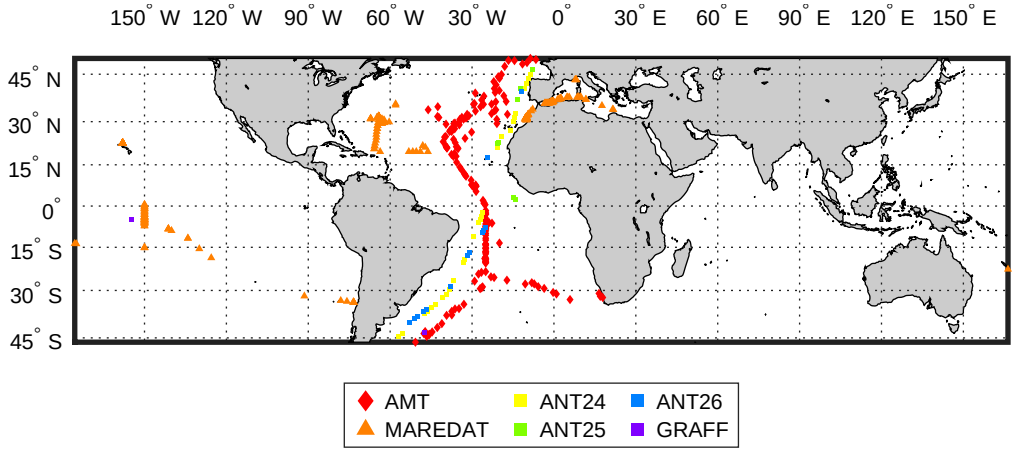

Figure E3. Picophytoplankton Cruises transects

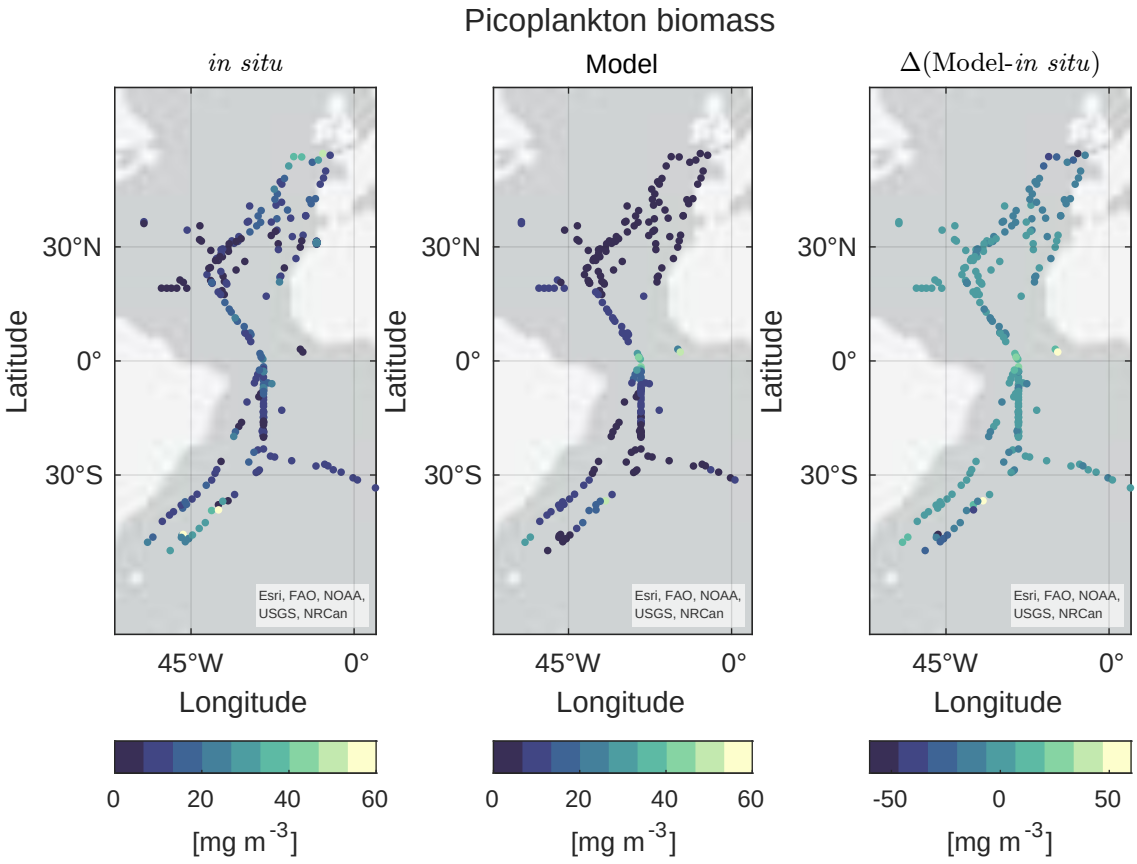

Figure E4. Comparison of picophytoplankton *in situ* data with model picoplankton biomass.

Appendix F: Additional figures

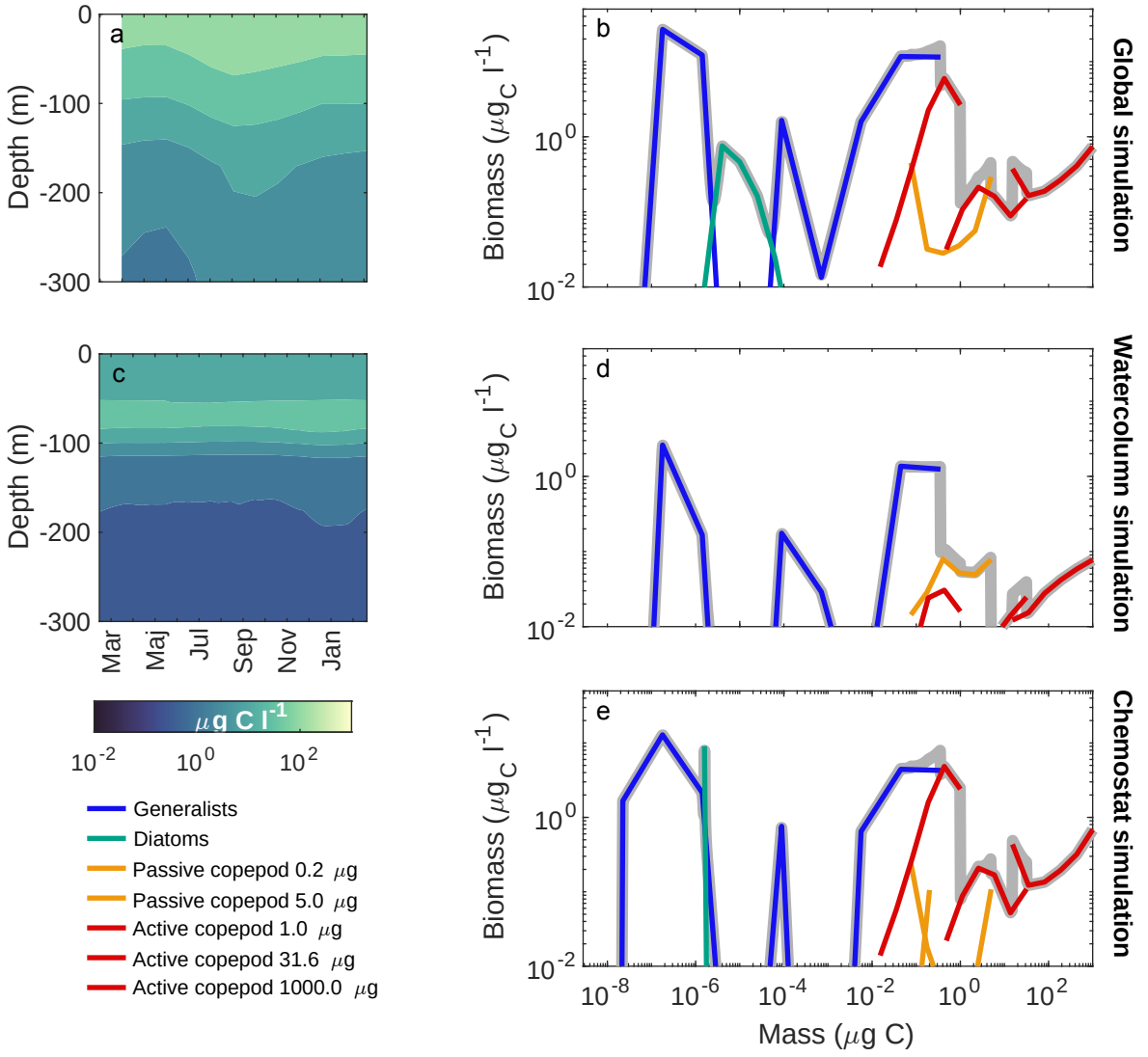

Figure F1. Model output for a 10-year simulation in the global, water column and chemostat model. (a) Total biomass in a water column extracted from the global simulation from Fig. 7 at an upwelling location (5°S, 5°E). (c) Same as (a) but extracted from the watercolumn model. (b),(d),(e): Sheldon spectrum of the community from global, water column and chemostat models at 5 meters depth and averaged over the last year.

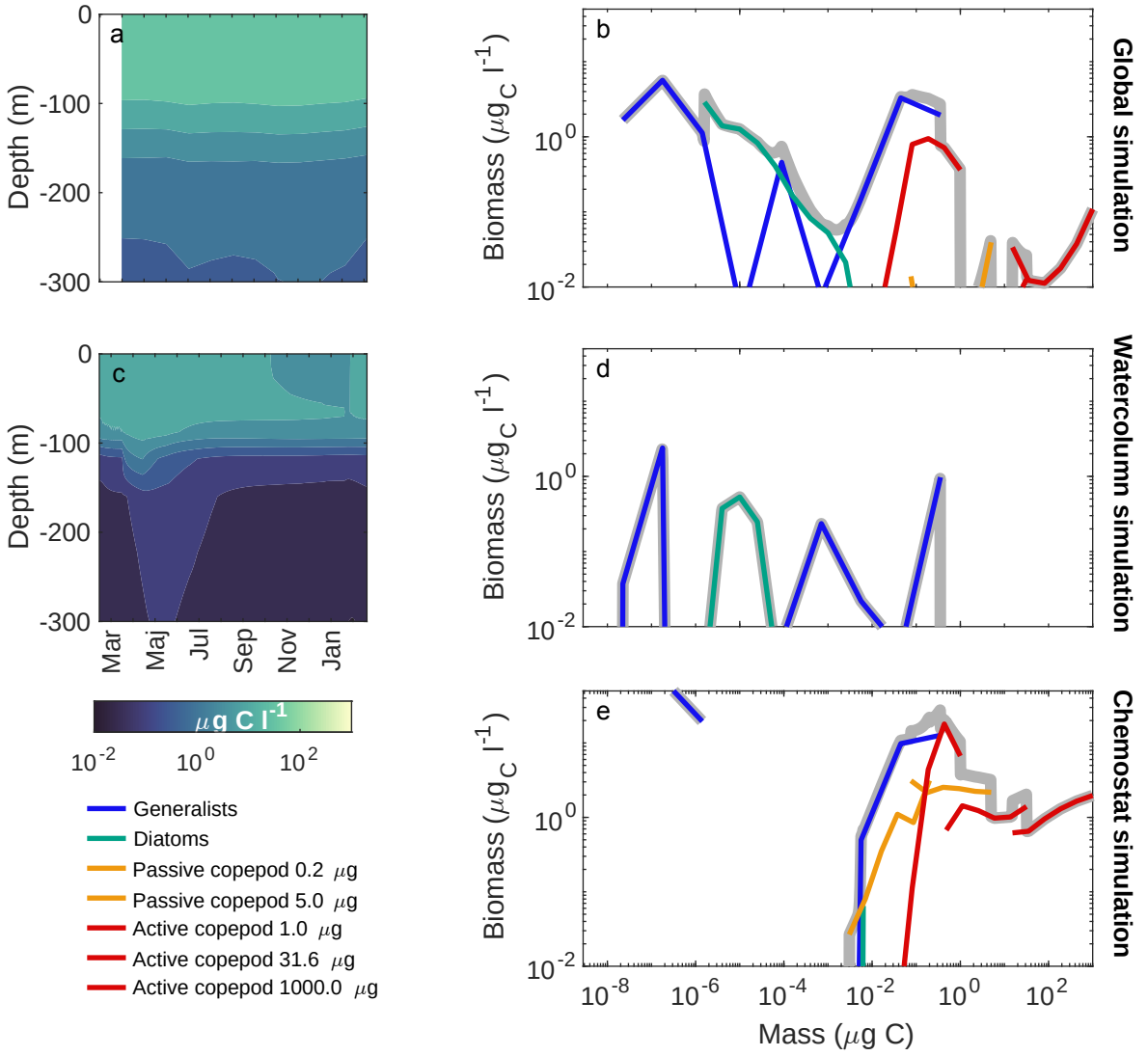

Figure F2. Same as in fig. F1, but at an oligotrophic location (24°N, 158°W).

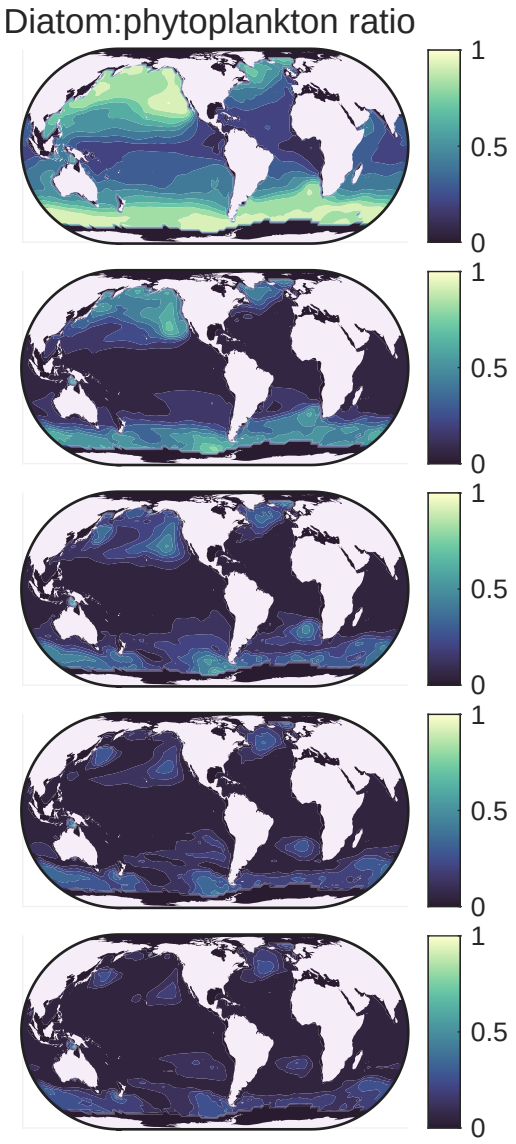

Figure F3. Annual mean surface diatom:phytoplankton ratio for values of the diatom vulnerability ranging from 0 (top) to 1 (bottom).

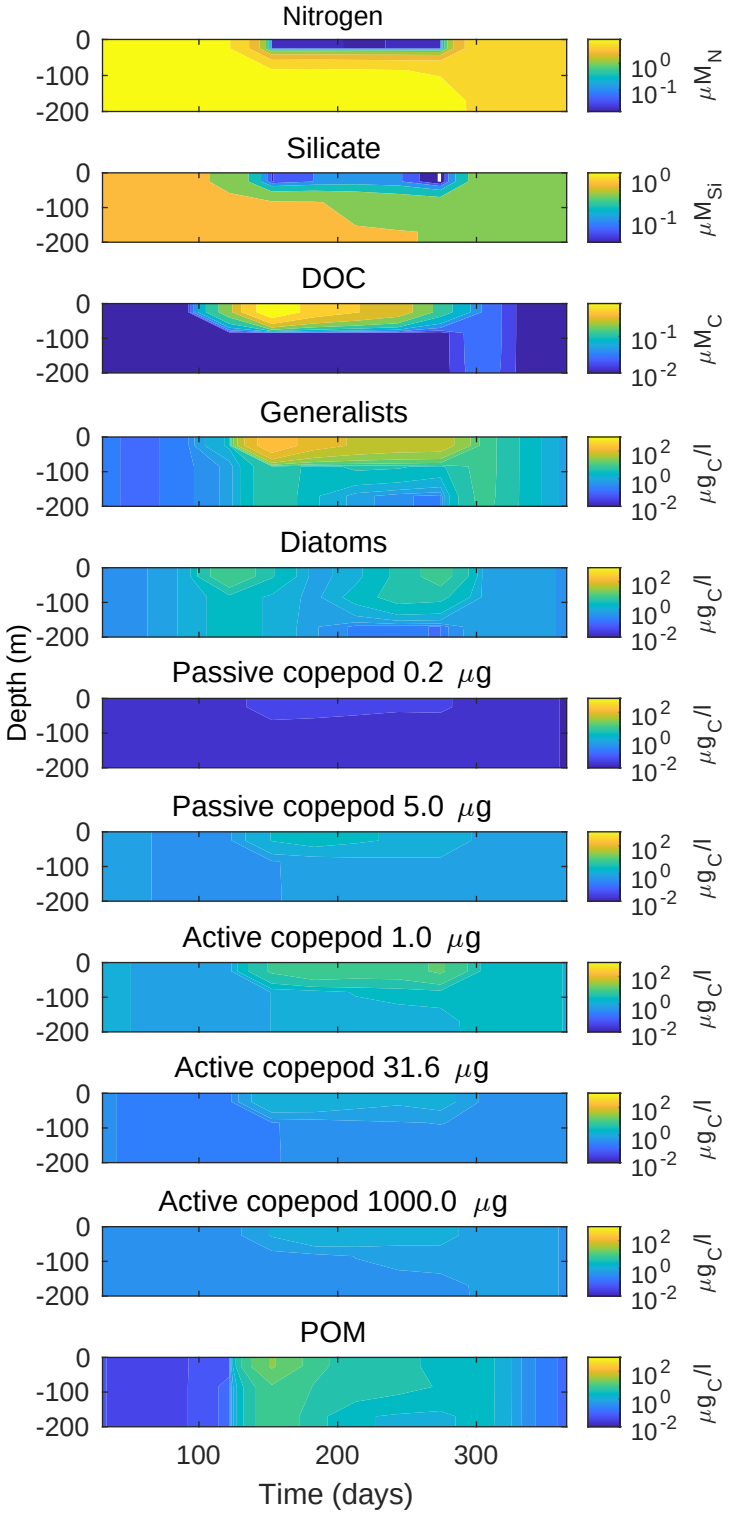

Figure F4. Output of a water-column simulation corresponding to Fig. 15c: (from top to bottom) (a) surface nitrogen ($\mu g_N/l$), (b) silicate ($\mu g_{Si}/l$), and (c) DOC ($\mu g_C/l$); (d) total generalist biomass ($\mu g_C/m^2$), (e) diatom biomass, (f) passive copepod $0.2\,\mu g_C$, (g) passive copepod $5\,\mu g_C$, (h) active copepod $1.0\,\mu g_C$, (i) active copepod $10\,\mu g_C$, (j) active copepod $100\,\mu g_C$, (k) active copepod $1000\,\mu g_C$, and (l) POM.

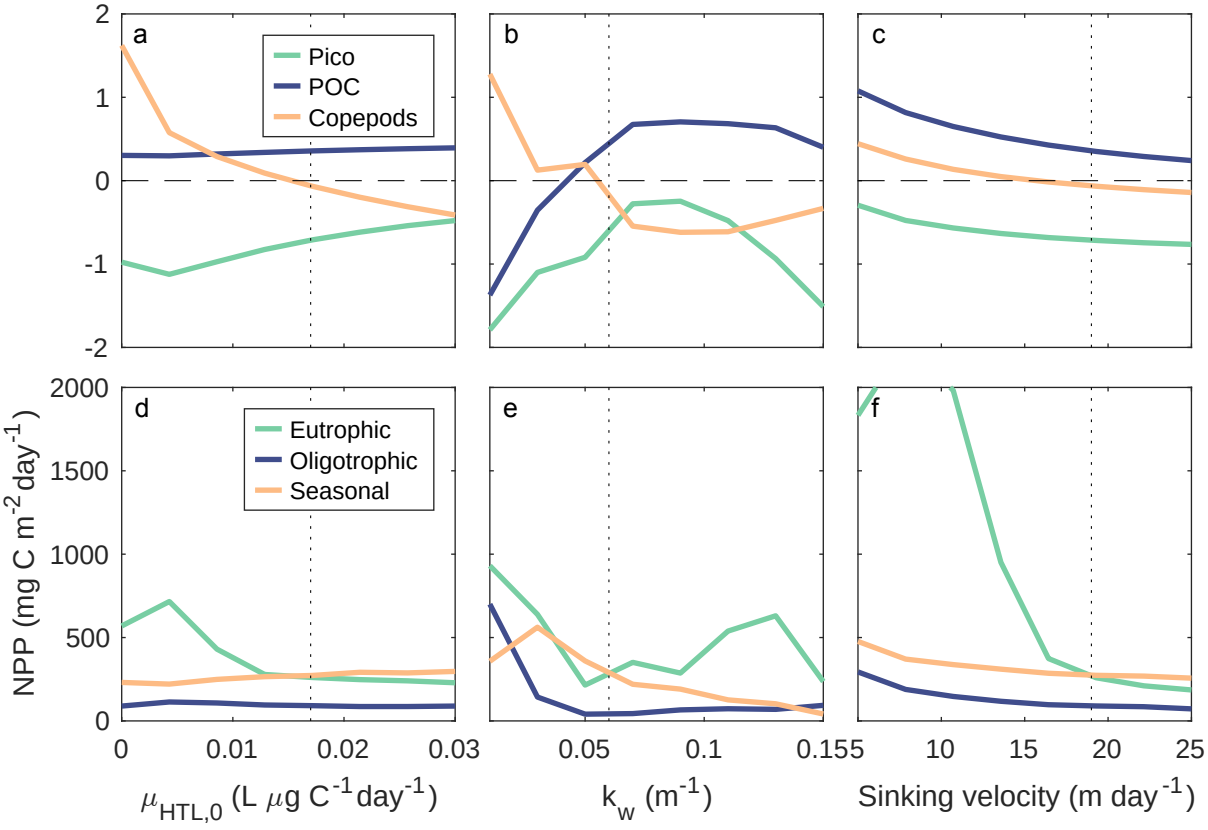

Figure F5. Sensitivity analysis of the parameters: higher trophic level mortality μ_{HTL}, light extinction coefficient k_w, and sinking velocity u. Top row: the objective function for comparison with pico-plankton, POC, and copepods (see Fig. 6. Bottom row: NPP at three characteristic water columns extracted from global simulations.

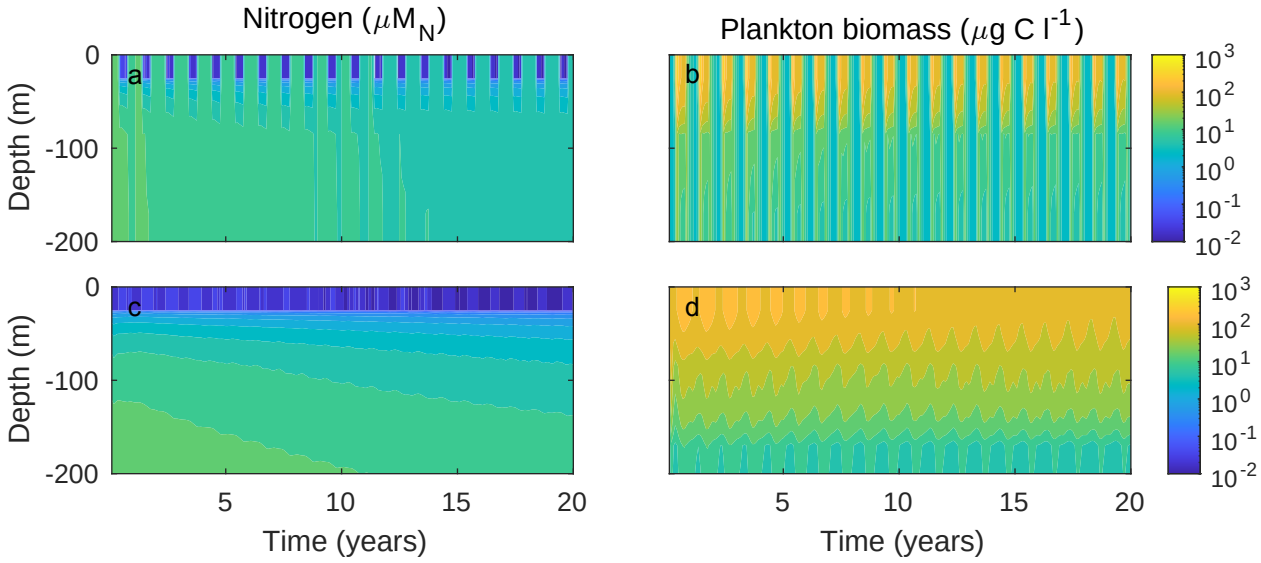

Figure F6. Convergence of a global simulation in a seasonal environment (60°N, 40°W; a,b) and at equator (0°N, 15°W; c,d).

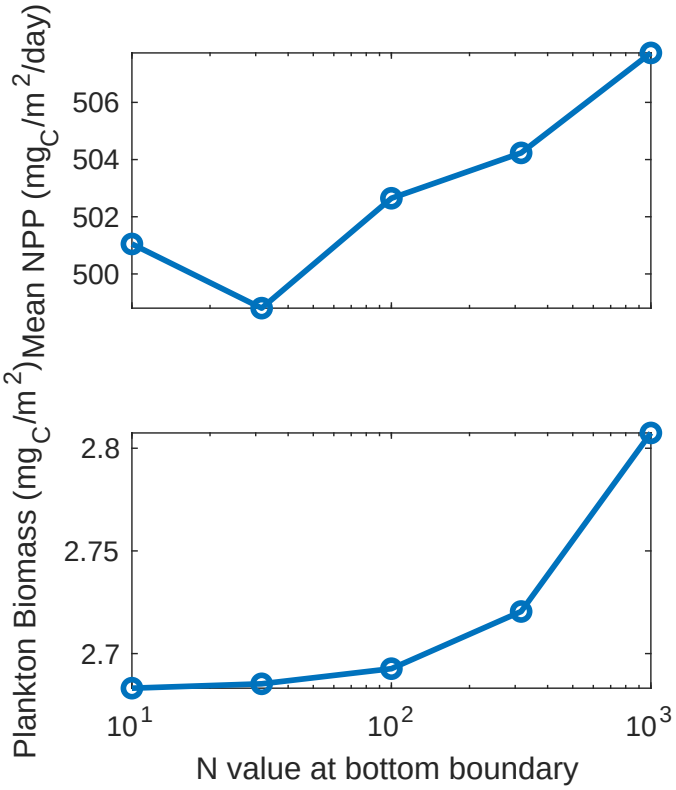

Figure F7. Exploration of the influence of the bottom condition on nutrient on net primary production (top) and plankton biomass (bottom). The values are the last year from a 10 year simulation of a seasonal water column at 60°N and 40°W.