Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-102, 2016
Manuscript under review for journal Geosci. Model Dev.
Published: 30 May 2016

Geoscientific Model Development
Discussions
Open Access

# PhytoSFDM version 1.0.0: Phytoplankton Size and Functional Diversity Model

Esteban Acevedo-Trejos[1], Gunnar Brandt[1,2], S. Lan Smith[3], and Agostino Merico[1,4]

[1]Systems Ecology Group, Leibniz Center for Tropical Marine Ecology, Fahrenheitstrasse 6, 28359 Bremen, Germany
[2]Current address: Brockmann Consult, Max-Planck-Str. 7, 21052 Geesthacht, Germany
[3]Marine Ecosystem Dynamics Research Group, Research and Development Center for Global Change, Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan
[4]Faculty of Physics & Earth Sciences, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany

*Correspondence to:* Esteban Acevedo-Trejos (esteban.acevedo@leibniz-zmt.de)

**Abstract.** Biodiversity is one of the key mechanisms that facilitate the adaptive response of planktonic communities to a fluctuating environment. How to allow for such a flexible response in marine ecosystem models is, however, not entirely clear. One particular way is to resolve the natural complexity of phytoplankton communities by explicitly incorporating a large number of species or plankton functional types. Alternatively, models of aggregate community properties focus on macroecological quantities such as total biomass, mean trait, and trait variance (or functional trait diversity), thus reducing the observed natural complexity to a few mathematical expressions. We developed the modelling tool PhytoSFDM, which can resolve species discretely and can capture aggregate community properties. The tool also provides a set of methods for treating diversity under realistic oceanographic settings. This model is coded in Python and is distributed as an open-source software. PhytoSFDM is implemented in a 0D physical scheme and can be applied to any location of the world oceans. We show that aggregate-community models reduce computational complexity while preserving relevant macroecological features of phytoplankton communities. Compared to species-explicit models, aggregate models are more manageable in terms of number of equations and have faster computational times. Further developments of this tool should address the caveats associated with the assumptions of aggregate community models and on implementations into spatially resolved physical settings (1D and 3D). With PhytoSFDM we embrace the idea of promoting open source software and encourage scientists to build on this modelling tool to further improve our understanding of the role that biodiversity plays in shaping marine ecosystems.

## 1 Introduction

Numerical models are simplified abstractions of complex phenomena. They are engineered for the problem at hand and inclined towards one end of Levins' triad (Levins, 1966): generality, realism and precision. Marine ecosystem models are no exceptions, and the scientific community has questioned the trend towards increasing model complexity, in terms of large numbers of state variables and parameters (Fulton et al., 2003; Anderson, 2005; Hood et al., 2006; Anderson, 2010). Alternatives such as trait-based models have been put forward as a way to simplify overly parameterised ecosystem models (Follows and Dutkiewicz, 2011).

In the past two decades, trait-based models of planktonic ecosystems have become important tools for elucidating the fundamental mechanisms behind emergent patterns of community structure and diversity. Most of these models describe the phytoplankton community by a *discrete* representation of many species or functional groups (Baird and Suthers, 2007; Follows et al., 2007; Bruggeman, 2007; Barton et al., 2010; Banas, 2011; Ward, 2012; Smith et al., 2015). Alternatively, models have been developed that treat the whole phytoplankton species assemblage as a single entity (Wirtz and Eckhardt, 1996; Norberg et al., 2001; Merico et al., 2009; Bruggeman, 2009; Wirtz, 2013; Wirtz and Sommer, 2013; Terseleer et al., 2014; Acevedo-Trejos et al., 2015). These models use *aggregate* community properties such as total biomass, mean trait, and trait variance to describe changes in phytoplankton community composition. Hence, by approximating the full spectrum of species or functional types with just a few macroecological properties, these models present a way of reducing the complexity of natural communities (Merico et al., 2009).

The simplification of both types of trait-based models (i.e. discrete and aggregate) relies on the use of a key trait, for which relationships with other traits can be formulated. Cell size is recognised as one of the most important traits for characterising phytoplankton communities (Litchman et al., 2008; Finkel et al., 2010; Litchman et al., 2010; Marañón, 2015), and it has been commonly used in plankton ecosystem models (Baird and Suthers, 2007; Banas, 2011; Ward, 2012; Wirtz, 2013; Wirtz and Sommer, 2013; Terseleer et al., 2014; Acevedo-Trejos et al., 2015; Smith et al., 2015). This morphological trait affects trophic organisation of foodwebs and the sequestration of $CO_2$ into the ocean interior (Chisholm, 1992). Phytoplankton size also impacts on many ecological and physiological functions and is linked to other relevant traits via tradeoff relationships (see reviews by Litchman et al., 2008; Finkel et al., 2010; Litchman et al., 2010). Therefore, studies on how cell size is associated to ecological and physiological processes and on the impact that these associations have on the structure and functioning of planktonic communities are of fundamental importance (Marañón, 2015; Andersen et al., 2015).

Here we present a new phytoplankton size and functional diversity model (called PhytoSFDM) that allows for five different ways of describing the size composition of phytoplankton communities in the upper mixed layers of the world oceans. In the first variant, the phytoplankton community is described according to the classical approach that resolves the discrete assemblage of many different species and then we present four alternative ways to express aggregate community properties of phytoplankton based on four different ways of treating size diversity. We provide this model as open-source so that it can be used, modified and redistributed freely with the aims of fostering reproducibility and encouraging investigations about the impact of environmental conditions on properties of phytoplankton community structure and diversity.

## 2 Model description

PhytoSFDM is developed from the study of Acevedo-Trejos et al. (2015), which used a size-based model of aggregate community properties to investigate the phytoplankton size structure and size diversity in two environmentally contrasting regions of the Atlantic Ocean. In this model, the phytoplankton community self-assembles according to a trade-off emerging from relationships between cell size and (1) nutrient uptake, (2) zooplankton grazing, and (3) phytoplankton sinking. In PhytoSFDM we have extended this work by providing four ways of treating size diversity using a moment-based approximation (see Smith et al., 2011; Bonachela et al., 2015, and section 2.1.3 in this study). In addition, we include a discrete version of the model (hereafter referred to as the full model) to better illustrate the potential of using aggregate models as compared to the equivalent discrete version. In the following, we present the mathematical equations, a description of the code structure, and easy-to-follow examples of how to use the model.

### 2.1 Mathematical formulations

### 2.1.1 Mixed layer scheme

The zero-dimensional physical setup consists of two vertical layers, the upper-mixed layer containing the pelagic ecosystem and the abiotic bottom layer with nutrient concentration as forcing. Following Evans and Parslow (1985) and Fasham et al. (1990), we describe material mixing between the two layers (K) as a function of the mixed-layer depth (M),

$$\mathrm{K} = \frac{\kappa + \mathrm{h}^{+}(\mathrm{t})}{\mathrm{M}(\mathrm{t})}, \qquad (1)$$

where $\kappa$ is a constant that parameterises diffusive mixing across the thermocline and $\mathrm{h}^{+}(\mathrm{t})$ is a function that describes entrainment and detrainment of material. The latter is given by $\mathrm{h}^{+}(\mathrm{t}) = max[\mathrm{h}(\mathrm{t}), 0]$, with $\mathrm{h}(\mathrm{t}) = \mathrm{dM}(\mathrm{t})/\mathrm{dt}$.

Zooplankton are considered capable of maintaining themselves within the upper mixed layer, thus, their mixing term simplifies to $\mathrm{K_Z} = \mathrm{h}(\mathrm{t})$.

### 2.1.2 Dynamics of the full phytoplankton community

The description of the phytoplankton community is a trait-based variant of the classical Nutrient-Phytoplankton-Zooplankton-Detritus (NPZD) model (Fasham et al., 1990). We define $n$ morphologically distinct phytoplankton types (hereafter referred to as morphotypes). Each morphotype is characterised by a biomass $P_i$ and a cell size $S_i$, in units of $\mu$m Equivalent Spherical Diameter. The distribution of phytoplankton cell size is known to be positively skewed (i.e. an asymmetrical size distribution with a pronounced right tail compared to its left tail), due to physiological, morphological and ecological constrains that limits phytoplankton from a minimum size of around 0.15 $\mu$m to a maximum size of more than 100 $\mu$m (Marañón, 2015; Andersen et al., 2015). Consequently, we assume a log-normal distribution of size to represent the size of each morphotype, thus transforming the cell size $S_i$ as follows $L_i = \ln(S_i)$. The net growth rate of the whole phytoplankton community (P) is then given by:

$$\frac{dP}{dt} = \sum_{i=1}^{n} f_i(L_i, E) \cdot P_i \,, \tag{2}$$

where $f_i(L_i, E)$ is the net growth rate of size class i, which we assume to be a proxy for fitness (Smith et al., 2011). Hence $f_i$ accounts for the gains and losses of each morphotype as a function of cell size ($L_i$) and environment E. The latter includes changes in nutrients, irradiance, temperature, and grazing. The equation describing the fitness functions of each size class i is thus given by:

$$f_i = \mu_P \cdot F(T) \cdot H(I) \cdot U(L_i, N) \; - \; \mu_Z \cdot G(L_i, P_i) \cdot Z \; - \; V(L_i, M) \; - \; m_P \; - \; K \,, \tag{3}$$

where $\mu_P$ indicates the maximum growth rate and $F(T) = e^{0.063 \cdot T}$ is Eppley's formulation for temperature-dependent growth (Eppley , 1972). The light limiting term, $H(I)$, represents the average light I available in the upper mixed layer. According to Steele's formulation (Steele , 1962):

$$H(I) = \frac{1}{M(t)} \int_0^M \left[ \frac{I(z)}{I_s} \cdot e^{\left(1 - \frac{I(z)}{I_s}\right)} \right] dz \,, \tag{4}$$

where $I_s$ is the light level at which photosynthesis saturates and $I(z)$ is the irradiance at depth $z$. The exponential decay of light with depth is computed according to the Beer-Lambert law with a generic extinction coefficient $k_w$

$$I(z) = I_0 \cdot e^{-k_w \cdot z} \,. \tag{5}$$

The nutrient limiting term U in equation 3 is determined by a Monod function with a half-saturation constant $\mathrm{K_N}$, which scales allometrically with phytoplankton cell size L (Litchman et al., 2007),

$$\mathrm{U(L_i,N)} = \frac{\mathrm{N}}{\mathrm{N+K_N}} = \frac{\mathrm{N}}{\mathrm{N}+(\beta_\mathrm{U}\cdot \mathrm{e}^{\mathrm{L_i}\cdot\alpha_\mathrm{U}})}\,, \tag{6}$$

with $\beta_{\mathrm{K_N}}$ and $\alpha_{\mathrm{K_N}}$, respectively, intercept and slope of the $\mathrm{K_N}$ allometric function. This empirical relationship is based on observations of different phytoplankton groups (see figure 3b in Litchman et al., 2007), with the regression parameters rescaled from cell volume to ESD.

The loss term $\mathrm{G(L_i,P_i)}$ in Equation 3, represents zooplankton grazing, which is a Monod-type function:

$$\mathrm{G(L_i,P_i)} = \frac{\mathrm{e}^{\mathrm{L_i}\cdot\alpha_\mathrm{G}}}{\sum_{\mathrm{i=1}}^{\mathrm{n}}\mathrm{P_i}\cdot\mathrm{e}^{\mathrm{L_i}\cdot\alpha_\mathrm{G}}+\mathrm{K_P}}\,, \tag{7}$$

where $\alpha_\mathrm{G}$ is the slope for size-dependent grazing and $\mathrm{K_P}$ is the half saturation constant. This formulation is based on observational evidence by Hansen et al. (1994, 1997).

The loss term $\mathrm{V(L_i,M)}$ in Equation 3 represents the sinking of phytoplankton as a function of size and depth of the mixed layer,

$$\mathrm{V(L_i,M)} = \frac{\beta_\mathrm{V}\cdot\mathrm{e}^{\mathrm{L_i}\cdot\alpha_\mathrm{V}}}{\mathrm{M(t)}}\,, \tag{8}$$

where the constants $\alpha_V$ and $\beta_V$ are the parameters of the function relating phytoplankton cell size to sinking velocity according to Stokes' law (Kiørboe, 1993). These parameters are expressed here in units of meters per day.

The loss term $\mathrm{m_P}$ in Equation 3 accounts for all phytoplankton losses other than those from grazing and mixing.

Differential equations for nutrients (N), zooplankton (Z), and detritus (D) complete the model system:

$$\frac{\mathrm{dN}}{\mathrm{dt}} = -\sum_{\mathrm{i=1}}^{\mathrm{n}}\mu_\mathrm{P}\cdot\mathrm{F(T)}\cdot\mathrm{H(I)}\cdot\mathrm{U(L_i,N)}\cdot\mathrm{P_i}\;+\;\delta_\mathrm{D}\cdot\mathrm{D}\;+\;\mathrm{K}\cdot(\mathrm{N_0}-\mathrm{N})\,, \tag{9}$$

$$\frac{\mathrm{dZ}}{\mathrm{dt}} = \sum_{\mathrm{i=1}}^{\mathrm{n}}\delta_\mathrm{Z}\cdot\mu_\mathrm{Z}\cdot\mathrm{Z}\cdot\mathrm{G(L_i,P_i)}\cdot\mathrm{P_i}\;-\;\mathrm{m_Z}\cdot\mathrm{Z}^2\;-\;\mathrm{K_Z}\cdot\mathrm{Z}\,, \tag{10}$$

$$\frac{\mathrm{dD}}{\mathrm{dt}} = \sum_{\mathrm{i=1}}^{\mathrm{n}}(1-\delta_\mathrm{Z})\cdot\mu_\mathrm{Z}\cdot\mathrm{Z}\cdot\mathrm{G(L_i,P_i)}\cdot\mathrm{P_i}\;+\;\sum_{\mathrm{i=1}}^{\mathrm{n}}\mathrm{m_P}\cdot\mathrm{P_i}\;+\;\mathrm{m_Z}\cdot\mathrm{Z}^2-\;\delta_\mathrm{D}\cdot\mathrm{D}\;-\;\mathrm{K}\cdot\mathrm{D}\,, \tag{11}$$

where $\delta_\mathrm{D}$ is the mineralization rate and $\mathrm{N_0}$ is the nutrient concentration below the upper mixed layer. $\mu_\mathrm{Z}$, $\delta_\mathrm{Z}$ and $\mathrm{m_Z}$ are, respectively, maximum growth rate, prey assimilation coefficient, and mortality rate of zooplankton. All parameter values and their units are reported in Table 1.

#### 2.1.3 Dynamics of the aggregate phytoplankton community

The phytoplankton community comprising many distinct morphotypes (Equations 2 to 8) can be approximated with the so-called moment-based approach (Wirtz and Eckhardt, 1996; Norberg et al., 2001; Merico et al., 2009; Terseleer et al., 2014; Acevedo-Trejos et al., 2015) by three macroscopic properties, namely: total biomass, mean trait, and trait variance (refer to Wirtz and Eckhardt, 1996; Norberg et al., 2001; Merico et al., 2009, for further details about the use of different moment closure techniques). Here the whole phytoplankton community is characterised by the morphological trait cell size and by a trade-off that emerges from three allometric relationships described by Equations 6-8). The equations for the respective three macroscopic properties are:

$$\frac{\mathrm{dP}}{\mathrm{dt}} \approx \mathrm{P} \cdot (\mathrm{f} + \frac{1}{2} \cdot \mathrm{f}^{(2)} \cdot \mathrm{V}), \tag{12}$$

$$\frac{\mathrm{d\bar{L}}}{\mathrm{dt}} \approx \mathrm{f}^{(1)} \cdot \mathrm{V}, \tag{13}$$

$$\frac{\mathrm{dV}}{\mathrm{dt}} \approx \mathrm{f}^{(2)} \cdot \mathrm{V}^2, \tag{14}$$

where f is the net growth rate (or the fitness function, see equation 3), and $\mathrm{f}^{(\mathrm{n})}$ is the nth derivative of the net growth with respect to the trait. Due to competitive exclusion, however, the phytoplankton community loses functional diversity over time, i.e. the variance declines to zero with time, in both full and aggregate model formulations (Merico et al., 2014). We name this standard formulation "Unsustained Variance".

Alternatively, one can use the approximated model to focus only on changes in the mean trait, thus ignoring changes in the variance by fixing it to an arbitrary constant value:

$$\frac{\mathrm{dP}}{\mathrm{dt}} \approx \mathrm{P} \cdot (\mathrm{f} + \frac{1}{2} \cdot \mathrm{f}^{(2)} \cdot \mathrm{V}), \tag{15}$$

$$\frac{\mathrm{d\bar{L}}}{\mathrm{dt}} \approx \mathrm{f}^{(1)} \cdot \mathrm{V}, \tag{16}$$

$$\frac{\mathrm{dV}}{\mathrm{dt}} = 0. \tag{17}$$

While using these two formulations (i.e. Unsustained and Fixed Variance) can be acceptable in some special cases (e.g. in experiments that lead to competitive exclusion or where diversity is being manipulated), it is clear that they fail to account for changes in the adaptive capacity of the community, which requires allowing the size variance, and thereby functional diversity, to vary over time (Merico et al., 2014).

Within our modelling tool we also provide two alternative ways of treating the size variance: immigration (following Norberg et al., 2001) and trait diffusion (following Merico et al., 2014). The treatment with immigration considers the introduction of biomass and new trait values from hypothetical adjacent communities into the resident community. The addition of incoming amount of biomass per day is named immigration I,

$$\frac{\mathrm{dP}}{\mathrm{dt}} \approx \mathrm{P}[\mathrm{f} + \frac{1}{2} \cdot \mathrm{f}^{(2)} \cdot \mathrm{V}] + \mathrm{I}\,, \tag{18}$$

$$\frac{\mathrm{d\overline{L}}}{\mathrm{dt}} \approx \mathrm{f}^{(1)} \cdot \mathrm{V} + \frac{\mathrm{I}}{\mathrm{P}}(\mathrm{L_I} - \overline{\mathrm{L}})\,, \tag{19}$$

$$\frac{\mathrm{dV}}{\mathrm{dt}} \approx \mathrm{f}^{(2)} \cdot \mathrm{V}^2 + \frac{\mathrm{I}}{\mathrm{P}}[(\mathrm{V_I} - \mathrm{V}) + (\mathrm{L_I} - \overline{\mathrm{L}})^2]\,, \tag{20}$$

where $\mathrm{L_I}$ and $\mathrm{V_I}$ are, respectively, the mean size and the size variance of the immigrating community. As implemented by Acevedo-Trejos et al. (2015), we treat I as a density-dependent process (i.e. $\mathrm{I} = \delta_\mathrm{I} \cdot \mathrm{P}$), and set $\mathrm{L_I}$ equal to the mean size of the resident community (i.e. $\mathrm{L_I} = \overline{\mathrm{L}}$).

The treatment of the size variance based on trait diffusion (Merico et al., 2014) gives:

$$\frac{\mathrm{dP}}{\mathrm{dt}} \approx \mathrm{P} \cdot [\mathrm{f} + \frac{1}{2} \cdot \mathrm{f}^{(2)} \cdot \mathrm{V} + \frac{1}{2} \cdot \nu(r_4 \cdot \mathrm{V} - 3 \cdot r_2)]\,, \tag{21}$$

$$\frac{\mathrm{d\overline{L}}}{\mathrm{dt}} \approx \mathrm{f}^{(1)} \cdot \mathrm{V} + \nu(r_3 \cdot \mathrm{V} - 3 \cdot r_1)\,, \tag{22}$$

$$\frac{\mathrm{dV}}{\mathrm{dt}} \approx \mathrm{f}^{(2)} \cdot \mathrm{V}^2 + \nu(r_4 \cdot \mathrm{V}^2 - 5 \cdot \mathrm{V} \cdot r_2 + 2 \cdot r)\,, \tag{23}$$

where $\nu$ is the trait diffusivity parameter, $r$ is the reproduction rate (or gross growth), and $r_n$ is the nth derivative of gross growth with respect to the trait.

The system of differential equations for all variance treatments is completed by equations describing gains and losses in nutrients (N), zooplankton (Z), and detritus (D):

$$\frac{\mathrm{dN}}{\mathrm{dt}} = -\mu_\mathrm{P} \cdot \mathrm{F(T)} \cdot \mathrm{H(I)} \cdot \mathrm{U}(\overline{\mathrm{L}},\mathrm{N}) \cdot \mathrm{P} \;+\; \delta_\mathrm{D} \cdot \mathrm{D} \;+\; \mathrm{K} \cdot (\mathrm{N_0} - \mathrm{N})\,, \tag{24}$$

$$\frac{\mathrm{dZ}}{\mathrm{dt}} = \delta_\mathrm{Z} \cdot \mu_\mathrm{Z} \cdot \mathrm{Z} \cdot \mathrm{G}(\overline{\mathrm{L}},\mathrm{P}) \cdot \mathrm{P} \;-\; \mathrm{m_Z} \cdot \mathrm{Z}^2 \;-\; \mathrm{K_Z} \cdot \mathrm{Z}\,, \tag{25}$$

$$\frac{\mathrm{dD}}{\mathrm{dt}} = (1 - \delta_\mathrm{Z}) \cdot \mu_\mathrm{Z} \cdot \mathrm{Z} \cdot \mathrm{G}(\overline{\mathrm{L}},\mathrm{P}) \cdot \mathrm{P} \;+\; \mathrm{m_P} \cdot \mathrm{P} \;+\; \mathrm{m_Z} \cdot \mathrm{Z}^2 - \;\delta_\mathrm{D} \cdot \mathrm{D} \;-\; \mathrm{K} \cdot \mathrm{D}\,. \tag{26}$$

## 2.2 Environmental forcing

We compiled monthly climatological forcing data for mixed-layer depth (MLD), photosynthetic active radiation (PAR), sea surface temperature (SST) and concentration of nutrients below the upper mixed layer ($N_0$). The MLD data were obtained from Monterey and Levitus (1997) using the variable density criterion and are openly accessible from https://www.nodc.noaa.gov/OC5/WOA94/mix.html. The PAR data were obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS), for the time period 2002-2011. This dataset is managed and distributed by the NASA's Ocean Biology Processing Group (http://oceancolor.gsfc.nasa.gov/cms/). SST and $N_0$ were obtained from the World Ocean Atlas 2009 (WOA09), which is maintained and distributed by NOAA (https://www.nodc.noaa.gov/OC5/WOA09/pr_woa09.html). For consistency and efficiency, all data were transformed from their original formats (e.g. TXT and HDF) to NetCDF. All monthly forcing were spatially averaged over the selected location (square boxes in Figure 1) and then interpolated to obtain daily values (Figure 2).

## 3 Test-case simulation

A test-case model configuration is provided for a location of the North Atlantic ocean at 47.5° N 15.5° W (Figure 1), a region where seasonal changes in mean size and size diversity are well known (Acevedo-Trejos et al., 2015). This region presents the typical oceanographic conditions of a temperate environment (Figure 2). The environmental conditions produce a pronounced phytoplankton bloom in spring, which stimulates secondary production and almost the full depletion of nutrients (Figure 3). Overturning of the water column in autumn restocks the pool of nutrient and light limitation halts primary production (Figures 2 and 3).

### 3.1 Comparison of full and aggregate models

Within PhytoSFDM, we provide a practical example of how to implement and compare phytoplankton community models that aim to describe a) a full assemblage of species or morphotypes (see section 2.1.2), and b) an aggregate community (see section 2.1.3). The aggregate community model is an approximation of the full assemblage of species or morphotypes(Wirtz and Eckhardt, 1996; Norberg et al., 2001; Merico et al., 2009).

Figures 3 and 4 show the results of, respectively, the full model and the aggregate model for the Unsustained Variance case. N, P, Z, and D are unaffected by the type of model considered. As expected, the dynamics of P, $\overline{\mathrm{L}}$, and V produced by the aggregate model are good approximations of those produced by the full model. Both models exhibit competitive exclusion, as indicated by the reduction in the number of morphotypes and consequently in the loss of size variance over time (Figure 4). The phytoplankton community evolves towards the optimal trait value, which is expressed by the fittest few morphotypes for the chosen parameterisation and the prevailing environmental conditions. Although competitive exclusion is well established theoretically (Hardin, 1960), natural communities of phytoplankton are typically very diverse, hence we will explore in the following the effects of different ways of sustaining the variance.

Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-102, 2016
Manuscript under review for journal Geosci. Model Dev.
Published: 30 May 2016

### 3.2 Comparison of variance treatments

The key aspect of trait-based models is their ability to describe the phytoplankton community in terms of mean trait and trait variance. Figures 5 and 6 show the results of one-year simulation after an initial spin-up phase of four years. While the four treatments produce very similar, if not identical, dynamics for N, P, Z, and D (Figure 5), the results for the mean size and the size variance differ considerably among treatments (Figure 6).

As already discussed, the system loses diversity over time when variance is unsustained. The loss of diversity reduces the capacity of the community to adapt to changing environmental conditions via shifts in species composition, as a flat year-round mean trait shows (Figure 6, grey lines). Under Fixed Variance, size diversity is locked at an arbitrary value. If this value is high enough, the mean size can adapt in response to changes in nutrient availability and grazing regimes (Figure 6). This treatment can be useful for studies focusing only on the size structure of the community but it is otherwise based on an arbitrarily fixed level of diversity and cannot offer meaningful insights, for example about biodiversity and ecosystem functioning relationships.

Trait Diffusion and Immigration show similar results for the mean size but not for the size variance (Figure 6). Since the mechanism of Trait Diffusion depends on reproduction, i.e. gross growth (see equation 23), the highest diversity of the community is reached in spring under high growth rates and declines when moving towards winter. Size diversity also peaks in spring for the case of Immigration because this mechanism is density-dependent (see equation 20), but the variances predicted in autumn and winter are, respectively, lower and higher than those obtained with Trait Diffusion (Figure 6). As highlighted before, the differences between Trait Diffusion and Immigration treatments are associated with their assumptions of endogenous (i.e. phytoplankton gross growth) and exogenous (i.e. density-dependent immigration) sources of biomass and trait values.

### 3.3 Sensitivity to changes in parameter values

We tested the sensitivity of the annual mean in P, $\overline{\mathrm{L}}$, and V to variations of +/- 25 % in parameter values. To quantify this sensitivity, we formulated an index S that accounts for relative changes in model results:

$$\mathrm{S} = \frac{\mathrm{X(p)} - \mathrm{X(p')}}{\mathrm{X(p)}} \cdot 100\,, \tag{27}$$

where X(p) is the result of the state variable X obtained with the standard parameter p and $\mathrm{X(p')}$ is the result of the state variable X obtained with the modified parameter $\mathrm{p'}$=p±25 %.

The four treatments of size variance respond similarly to changes in parameter values (Figure 7). The annual means of all three state variables (P, $\overline{\mathrm{L}}$, and V) are sensitive to changes in the parameters controlling zooplankton grazing (i.e. $\mu_{\mathrm{Z}}$,$\mathrm{m_Z}$,$\mathrm{K_P}$,$\delta_{\mathrm{Z}}$). However, P also shows a sensitive response to parameters affecting phytoplankton gross growth, such as $\mathrm{k_w}$,$\mathrm{I_s}$,$\mu_{\mathrm{P}}$, and $\mathrm{m_P}$. Mean size is the most robust variable with less than 10 % relative change compared to the standard run. The size variance treatments for Immigration and Trait Diffusion are affected by the parameters controlling the input of exogenous (i.e. $\delta_\mathrm{I}$ for Immigration) or endogenous variance (i.e. $\nu$, for Trait Diffusion). The results of the Unsustained Variance model are very sensitive to changes in $\mu_\mathrm{Z}$ and the case of Fixed Variance shows a sensitivity that is similar to the other cases, except for the variance itself.

### 3.4 Computational efficiency

Trait-based models that aim at resolving the complexity of natural communities by incorporating many different species or functional types can be expensive in terms of computational time (Baird and Suthers, 2007; Follows et al., 2007; Bruggeman, 2007; Banas, 2011; Ward, 2012). Alternatively, trait-based models that focus on aggregate community properties such as total biomass, mean trait, and trait variance can be more computationally efficient. In table 2 we report a quantification of the computation time required for calculating the full and aggregate models presented here. We obtained more than 10-fold longer computation time for the full model than for the aggregate model. In addition, when we increase the resolution of the full model from 10 to 100 morphotypes, the difference in computation time increases by more than 20-fold. Thus increasing the realism, in terms of number of species or morphotypes comes at a significant computational cost.

## 4 Strength and weakness of moment-based approximations

Models are simplifications of reality and, as such, are based on assumptions. For example, the simple exponential growth model is based on a number of assumptions that do not hold in all circumstances (many factors affect the intrinsic growth rate, which is often not time-invariant, not all individuals within a population are identical, nothing can grow indefinitely, etc.). However, this model can be used and, in fact, it is used within its range of validity. Likewise, the approximation of full models with moment-based approaches requires an assumption about the shape of the phytoplankton trait distribution (Wirtz and Eckhardt, 1996; Norberg et al., 2001; Merico et al., 2009). Typically, unimodal distributions, e.g. normal or log-normal, are assumed. However, depending on how the fitness function (i.e. the net growth rate of the phytoplankton community) is constructed and parameterised, the value of $\mathrm{f}^{(2)}$, that is the rate of change of the variance (Equations 14, 17, 20 23), can be positive, implying a disruptive selection or branching. This represents an indication that the unimodality assumption does not hold (Bonachela et al., 2015). Alternatively, $\mathrm{f}^{(2)}$ can remain negative over time, implying that the community continually loses variance, thus constituting a strong indication against the occurrence of disruptive selection. Therefore, models based on moment approximations require careful checks about the validity of the unimodality assumption throughout the time of the simulations. Figure 8 shows, for our test case, the predicted variance V, $\mathrm{f}^{(2)}$ and its components for the four variance treatments. In our test case, $\mathrm{f}^{(2)}$ is negative for all treatments and its changes are jointly driven by bottom-up ($\mathrm{f}^{(2)}\,\mathrm{U}(\overline{\mathrm{L}},\mathrm{N})$) and top-down processes ($\mathrm{f}^{(2)}\,\mathrm{G}(\overline{\mathrm{L}},\mathrm{P})$), i.e. the second derivatives with respect to trait for nutrient uptake (equation 6) and grazing (equation 7) terms. Sinking plays a role mainly during spring, but its influence is minor compared to the effects of nutrient uptake and grazing.

It is unclear whether unimodality in size distributions is a robust feature in the oceans. Observational evidence from recent work (Downing et al., 2014) suggests that at large temporal scales, i.e. from 5 to 20 years, unimodality of size distributions is a consistent feature of phytoplankton communities of the North Sea. By contrast, multimodality is typically observed on temporal scales of less than one year (Downing et al., 2014). We believe that the observational evidence available in favour or against this case is still insufficient. However, the ocean is a highly variable environment and it is more likely that multimodality, for example because of size-selective grazing events, is a short-term, transient situation rather than the norm, because mixing would continuously reshuffle plankton assemblages and restore homogeneous conditions.

## 5 Concluding remarks

Biological communities are complex adaptive systems (Levin, 1998) characterised by many components and interconnections that lead to emergent properties and nonlinear responses. Models help us to formalise and simplify the complexity we observe in nature. This simplification allows us to render natural phenomena treatable and testable (Levins, 1966; Anderson, 2005, 2010). Over time, however, phytoplankton models have grown more complex, computationally more complicated, and often unavailable to the wider scientific community, aspects that can all hamper advancements in the field. To help reverse this trend we developed PhytoSFDM as a tool to promote the use of trait-based models (whether species-explicit or aggregate models) in marine ecosystems.

A key decision in modelling is choosing an appropriate level of detail for the problem at hand. For example, a species-explicit model offers obvious advantages, which aggregate models cannot offer, when the interest lies in understanding the relative importance of particular species in providing certain ecological services or in quantifying the effect of disruptive selection. Aggregate models, instead, can be more useful at a higher level of abstraction, when the interest lies on macroecological properties. In addition, as we have shown, aggregate models present an advantage with respect to computation time when compared to full models. The advantages in terms of reducing complexity and computation time remain unproven in spatially explicit settings (e.g. in 1D and 3D), although preliminary applications have shown promising results (Bruggeman, 2009).

PhytoSFDM provides a set of methods, under the open source concept, to quantify macroecological properties of phytoplankton communities, as an alternative to the traditional discrete, species-explicit approach. This effort, we hope, will foster our understanding about the role that biodiversity plays in shaping marine ecosystems.

## 6 Code availability

PhytoSFDM is written in Python (version 2.7.x) as a lightweight and user-friendly package to facilitate use and re-distribution. We provide PhytoSFDM as free software under the GNU General Public License version 2. The python package is hosted in: a) GitHUB (https://github.com/SEGGroup/PhytoSFDM) a software repository that allows for version control, b) Zenodo (https://zenodo.org/record/49849) an open scientific repository, and c) PyPI (https://pypi.python.org/pypi/PhytoSFDM) one of the most popular python package repository. To be able to install and operate the package, the user should be familiar with the Python language and should have a running python distribution (preferably version 2.7.x) that includes the latest versions of the libraries *pip* and *setuptools*. Additional required libraries are *matplotlib*, *numpy*, *scipy* and *sympy*. PhytoSFDM can then be conveniently installed by typing the following command from a terminal window:

```
$ pip install PhytoSFDM
```

or by downloading the tarball from the GitHub repository. This is installed using the source file *setup.py* contained in the PhytoSFDM folder by typing:

```
$ python setup.py install
```

The package consists of three main modules: *Example*, *SizeModels*, and *EnvForcing*. *Example* is the entry point: it computes and compares full and aggregate models with the four treatments of variance (unsustained, fixed, trait diffusion, and immigration) at the testing location in the north Atlantic Ocean (centred at 47.5° N and 15.5° W). The example is run from a terminal by typing:

```
$ PhytoSFDM_example
```

or from an interactive python shell by typing:

```
>>> import phytosfdm.Example.example as exmp
>>> exmp.main()
```

The module *SizeModels*, contains the model variants. Here the user can a) modify the default parameters, b) symbolically solve the derivatives with respect to the trait, and c) log-transform mean trait and trait variance. To run the model at a specific location in an interactive python shell one should type:

```
>>> from phytosfdm.SizeModels.sizemodels import SM
>>> Lat=47.5
>>> Lon=344.5
>>> RBB=2.5
>>> SM1= SM(Lat,Lon,RBB,"Imm")
```

In the above example, the model is executed at a location in the North Atlantic Ocean centred at 47.5°N and 15.5°W (here transformed to a scale of 0° to 360°). RBB specify the range of the bounding box (in degrees) for averaging the environmental forcing variables. The fourth argument SM1 is an object that contains the call of the function SM, which runs the size model at the specified location and with the desired treatment for the size variance, in this case Immigration.

The last module, *EnvForcing*, consists of a class containing spatially averaged forcing data. The climatological data have monthly resolution but we include a method to interpolate the data to a daily time step. Spatially averaged and temporally interpolated forcing at a specific location can be extracted by typing:

```
>>> MLD=ExtractEnvFor(Lat,Lon,RBB,'mld')
```

Additional information on the usage of the package is contained in the Readme file.

*Acknowledgements.* We would like to thank Jorn Bruggeman for his support on earlier versions of the model and for his suggestions while we were preparing the draft of this manuscript. EA-T and AM are supported by the German Research Foundation (DFG), through the priority programme DynaTrait (DFG-Schwerpunktprogramm 1704, subproject 19). SLS received support from the Japan Science and Technology Agency (JST) through a CREST Project.

## References

Acevedo-Trejos, E., Brandt, G., Bruggeman, J. and Merico, A.: Mechanisms shaping phytoplankton community structure and diversity in the ocean. Sci. Rep., 5, 8918, 2015.

Andersen, K. H., Berge, T., Gon̨alves, R. J., Hartvig, M., Heuschele, J., Hylander, S., Jacobsen, N. S., Lindemann, C., Martens, E. A., Neuheimer, A. B., Olsson, K., Palacz, A., Prowe, F., Sainmont, J., Traving, S. J., Visser, A. W., Wadhwa, N. and Kiørboe, T.: Characteristic Sizes of Life in the Oceans, from Bacteria to Whales. Ann. Rev. Mar. Sci. 8, 1-25, 2015.

Anderson, T. R.: Plankton functional type modelling: running before we can walk? J. Plankton Res. 27, 2005.

Anderson, T. R.: Progress in marine ecosystem modelling and the unreasonable effectiveness of mathematics. J. Mar. Syst. 81, 4-11, 2010.

Baird, M. E. and Suthers, I. M.: A size-resolved pelagic ecosystem model. Ecol. Modell. 203, 185-203, 2007.

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

Barton, A. D., Dutkiewicz, S., Flierl, G., Bragg, J. G. and Follows, M. J.: Patterns of diversity in marine phytoplankton. Science 327, 1509-11, 2010.

Bonachela, J. A., Klausmeier, C. A., Edwards, K. F., Litchman, E. and Levin, S. A.: The role of phytoplankton diversity in the emergent oceanic stoichiometry. J. Plankton Res., 2015.

Bruggeman, J.: Succession in plankton communities: A trait-based perspective. 2009.

Bruggeman, J. and Kooijman, S. A. L. M.: A biodiversity-inspired approach to aquatic ecosystem modeling. Limnol. Oceanogr. 52, 1533-1544, 2007.

Chisholm, S. W.: Phytoplankton Size in Primary productivity and biogeochemical cycles in the sea (eds. Falkowski, P. G. and Woodhead, A. D.). Plenum Press, pages 213-237, 1992.

Downing, A. S. et al.: Zooming in on size distribution patterns underlying species coexistence in Baltic Sea phytoplankton. Ecol. Lett. 17, 1219-1227, 2014.

Eppley, R.: Temperature and phytoplankton growth in the sea. Fish. Bull. 70, 1063-1085, 1972

Evans, G. and Parslow, J.: A model of annual plankton cycles. Biol. Oceanogr. 3, 327-347, 1985.

Fasham, M., Ducklow, H. W. and Mckelvie, S. M.: A nitrogen-based model of plankton dynamics in the oceanic mixed layer. J. Mar. Res. 48, 591-639, 1990.

Finkel, Z. V., Beardall, J., Flynn, K., Quigg, A., Rees, T. A. V., and Raven, J. A.: Phytoplankton in a changing world: cell size and elemental stoichiometry. J. Plankton Res. 32, 119-137, 2010.

Follows, M. J. and Dutkiewicz, S.: Modeling diverse communities of marine microbes. Ann. Rev. Mar. Sci. 3, 427-451, 2011.

Follows, M. J., Dutkiewicz, S., Grant, S. and Chisholm, S. W.: Emergent biogeography of microbial communities in a model ocean. Science 315, 1843-1846, 2007.

Fulton, E. A., Smith, A. D. M. and Johnson, C. R.: Effect of complexity on marine ecosystem models. Mar. Ecol. Prog. Ser. 253, 1-16, 2003.

Irigoien, X., Huisman, J. and Harris, R. P.: Global biodiversity patterns of marine phytoplankton and zooplankton. Nature 429, 863-867, 2004.

Hansen, B., Bjørnsen, P. K. and Hansen, P. J.: The size ratio between planktonic predators and their prey. Limnol. Oceanogr. 39, 395-403, 1994.

Hansen, P. J., Bjørnsen, P. K. and Hansen, B. W.: Zooplankton grazing and growth: Scaling within the 2-2,000-$\mu$m body size range. Limnol. Oceanogr. 42, 687-704, 1997.

Hardin, G.: The competitive exclusion principle. Science 131, 1292-7, 1960.

Hood, R. R. et al.: Pelagic functional group modeling: Progress, challenges and prospects. Deep. Res. Part II Top. Stud. Oceanogr. 53, 459-512, 2006.

Kiørboe, T.: Turbulence, phytoplankton cell size, and the structure of pelagic food webs. Adv. Mar. Biol. 29, 1-72, 1993.

Levin, S. A.: Ecosystems and the biosphere as complex adaptive systems. Ecosystems 1, 431-436, 1998.

Levins, R.: The strategy of model building in population biology. Am. Sci. 54, 421-431, 1966.

Litchman, E., Klausmeier, C. A., Schofield, O. and Falkowski, P. G.: The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level. Ecol. Lett. 10, 1170-1181, 2007.

Litchman, E. and Klausmeier, C. A.: Trait-based community ecology of phytoplankton. Annu. Rev. Ecol. Evol. Syst. 39, 615-639, 2008.

Litchman, E., de Tezanos Pinto, P., Klausmeier, C. A., Thomas, M. K. and Yoshiyama, K.: Linking traits to species diversity and community structure in phytoplankton. Hydrobiologia 653, 15-28, 2010.

Norberg, J., Swaney, D. P., Dushoff, J., Lin, J., Casagrandi, R., and Levin, S. A.: Phenotypic diversity and ecosystem functioning in changing environments: a theoretical framework. Proc. Natl. Acad. Sci. 98, 11376-81, 2001.

Marañón, E.: Cell Size as a Key Determinant of Phytoplankton Metabolism and Community Structure. Ann. Rev. Mar. Sci. 7, 1-24, 2015.

Merico, A., Brandt, G., Smith, S. L. and Oliver, M.: Sustaining diversity in trait-based models of phytoplankton communities. Front. Ecol. Evol. 2, 1-8, 2014.

Merico, A., Bruggeman, J. and Wirtz, K.: A trait-based approach for downscaling complexity in plankton ecosystem models. Ecol. Modell. 220, 3001-3010, 2009.

Monterey, G. I. and Levitus, S.: Seasonal variability of the mixed layer depth for the world ocean. U.S. Gov. Printing Office, 1997.

Smith, S. L., Pahlow, M., Merico, A. and Wirtz, K. W.: Optimality-based modeling of planktonic organisms. Limnol. Oceanogr. 56, 2080-2094, 2011.

Smith, S. L. et al.: Flexible phytoplankton functional type (FlexPFT) model: size-scaling of traits and optimal growth. J. Plankton Res., 1-16, 2015.

Steele, J.: Environmental control of photosynthesis in the sea. Limnol. Oceanogr. 7, 137-150, 1962.

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. Limnol. Oceanogr. 59, 1-16, 2014.

Ward, B. A., Dutkiewicz, S., Jahn, O. and Follows, M. J.: A size-structured food-web model for the global ocean. Limnol. Oceanogr. 57, 1877-1891, 2012.

Wirtz, K. W. and Eckhardt, B.: Effective variables in ecosystem models with an application to phytoplankton succession. Ecol. Modell. 92, 33-53, 1996.

Wirtz, K. W.: Mechanistic origins of variability in phytoplankton dynamics: Part I: niche formation revealed by a size-based model. Mar. Biol. 160, 2319-2335, 2013.

Wirtz, K. W. and Sommer, U.: Mechanistic origins of variability in phytoplankton dynamics. Part II: analysis of mesocosm blooms under climate change scenarios. Mar. Biol. 160, 2503-2516, 2013.

## Figures

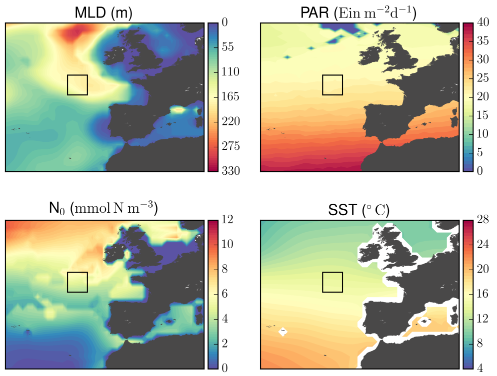

**Figure 1.** Environmental forcing variables considered in PhytoSFDM. The data shown are annual average of mixed-layer depth (MLD), photosynthetic active radiation (PAR), sea surface temperature (SST), and nutrient concentration below the mixed-layer ($N_0$). The square boxes mark the location of the test-case simulation.

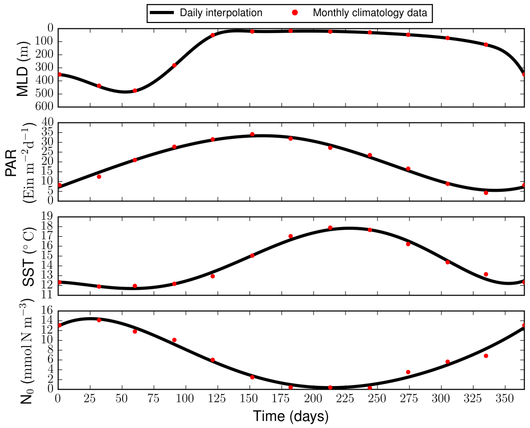

**Figure 2.** Temporal variation of the environmental variables. The monthly climatology data (red dots) are spatially averaged over the test location (square boxes in Figure 1). The interpolation (continuous line) is obtained with a $3^{rd}$ (MLD and PAR) and a $5^{th}$ (SST and $N_0$) order polynomial.

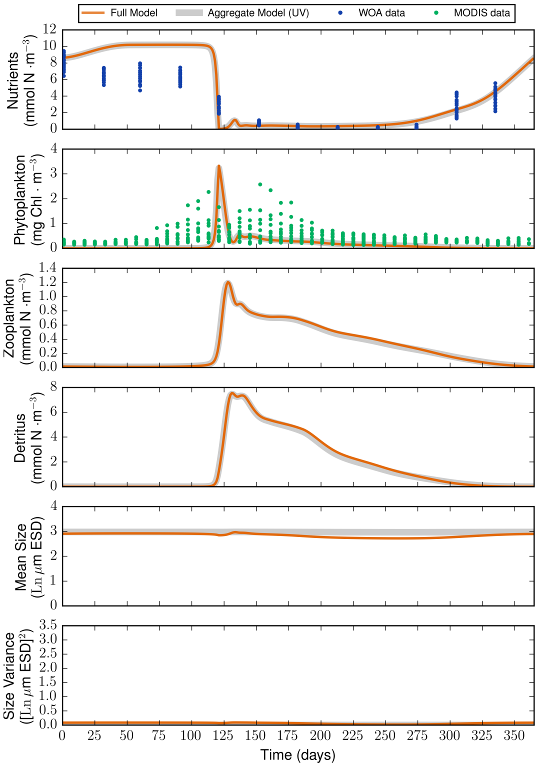

**Figure 3.** NPZD dynamics of the full model (Section 2.1.2 ) and of its equivalent aggregate model (Unsustained Variance, Section 2.1.3) for the last year of the simulations. The total phytoplankton in the full model corresponds to the sum of all $\mathrm{P_n}$. The red dots are nutrient observations (monthly data obtained from the World Ocean Atlas) and the green dots are remotely sensed Chl-a data (8-day composite obtained from MODIS).

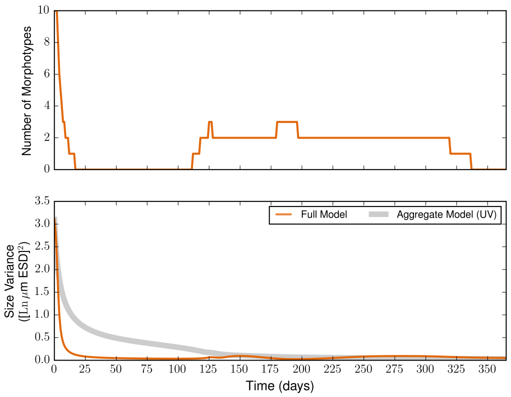

**Figure 4.** Number of morphotypes and size variance over the first year of the simulation. Here we included the morphotypes with a biomass greater than 0.01 $\mathrm{mmol\,N\,m^{-3}}$. Models that do not consider a mechanism to sustain variance exhibit competitive exclusion of morphotypes and a rapid decline of size diversity.

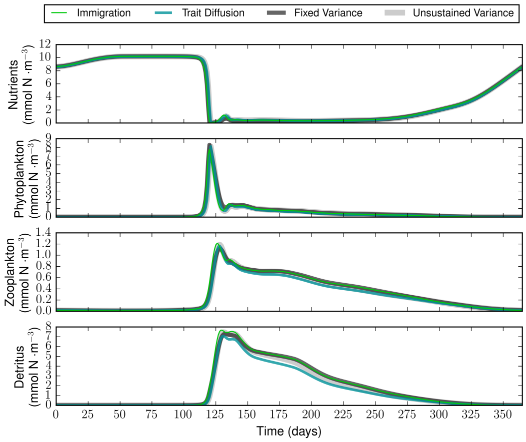

**Figure 5.** Nutrient, Phytoplankton, Zooplankton and Detritus dynamics over a seasonal cycle for the four variants of the aggregate model (see section 2.1.3), named Unsustained and Fixed Variance, Trait Diffusion, and Immigration.

Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-102, 2016
Manuscript under review for journal Geosci. Model Dev.
Published: 30 May 2016

Geoscientific
Model Development
Discussions
Open Access

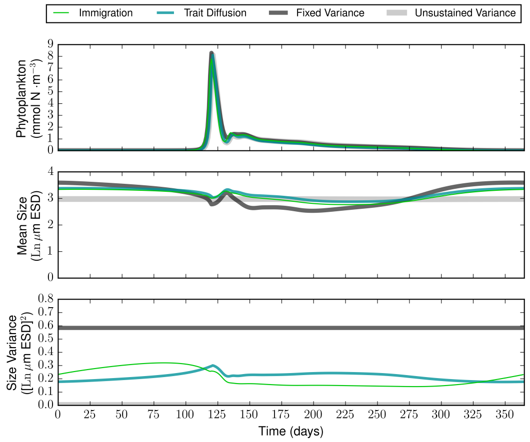

**Figure 6.** Dynamics of the size structured phytoplankton community and its functional size diversity for the four variance treatments (see section 2.1.3), named Unsustained and Fixed Variance, Trait Diffusion, and Immigration.

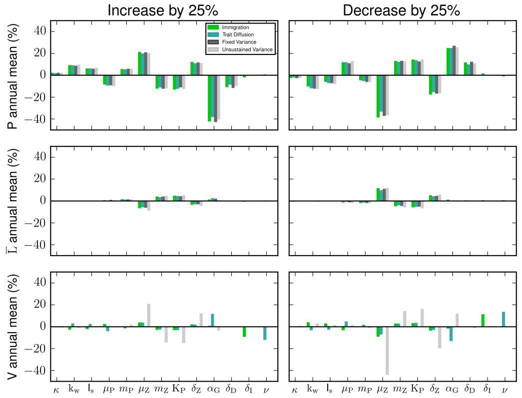

**Figure 7.** Sensitivity of four variance treatments to an increase and a decrease by 25 % in the default parameter values. The values and definitions of all parameters are given in table 1.

Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-102, 2016
Manuscript under review for journal Geosci. Model Dev.
Published: 30 May 2016

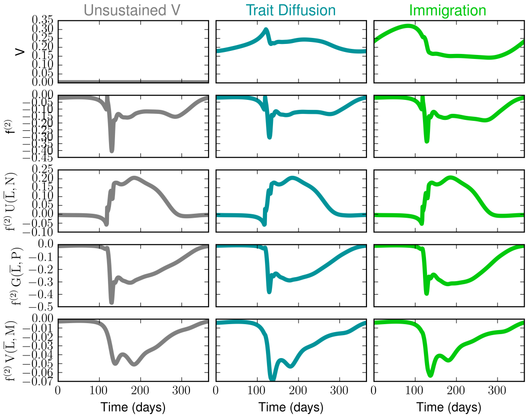

**Figure 8.** Components of the size variance (V), where $\mathrm{f}^{(2)}$ is the second derivative of the fitness function with respect to the trait, $\mathrm{f}^{(2)}\,\mathrm{U}(\overline{\mathrm{L}},\mathrm{N})$, $\mathrm{f}^{(2)}\,\mathrm{G}(\overline{\mathrm{L}},\mathrm{P})$ and $\mathrm{f}^{(2)}\,\mathrm{V}(\overline{\mathrm{L}},\mathrm{M})$ are, respectively, nutrient uptake, zooplankton grazing and phytoplankton sinking components of $\mathrm{f}^{(2)}$.

## Tables

**Table 1.** Parameters definitions, their units and their default values as provided in PhytoSFDM.

| Definition | Symbol(Units) | Value |
|---|---|---|
| Diffusive mixing across the thermocline | $\kappa$ (m·d$^{-1}$) | 0.1 |
| Light attenuation constant | $\mathrm{k_w}$ (m$^{-1}$) | 0.1 |
| Optimum irradiance | $\mathrm{I}_s$ (E m$^{-2}$d$^{-1}$) | 30 |
| P max growth rate | $\mu_\mathrm{P}$ (d$^{-1}$) | 1.5 |
| P mortality rate | $\mathrm{m_P}$ (d$^{-1}$) | 0.05 |
| Z grazing rate | $\mu_\mathrm{Z}$ (d$^{-1}$) | 1.35 |
| Z mortality rate | $\mathrm{m_Z}$ (d$^{-1}$) | 0.3 |
| P half-saturation | $\mathrm{K}_P$ (mmol N m$^{-3}$) | 0.1 |
| P assimilation coefficient | $\delta_\mathrm{Z}$ (-) | 0.31 |
| Mineralization rate | $\delta_\mathrm{D}$ (d$^{-1}$) | 0.1 |
| Immigration rate | $\delta_\mathrm{I}$ (mmol N d$^{-1}$) | 0.008 |
| Trait diffusivity parameter | $\nu$ (-) | 0.008 |
| Slope for allometric grazer preference | $\alpha_G$ ([$\mu$m ESD]$^{-1}$) | -0.75 |
| Intercept of the $\mathrm{K}_N$ allometric function | $\beta_U$ (mmol N m$^{-3}$) | 0.14257 |
| Slope of the $\mathrm{K}_N$ allometric function | $\alpha_U$ (mmol N m$^{-3}$ [$\mu$m ESD]$^{-1}$) | 0.81 |
| Intercept of the V allometric function | $\beta_V$ (m·d$^{-1}$) | 0.01989 |
| Slope of the V allometric function | $\alpha_V$ (m·d$^{-1}$ [$\mu$m ESD]$^{-1}$) | 1.17 |
| Size variance of immigrating P | $\mathrm{V_0}$ (Ln [$\mu$m ESD]$^2$) | 0.58 |

**Table 2.** Computation time in seconds for the full model with 10 and 100 morphotypes and the four variants of the aggregate model.

| System | $\mathrm{Full}_{10}$ | $\mathrm{Full}_{100}$ | Unsustained | Fixed | Trait Diffusion | Immigration |
|---|---|---|---|---|---|---|
| MacOS 2.8 GHz Intel i7 | 282.515 | 6696.503 | 16.012 | 20.374 | 15.177 | 16.644 |
| Windows 3.0 GHz Intel i5 | 369.597 | 8722.753 | 20.236 | 22.485 | 20.736 | 20.298 |