Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Geoscientific Model Development Discussions

CITRATE 1.0: Phytoplankton continuous trait-distribution model with one-dimensional physical transport applied to the Northwest Pacific

Bingzhang Chen, S. Lan Smith

Research Center for Global Change Research, JAMSTEC (Japan Agency for Marine-Earth Science and Technology), 3173-25 Showa-machi, Kanazawa-ku, Yokohama 236-0001, Japan

Correspondence to: Bingzhang Chen (bingzhang.chen@gmail.com)

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Geoscientific Model Development Discussions
Open Access
EGU

Abstract. Diversity plays critical roles in ecosystem functioning, but it remains unclear how best to model phytoplankton diversity in order to better understand those roles and reproduce consistently observed patterns in the ocean. In contrast to the typical approach of resolving distinct species or functional groups, we present a ContInuous TRAiT-basEd phytoplankton model (**CITRATE**) that focuses on macroscopic properties such as total biomass, mean trait values, and trait variance. This phytoplankton component is embedded within a Nitrogen-Phytoplankton-Zooplankton-Detritus-Iron model that itself is coupled with a simplified one-dimensional ocean model. Size is used as the master trait for phytoplankton. CITRATE also incorporates "trait diffusion" for sustaining diversity, as well as simple representations of physiological acclimation, i.e. flexible chlorophyll-to-carbon and nitrogen-to-carbon ratios. We implemented CITRATE 1.0 at two contrasting stations in the Northwest Pacific where several years of observational data are available. The model is driven by physics forcing including vertical eddy diffusivity imported from three-dimensional ocean circulation models. One common set of model parameters for the two stations was optimized using the Delayed Rejection Adaptive Metropolis-Hasting Monte Carlo (DRAM) algorithm. The model faithfully reproduced most of the observational patterns and gave robust predictions on phytoplankton mean size and size diversity. With proper physical forcing, CITRATE 1.0 can be applied to any oceanic station where either nitrogen or iron limits phytoplankton growth.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

1 Introduction

Both species identity and diversity play critical roles in ecosystem functioning (Tilman et al., 1997, 2014). Phytoplankton are a polyphyletic group of oxygenic organisms that account for nearly half of the global primary production (Fields et al., 1998) and also play indispensable roles in other biogeochemical cycles in the Earth system (Falkowski, 2012). They have astonishingly high diversity, with several thousand species already documented and many remaining to be explored (Sournia et al., 1991; Moon-van der Staay et al., 2001). Their sizes range from less than one micron for the cyanobacteria such as *Prochlorococcus* (Chisolmn et al., 1988) to more than 1 mm for some giant diatoms (Villareal, 1993). Furthermore, physiology differs substantially even within the same genera or species (Strzepek and Harrison, 2004; Johnson et al., 2006; Palenik et al., 2006; Kooistra et al., 2008).

Although various ocean models have been developed by accounting for various functional groups or categories of phytoplankton (e.g., Le Quéré et al., 2005; Hashioka et al., 2013), the finite number of such distinct types included limits their ability to resolve the vast diversity of trait values. Some pioneering studies have considered greater numbers of species, each of which has a particular set of multivariate trait axes that constitute a hyper-volume niche space (Follows et al., 2007; Barton et al., 2010; Follows and Dutkiewicz, 2011; Matsuda et al., 2016). “Functional traits” are the key to linking phytoplankton diversity, environmental conditions, and ecosystem functioning. Important phytoplankton traits include maximal growth rate, the light absorption and nutrient uptake affinities, optimal growth temperature, and edibility (i.e., susceptibility to grazing), etc (Litchman et al., 2007; Litchman and Klausmeier, 2008; Edwards et al., 2011, 2012, 2015; Merico et al., 2009; Thomas et al., 2012; Chen, 2015). The total species pool in these modelling studies should ideally cover the entire multi-dimensional trait space of all possible species, which is practically impossible. Although that approach has effectively generated large-scale patterns of plankton diversity, such models generally underestimate local diversity, for two reasons: 1) lack of appropriate mechanisms for sustaining diversity, and 2) insufficient resolution of the trait space so that fitness differences between species are too large to allow coexistence (i.e. insufficient equalizing effect; Chesson, 2000). In any case, a substantial proportion of the idealized species so modelled cannot survive under realistic oceanic conditions, and therefore the models do not capture the functions associated with many species.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Continuous trait-based models have been developed to address the above questions (Wirtz and Eckhardt, 1996; Norberg et al., 2001; Merico et al., 2009, 2014; Terseleer et al., 2014; Acevedo-Trejos et al., 2016; Smith et al. 2016). Instead of modeling the dynamics of individual species, continuous trait-based models or so-called "adaptive dynamics" models focus on macroscopic or aggregate properties of a community such as total biomass, average trait, and trait variance by assuming that phytoplankton traits follow some distribution (usually Gaussian). These models do not have the problem of inadequate trait resolution, because they have infinitesimally fine trait resolution. The trait variance, treated as a tracer in the model, serves as a measure of trait diversity. Thus, the continuous trait-based model has the advantage that the factors controlling diversity can be directly quantified and better understood. In addition, these models are computationally much more efficient than classic species-based models. For example, assuming two independent traits for the phytoplankton community, a continuous trait-based model only requires 1 (biomass) + 2x2 (trait mean and variance) = 5 tracers for the phytoplankton community, while a discrete species-based model requires 2x10 = 20 tracers if assuming ten discrete values in each trait dimension, which still provides only coarse trait resolution. Furthermore, this difference increases exponentially with trait dimension.

Relatively few continuous trait-based models have been coupled with physics transport and validated against oceanic observations. Here we describe a new one-dimensional model, **CITRATE 1.0,** built upon the classic nitrogen-phytoplankton-zooplankton-detritus (NPZD) model with a phytoplankton community represented using a continuous distribution of size, taken as a master trait (Fig. 1). In this way, not only total phytoplankton biomass, but also phytoplankton mean size and size variance are explicitly modeled. The distributions of other important functional traits are implicitly modeled via well-established scaling power laws. For the model to be implemented in the subarctic North Pacific, a well-known high nitrate low chlorophyll (HNLC) region, **CITRATE** also incorporates an iron limitation module. We optimized the model parameters against the extensive observational data at two contrasting stations (K2: 160 ºE, 47 ºN; S1: 145 ºE, 30ºN) in the North Pacific (Fig. 2a). The station K2 is located within the western subarctic North Pacific gyre and is characterized by low temperature, high nitrate, and high carbon export. Iron limitation on phytoplankton growth has been suggested at this station (Fujiki et al., 2014). The station S1 is located within the western subtropical

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Geoscientific Model Development Discussions

North Pacific and is characterized by high sea surface temperature, low levels of nitrate and carbon export efficiency (Matsumoto et al., 2016; Sasai et al., 2016; Wakita et al., 2016).

In the following sections, we first describe the details of the model structure and the parameter optimization subroutine. Then we show the results of parameter optimization and modeled patterns of nutrients, phytoplankton biomass, mean size and size diversity. We also discuss the merits and limitations of the model and of the continuous trait-distribution approach. **CITRATE** 1.0 is intended to be a starting model for later incorporation into three-dimensional (3D) ocean circulation models and for further development of more comprehensive trait-based models.

2 Model description

For simplicity, we assume that phytoplankton size is the master trait that determines all physiological functions (Litchman et al., 2007; Finkel et al., 2010; Edwards et al., 2011, 2012, 2015). In addition to phytoplankton total biomass (***P***), we also include phytoplankton mean log size ($\bar{\boldsymbol{l}}$, ln μm^3) and log size variance ($\boldsymbol{v}$, (ln μm^3)^2) as independent tracers in the model. We use the product of $\boldsymbol{P\bar{l}}$ and $\boldsymbol{P^2v}$, respectively, for the tracers of mean and variance involved in diffusion. Note that we prefer $\boldsymbol{P^2v}$ over $\boldsymbol{Pv}$ because we found that the use of $\boldsymbol{P^2v}$ slightly improved the model fitting. Because in certain parts of the North Pacific iron instead of nitrogen is the main limiting nutrient for phytoplankton growth (Fujiki et al., 2014), we also include dissolved iron (***fer***) as another independent tracer. Hence the model includes 7 tracers in total (Fig. 1).

2.1 Continuous trait-based phytoplankton model

Following the moment closure techniques in Merico et al. (2009, 2014), the differential equations for P, l, and v can be written as:

$$\frac{dP}{dt} \approx P\left(\mu(l) + \frac{v}{2}\left(\frac{d^2\mu(l)}{dl^2} + u\frac{d^4\mu(l)}{dl^4}\right) - 3u\frac{d^3\mu(l)}{dl^3}\right)\Big|_{l=\bar{l}} + \frac{d}{dz}\left(K_v\frac{dP}{dz}\right) \quad (1a)$$

$$\frac{d\bar{l}}{dt} \approx \left[v\left(\frac{d\mu(l)}{dl} + u\frac{d^3\mu(l)}{dl^3}\right) - 3u\frac{d\mu(l)}{dl}\right]\Big|_{l=\bar{l}} \quad (1b)$$

$$\frac{dv}{dt} \approx \left\{v\left[v\left(\frac{d^2\mu(l)}{dl^2} + u\frac{d^4\mu(l)}{dl^4}\right) - 5u\frac{d^2\mu(l)}{dl^2}\right] + 2u\mu(l)\right\}\Big|_{l=\bar{l}} \quad (1c)$$

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Geoscientific Model Development Discussions

where $\mu(\boldsymbol{l})$ is the phytoplankton growth rate (d^{-1}) at mean size $\boldsymbol{l}$, u is the trait diffusion parameter, K_v is the vertical eddy diffusivity ($m^2\ s^{-1}$). Eqs. (1a-c) are approximations because the higher-order moments such as the skewness and kurtosis have been ignored and a Gaussian distribution needs to be assumed for $\boldsymbol{l}$. The differential equations of $P\bar{l}$ and P^2v follow:

$$\frac{d(P\bar{l})}{dt} = P\frac{d\bar{l}}{dt} + \bar{l}\frac{dP}{dt} + \frac{d}{dz}\left(K_v\frac{d(P\bar{l})}{dz}\right) \tag{1d}$$

$$\frac{d(P^2v)}{dt} = P^2\frac{dv}{dt} + 2Pv\frac{dP}{dt} + \frac{d}{dz}\left(K_v\frac{d(P^2v)}{dz}\right) \tag{1e}$$

Phytoplankton growth rate depends on temperature (T, K), light (I, W m^{-2}), dissolved inorganic nitrogen (N, μmol L^{-1}) and iron concentrations (fer, nmol L^{-1}):

$$\mu = \mu_m\frac{N}{N+K_N}\frac{fer}{fer+K_{fer}}\left(1-e^{-\frac{\alpha_I l}{\mu_m}}\right) \tag{2}$$

in which μ_m is a function of T:

$$\mu_m = \mu'_m e^{\frac{E_p}{k}\left(\frac{1}{T_0}-\frac{1}{T}\right)} \tag{3}$$

The parameters μ'_m, K_N, K_{fer}, and α_c are all dependent on cell size:

$$\mu'_m = \mu'_{0,m}e^{\alpha_\mu l+\beta_\mu l^2} \tag{4a}$$

$$K_N = K_{0,N}e^{\alpha_K l} \tag{4b}$$

$$K_{fer} = K_{0,fer}e^{\alpha_{fer} l} \tag{4c}$$

$$\alpha_c = \alpha_{0,c}e^{\alpha_I l} \tag{4d}$$

Eq. (5a) follows that maximal phytoplankton growth rate is a unimodal function of phytoplankton size (Chen et al., 2010, 2011; Marañón et al., 2013).

Following Flynn (2003), we are also able to directly estimate phytoplankton chlorophyll-to-carbon (θ) and nitrogen-to-carbon (Q_N) ratios directly from ambient light and nutrient levels:

$$\theta = \theta_{min} + \frac{\mu}{I\alpha_C}(\theta_{max}-\theta_{min}) \tag{5a}$$

$$Q_N = \frac{Q_{min}}{1-\left(1-\frac{Q_{min}}{Q_{max}}\right)\frac{N}{N+K_N}} \tag{5b}$$

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

where θ_{min} and θ_{max} are minimal and maximal Chl:C ratios, respectively. Q_{min} and Q_{max} are minimal and maximal N:C ratios, respectively. The total Chl *a* concentrations (*Chl*, μg L^{-1}) and net primary production (*NPP*, μgC L^{-1} d^{-1}) can be calculated as:

$$Chl = P\left(\frac{\theta}{Q_N} + \frac{v}{2}\frac{d^2\left(\frac{\theta}{Q_N}\right)}{dl^2}\right)\Bigg|_{l=\bar{l}} \quad (5c)$$

$$NPP = P\left(\frac{\mu}{Q_N} + \frac{v}{2}\frac{d^2\left(\frac{\mu}{Q_N}\right)}{dl^2}\right)\Bigg|_{l=\bar{l}} \quad (5d)$$

2.2 General description of the 1D model (nutrient, zooplankton (*Z*), detritus (*D*), iron (*fer*), and light)

The 1D model focuses on the upper 250 meters of the ocean. The vertical grid, a total of 30 layers, follows a stretched vertical coordinate with increasing resolution towards the sea surface (surface stretching parameter = 2.0), similar to that used in the Regional Ocean Modelling System (ROMS) (Shchepetkin and McWilliams, 2005). For computational efficiency, the 1D model contains only biological tracers. The physics variables are imported as external data (*see* Sect. 2.3 for details).

For simplicity, phytoplankton cells are assumed not to excrete inorganic nitrogen or to have any natural mortality to be converted into detritus. We assume a generic zooplankton compartment that feeds only on phytoplankton following a Holling Type III functional response without any size-dependent feeding preference:

$$Zoo\ ingestion\ per\ capita = g_{max}\frac{P^2}{P^2+K_P^2} \quad (6)$$

where g_{max} is the maximal zooplankton ingestion rate, K_P is the half-saturation constant for zooplankton grazing.

The total amount of phytoplankton ingested by zooplankton is divided among three fates: zooplankton net growth, excretion into the inorganic nitrogen pool, and defecation of unassimilated

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

food into the detritus pool (Buitenhuis et al., 2010). The mortality rate of zooplankton is set to be proportional to the squares of zooplankton biomass. The dynamics of zooplankton follow:

$$\frac{dZ}{dt} = e^{\frac{E_Z}{k}\left(\frac{1}{T_0}-\frac{1}{T}\right)}\left(Zg_{max}\frac{P^2}{P^2+K_P^2}NGE - m_z Z^2\right) + \frac{d}{dz}\left(K_v\frac{dZ}{dz}\right) \quad (7a)$$

where E_z is the activation energy for heterotrophic processes. *NGE* is the net growth efficiency of zooplankton. m_z is the zooplankton mortality coefficient. z is water depth (m).

Detritus is converted to inorganic nitrogen at a rate (R_{dn}, d^{-1}) that has the same temperature sensitivity with zooplankton grazing. Detritus is also assumed to have a constant sinking rate (W_d).

$$\frac{dD}{dt} = e^{\frac{E_Z}{k}\left(\frac{1}{T_0}-\frac{1}{T}\right)}\left(Zg_{max}\frac{P^2}{P^2+K_P^2}unass + m_z Z^2 - -W_d\frac{dD}{dz} + \frac{d}{dz}\left(K_v\frac{dD}{dz}\right)\right. \quad (7b)$$

The dynamics of inorganic nutrients follow:

$$\frac{dN}{dt} = -P\left(\mu(l) + \frac{v}{2}\left(\frac{d^2\mu(l)}{dl^2} + u\frac{d^4\mu(l)}{dl^4}\right) - 3u\frac{d^3\mu(l)}{dl^3}\right)\Big|_{l=\bar{l}} + e^{\frac{E_Z}{k}\left(\frac{1}{T_0}-\frac{1}{T}\right)}\left(Zg_{max}\frac{P^2}{P^2+K_P^2}(1 - NGE - unass) + DR_{dn}\right) + \frac{d}{dz}\left(K_v\frac{dN}{dz}\right) \quad (7c)$$

The dynamics of dissolved iron largely follow Aumont et al. (2003), Nikelsen et al. (2015), and the PlankTOM10 model (Buitenhuis et al., 2010):

$$\frac{dfer}{dt} = \left[-P\left(\mu(l) + \frac{v}{2}\left(\frac{d^2\mu(l)}{dl^2} + u\frac{d^4\mu(l)}{dl^4}\right) - 3u\frac{d^3\mu(l)}{dl^3}\right)\Big|_{l=\bar{l}} + e^{\frac{E_Z}{k}\left(\frac{1}{T_0}-\frac{1}{T}\right)}\left(Zg_{max}\frac{P^2}{P^2+K_P^2}(1 - NGE - unass) + DR_{dn}\right)\right]R_{fer\_N} + \text{dust deposition} - fer_{scav} + \frac{d}{dz}\left(K_v\frac{dfer}{dz}\right) \quad (7d)$$

To translate between nitrogen and iron in particles, a constant *fer*:*N* ratio (R_{fer_N}) of 0.0265 is assumed. Monthly atmospheric dust deposition data are extracted from Tegen and Fung (1995). Following the PlankTOM10 model, iron scavenging rate (fer_{scav}) is composed of both background scavenging rate (k_{scm}) and particle absorption rate (k_{sc}):

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

$$fer_{scav} = k_{scm} + k_{sc} D e^{\frac{E_Z}{k}\left(\frac{1}{T_0}-\frac{1}{T}\right)} \frac{\left(-(1+(l_{fe}-fer)k_{eq})+\sqrt{4ferk_{eq}+(1+(l_{fe}-fer)k_{eq})^2}\right)}{2k_{eq}} \quad (8a)$$

where k_{eq} is the equilibrium constant between free iron and ligands and organic complexes:

$$k_{eq} = 10^{\left(17.27-\frac{1565.7}{T}\right)} \quad (8b)$$

Note that T is in absolute temperature. l_{fe} is the total iron ligand concentration that is assumed constant (0.6 nM).

Light levels (I_z) at depth z were calculated based on PAR_0 and Chl a concentrations following the Beer-Lambert law:

$$I_z = PAR_0 e^{-z\left(K_w + K_{chl}\int_z^0 Chl(x)dx\right)} \quad (9)$$

in which K_w and K_{chl} are the attenuation coefficients for seawater and Chl a, respectively. To realistically estimate the average light field that a phytoplankton cell should experience in a mixing water column (Franks, 2015), the ambient light level for phytoplankton within the surface mixed layer (ML) is calculated as the average light throughout the surface ML, which is defined as the deepest depth with $K_v > 10^{-3}$ m^2 s^{-1}. This calculation is based on eq. (1) in Franks (2015), which gives that the average time for a phytoplankton cell to move 100 m at the local diffusivity of 10^{-3} m^2 s^{-1} is roughly half a day. However, to compare with *in situ NPP* estimates that were calculated from incubation bottles without continuous mixing, phytoplankton μ, θ, and Q_N are recalculated from I_z based on the Beer-Lambert law.

The initial condition of inorganic nitrogen is set to the vertical profile of nitrate in January of the World Ocean Atlas (WOA) 2013 monthly climatology. Initial phytoplankton, zooplankton, and detritus biomass are all set to 0.1 μmol L^{-1} in each grid. Initial phytoplankton mean log size ($\bar{l}$) and log size variance (v) are set to be 1. Initial dissolved iron concentration is set to the vertical profile of iron in January from a 3D global biogeochemical model output (Aumont et al., 2003). The time step of the model is 10 minutes. All the fixed model parameters are shown in Table 1 and the model parameters that are optimized to match observational data are shown in Table 2.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

2.3 External physics data for one-dimensional (1D) model

Four types of external physics forcing data were imported into the 1D model: vertical eddy diffusivity (K_v), surface photosynthetic available radiation (PAR_0), atmospheric dust deposition, and vertical temperature profiles. Vertical advection of water was neglected, which had been shown relatively unimportant (Fernández-Castro et al., 2016). The most important physics forcing data, K_v, determined the upward nutrient flux to the upper euphotic zone and were imported from the output of a three dimensional (3D) eddy-permitting model targeted for North Pacific (Hashioka et al., 2009). This 3D model was able to faithfully simulate the Kuroshio Current and the spatial distributions of the Chl *a* fields. The extracted vertical profiles of K_v were also consistent with the *in situ* estimated mixed layer depths (MLD) at the two stations (Fig. 2). Practically, any reasonable outputs of seasonal K_v profiles can be used. PAR_0 were imported from SeaWIFS satellite monthly climatology products. Seasonal temperature vertical profiles were imported from WOA2013 monthly climatology.

2.4 Delayed Rejection Adaptive Metropolis-Hasting Monte Carlo (DRAM) algorithm

The DRAM algorithm, built upon the classic Metropolis-Hasting Monte Carlo (MHMC) algorithm, incorporates the merits of both adaptive and delayed-rejection MHMC algorithm to increase the acceptance rate and thus more efficiently find the target distribution of parameter values (Haario et al., 2006; Laine, 2008). The adaptive component uses information from previous model runs to approximate the distribution of parameter values as a multivariate Gaussian, which it periodically updates as more simulations are conducted. Based on this, the algorithm alters the magnitude and direction of proposed 'jumps' in parameter space (i.e., the proposal covariance matrix (P_{cvm})) in order to more efficiently explore the parameter space. Specifically, the P_{cvm} is tuned based on the covariance matrix (C_{vm}) of the already accepted parameter sets after a fixed number of iterations following Gelman et al. (2014) (i.e. $P_{cvm} = C_{vm} \bullet 2.4^2/d$, where d is the dimension of the target parameter vector).

With the delayed rejection MCMC, when a newly proposed set of parameters is rejected, P_{cvm} is temporarily downscaled (to 1% of the original P_{cvm} in our case) and a second set of parameters is proposed based on the rejected parameters and the downscaled P_{cvm}. DRAM has the advantage of speeding up the mixing properties of MHMC and has been shown to better explore the parameter space

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

than some other algorithms such as the families of Simulated Annealing possibly due to the use of two-stage proposal covariance matrices (Villagran et al., 2008). Compared with the widely used ensemble Kalman filter, DRAM is perhaps more suitable for the nonlinear ecosystems (Annan and Hargreaves, 2007).

Preliminary model runs suggested that from the third year, the model reached a quasi-steady state, exhibiting regular seasonal cycles under the climatological forcing (Fig. 3). As such, we ran the model for four years and the output of the final year was used for validation against observational data. The model outputs were linearly interpolated to the observational depths and both model. To allow fair comparisons among different data types and downplay the effects of extreme values, both the model outputs and observational data were transformed to their 1/4 power and normalized between 0 and 1 to achieve a quasi-normal distribution before calculating sum of squared errors (*SSqE*):

$$SSqE_{k,i} = \sum_{j=1}^{n_{k,i}} \left(\frac{m_{k,i,j}^{0.25} - o_{k,i,min}^{0.25}}{o_{k,i,max}^{0.25} - o_{k,i,min}^{0.25}} - \frac{o_{k,i,j}^{0.25} - o_{k,i,min}^{0.25}}{o_{k,i,max}^{0.25} - o_{k,i,min}^{0.25}} \right)^2 \quad (10)$$

where $SSqE_{k,i}$ is the sum of squared errors of data type i at station k. $n_{k,i}$ is the number of observations for data type i at station k. $o_{k,i,j}$ is the observed j^{th} value for data type i at station k. $o_{k,i,min}$ and $o_{k,i,max}$ are minimal and maximal observed values for data type i at station k, respectively (Note that for all size-fractions of Chl a, we intentionally set $o_{k,i,min} = 0$ and $o_{k,i,max} = 1$ to minimize the effects of the large measurement variability). $m_{k,i,j}$ is the value linearly interpolated from model outputs to the same depth and date of $o_{k,i,j}$.

Following Laine (2008), the likelihood function for the eight types of observational data (total dissolved inorganic nitrogen (TIN), total Chl a concentration, *NPP*, particulate organic nitrogen (PON), and the fractions of four size-fractionated (> 10 µm, 3–10 µm, 1–3 µm, <1 µm) Chl a; Table 3) at two stations is calculated as:

$$p = \prod_{k=1}^{2} \left[\prod_{i=1}^{8} (2\pi)^{-\frac{n_{k,i}}{2}} \sigma_{k,i}^{-n_{k,i}} e^{-\frac{SSqE_{k,i}}{2\sigma_{k,i}^2}} \right] \quad (11)$$

in which $\sigma_{k,i}$ is the standard deviation of the Gaussian errors of data type i at station k.

Following Laine (2008), we assume that the prior of $1/\sigma_{k,i}$ follows a gamma distribution, with the prior mean as S_0^2 and prior accuracy as n_0. At each step the value of $1/\sigma_{k,i}$ is sampled from a conditional

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

gamma distribution $\Gamma\left(\frac{n_0+n_{k,i}}{2}, \frac{n_0 S_0^2+SSqE_{k,i}}{2}\right)$. The model parameters are assumed to follow multivariate normal distributions. The likelihood function contributed by the priors of the parameters is:

$$p_{pri} = (2\pi)^{-\frac{n_p}{2}}\left(\prod_{i=1}^{n_p}\eta_i^{-1}\right)e^{-\sum_{i=1}^{n_p}\left(\frac{\theta_i-\gamma_i}{\eta_i}\right)^2} \quad (12)$$

in which n_p is the number of parameters to be estimated, γ_i and η_i are the prior estimates of the i^{th} parameter and its standard deviation, respectively (Table 2). Values of η_i were calculated as one-sixth of the difference between the preset maximal and minimal parameter boundaries. θ_i is the current parameter value. The MCMC chain was run for an ensemble of 10000 simulations.

2.5 **Observational data**

The observational data including MLD and eight types of data used in DRAM (*see* Sect. 2.4) were obtained from the K2S1 project (https://ebcrpa.jamstec.go.jp/k2s1/en/index.html; Honda, 2016). The observations spanned from 2010 to 2013. All the data were pooled together to generate a quasi-climatological seasonal pattern and inter-annual variations were treated as random noise. Part of the data have been published in Wakita et al. (2016), Fujiki et al. (2016), Matsumoto et al. (2016), and Sasai et al. (2016). To improve data coverage, we also included the nitrate data of World Ocean Atlas (WOA) 2013 for observed TIN. MLD was calculated as the first depth that the seawater density exceeds surface density by 0.125 kg m^{-3} (Shigemitsu et al., 2012).

3 Results

3.1 External physics forcing

The validity of external physics forcing data, particularly vertical mixing that determines upward nutrient diffusive supply to the surface mixed layer, is essential for correct results and parameter optimization with the ecosystem model. Here we show in Fig. 2 a representative year of seasonal variations of K_v, temperature, surface PAR, and dust deposition. Vigorous winter mixing preceded summer water column stratification at both stations, with the model estimates of mixed layer depths consistent with those measured from *in situ* temperature and salinity profiles. Water temperatures and surface PAR values at the subarctic station K2 were significantly lower than at the subtropical station

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

S1. Compared to station S1, the station K2 was characterized by a pronounced spring peak of atmospheric dust deposition.

3.2 Optimization efficiency and optimized parameter distributions

As expected, the log-likelihood continued to increase with the number of model runs and reached a plateau after 2000 iterations (Fig. 4). Except that the *SSqE* of NPP and the largest size fraction (> 10 μm) increased with time at S1, all the other model–data mismatches consistently decreased. Comparing the two stations, the model fits to the TIN, Chl, NPP and large size fractions (> 3 μm) of Chl were better at station K2 than S1. The model fits to the small size fractions (< 3 μm) were better at S1 than K2.

Most values of the optimized parameters fell into reasonable ranges (Table 2; Fig. 5). For some of the parameters such as $K_{0,N}$, W_d, $\alpha_{0,C}$, and $K_{0,fer}$, the final optimized value differed substantially from initial estimates, an expected outcome of the algorithm striving to match with the eight different types of observations at both stations with contrasting environments. Note that the optimized u value was much higher than in Acevedo-Trejos et al. (2016), which might be a necessary condition to match with the observed distributions of size-fractionated chlorophyll. We also note that increasing u beyond some threshold (e.g. 0.2) could drive the model to unstable conditions in which the size variance kept increasing.

3.3 Comparison between best model outputs and observation

The best model outputs in terms of the highest likelihood could capture most of the observational patterns quantitatively (Figs. 6–9). At both stations, the model could reproduce the vertical increasing trend of TIN with depth and the higher concentrations of TIN during winter than summer and autumn. It is noteworthy that the model could also successfully reproduce the relatively abundant summer TIN concentrations at surface at station K2 due to the incorporation of iron limitation. At station S1, the depth of the nutricline seems to be underestimated, mostly related to the shallow mixed layer depth from spring to summer (Fig. 2b).

The vertical and seasonal patterns of Chl *a* concentrations, including the deep chlorophyll maximum at the subtropical station S1, could also be well reproduced at both stations. The patterns of NPP at station K2 were also well reproduced. One noticeable problem at station S1 is that the surface peak of NPP during summer and fall could not be well reproduced, probably due to the lack of N_2

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

fixation in the model that is expected to enhance NPP within the surface mixed layer. The model fits of PON were less satisfying at station K2, probably due to the use of a single parameter for the sinking velocity of detritus at both stations.

Validation against observed phytoplankton size data is critical for testing **CITRATE** 1.0 in which phytoplankton size structure is the core component. The model could reproduce most patterns of the proportions of size-fractionated Chl at both stations (Figs. 8, 9). Although with some deviations, the model could show the relative dominance of larger sizes at K2 and smaller sizes at S1. At station S1, the model also reproduced the relatively invariant proportions of size-fractionated Chl with depth during stratified seasons. This is not expected because of a steep vertical gradient of TIN around 100 m, which was expected to enhance the growth of large phytoplankton. The relative constancy of phytoplankton size structure with depth suggests that large phytoplankton are more susceptible to light limitation as previously suggested (Finkel, 2001; Edwards et al., 2015). We also note some deficiencies of the model, including the overestimation of the 3~10 μm fractions at station K2. The model also tended to underestimate >10 μm fractions at station S1. All these problems relate to the assumption of a fixed trait distribution as discussed later.

3.4 Modelled seasonal patterns of nutrients, phytoplankton biomass, mean size, and size diversity

At both stations, TIN concentrations were higher during winter in the surface mixed layer due to more vigorous mixing (Figs. 10, 11). Significant drawdown of TIN occurred in surface water following water column stratification. At station K2, the patterns of dissolved iron concentration mirrored those of TIN, suggesting that the spring peak of dust deposition was counteracted by stimulated phytoplankton growth. By contrast, surface iron concentrations accumulated from late summer to fall due to nitrogen limitation at station S1.

In accordance with the TIN patterns, higher concentrations of Chl *a* were found during winter at station S1, which results from both increased phytoplankton biomass and chlorophyll-to-carbon ratios (Fig. 11). Starting from spring to fall, subsurface maximal layers of Chl *a* formed and progressively deepened with time. By contrast, at station K2, Chl *a* concentrations peaked from spring to early summer and subsurface chlorophyll maximum layers were not evident even during the summer (Fig. 10), suggesting light limitation played a stronger role in limiting phytoplankton growth at K2 than S1.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

At both stations, in spite of the nutrient increases in winter, phytoplankton mean size peaked around spring. This is because wintertime conditions, with high nutrient concentrations, favour large sizes, and phytoplankton mean size continues to increase, i.e. $\frac{d\mu(l)}{dl}$ values are usually positive throughout the winter, until spring. Thus, the spring peak of mean size results from winter mixing that favours large size. This scenario is similar to the spring bloom cycles described by Behrenfeld (2010) that spring bloom is also a result of winter mixing that favours phytoplankton growth.

Light also played a role in affecting size dynamics. At station S1, even though phytoplankton mean size also increased during winter, it increased most dramatically in early spring when light level increased and there were still nutrients left from winter mixing (Fig. 11e,h). The increases in light were contributed by both increases in surface PAR and shallower mixing. Nutrient (dissolved iron in the case of K2) depletion led to negative values of $\frac{d\mu(l)}{dl}$ since late spring or summer at both stations, resulting in the decrease in mean size. In general, the modelled mean sizes were significantly larger at station K2 than S1, mainly due to less severe nutrient limitation.

Modelled phytoplankton size variances largely followed the patterns of nutrients. Similar to mean size, size variance also increased from fall when mixing started to intensify. However, size variance decreased with nutrient depletion appreciably faster than mean size because the rate of change of size variance is a function of the square of size variance, while the rate of change of mean size is only a function of size variance (Eq. 1). This led to a temporal decoupling of mean size and size diversity. We also observed a temporal decoupling of μ and size variance. The growth rate μ peaked during spring and summer at S1 and K2, respectively, due to the improved light condition and high temperature. In other words, although spring or summer blooms can be characterized by high biomass, growth rate, and large size, the size diversity could be higher during winter.

The values of $\frac{d^2\mu(l)}{dl^2}$, a surrogate for the intensity of competition, were always negative in all times at both stations, indicative of the effect of resource competition that reduces diversity. The values of $\frac{d^2\mu(l)}{dl^2}$ were less negative in nutrient-replete wintertime and deep waters, suggesting that enhanced nutrient supply relieved the nutrient competition among phytoplankton species, providing a window for large

species to thrive, and thus increasing diversity. The absolute magnitude of $\frac{d^2\mu(l)}{dl^2}$ positively correlated to μ, indicating that higher growth rates induced greater resource competition. This agrees well with the "dynamic equilibrium theory" proposed by Huston (1979).

The patterns of trait diffusion, calculated as the sum of all components containing u in the right side of Eq. (1c) to sustain diversity, mirrored those of growth rates. From late fall to early spring, because the values of $\frac{d^2\mu(l)}{dl^2}$ were less negative and trait diffusion estimates were also relatively high, phytoplankton size diversity kept increasing. With stratifying water column starting from spring, the values of $\frac{d^2\mu(l)}{dl^2}$ became more negative and trait diffusion became critically important for sustaining diversity in the system by counteracting the effect of competitive exclusion in nutrient depleted waters.

4 Discussion

4.1 Model merits

The most important advantage of the continuous trait-based or the adaptive dynamics approach can be quoted from Bak (1996): "If, following traditional scientific methods, we concentrate on an accurate description of the details, we lose perspective" (p. 10) and also "It is a futile endeavour to try to explain most natural phenomena in detail by starting from particle physics and following the trajectories of all particles" (p. 5). We believe that this merit overweighs other technical advantages such as computational efficiency (Acevedo-Trejos et al., 2016) because the model equations themselves (Eq. 1) have already provided the genuine insights for the mechanisms regulating phytoplankton diversity, size structure and biomass. For example, the second derivative of the growth rate at mean size, $\frac{d^2\mu(l)}{dl^2}$, can be conveniently perceived as a proxy for the intensity of resource competition. In typical NPZD-type models in which phytoplankton species compete for the same set of nutrients but do not directly confront each other, the parameters representing the direct interactions among species (such as in the Lotka–Volterra equations) are usually not available. Thus, in traditional approaches that model the trajectories of a number of species, the competition is not easily quantifiable and therefore is rarely

quantified. The ability to easily quantify competition intensity makes it easy to test ecological theories such as Huston's "general hypothesis of species diversity" (Huston, 1979).

The set of Eq. 1 also provides an excellent platform to investigate the underlying mechanisms for the relationship between biodiversity and ecosystem functioning (productivity in this case), which have been extensively studied (Tilman et al., 2014). One insight from the results of the temporal decoupling of growth rate and size diversity on the seasonal scale is that productivity and diversity are not necessarily correlated because the productivity can strongly depend on *current* environmental condition, while diversity emerges over time, depending on standing stocks, which themselves are contingent upon *historical* events.

The incorporation of trait diffusion originally developed for continuous trait-based models (Merico et al., 2014) also provides a mechanism similar to speciation for sustaining diversity, linking ecological and evolutionary processes (Rosenzweig, 1995). The increasing effect of trait diffusion with growth rate is consistent with evolutionary theory that metabolic rates, closely coupled with growth rates, are expected to correlate with speciation rates (Allen et al., 2006).

By solving these theoretical differential equations under quasi-realistic ocean conditions, we can evaluate the relative roles of each mechanism that structures the community properties mentioned above. Although any model is necessarily an abstraction of the real world, to avoid being too far away from reality, we also considered realistic phytoplankton physiology and optimized model parameters guided by real data. For example, our model has incorporated some features of phytoplankton acclimatory plasticity such as variable Chl:C ratio and N:C ratio. Although, for the sake of simplicity, these variable ratios do not directly influence phytoplankton specific growth rate as in Geider et al. (1997), they are able to reproduce the high Chl:C ratios in the DCM layer, thus providing a more realistic mechanism for the formation of the DCM layer than the models with fixed ratios (Fennel and Boss, 2003). Similarly, the variable N:C ratio also allows phytoplankton cells to achieve higher carbon-based NPP in surface waters than the models with fixed N:C ratios (Christian, 2005). Although cellular chlorophyll and nitrogen quota are not calculated as independent tracers, model comparisons suggest that more complex models do not yield better fits to the data (Flynn, 2003).

Compared to previous continuous trait-based models (Terseleer et al., 2014; Acevedo-Trejos et al., 2015, 2016), **CITRATE** 1.0 does not impose a size-dependent feeding preference of zooplankton to affect the average size of phytoplankton. Instead, we employ an observation-based unimodal relationship between maximal growth rate and size to give the large phytoplankton the advantage under nutrient-replete conditions (Chen and Liu, 2010, 2011; Marañón et al., 2013), thus allowing a tradeoff between nutrient affinity and maximal growth rate. One reason is that, in the open ocean, such size-dependent feeding preference is either very weak or does not have a robust pattern for the dominant grazers, microzooplankton (Hansen and Hansen, 1994; Chen et al., 2009, 2010). While previous studies usually assume that the specific clearance rates of zooplankton on phytoplankton decrease with increasing phytoplankton size, which is usually based on the copepod data, field dilution experiments often suggest that the grazing rates of microzooplankton are higher on fast-growing diatoms (Latasa et al., 1997; Zhou et al., 2015). Because our model incorporates a generic zooplankton compartment, which should have an adaptive feeding behaviour on phytoplankton (Behrenfeld and Boss, 2014), we feel that it is justifiable not to apply any size-related feeding preference in **CITRATE** 1.0. Thus, phytoplankton mean size and size diversity are controlled only by bottom-up factors, although zooplankton grazing plays an indirect role in affecting nutrient regeneration and phytoplankton mortality. Whether or not size-dependent zooplankton grazing plays an important role in determining patterns of phytoplankton mean size and size diversity (Terseleer et al., 2014) deserves to be further investigated in future work.

4.2 Model limitations

4.2.1 Assumption of trait distribution

To facilitate calculation of trait moments, a certain distribution has to be assumed for the trait (Merico et al. 2009; 2014). A lognormal distribution is usually appropriate for phytoplankton size (Finkel, 2007; Terseleer et al., 2014; Smith et al., 2016). However, this does not guarantee that a fixed type of probability distribution can hold for all situations (Coutinho et al., 2016). In oligotrophic waters where picophytoplankton, particularly the unicellular cyanobacteria *Prochlorococcus* and *Synechococcus*, dominate (Campbell et al., 1994; Liu et al., 1997), the distribution of phytoplankton log size is more

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

likely right skewed. In other words, abundances of large species are higher than expected from a pure lognormal distribution, which is consistent with the observation that some large diatoms, with significant contributions to new production, can be found in the oligotrophic gyres (Villareal et al., 1999). This is probably one of the major reason that our model tends to underestimate the fraction of > 10 μm size. This is an inevitable consequence of aggregating the description of the entire community into only the three descriptors (i.e. total biomass, mean and variance), which reduces the degrees of freedom, thus sacrificing detailed accuracy for generality and perspective.

One remedy for this problem might be to assign more functional groups in phytoplankton and assume a lognormal distribution for each group (Terseleer et al., 2014). Having a number of functional groups also circumvents the problem of size-independent functional differences among phytoplankton, such as the differences of maximal growth rates between diatoms and dinoflagellates that may have similar sizes (Chen and Laws, 2017).

4.2.2 Transport of moments

Another potential problem is the transport of trait moments in ocean circulation models. Unlike nutrients or plankton biomass, trait moments are not "concentrations" that can be directly involved in advection and diffusion. The immediate summation of two Gaussian curves with different areas (representing the total biomass), mean, and variance is certainly not another perfect Gaussian curve. However, if we consider that during one short time step, the intruded biomass of the external community from an adjacent grid represents only a minor fraction of the local community, then we can use the products of the biomass and the moments such as $\boldsymbol{P\bar{l}}$ and $\boldsymbol{P^2 v}$ as tracers being transported as other normal tracers to approximate the moments of the mixed community. This approach has already been suggested by Norberg (2004). A schematic diagram is shown in Fig. 12 to represent one example of such approximation, which seems to work well to the first order. Our tentative conclusion is that this approximation should work well in most cases and may potentially be a concern only under the presence of large gradients of phytoplankton biomass or trait moments.

4.2.3 Lack of multiple traits

As a first step, we incorporated only size as the master trait that affects all physiological functions of phytoplankton. In reality, many phytoplankton functional traits, such as optimal temperature, N_2

fixation, and mixotrophy, are independent of size. For example, the optimal growth temperature of phytoplankton is closely related to environmental temperature, but only weakly relates to size (Thomas et al., 2012; Chen, 2015). The optimal growth temperature and irradiances are certainly function traits that deserve to be incorporated into trait-based models (Follows et al., 2007; Norberg, 2004; Edwards et al., 2015) and are expected to strongly affect phytoplankton functional identity and diversity at large scales.

4.2.4 Difficulty in modelling surface peaks of NPP at station S1

The near-surface peak of NPP at the oligotrophic station S1 during summer is not expected if we assume that the source of nutrients comes from below the euphotic zone. Even if variable N:C ratios are used in the model to allow more carbon to be fixed given the same amount of nitrogen near surface waters, surface NPP is still likely to be underestimated even with the presence of N_2 fixation because of phosphorus limitation (Christian, 2005). It is possible that other mechanisms such as vertical migration of phytoplankton need to be taken into account (Villareal et al., 1999; Chavez et al., 2012). Therefore, this problem is not only restricted to **CITRATE** 1.0.

4.2.5 Deficiency of the external physics driver

We note that various aspects of the external environmental drivers such as vertical mixing coefficients (K_v) and iron deposition rates may still have some deficiencies that can bias model results and parameter optimization. For example, it is puzzling that although the MLDs derived from K_v seemed consistent with *in situ* temperature and salinity profiles (Fig. 2b), they were usually shallower than the observed nutricline from spring to fall at station S1. It is also possible that this problem arose because certain aspects of phytoplankton biology such as the varying adaptive response to light (Moore et al., 1998) were not included in the model. For example, we expect that including another functional group adapted to low light environments would deepen the nutricline.

4.3 Future directions

Considering the above limitations, one future direction is to increase the number of traits in the model to generate more realistic phytoplankton diversity patterns, which requires both an “envelope” function relating the maximal growth rate with the optimal trait value and a relationship between growth rate and trait value for each species (Norberg, 2004). Another refinement as noted above is to model

each functional group with a continuous trait-based model, thus mixing the continuous trait-distribution and discrete trait approaches to better capture trait distributions that deviate from normality.

It is relatively easy to couple the one-dimensional **CITRATE** model with 3D global or regional ocean models, thus providing a means to model the large-scale patterns of phytoplankton size and diversity. In particular, by including both trait diffusion and competitive exclusion it may be possible to begin to untangle the relative roles of ecological versus evolutionary processes in shaping global phytoplankton diversity patterns.

5. Conclusions

- We present a 1D model with continuous size distribution for phytoplankton (**CITRATE**). The dynamics of phytoplankton mean size and size variance are directly linked to environmental factors and moments of the size distribution in Eq. (1), facilitating understanding of the underlying mechanisms controlling phytoplankton size and diversity. **CITRATE** 1.0 also incorporates "trait diffusion" as an eco-evolutionary process to sustain phytoplankton diversity.
- We optimized the parameters of **CITRATE** using the DRAM algorithm, which revealed that the model can faithfully reproduce observed seasonal patterns of inorganic nitrogen, Chl *a*, and phytoplankton size structure. The model structure and associated parameters obtained herein can be useful for 3D regional and global ocean models.
- The shortcomings of **CITRATE** include its assumption of a lognormal distribution for phytoplankton size, which to some extent limits the precision with which it can reproduce large size classes of phytoplankton.
- Future continuous trait-based models can be built on **CITRATE** 1.0 to reproduce more realistic patterns of phytoplankton diversity, for example by accounting for multi-dimensional trait distributions.

6. Code and data availability

The code and data of **CITRATE** 1.0 are freely available at: https://github.com/BingzhangChen/NPZDFeCONT.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

6.1 General instruction

The code for **CITRATE** 1.0 is written in Fortran90 with the Intel Fortran compiler used. The running environment is Linux or Unix. The user is supposed to be familiar with the Fortran language.

For each station, four different physical forcing data including vertical profiles of eddy diffusive coefficients and temperatures, surface PAR and atmospheric dust deposition. We already provided the relevant data for stations S1 and K2. The temporal resolution is one day for the vertical eddy diffusivity and one month for three other types of data.

After downloading the folder *FlexEFT1D*, go to the directory *FlexEFT1D/DRAM_0.9/NPZDcont/BOTH* and type *vi run*. The user can change the value of *Test* inside the bash script *run*. The *Test* value being 0 means a fast run, usually for a formal model run for a large number of iterations. The *Test* value being 1 means running a model for debugging, which is much slower than the fast run. The user can also modify the compiler flags depending on the purpose in the script. Having chosen the right compiler option, type *./run*, the model will compile and an executable file (NPZDCONT) will be generated.

The next step is to type *vi Model.nml*, the file containing two namelists. The namelist *&Model* contains the options for station names, the type of ecological model, the type of nutrient uptake function (1 only for **CITRATE**), and the type for grazing function (four different grazing functions including the three Holling type functions and the Ivlev function). The station name determines the right physics files to be read and the filenames for model output. For now we only allow three possible stations: S1, K2, and HOT. Other station names will generate an error. If the user wants to add more station names, the subroutine *Setup_OBSdata* within *MOD_1D.f90* is the place to be modified. A number of ecological models besides CITRATE have been developed. It is beyond the scope of the present study to describe all of them in detail. Just note that the model lists are in the fortran file bio_MOD.f90 and some other details are in choose_model.f90 and MOD_1D.f90.

The namelist *&MCMCrun* contains the options for defining the total length of the MCMC chain which is at least 2, the number of the ensemble runs, the number of days for each model run, whether the model should start from previous runs (*Readfile* = 1) or start a new run (*Readfile* = 0), and the number of runs in the historical files (*enssig* and *enspar*).

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

After defining all the model settings, type "./NPZDCONT" and then the model will run and some outputs will be shown on the screen. Type "./NPZDCONT > out" to make the model outputs stored in the "out" file.

For each model run, the model saves the current parameters into the "enspar" file and the current values of σ and SSqE into the 'enssig' file. In this way, even if the model crashes, the user can pick up the current parameter position and updated parameter covariance matrix. The model also generates the files of best parameters, best σ and SSqE files, best model output files that correspond to observational data, and model output files at daily resolution at each grid after an ensemble run.

6.2 Code structure

All the source files including the makefile are stored in the *src* folder. Here we briefly describe the functions of the most important source files:

- Main.f90: The main program for DRAM that calls each subroutine in serial.

- MOD_1D.f90: The major module that sets up and runs the 1D model. The module also generates model output that matches with the observational data.

- Interface_MOD.f90: the module that initializes the absolute and normalized parameter vectors, the covariance matrix of the parameters, the prior parameter values, and the upper and lower parameter boundaries.

- SUB_MOD.f90: the module that calculates sum of squared errors (SSqE) between model outputs and observational data. This module also contains the I/O subroutines that save the parameters, σ, and SSqE for each iteration. It also contains the major subroutine MCMC_adapt that determines whether to accept new parameters, updates covariance matrix, proposes new parameter vectors and calls the subroutine that runs the 1D model with the newly proposed parameters.

- choose_model.f90: the subroutine that defines the number and indices of tracers and the model outputs that need to be written into the output file.

- NPZD_cont.f90: the major biological subroutine for the CITRATE model.

- bio_MOD.f90: the module that declares most of the model names, indices for model input and output variables and parameters.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Geoscientific Model Development Discussions
Open Access
EGU

Table 1. Fixed parameters including units and values of the **CITRATE** 1.0 model.

Symbol	Description	Value	Unit
K_w	Light attenuation coefficient of seawater	0.04[a]	m^{-1}
K_{chl}	Light attenuation coefficient of chlorophyll	0.025[a]	(mg Chl *a* m^2) $^{-1}$
E_p	Activation energy of phytoplankton rates	0.5[b]	eV
E_z	Activation energy of heterotrophic rates	0.6[b]	eV
θ_{min}	Minimal chlorophyll-to-carbon ratio	0.02[c]	gChl $molC^{-1}$
θ_{max}	Maximal chlorophyll-to-carbon ratio	0.5	gChl $molC^{-1}$
unass	Fraction of unassimilated food by zooplankton	0.24[d]	dimensionless
NGE	Net growth efficiency of zooplankton	0.3[d]	dimensionless
K_p	Grazing half-saturation constant of zooplankton	0.5[e]	μM N
R_{dn}	Conversion rate of detritus to inorganic nitrogen	0.1	d^{-1}
l_{fe}	Total iron ligand concentration	0.6 [d]	nM
K_{scm}	Minimal iron scavenging rate	3 x 10^{-5} [d]	d^{-1}
K_{sc}	Particle dependent scavenging rate	0.03[d]	(μM N)$^{-1}$ d^{-1}
R_{Fe_N}	Plankton iron-to-nitrogen ratio	0.0265	nM:μM

[a]Fennel et al. (2006); [b]Chen and Laws (2017); [c]Flynn (2003); [d]Buitenhuis et al. (2010); [e]Chai et al. (2002).

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Table 2. Parameters optimized by the DRAM algorithm. The values inside the parentheses of the initial values indicate the "hard" boundaries for the parameters. The numbers inside the parentheses of the optimized values indicate the standard deviation after the "burn-in" period (2000 iterations) has been removed.

Symbol	Description	Initial	Optimized	Unit
W_d	Sinking rate of detritus	1 (0.01, 10)	3.3 (0.3)	m d^{-1}
Q_{0N}	Phytoplankton minimal N:C ratio	0.076[a] (0.05, 0.13)	0.093 (0.0007)	mol: mol
g_{max}	Maximal zooplankton grazing rate at 15 ºC	1 (0.5, 2)	0.90 (0.024)	d^{-1}
m_z	Coefficient of mortality rate of zooplankton	0.1 (0.05, 0.3)	0.094 (0.002)	(μM N)$^{-1}$ d^{-1}
$\alpha_{0,c}$	Initial slope of photosynthesis versus light at 1 μm^3	0.055[b] (0.01, 0.1)	0.019 (0.0013)	(W m^{-2})$^{-1}$ d^{-1}
α_I	Size scaling exponent for α_c	0.08 (–0.3, 0.2)	0.097 (0.0017)	dimensionless
$K_{0,N}$	Growth half-saturation constant for nitrogen for a phytoplankton cell of 1 μm^3	0.5 (0.001, 3)	0.89 (0.01)	μM N
α_K	Size scaling exponent for K_N	0.27[c] (0.1, 0.3)	0.24 (0.001)	dimensionless
$\mu'_{0,m}$	Phytoplankton maximal growth rate at 1 μm^3 at 15 ºC	1.2[d] (0.3, 3)	1.34 (0.018)	d^{-1}
α_μ	First-order size scaling component for μ_m	0.25[e] (0.1, 0.4)	0.27 (0.005)	dimensionless
β_μ	Second-order size scaling component for μ_m	–0.025[e] (–0.05, 0)	–0.013 (0.0002)	dimensionless
$K_{0,fer}$	Growth half-saturation constant for iron of phytoplankton with 1 μm^3	0.08[f] (0.04, 0.2)	0.058 (0.001)	nM Fe
α_{fer}	Size scaling exponent for $K_{0,fer}$	0.27[c] (0.1, 0.3)	0.30 (0.001)	dimensionless
u	Trait diffusion parameter	0.08[g] (0.01, 0.12)	0.092 (0.0016)	dimensionless
dustso	Dust iron solubility	0.02[h] (0.01, 0.05)	0.022 (0.0007)	dimensionless

[a]Marañón et al., (2013); [b]Fennel et al., (2006); [c]Ward et al., (2012); [d]Flynn et al., (2016); [e]Chen and Liu, (2010); [f]Gregg et al., (2003); [g]Merico et al., (2014); [h]Aumont et al., (2003).

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Table 3. Observational data at stations S1 and K2. N: number of observations. Min and Max are minimal and maximal values used in data normalization (*see* Sect. 2.4 for details). TIN: dissolved inorganic nitrogen (μmol L^{-1}). Chl *a*: total chlorophyll a concentration (μg L^{-1}). NPP: net primary production measured by ^{13}C uptake (μgC L^{-1} d^{-1}). PON: particulate organic nitrogen (μmol L^{-1}). SF Chl: percentages of four size fractionated Chl *a*.

	K2			S1		
Type	N	Min	Max	N	Min	Max
TIN	974	4.1	45.7	902	0	11.2
Chl *a*	470	0	3.4	426	0	1.0
NPP	112	0.1	37.1	128	0.1	34.9
PON	29	0.1	2.2	32	0.1	1.0
SF Chl	143 x 4	0	1.0	166 x 4	0	1.0

Author contribution

B. Chen and S. L. Smith conceived and designed the study. S. L. Smith wrote the initial MCMC code. B. Chen acquired and organized the observational data, did subsequent coding, and wrote the first draft of the manuscript. Both authors contributed to later revision of the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

We sincerely thank the K2S1 project personnel, particularly M. C. Honda, T. Fujiki, and K. Matsumoto, for sharing the cruise data. We also sincerely thank T. Hashioka for sharing the data of vertical eddy diffusivity. This study is supported by CREST (Grant Number JPMJCR12A3; P.I. SLS) funded by the Japan Science and Technology (JST) Agency and a Grants-in-Aid for Scientific Research (KAKENHI) (Grant Number JP16K21701; P.I. BC) funded by Japan Society for the Promotion of Science (JSPS).

References

Acevedo-Trejos, E., Brandt, G., Bruggeman, J. and Merico, A., Mechanisms shaping size structure and functional diversity of phytoplankton communities in the ocean, Sci. Rep., 5, 8918, doi: 10.1038/srep08918, 2015.

Acevedo-Trejos, E., Brandt, G., Smith, S. L. and Merico, A., PhytoSFDM version 1.0. 0: Phytoplankton Size and Functional Diversity Model, Geosci. Model Dev., 9, 4071–4085, doi: 10.5194/gmd-9-4071-2016, 2016.

Allen, A.P., Gillooly, J.F., Savage, V.M. and Brown, J.H.: Kinetic effects of temperature on rates of genetic divergence and speciation. Proc. Nat. Acad. Sci., 103, 9130-9135, 2006.

Annan, J. D., and Hargreaves, J. C.: Efficient estimation and ensemble generation in climate modelling. Phil. Trans. R. Soc. A, 365, 2077–2088, 2007.

Aumont, O., Maier‐Reimer, E., Blain, S., and Monfray, P.: An ecosystem model of the global ocean including Fe, Si, P colimitations. Global Biogeochem. Cycles, 17, 1060, doi:10.1029/2001GB001745, 2003.

Bak, P.: How nature works: the science of self-organized criticality, New York: Copurnicus, 1996.

Behrenfeld, M. J.: Abandoning Sverdrup's critical depth hypothesis on phytoplankton blooms, Ecology, 91, 977–989, 2010.

Behrenfeld, M. J. and Boss, E. S.: Resurrecting the ecological underpinnings of ocean plankton blooms, Ann. Rev. Mar. Sci., 6, 167–194, 2014.

Buitenhuis, E. T., Rivkin, R. B., Sailley, S., and Le Quéré, C.: Biogeochemical fluxes through microzooplankton, Global Biogeochem. Cycles, 24, GB4015, doi:10.1029/2009GB003601, 2010.

Campbell, L., Nolla, H. A. and Vaulot, D., The importance of *Prochlorococcus* to community structure in the central North Pacific Ocean. Limnol. Oceanogr., 39, 954–961, 1994.

Chai, F., Dugdale, R. C., Peng, T. H., Wilkerson, F. P. and Barber, R. T.: One-dimensional ecosystem model of the equatorial Pacific upwelling system. Part I: model development and silicon and nitrogen cycle, Deep Sea Res. II, 49, 2713–2745, 2002.

Chavez, F. P., Messié, M. and Pennington, J. T.: Marine primary production in relation to climate variability and change, Ann. Rev. Mar. Sci., 3, 227–260, 2011.

Chen, B., and Laws, E. A.: Is there a difference of temperature sensitivity between marine phytoplankton and heterotrophs? Limnol. Oceanogr., 62, 806–817, doi: 10.1002/lno.10462, 2017.

Chen, B., and Liu, H.: Comment: Unimodal relationship between phytoplankton mass–specific growth rate and size: A reply to the comment by Sal and López-Urrutia (2011), Limnol. Oceanogr., 56, 1956–1958, 2011.

Chen, B., and Liu, H.: Relationships between phytoplankton growth and cell size in surface oceans: Interactive effects of temperature, nutrients, and grazing, Limnol. Oceanogr., 55, 965–972, 2010.

Chen, B.: Patterns of thermal limits of phytoplankton, J. Plankton Res., 37, 285–292, 2015.

Chen, B., Liu, H., Landry, M. R., Dai, M., Huang, B. and Sun, J.: Close coupling between phytoplankton growth and microzooplankton grazing in the western South China Sea, Limnol. Oceanog., 54, 1084-1097, 2009.

Chesson, P.: Mechanisms of maintenance of species diversity, Ann. Rev. Ecol. Syst., 31, 343–366, 2000.

Chisholm, S.W., Olson, R. J., Zettler, E. R., Goericke, R., Waterbury, J. B. and Welschmeyer, N.A.: A novel free-living prochlorophyte abundant in the oceanic euphotic zone. Nature, 334, 340–343, 1988.

Christian, J. R.: Biogeochemical cycling in the oligotrophic ocean: Redfield and non-Redfield models, Limnol. Oceanogr., 50, 646–657, 2005.

Coutinho, R.M., Klauschies, T. and Gaedke, U.: Bimodal trait distributions with large variances question the reliability of trait-based aggregate models, Theor. Ecol., 9, 389–408, 2016.

Edwards, K.F., Klausmeier, C.A. and Litchman, E.: Evidence for a three-way trade-off between nitrogen and phosphorus competitive abilities and cell size in phytoplankton. Ecology, 92, 2085–2095, 2011.

Edwards, K. F., Thomas, M. K., Klausmeier, C. A. and Litchman, E.: Light and growth in marine phytoplankton: allometric, taxonomic, and environmental variation, Limnol. Oceanogr., 60, 540–552, 2015.

Edwards, K. F., Thomas, M. K., Klausmeier, C.A. and Litchman, E.: Allometric scaling and taxonomic variation in nutrient utilization traits and maximum growth rate of phytoplankton. Limnol. Oceanogr., 57, 554–566, 2012.

Falkowski, P.: Ocean science: the power of plankton, Nature, 483, S17–S20, 2012.

Fennel, K. and Boss, E.: Subsurface maxima of phytoplankton and chlorophyll: Steady-state solutions from a simple model. Limnol. Oceanogr., 48, 1521–1534, 2003.

Fennel, K., Wilkin, J., Levin, J., Moisan, J., O'Reilly, J. and Haidvogel, D., Nitrogen cycling in the Middle Atlantic Bight: Results from a three-dimensional model and implications for the North Atlantic nitrogen budget, Global Biogeochem. Cycles, 20, GB3007, doi:10.1029/2005GB002456, 2006.

Fernández-Castro, B., Pahlow, M., Mouriño-Carballido, B., Marañón, E. and Oschlies, A.: Optimality-based *Trichodesmium* diazotrophy in the North Atlantic subtropical gyre, J. Plankton Res., 38, 946–963, doi: 10.1093/plankt/fbw047, 2016.

Finkel, Z. V.: Light absorption and size scaling of light-limited metabolism in marine diatoms, Limnol. Oceanogr., 46, 86–94, 2001.

Finkel, Z. V.: Does phytoplankton cell size matter? The evolution of modern marine food webs (pp. 333-350). Boston, Elsevier, 2007.

Finkel, Z. V., Beardall, J., Flynn, K. J., 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.

Flynn, K. J.: Do we need complex mechanistic photoacclimation models for phytoplankton? Limnol. Oceanogr., 48, 2243–2249, 2003.

Flynn, K.J. and Raven, J.A.: What is the limit for photoautotrophic plankton growth rates? J. Plankton Res., doi:10.1093/plankt/fbw067, 2016.

Franks, P.J.: Has Sverdrup's critical depth hypothesis been tested? Mixed layers vs. turbulent layers. ICES J. Mar. Sci., 72, 1897–1907, 2015.

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.

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

Fujiki, T., Matsumoto, K., Mino, Y., Sasaoka, K., Wakita, M., Kawakami, H., Honda, M. C., Watanabe, S., and Saino, T.: Seasonal cycle of phytoplankton community structure and photophysiological state in the western subarctic gyre of the North Pacific. Limnol. Oceanogr., 59, 887–900, 2014.

Fujiki, T., Sasaoka, K., Matsumoto, K., Wakita, M., and Mino, Y.: Seasonal variability of phytoplankton community structure in the subtropical western North Pacific, J. Oceanogr., 72, 343–358, 2016.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B.: Bayesian data analysis (Vol. 2), Boca Raton, FL, USA, Chapman & Hall/CRC, 2014.

Geider, R. J., MacIntyre, H. L. and Kana, T. M.: Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a: carbon ratio to light, nutrient-limitation and temperature. Mar. Ecol. Prog. Ser., 148, 187–200, 1997.

Gregg, W. W., Ginoux, P., Schopf, P. S. and Casey, N.W.: Phytoplankton and iron: validation of a global three-dimensional ocean biogeochemical model, Deep Sea Res. II, 50, 3143–3169, 2003.

Haario, H., Laine, M., Mira, A. and Saksman, E.: DRAM: efficient adaptive MCMC. Stat. Comput., 16, 339–354, 2006.

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

Hashioka, T., Sakamoto, T.T., and Yamanaka, Y.: Potential impact of global warming on North Pacific spring blooms projected by an eddy-permitting 3-D ocean ecosystem model. Geophys. Res. Let., 36, L20604, doi:10.1029/2009GL038912, 2009.

Hashioka, T., Vogt, M., Yamanaka, Y., Le Quere, C., Buitenhuis, E.T., Aita, M.N., Alvain, S., Bopp, L., Hirata, T., Lima, I.D. and Sailley, S. F.: Phytoplankton competition during the spring bloom in four plankton functional type models. Biogeosciences, 10, 6833–6850, doi: 10.5194/bg-10-6833-2013, 2013.

Honda, M. C.: Short introduction to the K2S1 project, J. Oceanogr., 72, 341–342, 2016.

Huston, M.: A general hypothesis of species diversity, Am. Nat., 113, 81–101, 1979.

Johnson, Z. I., Zinser, E. R., Coe, A., McNulty, N. P., Woodward, E. M. S., and Chisholm, S. W.: Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients, Science 311, 1737–1740, 2006.

Kooistra, W. H., Sarno, D., Balzano, S., Gu, H., Andersen, R. A., and Zingone, A.: Global diversity and biogeography of *Skeletonema* species (Bacillariophyta), Protist, 159, 177-193, 2008.

Laine, M.: Adaptive MCMC methods with applications in environmental and geophysical models, Finnish Meteorological Institute, 2008.

Latasa, M., Landry, M. R., Louise, S. and Bidigare, R. R.: Pigment specific growth and grazing rates of phytoplankton in the central equatorial Pacific, Limnol. Oceanogr., 42, 289–298, 1997.

Large, W. G., McWilliams, J. C., and Doney, S. C., Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization, Rev. Geophys., 32, 363–403, 1994.

Le Quéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., et al.: Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. Global Change Biol., 11, 2016–2040, 2005.

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

Liu, H., Nolla, H. A., and Campbell, L.: *Prochlorococcus* growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean, Aquat. Microb. Ecol., 12, 39–47, 1997.

Marañón, E., Cermeño, P., López‐Sandoval, D. C., Rodríguez‐Ramos, T., Sobrino, C., Huete‐Ortega, M., Blanco, J. M., and Rodríguez, J.: Unimodal size scaling of phytoplankton growth and the size dependence of nutrient uptake and use. Ecol. Let. 16, 371–379, 2013.

Masuda, Y., Yamanaka, Y., Hirata, T. and Nakano, H.: Competition and community assemblage dynamics within a phytoplankton functional group: Simulation using an eddy-resolving model to disentangle deterministic and random effects, Ecol. Mod., 343, 1–14, 2017.

Matsumoto, K., Abe, O., Fujiki, T., Sukigara, C. and Mino, Y., Primary productivity at the time-series stations in the northwestern Pacific Ocean: is the subtropical station unproductive? J. Oceanogr., 72, 359–371, 2016.

Merico, A., Brandt, G., Smith, S. L., and Oliver, M.: Sustaining diversity in trait-based models of phytoplankton communities, Front. Ecol. Evol., 2, 59, doi: 10.3389/fevo.2014.00059, 2014.

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

Moon-van der Staay, S. Y., De Wachter, R., and Vaulot, D.: Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity, Nature, 409, 607–610, 2001.

Moore, L. R., Rocap, G., Chisholm, S. W.: Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes, Nature, 393, 464–467, 1998.

Nickelsen, L., Keller, D. and Oschlies, A.: A dynamic marine iron cycle module coupled to the University of Victoria Earth System Model: the Kiel Marine Biogeochemical Model 2 for UVic 2.9, Geosci. Model Dev., 8, 1357–1381, doi:10.5194/gmd-8-1357-2015, 2015.

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. Nat. Acad. Sci., 98, 11376–11381, 2001.

Norberg, J.: Biodiversity and ecosystem functioning: a complex adaptive systems approach, Limnol. Oceanogr., 49, 1269–1277, 2004.

Palenik, B., Ren, Q., Dupont, C. L., Myers, G. S., Heidelberg, J. F., Badger, J. H., Madupu, R., et al.: Genome sequence of *Synechococcus* CC9311: insights into adaptation to a coastal environment, Proc. Nat. Acad. Sci., 103, 13555–13559, 2006.

Rosenzweig, M. L.: Species diversity in space and time, Cambridge University Press, 1995.

Sasai, Y., Yoshikawa, C., Smith, S. L., Hashioka, T., Matsumoto, K., Wakita, M., Sasaoka, K. and Honda, M. C., Coupled 1-D physical–biological model study of phytoplankton production at two contrasting time-series stations in the western North Pacific, J. Oceanogr., 72, 509–526, 2016.

Shchepetkin, A. F., and McWilliams, J. C., The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model, Ocean Mod., 9, 347–404, 2005.

Shigemitsu, M., Okunishi, T., Nishioka, J., Sumata, H., Hashioka, T., Aita, M. N., Smith, S. L., Yoshie, N., Okada, N. and Yamanaka, Y., Development of a one-dimensional ecosystem model including the iron cycle applied to the Oyashio region, western subarctic Pacific, J. Geophys. Res., 117, C06021, doi:10.1029/2011JC007689, 2012.

Smith, S. L., Vallina, S. M., and Merico, A.: Phytoplankton size-diversity mediates an emergent trade-off in ecosystem functioning for rare versus frequent disturbances. Sci. Rep., 6, 34170, doi: 10.1038/srep34170, 2016.

Sournia, A., Chrdtiennot-Dinet, M-J., Ricard, M.: Marine phytoplankton: how many species in the world ocean? J. Plankton Res. 13, 1093–1099, 1991.

Strzepek, R. F., and Harrison, P. J.: Photosynthetic architecture differs in coastal and oceanic diatoms, Nature, 431, 689–692, 2004.

Tegen, I., and Fung, I., Contribution to the atmospheric mineral aerosol load from land surface modification, J. Geophys. Res.: Atmospheres, 100(D9), 18707-18726, 1995.

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, 1958–1972, 2014.

Thomas, M. K., Kremer, C. T., Klausmeier, C. A., and Litchman, E.: A global pattern of thermal adaptation in marine phytoplankton, Science, 338, 1085–1088, 2012.

Tilman, D., Isbell, F., and Cowles, J. M., Biodiversity and ecosystem functioning. Ann. Rev. Ecol. Evol. Syst., 45, 471–493, 2014.

Tilman, D., Lehman, C. L., and Thomson, K. T.: Plant diversity and ecosystem productivity: theoretical considerations. Proc. Nat. Acad. Sci., 94, 1857–1861, 1997.

Villagran, A., Huerta, G., Jackson, C. S. and Sen, M. K.: Computational methods for parameter estimation in climate models, Bayesian Analysis, 3, 823-850, 2008.

Villareal, T. A.: Abundance of the giant diatom *Ethmodiscus* in the southwest Atlantic Ocean and central Pacific gyre, Diatom Res., 8, 171–177, 1993.

Villareal, T.A., Pilskaln, C., Brzezinski, M., Lipschultz, F., Dennett, M. and Gardner, G.B.: Upward transport of oceanic nitrate by migrating diatom mats, Nature, 397, 423–425, 1999.

Wakita, M., Honda, M. C., Matsumoto, K., Fujiki, T., Kawakami, H., Yasunaka, S., Sasai, Y., Sukigara, C., Uchimiya, M., Kitamura, M. and Kobari, T.: Biological organic carbon export estimated from the annual carbon budget observed in the surface waters of the western subarctic and subtropical North Pacific Ocean from 2004 to 2013, J. Oceanogr., 72, 665-685, 2016.

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. Mod., 92, 33–53, 1996.

Zhou, L., Tan, Y., Huang, L. and Li, G.: Does microzooplankton grazing contribute to the pico-phytoplankton dominance in subtropical and tropical oligotrophic waters? Acta Ecol. Sin., 35, 29–38, 2015.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Figure captions

Fig. 1. Schematic description of the CITRATE model. Thick arrows indicate nitrogen flows and dashed lines indicate the simplified iron cycle.

Fig. 2. (a) Locations of the two stations, K2 and S1, overlaid on annual Chl *a* climatology of the North Pacific. (b-e) Seasonal forcing of vertical eddy diffusivity (K_v), temperature, surface PAR, and atmospheric dust deposition, respectively, at station S1. (d-i) The same as (b-e), but for station K2.

Fig. 3. An example of modelled patterns of total inorganic nitrogen (TIN), Chl *a* (Chl), mean size, and ln size variance for four years at stations K2 (a-d) and S1 (e-h).

Fig. 4. (a) Time evolution of log-likelihoods of the MHMC chain. (b-i) Time evolution of sum of squared errors (SSqE) for TIN, Chl, net primary production (NPP), particulate organic nitrogen (PON), and fractions of size-fractionated Chl *a* concentrations of > 10 μm (P10), 3–10 μm (P03), 1–3 μm (P01), and <1 μm (P_1). (j-q) The same as (b-j), but for station S1.

Fig. 5. Time evolution of fitted model parameters.

Fig. 6. Model fittings to vertical profiles of TIN, CHL, NPP, and PON at four seasons at station K2. Black dots represent observational data and red thick solid lines represent the averaged seasonal values predicted by the model. Thin dashed lines represent 95% percentiles of the seasonal data.

Fig. 7. The same as Fig. 6, but for station S1.

Fig. 8. Model fittings for the percentages of the four size fractions of Chl *a* at station K2.

Fig. 9. The same as Fig. 8, but for station S1.

Fig. 10. Modelled seasonal patterns at station K2: (a) TIN, (b) dissolved iron, (c) Chl *a*, (d) phytoplankton biomass in terms of nitrogen (PHY), (e) phytoplankton mean size, (f) ln size variances, (g) community-based specific growth rate, (h) first derivative of phytoplankton growth rate against ln size evaluated at mean size, (i) second derivative of phytoplankton growth rate evaluated at mean size, and (j) trait diffusion effects.

Fig. 11. The same as Fig. 10, but for station S1.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Fig. 12. A schematic diagram for the transport of phytoplankton mean size of two idealized communities, each having a lognormal size distribution. The hypothetical biomass, log mean size, and log size variance are 1, 5, 1 for community A, and 0.2, 4, 0.8 for B, respectively.

5

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

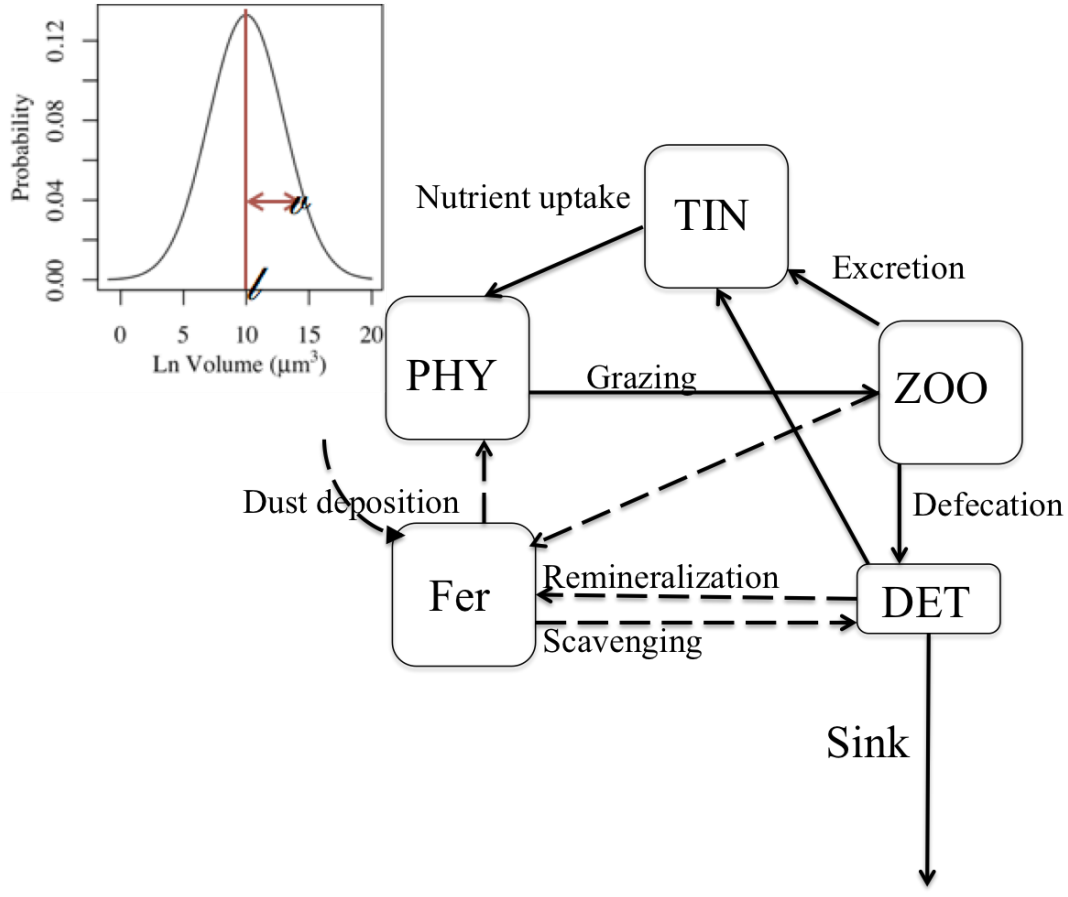

Fig. 1

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

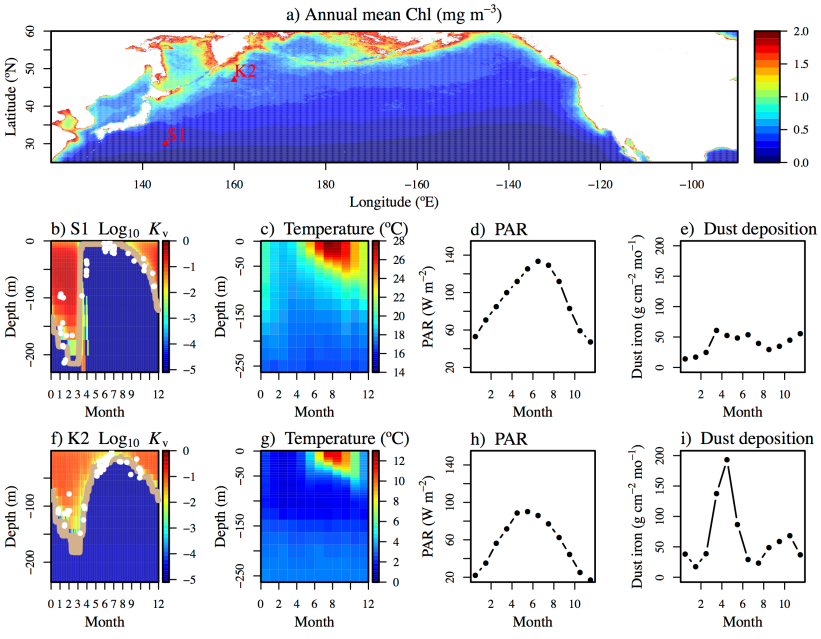

Fig. 2

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

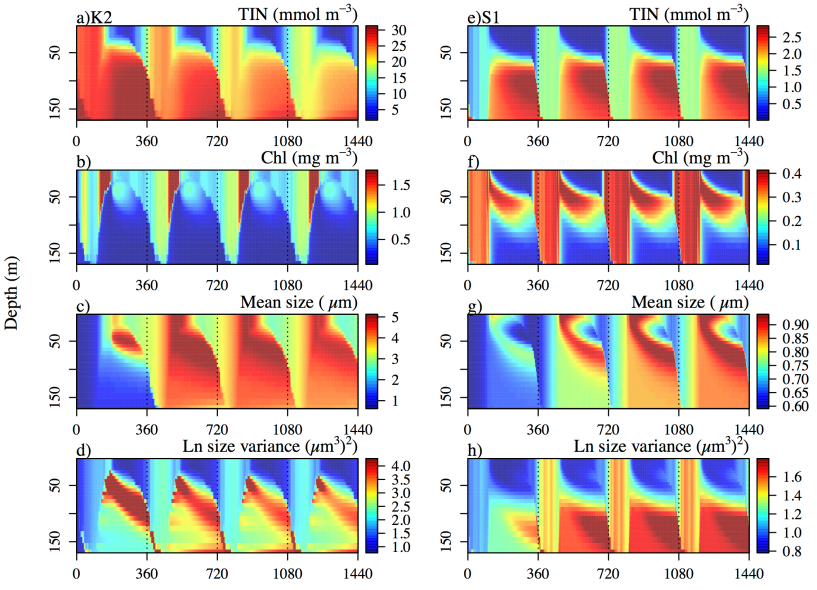

Fig. 3 An example of modelled 4 year patterns at K2 and S1

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Geoscientific Model Development Discussions

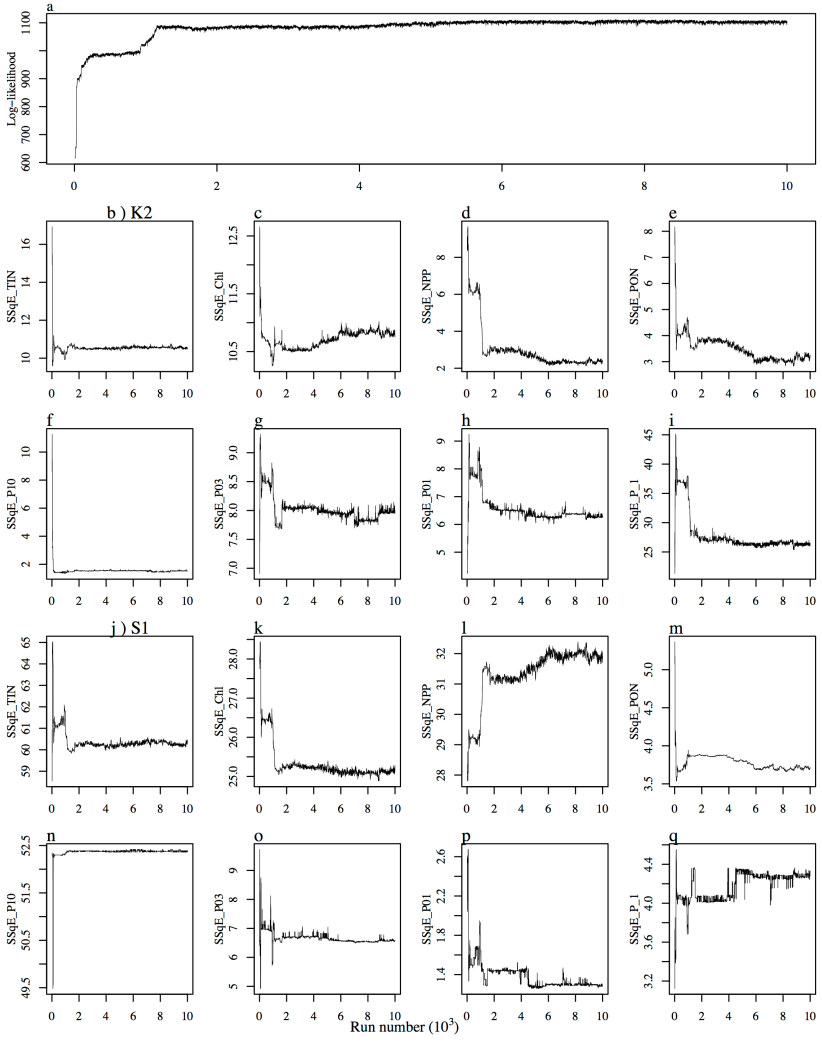

Fig. 4

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

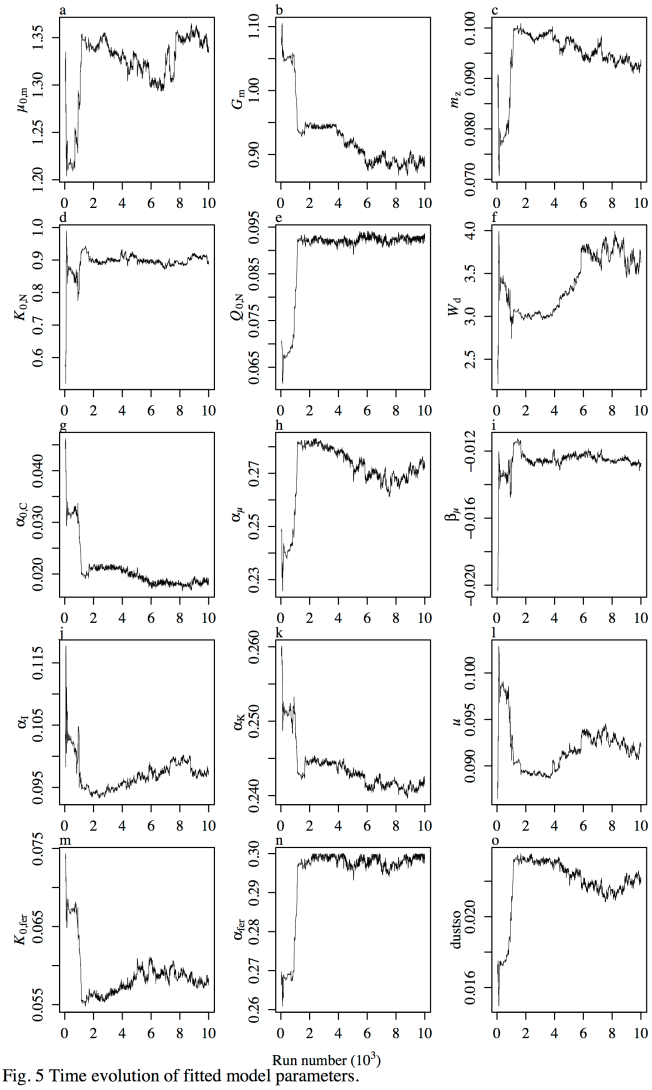

Fig. 5 Time evolution of fitted model parameters.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

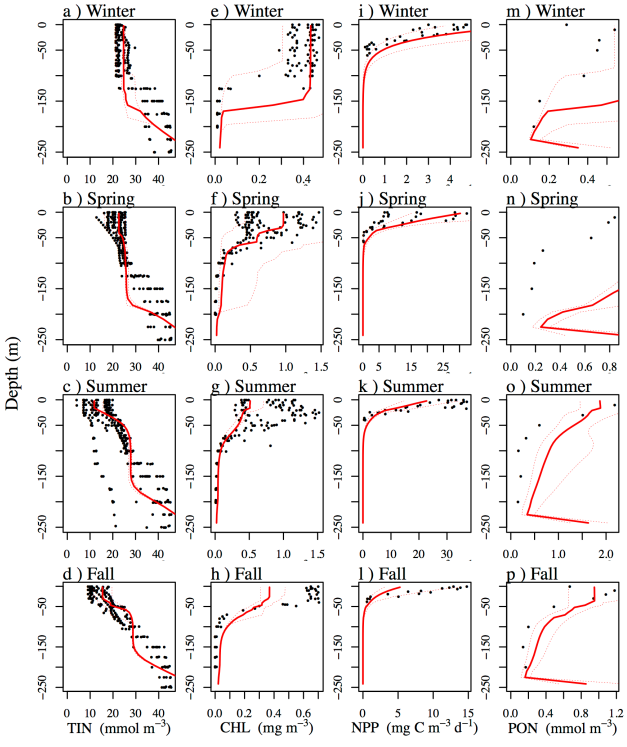

Fig. 6 . Model fittings to vertical profiles of TIN, CHL, NPP, and PON at K2

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

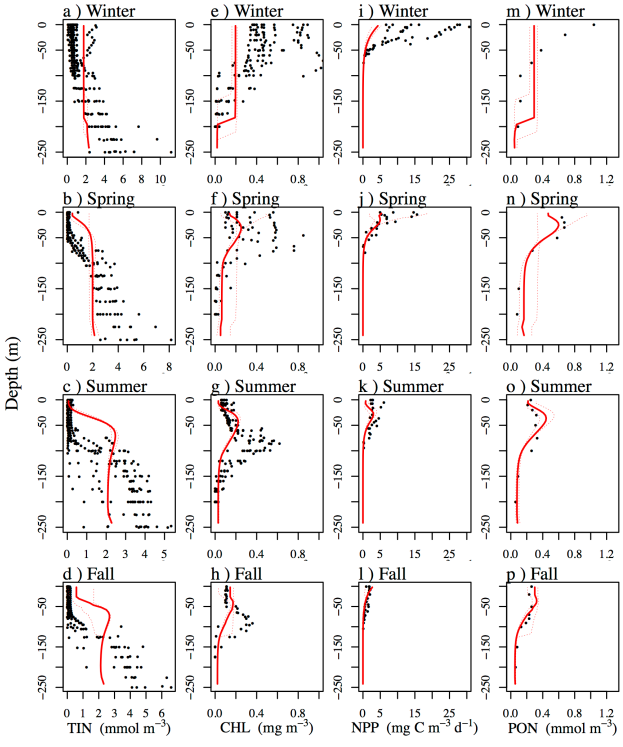

Fig. 7 . Model fittings to vertical profiles of TIN, CHL, NPP, and PON at S1

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

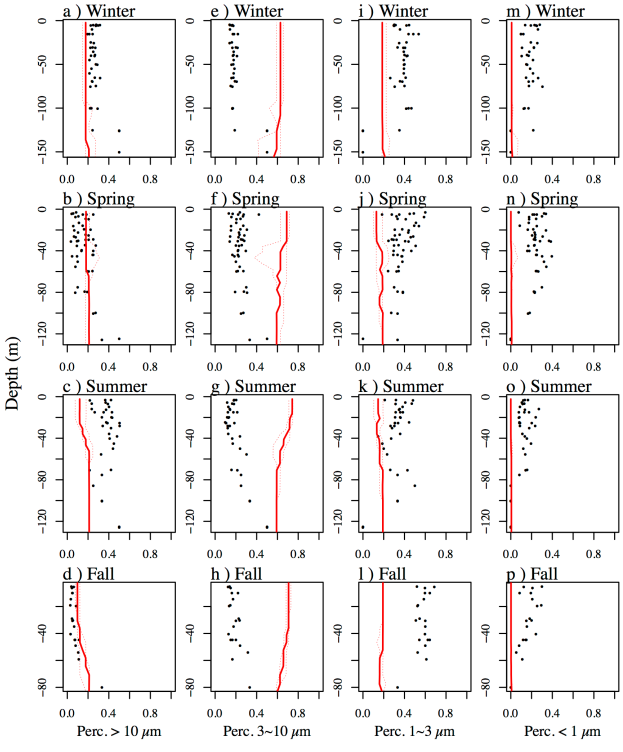

Fig. 8 . Model fittings to vertical profiles of four size fractions at K2

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

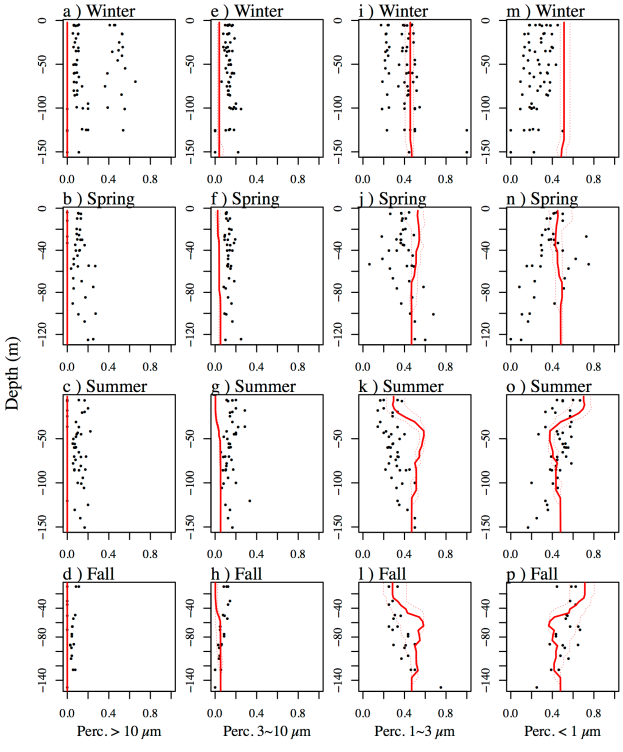

Fig. 9 . Model fittings to vertical profiles of four size fractions at S1

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

Geoscientific Model Development Discussions

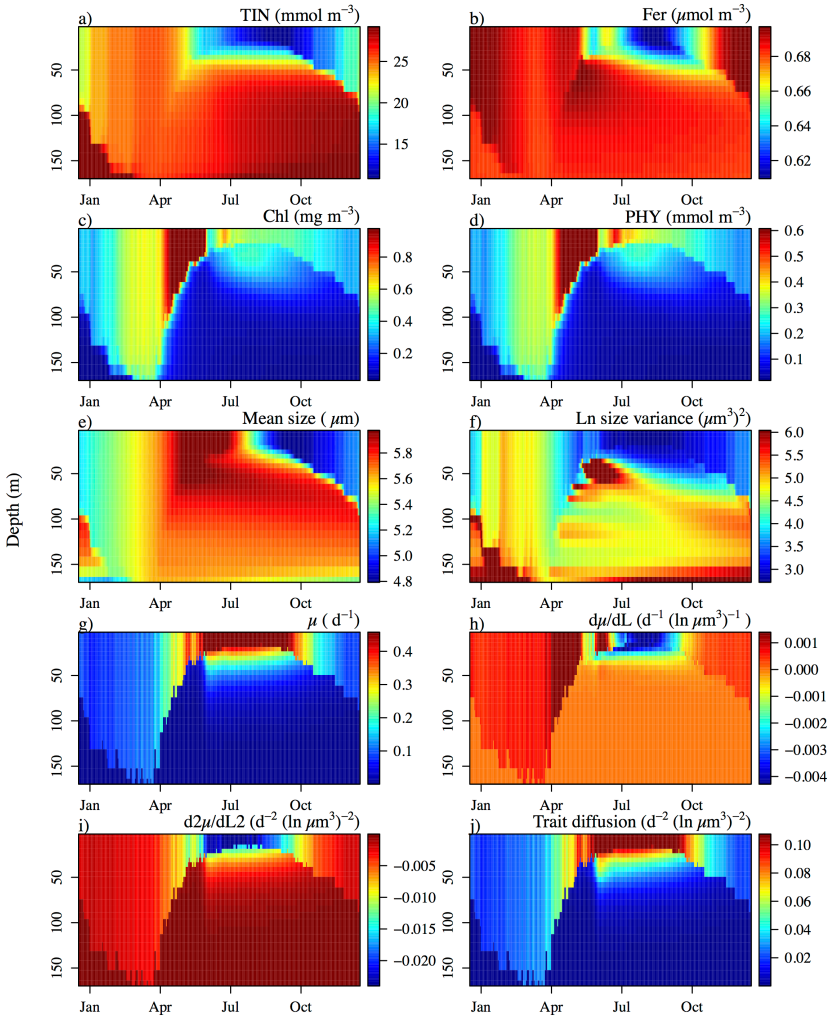

Fig. 10 . Modelled seasonal patterns at K2

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

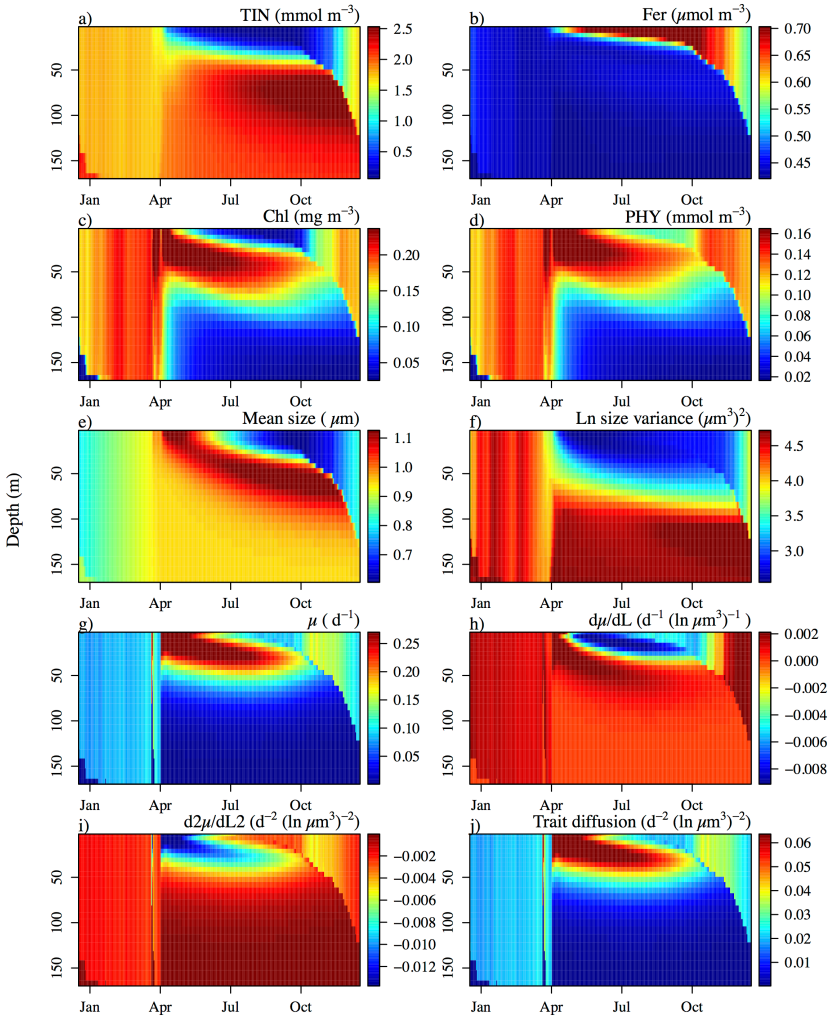

Fig. 11 . Modelled seasonal patterns at S1

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-104
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 June 2017

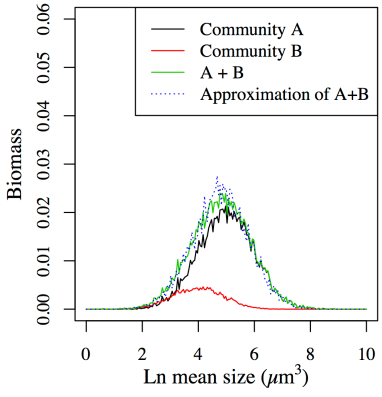

Fig. 12