WAP-1D-VAR v1.0: Development and Evaluation of a One-Dimensional Variational Data Assimilation Model for the Marine Ecosystem Along the West Antarctic Peninsula

Hyewon Heather Kim[1,2], Ya-Wei Luo[3], Hugh W. Ducklow[4], Oscar M. Schofield[5], Deborah K. Steinberg[6], Scott C. Doney[2]

[1]Woods Hole Oceanographic Institution, Woods Hole, MA 02543, United States
[2]University of Virginia, Charlottesville, VA 22904, United States
[3]Xiamen University, Xiamen, Fujian 361102, China
[4]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964, United States
[5]Rutgers University, New Brunswick, NJ 80901, United States
[6]Virginia Institute of Marine Science, Gloucester Point, VA 23062, United States

Correspondence to: Hyewon Heather Kim (hkim@whoi.edu)

Abstract

The West Antarctic Peninsula (WAP) is a rapidly warming region, with substantial ecological and biogeochemical responses to the observed change and variability for the past decades, revealed by multi-decadal observations from the Palmer Antarctica Long-Term Ecological Research (LTER) program. The wealth of these long-term observations provides an important resource for ecosystem modelling, but there has been a lack of focus on the development of numerical models that simulate time-evolving plankton dynamics over the Austral growth season along the coastal WAP. Here we introduce a one-dimensional, data assimilation planktonic ecosystem model (i.e., the WAP-1D-VAR model v1.0) equipped with a variational adjoint and model parameter optimization scheme. We first demonstrate the modified and newly added model schemes to the pre-existing food-web and biogeochemical components of the other ecosystem models that WAP-1D-VAR model was adapted from, including diagnostic sea-ice forcing and trophic interactions specific to the WAP region. We then present the results from model experiments where we assimilate eleven different data types from an example Palmer LTER growth season (October 2002 - March 2003) directly related to corresponding model state variables and flows between these variables. The iterative, data assimilation procedure reduces by 58% the misfits between observations and model results, compared to before optimization, via an optimized set of 12 parameters out of total 72 free parameters. The optimized model results capture key WAP ecological features, such as blooms during seasonal sea-ice retreat, the lack of macronutrient limitation, and modelled variables and flows comparable to other studies in the WAP region, as well as several important ecosystem metrics. One exception is that the model slightly underestimates particle export flux, for which we discuss potential underlying reasons. The data assimilation scheme of the WAP-1D-VAR model enables the available observational data to constrain previously poorly understood processes, including the partitioning of primary production by different phytoplankton groups, the optimal chlorophyll to carbon ratio of the WAP phytoplankton community, and the partitioning of dissolved organic carbon pools with different lability. The WAP-1D-VAR model can be successfully employed to link the snapshots collected by the available data sets together to explain and understand the observed dynamics along the coastal WAP.

1 Introduction

The West Antarctic Peninsula (WAP) has experienced significant atmospheric and surface ocean warming since the 1950s, resulting in decreased winter sea-ice duration, the retreat of glaciers, and changes in upper ocean dynamics (Clarke et al., 2009; Cook et al., 2005; Henley et al., 2019; King, 1994; Meredith & King, 2005; Stammerjohn et al., 2008; Vaughan et al., 2003, 2006; Whitehouse et al., 2008). These climate-driven changes propagate through marine food-webs by affecting physiology of individual organisms and the whole communities (Ducklow et al., 2007). Long-term observational efforts through the Palmer Antarctica Long-Term Ecological Research program (LTER; since 1991) have demonstrated a range of ecological and biogeochemical responses to changing environments, including phytoplankton (Montes-Hugo et al., 2009; Saba et al., 2014; Schofield et al., 2017), marine heterotrophic bacteria (Bowman & Ducklow, 2015; Ducklow et al., 2012; Kim & Ducklow, 2016; Luria et al., 2014, 2017), nutrient drawdown (Kim et al., 2016), and micro- and macrozooplankton (Garzio & Steinberg, 2013; Steinberg et al., 2015; Thibodeau et al., 2019).

The wealth of Palmer LTER time-series observations provides an important resource for ecological and biogeochemical modelling, and different types of modelling approaches have been developed to explore the WAP responses to climate change and variability. For instance, an inverse modelling study estimated the steady-state dynamics of the WAP food-web by deriving snapshots of flows among different plankton functional types and higher trophic levels (Sailley et al., 2013). However, there has been a less focus on numerical ecosystem models that simulate time-evolving plankton dynamics over the full Austral growth season along the coastal WAP. Numerical ecosystem models provide estimates of key rate processes for which observations have been less frequently or seldom made compared to frequently measured stocks and rates. Despite its strengths, constructing an ecosystem model is a challenge due to the lack of *a priori* knowledge on model parameter values and incomplete understanding of ecological processes that should be explicitly presented in the model structure (Ducklow et al., 2008; Murphy et al., 2012). Owing to many observational studies, a more robust, yet still incomplete, data-based picture is emerging of WAP food-web interactions and ecosystem dynamics, which could guide a development of the WAP-specific numerical ecosystem model.

Here we introduce a one-dimensional (1-D) variational data assimilation model specific to the coastal WAP (i.e., the WAP-1D-VAR model v1.0) that we develop by adapting an existing biogeochemical-planktonic model of different ocean basins (Friedrichs, 2001; Friedrichs et al., 2006, 2007; Luo et al., 2010). The WAP-1D-VAR model is compared against the roughly semi-weekly, bio-physical observations over the Austral growth season near Palmer Station on Anvers Island, Antarctica (64.77°S, 64.05°W). The field data record the seasonal variations in the initiation, peak, and termination of phytoplankton blooms and other biogeochemical processes modulated by variations in surface light, mixed layer depth, and sea-ice cover. In the present study, we 1) describe the structure and schemes of the WAP-1D-VAR in great detail, 2) evaluate the model performance and robustness using a variety of quantitative metrics, and 3) discuss the model applicability with regard to capturing the key WAP ecological and biogeochemical features using the data from an example growth season.

2 Model development and implementation

2.1 Model state variables

The WAP-1D-VAR model v1.0 (Figure 1) is originally derived and modified from data-assimilative, ocean regional test-bed models of the Arabian Sea, the Equatorial Pacific, and the Hawaii Ocean Time-series Station ALOHA (Friedrichs, 2001; Friedrichs et al., 2006, 2007; Luo et al., 2010). The WAP-1D-VAR model simulates stocks and flows of C, N, and P through 11 different model prognostic state variables. The two size-fractionated phytoplankton compartments (i.e., diatoms and cryptophytes) and the two different zooplankton compartments (i.e., microzooplankton and krill) are separately simulated

following the plankton functional types as in Sailley et al (2013) and the observations of the phytoplankton community structure along the coastal WAP. Typically, the coastal WAP is associated with large phytoplankton accumulations dominated by large (> 20 μm) diatoms, but nanoflagellates (< 20 μm) or cryptophytes are also an important component of the food web (Schofield et al., 2017). Mixed flagellates, prasinophytes, and type-4 haptophytes are also found in the region, but we choose to model only diatoms and cryptophytes, in order to avoid too many free (optimizable) parameters associated with each phytoplankton group. The third most dominant species is mixed flagellate but little is known about this taxa in the region and this taxa generally exhibit low interannual variability (Schofield et al. 2017). Functional grazing relationships are defined in which diatoms are consumed by both krill (*Euphausia superba*) and microzooplankton (mostly ciliates and other protozoa), cryptophytes are consumed by microzooplankton, and microzooplankton are grazed by krill. Other abundant zooplankton taxa in the WAP, such as salps, pteropods, and copepods (Steinberg et al., 2015), are not explicitly simulated in the WAP-1D-VAR model, in part to limit the model complexity and in part because of the limited data constraints on these groups, especially feeding. Higher trophic levels are implicitly represented to close the model. The WAP-1D-VAR model allows for the partitioning between labile dissolved organic matter (LDOM) and semi-labile DOM (SDOM) such that the entire LDOM pool is available but only a limited portion of the SDOM is available for bacterial utilization to account for lower lability of SDOM. Refractory DOM (RDOM) is not explicitly modelled due to its much longer turnover time than labile and semi-labile pools, but some mass flows are included to RDOM from other prognostic model compartments, such as bacteria, krill, and SDOM, to account for loss terms for those state variables. Detritus represents an average particulate organic matter (POM) pool after removing living phytoplankton and bacterial biomasses, and sinking of the detritus pool contributes to particle export flux. The WAP-1D-VAR model explicitly simulates NO_3, NH_4, and PO_4 for inorganic (macro)nutrient compartments, but there is not a separate Fe model compartment or Fe uptake processes, given that even during the peak of the blooms iron is still measurable and iron limitation is absent or occurs only minimally and seasonally in the nearshore Palmer Station area (Carvalho et al., 2016; Sherrell et al., 2018).

2.2 Model equations

Here we demonstrate key model processes that are either based on the existing schemes or built as new schemes for the coastal WAP region. The original model schemes are detailed in Supplementary Material of Luo et al (2010). The WAP-1D-VAR model v1.0 simulates biological-physical model processes for a 1-D vertical water column, solving numerically for a discretized version of the time-rate of change for each model state variable. For a generic tracer variable C the time-rate of change equation takes the form (Glover et al., 2011):

$$\frac{\partial C}{\partial t} = -\frac{\partial}{\partial z}(wC) + \frac{\partial}{\partial z}\left(K_z \frac{\partial C}{\partial z}\right) + J_C \qquad (1)$$

where z is the depth, w is the vertical velocity (the sum of water motion and gravitational particle sinking), K_z is the turbulent eddy diffusivity, and J_C is the biological and biogeochemical net source and sink term for C (Appendix A Equations; Eq. A.2.41-44, A.3.37-40, A.4.53-55, A.5.24-26, A.6.27-29, A.7.4-6, A.8.4-9, A.9.2-4). The physical advection and mixing terms are discussed below in section 2.3 and applied sequentially following the computation of the biological and biogeochemical terms J_C using a constant time step of 1 hour. The contributions of the source sink terms J_C to the full time rate of change equations are constructed as a series of coupled ordinary differential equations, detailed in Appendix A (sections A1-9), and solved using a second-order Runge-Kutta numerical integration scheme. The WAP-1D-VAR model simulates the dynamics of C, N, and P, but here we only focus on the presentation of the model C dynamics. The cellular molar (e.g., N/C, P/C) quota parameter values of most state variables are fixed (Table 1) and not submitted to the optimization and data assimilation procedure. To first order, most model physiological processes are affected by water temperature, including the maximum growth rates of phytoplankton, bacteria, and zooplankton and basal respiration rates of bacteria and zooplankton. The Arrhenius function is implemented to change these physiological rates as a function of water temperature (Eq. A.1.1).

The net change of phytoplankton (both diatoms and cryptophytes) C biomass is driven by gross growth, DOC excretion, particulate organic carbon (POC) production via aggregation, respiration, and grazing (Eq. A.2.41, A.3.37), the net change of their N and P biomass by gross growth, dissolved organic nitrogen (phosphorus) excretion, particulate organic nitrogen (phosphorus) production, and grazing (Eq. A.2.42-43, A.3.38-39). The net change of their chlorophyll *a* (Chl) by gross growth, DOM excretion, and grazing (Eq. A.2.44, A.3.40). The WAP-1D-VAR model adapts a phytoplankton growth scheme with flexible stoichiometry, in which phytoplankton cells are allowed to accumulate and store more nutrients under light stress (Bertilsson et al., 2003; Droop, 1974, 1983; McCarthy, 1980). The phytoplankton C growth rate is limited by their cellular nutrient quota (Eq. A.2.1-2, A.3.1-2). Modified from Geider et al. (1998), phytoplankton nitrogen uptake decreases when their cellular N/C quota is higher than their reference (Redfield) ratio, but not limited when lower than their reference ratio (Eq. A.2.4, A.2.8, A.3.4, A.3.8). The nitrogen consumption completely ceases when the phytoplankton cellular quota reaches their maximum allowable ratios and is additionally limited by the ambient NO_3 and NH_4 concentrations with a Monod function (Eq. A.2.10-11, A.3.10-11). NH_4 inhibition on NO_3 uptake of phytoplankton is modelled by assigning lower k^{NH4} compared to k^{NO3} (Table 1). The inhibition term does not exist for PO_4. The uptake scheme is similar for PO_4 (Eq. A.2.13, A.3.13), but PO_4 can be consumed in great excess of current needs (Armstrong 2006). Such luxury uptake is modelled by assigning smaller maximum and minimum P quota, which acts to alleviate P limitation. The maximum photosynthesis rate decreases when the phytoplankton cellular quota is lower than their reference ratio, and approaches zero near their minimum ratio (Eq. A.2.6, A.3.6). The Chl production decreases with lowering photosynthetic active radiation (PAR) and completely ceases in dark (Eq. A.2.14, A.3.14). Phytoplankton release LDOM via passive diffusion of the low molecular weights DOM (e.g., neutral sugars and dissolved free amino acids) with the same cellular elemental ratio as that of phytoplankton (Fogg 1966, Bjørnsen 1988, Biddanda & Benner 1997; Eq. A.2.16-18, A.3.16-18). Phytoplankton also release L- and SDOM actively, in the form of carbohydrate, as 75% of the labile (Eq. A.2.19, A.2.23, A.3.19, A.3.23) and 25% of the semi-labile pools (Eq. A.2.20, A.2.26, A.3.20, A.3.26). This active DOM production enables phytoplankton to adjust their stoichiometry to approach their reference ratio. If cellular organic C is in excess, DOC is released on a time scale of 2 days, and if excess nitrogen (phosphorus), DON (DOP) is released on a time scale of 8 days (Eq. A.2.21-22, A.3.21-22). Diatoms are grazed by both microzooplankton and krill (Eq. A.2.33-40), while cryptophytes are only grazed by microzooplankton (Eq. A.3.33-36). Microzooplankton grazing functions are altered by assigning grazing limitation terms (ϵ) to provide a limit on diatom grazing and route more cryptophytes to microzooplankton (Eq. A.2.33, A.3.33), based on initial modelling attempts where elevated diatom Chl was not simulated due to their much stronger removal by microzooplankton than cryptophytes. In principle, optimization should be able to capture the elevated diatom Chl by adjusting free parameters unless: 1) the right parameters are not adjusted and/or the baseline (non-optimized) parameters need significant adjusting, and/or 2) the model equations are not adequate even with the optimized parameters. In our initial modelling attempts, the model failed to simulate the elevated diatom Chl with varying sets of the model initial parameter values assigned to decouple diatoms from their grazers. Thus, grazing limitation terms (ϵ) are instead assigned to limit microzooplankton grazing on diatoms for modelling purposes, the implementation of which is not strictly based on the ecological evidence of prey switching, or of zooplankton mortality thresholds.

The net change of bacterial biomass is driven by their gross growth (via L- and SDOM uptake; Eq. A.4.12-14, A.4.15-16, A.4.21-22), respiration (Eq. A.4.25), S- and RDOM excretion (Eq. A.4.26-28, A.4.35-43), grazing (Eq. A.4.44-46), and mortality due to viral attack (Eq. A.4.47-49). The WAP-1D-VAR model allows both L- and SDOM as the substrate sources for bacteria, and bacterial nutrient quota lets the lability of SDOM variable for their selective utilization. All the LDOM pool is available, while only a limited portion of the SDOM pool is allowed for bacterial utilization, the degree of which is controlled by an optimizable parameter controlling the relative utilization of SDOM to LDOM, or SDOM lability (i.e., r_{SDOC}, Eq. A.4.11, Table 1). Bacterial C growth is determined by their cellular quota and available L- and SDOC concentration (Eq. A.4.12-13), in which the growth would be limited if bacterial cellular nitrogen (phosphorus) quota is smaller than their reference ratios (Eq. A.4.8-9). Bacteria take up LDOM in the way that the ratio of LDON (LDOP) to LDOC uptake is the same as the bulk

N/C (P/C) ratio of the LDOM (Eq. A.4.15, Eq. A.4.21). Bacteria take up SDOM with higher N/P ratios to reflect that SDOM with higher N/P ratios is more labile (Eq. A.4.13). The ratio of SDON to SDOC uptake by bacteria would vary between the bulk N/C of SDOM and the bacterial reference cellular quota (Eq. A.4.16, A.4.22). Bacteria are modelled to either take up or release NH_4 and PO_4 to maintain their stable and consistent stoichiometry (Kirchman, 2000). Bacteria take up NO_3 only if their cellular N/C ratio is smaller than their reference ratio (i.e., when bacteria are in short of nitrogen), in order to reflect higher energetic cost of NO_3 uptake than NH_4, but the amount of NO_3 uptake is modelled to be no more than 10% of N-specific bulk L- and SDOM uptake, and the sum of NO_3 and NH_4 uptake is modelled to be no more than N-specific bulk L- and SDOM uptake (Eq. A.4.17-20). These limit the maximum NO_3 uptake rate and set the inhibition of NH_4 uptake on NO_3 uptake. Bacteria excrete RDOM by transforming LDOM to RDOM (A4.26-28). Bacteria also adjust their cellular stoichiometry by remineralizing NH_4 and PO_4 if carbon is in short (i.e., N and P in excess; Eq. A.4.29-30) and by excreting SDOC if carbon is in excess (i.e., nitrogen and phosphorus are in short; Eq. A.4.38-43). Bacteria are grazed by microzooplankton (Eq. A.4.44-46), and a certain percentage of bacteria gets lost to LDOC pool due to viral attack (Eq. A.4.47-49).

The net change of zooplankton (both microzooplankton and krill) biomass is driven by their gross growth (via grazing on preys; Eq. A.5.3-5, A.6.3-5), L- and SDOM excretion (Eq. A.5.6-14, A.6.6-14), respiration (Eq. A.5.17, A.6.17), POM production (Eq. A.5.18-20, A.6.18-20), and grazing (Eq. A.5.21-23, A.6.24-26). Microzooplankton C growth is supported by consuming cryptophytes and bacteria (Eq. A.5.3-5), while krill carbon growth is supported by consuming diatoms and microzooplankton (Eq. A.6.3-5). Both zooplankton compartments follow the Holling Type 2 density-dependent grazing function with a preferential selection on different prey species (Eq. A.2.33, A.2.37, A.3.33, A.4.44, A.5.21). Both zooplankton groups release a portion of the organic matter that they ingest as DOM via sloppy feeding and excretion (Eq. A.5.6-8, A.5.9-11, A.6.6-8, A.6.9-11) such that the ratio of the released DON (DOP) to LDOC is equivalent to the N/C (P/C) ratio of zooplankton. The amount of SDOC excretion is a function of the total carbon growth (Eq. A.5.9, A.6.9), while the amount of SDON (SDOP) excretion is also a function of the zooplankton cellular N/C (P/C) ratio relative to their reference ratio (Eq. A.5.10-11, A.6.10-11). Zooplankton adjust their body cellular quota by either releasing SDOM if carbon is in excess, or by regenerating NH_4 or PO_4 if nitrogen or phosphorus is in excess (Eq. A.5.12-16, A.6.12-16), similar to the bacterial scheme. Respiration is formulated such that basal respiration is based on a portion of zooplankton biomass, while active respiration is based on a portion of their grazed C (Eq. A.5.17, A.6.17). Both zooplankton egest fecal matter as POM (Eq. A.5.18-20, A.6.18-20), but only krill additionally excrete RDOM with N/C and P/C similar to bacteria (Eq. A.6.21-23). Microzooplankton are grazed by krill (Eq. A.5.21-23), while krill are removed by implicit higher trophic levels (Eq. A.6.24-26), similarly calculated as a bacterial mortality term, rather than as an explicit grazing process.

The net change of detritus is driven by POM production by all phyto- and zooplankton compartments that is routed to detrital pool (Eq. A.2.29-32, A.3.29-32, A.5.18-20, A.6.18-20) and its dissolution (Eq. A.7.1-3). An optimizable vertical sinking speed is assigned to detritus to derive particle export flux (i.e., particle export flux = detrital concentration × particle sinking velocity, *wnsv,* Table 1). The detritus that is lost due to dissolution is incorporated to SDOM pool when it sinks (Eq. A.8.7-9) before regenerated to inorganic nutrients, rather than directly regenerated from as the particulate form. The net change of LDOM is driven by LDOM excretion by all phyto- and zooplankton compartments (Eq. A.2.16-19, A.3.16-19, A.5.6-8, A.6.6-8) and the amount of bacterial mortality that is incorporated to LDOM due to viral attack (Eq. A.4.47-49) and its uptake by bacteria (Eq. A.4.12). The net change of SDOM is driven by SDOM excretion by all organisms (Eq. A.2.20-22, A.3.20-22, A.5.9-14, A.6.9-14), the amount of detrital dissolution (Eq. A.7.1-3), uptake by bacteria (Eq. A.4.13-14), and conversion to RDOM pool (Eq. A.8.1-3). The conversion of SDOM to RDOM pool is a function of the stoichiometry of SDOM, in which the conversion process is slower for higher N/C and P/C of SDOM, to reflect that nitrogen- and phosphorus-enriched SDOM is more likely labile. A certain percentage of NH_4 is converted to NO_3 on a daily basis to represent a simple nitrification process in the model (Eq. A.9.1).

2.3 Physical forcing

The WAP-1D-VAR model v1.0 is forced by mixed layer depth (MLD), PAR at the ocean surface, sea-ice concentration, water-column temperature, vertical velocity, and vertical eddy diffusivity, at a temporal resolution of 1 day. Temperature, sea ice, and vertical eddy diffusivity are set up at every vertical grid (depth) point.

MLD is determined based on a finite difference density criterion with a threshold value of $\Delta\sigma_\theta = 0.03$ kg m^{-3} (Montégut et al., 2004) after calculating potential density of water mass from temperature and salinity Conductivity-Temperature-Density (CTD) observations. Vertical velocity, w, is assigned as zero because it is very weak in the surface waters of the study site and materials are transported vertically mostly by diffusion. The vertical eddy diffusivity scheme treats the rapid vertical mixing in the surface boundary layer by homogenizing model state variables instantaneously in the mixed layer (i.e., by averaging at every time step). Thus, K_z value above MLD is not required, and only K_z below MLD is calculated as follows:

$$K_z(z) = K_{z0} \times \exp\{ -\alpha \times (z - MLD) \} \quad (2)$$

where z is depth (m) below MLD and K_{Z0} is the vertical eddy diffusivity at the bottom of the mixed layer (1.1×10^{-4} m^2 s^{-1}) (Klinck, 1998; Smith et al., 1999), and α is 0.01 (m^{-1}).

Daily surface downward solar radiation flux (National Centers for Environmental Prediction reanalysis daily averages) is used to calculate sea surface PAR. PAR is estimated as 46% of the total solar radiation (Pinker & Laszlo 1992, Kirk 1994). The attenuation of PAR as a function of depth is calculated as follows:

$$PAR(z) = PAR_0 \times \exp\{ -(k_w + k_c \times CHL) \times z \} \quad (3)$$

where z is depth (m), PAR_0 is PAR level at sea surface (W m^{-2}), k_w is the attenuation coefficient for seawater (m^{-1}), k_c is the attenuation coefficient for Chl ((mg Chl)$^{-1}$ m^2), and CHL is the Chl concentration (mg Chl m^{-3}).

Sea-ice conditions in the coastal WAP do not necessarily represent solely local temperature and climate conditions, given that sea ice can be impacted by temperature, mixed layer, heat fluxes, regional winds, and other physical processes (Saenz et al. *in review*). We implement a sea-ice model scheme to account for light transmission through sea ice (5% of incident irradiance, as a typical transmittance value used in the Community Earth System Model) and non-linearities in the photosynthesis-irradiance (P-I) response under partial ice concentration (Long et al., 2015) using percent daily sea-ice concentration data (GSFC Bootstrap versions 2/3, derived from SMMR/SSMI satellite temperature brightness data binned into 25 by 25 km grid cells). In many previous models, the light-limitation term $\mathcal{L}(I)$ is calculated as a function of mean irradiance $\bar{I}$ averaged over both ice-covered and open-water conditions, so $\mathcal{L}(\bar{I})$; instead we compute the mean of light-limitation term ($\overline{\mathcal{L}(I)}$) as a function of fractional sea ice and open water and incident irradiance:

$$\mathcal{L}(I) = P^C / P^C_{MAX} = 1 - \exp(-I/I_k) \quad (4)$$

$$\overline{\mathcal{L}(I)} = f_i \times \mathcal{L}(I_i) + f_o \times \mathcal{L}(I_o) \quad (5)$$

where P^C is the C-specific photosynthetic rate (d^{-1}), P^C_{MAX} is the maximum photosynthetic rate (d^{-1}), I_k is the parameter describing the light-saturation behaviour of the PI-curve (W m^{-2}), I_o is the open-water irradiance, I_i is the under-ice irradiance (i.e., $I_i = 0.05 \times I_o$), f_i is the fraction of area covered with sea ice, and f_o is the fraction of open water (i.e., $f_o = 1 - f_i$).

2.4 Variational data assimilation

The WAP-1D-VAR model v1.0 is equipped with a built-in data assimilation scheme based on a variational adjoint method (Lawson et al., 1995). This method generates optimal model solutions that minimize the difference between model results and observations by objectively optimizing model parameter values (Friedrichs, 2001; Spitz et al., 2001; Ward et al., 2010). In detail, the assimilation scheme (Figure 2) consists of four steps (Glover et al., 2011): 1) starting with initial values of the model parameters (see below), the model is integrated forward in time from specified initial conditions to calculate the difference between the model simulation and the field data, or the model-observation misfit (i.e., cost function; section 2.5, Eq. 6); 2) an adjoint model constructed using the Tangent linear and Adjoint Model Compiler (TAPENADE) is integrated

backward in time to compute the gradients of the total cost with respect to the model parameters; 3) the computed gradients are then passed to a limited-memory quasi-Newton optimization software M1QN3 3.1 (Gilbert & Lemaréchal, 1989) to determine the direction and optimal step size by which the selected model parameters (see below) need to be modified in order to reduce the total cost; and 4) a new forward in time simulation is conducted using the new set of modified (optimized) parameter values. These four-step procedures are conducted in an iterative manner until the pre-set convergence criteria (i.e., low gradients of the total cost function with respect to optimized parameters and positive eigenvalues of the Hessian matrix) are satisfied to ensure that the optimized parameters converge and the total cost function reaches a local minimum.

Initial values of the model parameters (total of 72 free or optimizable parameters, Table 1) are assigned based on literature values (Caron et al., 2000, Luo et al., 2010, Garzio et al., 2013) without examining the effects of the initial parameter values on the model results prior to optimization. As is typical for many types of ecosystem models, a collection of what appear to be reasonable initial parameter estimates can result in relatively poor overall system behaviour because of system-level interactions of different model components. In most marine ecosystem models, these initial parameter values are subjectively and manually adjusted to improve the simulation, and the simulations with the initial, unadjusted parameter values are rarely shown. However, here with a more objective optimization approach that we conduct, the initial and optimized solutions can be explicitly compared (section 4). Optimization starts by submitting a subset of the 72 free model parameters rather than submitting all of them at once. This initial parameter subset consists of 10 different model parameters, the change of which yields the largest decrease in the total cost function (i.e., which also happens to be usually one per each state variable). These include α_{DA} (initial slope of photosynthesis vs. irradiance curve of diatoms, mol C (g Chl *a*)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$), α_{CR} (initial slope of photosynthesis vs. irradiance curve of cryptophytes, mol C (g Chl *a*)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$), Θ (maximum Chl/N ratio, g Chl *a* (mol N)$^{-1}$), μ_{BAC} (maximum bacterial growth rate, d^{-1}), $r^{A}_{max,BAC}$ (maximum bacterial active respiration rate, d^{-1}), g_{BAC} (half-saturation density of bacteria in microzooplankton grazing, mmol C m^{-3}), μ_{MZ} (maximum microzooplankton growth rate, d^{-1}), μ_{KR} (maximum krill growth rate, d^{-1}), and $remv_{KR}$ (krill removal rate by higher-trophic levels, (mmol C m^{-3})$^{-1}$ d^{-1}; Table 1).

When computed at the minimum of the cost function value, the inverse of the Hessian matrix provides the uncertainties of optimized parameters, cross-correlations among parameters, and sensitivities of the total cost function to each parameter (Matear, 1996; Tziperman & Thacker, 1989). High off-diagonal values in the inversed Hessian matrix indicate highly cross-correlated model parameters, so one of the highly cross-correlated parameters is removed from the optimization. The square root of a diagonal element in the inversed Hessian matrix is the logarithm of the relative uncertainty (σ_f) of the corresponding optimized parameter. The absolute uncertainty of the constrained parameter is calculated as $p_f \times e^{\pm\sigma f}$ where p_f is the value of the optimized parameter (Table 1). If optimized to ecologically unrealistic values, parameters are kept back to their respective initial values and removed from the next optimization cycle. Optimized parameters with σ_f larger than 50% are updated but removed from the next optimization cycle (i.e., defined as 'optimized' parameters), while optimized parameters with σ_f smaller than 50% are updated and kept for the next optimization cycle (i.e., defined as 'constrained parameters'). Constrained parameters are reported with the uncertainties, while optimized parameters are reported without the uncertainties (Table 1) because both changed parameters consist of an optimized model parameter set, but the parameters reported with the uncertainty ranges are the ones optimized with relatively small uncertainties and considered constrained. This way, a part of the initial parameter subset forms a final optimized parameter set. The gradients of the total cost function with respect to all 72 parameters are then evaluated, the parameters with large gradients (e.g., > 5) are re-submitted to optimization to further reduce the total cost, the gradients are evaluated again, and these cycles repeat until the termination of optimization. Optimization terminates when the gradients are reasonably low (e.g., < 10^{-2} for constrained parameters, < 5 for optimized parameters, and < 10 for unoptimized parameters). This final optimized model parameter set forms the basis of the results presented throughout this study (section 4). Additionally, in order to assess the sensitivity of the model optimization results with regard to the initial parameter choice, we perturb by ± 50% a subset of the initial parameter values used in the reference (original) optimization experiments to form different initial parameter sets (a total of 15 sets consisting of partially or fully perturbed 18 parameters, Tables B1-2) and conduct new optimization experiments from each set (section 4.1).

2.5 Cost function

To represent a misfit between observations and model output, a total cost function is calculated as follows (Luo et al., 2010):

$$J = \sum_{m=1}^{M} \frac{1}{N_m} \sum_{n=1}^{N_m} \left(\frac{\hat{a}_{m,n} - a_{m,n}}{\sigma_m} \right)^2 \quad (6)$$

where m and n represent assimilated data types and data points, respectively, M and N_m are the total number of assimilated data types and data points for data type m, respectively, σ_m is the target error for data type m, $a_{m,n}$ is observations, and $\hat{a}_{m,n}$ is model output. Given the high biological productivity of the WAP waters and the approximate log-normal distribution of many marine biological variables, the base-10 logarithms of Chl and primary production (PP) are used in the cost function calculation to capture phytoplankton dynamics (Campbell, 1995; Glover et al., 2018). The target error is calculated for each data type as follows:

$$\sigma_m = \overline{a_{m,n}} \times \mathrm{CV}_m \quad (7)$$

where $\overline{a_{m,n}}$ is the climatological mean (over the select 9 growth-seasons, see below) of the observations and CV_m is the averaged coefficient of variation (CV) of the observations of each data type in the mixed layer (due to observational error and seasonal and interannual variations) calculated using all of the observational data over 9 growth-season periods between 2002-2003 and 2011-2012, except the 2007-2008 growth season due to its missing data. These 9 growth seasons are chosen, instead of the multi-decadal observations available from Palmer LTER (since 1991), due to the relatively more complete data coverage in those seasons. The standard deviations are used as target errors of the log-converted data types. The CV of the log-converted data type is estimated as the average of ± 1 standard deviation in log space converted back into normal space (Doney et al., 2003; Glover et al., 2018). Hereafter, we present the total cost normalized by M (J equivalent to J/M hereafter) as it indicates the model-observation misfit equivalent to a reduced Chi-square estimate of model goodness of fit. We report the normalized total cost J along with normalized costs of individual data types throughout this study. $J = 1$ indicates a good fit, $J >> 1$ indicates a poor fit or underestimation of the error variance, and $J << 1$ indicates an overfitting of the data, fitting the noise, or overestimation of the error variance.

3 Model experiments

3.1 Modelling framework

To examine the applicability of the WAP-1D-VAR model v1.0 to the coastal WAP region, we select a nearshore Palmer LTER water-column time-series station, Station E (64.77°S, 64.05°W), as the modelling site that is ~200 m deep and situated approximately 3 km south of Palmer Station and 6.5 km northeast of the head of Palmer Deep (Sherrell et al., 2018). Physical forcing (Figure 3) and data types assimilated are derived from roughly semi-weekly physical, chemical and biological profiles collected from small boat via a profiling CTD and discrete water samples at Station E. When weather and ice conditions permit, water column sampling at the station has been conducted twice a week over the growth season. Seven upper-ocean layer depths (2.5, 10, 20, 30, 40, 50 and 60 m) are chosen for the model vertical grids. The model depth can be extended to as deep as needed, but this study is focused to upper 60-m water column to fully take advantage of the large data availability. Also, conceptually, the application of the 1-D model framework makes the most sense for the upper water column dominated by local seasonal processes, and extension of the model into deeper water well below the maximum seasonal mixed layer becomes more problematic because of the growing importance of lateral advective processes that are not well captured in the 1-D model framework. The vertical structure of the water column can be affected by growing sea ice due to reduced wind-driven turbulence and brine rejection during winter, but this is what a prognostic, coupled ocean-ice 1-D model can offer to

simulate, not our diagnostic forcing based model used in this study. Also, because our model simulates only the spring-summer growth season, the impact of winter sea ice on ecosystem dynamics is less of a concern.

Given the routine observations of Palmer LTER available over the growth season (October - March), we simulate one example growth season with the most complete data coverage, from October 2002 to March 2003 (2002-2003 growth season hereafter), instead of a series of different growth seasons in a continuous manner. The example growth season simulations utilize this year's specific observed physical forcing fields and assimilated biological and biogeochemical observations. Each Palmer LTER growth season should be modelled to have its own unique optimized parameter set, as well as initial conditions and physical forcing that together determine the model solution for that year; however, only 2002-2003 growth season simulations are modelled in this study for model analysis and evaluation.

3.2 Initial and boundary conditions

Model initial conditions are prescribed 135 days before the model start date for the growth season (October 15, 2002), so on June 1, 2002. This 135-day spin up is conducted to minimize the impact of initial conditions on the model output over the growth season. Initial conditions are prepared by optimizing the full growth seasonal cycle forced by climatological physics and assimilated with climatological observations and with the same bottom boundary conditions used in the optimization of the 2002-2003 growth season (i.e., climatological model; using climatological physics and observations averaged over 9 growth-season periods between 2002-2003 and 2011-2012 except the 2007-2008 growth season due to its missing data). For the first climatological model simulation initial conditions are prepared by adjusting manually following literature values (e.g., Luo et al. 2010). Due to strong interannual variability in the phytoplankton bloom phenology at Palmer Station, averaging across all these 9 years does not reflect distinct seasonal phytoplankton peaks, leading to underestimated phytoplankton values (not shown). To capture this non-linear aspect of the coastal WAP system, we construct the climatological year by applying a single time shift to all variables so that a seasonal PP peak of each year lines up with an average date of seasonal PP peaks from all years. Most biological initial conditions on June 1 are close to zero given the lack of active physiological processes in the very low light and the presence of sea ice during wintertime before the model growth season starts. All the data types are set to zero at the lower boundary (bottom) except for NO_3, PO_4, SDOC, SDON, and SDOP in which the climatological values at 65 m are used for lower boundary values (25.9 mmol m^{-3}, 1.9 mmol m^{-3}, 6.5 mmol m^{-3}, 0.6 mmol m^{-3}, and 0.03 mmol m^{-3}, respectively).

3.3 Assimilated data

We include the data types directly related to corresponding model outputs, including a mix of ecosystem stocks or state variables – NO_3, PO_4, Chl for diatoms and cryptophytes, bacterial biomass, microzooplankton biomass, SDOC, POC, and PON as well as carbon flows among model stocks – bulk net PP and bacterial production (BP). These data sets have been sampled semi-weekly at Palmer Station E (64.77°S, 64.05°W), the same location where our model is set up, and are available from the Palmer LTER data website (see Data availability). The distinction between diatoms and cryptophytes is established by assimilating phytoplankton taxonomic-specific Chl data for diatoms and non-diatom species derived from a High-Performance Liquid Chromatography (HPLC) and CHEMTAX analysis (Schofield et al., 2017), but given cryptophytes being the second dominant species in the water samples at the study site, cryptophytes are assumed to represent all non-diatom species for modelling purposes. Given that POC (PON) from bottle filtration may capture both living biomass and detrital material, we adjust the observed POC (PON) by subtracting phytoplankton and bacterial C (N) biomass to estimate the detrital pool, in order to only include non-living particles to detrital pool. When phytoplankton or bacterial biomass data are not available, we assign climatological (2002-2003 to 2011-2012) fractions of POC (PON) to detrital pool. Phytoplankton- and bacterial biomass accounts for 26% of total POC and 29% of total PON. In converting Chl to phytoplankton carbon (nitrogen)

biomass, the maximum Chl/C (Chl/N) ratio submitted for optimization is used along with other reference ratios (Table 1). Microzooplankton biomass data are not available for the full time-series, so their data from grazing experiments at Palmer Station (Garzio et al., 2013) are assimilated to at least provide constraints on bacterial and cryptophyte grazing processes. However, due to the discrepancy in the timing and location from model simulations of this study, the microzooplankton model-observation misfits are not analysed in the present study. Krill biomass data are not assimilated due to the strong patchiness of their distribution that may hinder proper model optimization. The vertical profiles of most of the data types are assimilated, whereas average NO_3 and PO_4 concentrations in the mixed layer are assimilated due to the difficulty of simulating depth-dependent nutrient concentrations and the fact that net PP is mostly determined by surface nutrient concentrations (Luo et al., 2010). BP (mmol C m^{-3} d^{-1}) is derived from the ^{3}H-leucine incorporation rate (pmol l^{-1} h^{-1}) data using the conversion factor of 1.5 kg C (mol leucine)$^{-1}$ incorporated (Ducklow, 2000). Bacterial biomass (mmol C m^{-3}) is estimated from bacterial abundance measured by flow cytometry with the conversion factor of 10 fg C $cell^{-1}$ (Fukuda et al., 1998). SDOC is calculated by subtracting the background concentration (41.2 mmol m^{-3} for the modelling site) from total DOC concentration.

3.4 Uncertainty analysis

Uncertainties of the optimized parameters are computed from a finite difference approximation of the complete Hessian matrix (i.e., second derivatives of the cost function with respect to the model parameters) during the iterative optimization process (details in section 2.4). We then conduct Monte Carlo experiments to calculate the impact of the optimized parameter uncertainties on the model results. We first create an ensemble of parameter sets (n = 1000) by randomly sampling values within the uncertainty ranges of the constrained parameters, and then perform a model simulation using each parameter set. 1000 Monte Carlo experiments were shown to be adequate from a series of tests with different numbers of Monte Carlo sampling (n = 500-2000), where standard deviations of model simulated values converged after >1000 Monte Carlo sampling (not shown). All uncertainty estimates are calculated following standard error propagation rules and presented as ± 1 standard deviation in the study.

4 Results and discussion

4.1 Model skill assessment

In the case of the example growth season (2002-2003) modelled in this study, the iterative data assimilation-parameter optimization procedure reduced by 58% the misfits between observations and model output compared to the misfits obtained from the initial parameter values (Table 2). The optimized model solution satisfied the pre-set convergence criteria, with the low gradients of the total cost with respect to the optimized parameters and positive eigenvalues of the Hessian matrix. Notably, this was achieved by optimizing a subset of 12 (9 constrained and 3 optimized) parameters among the total of 72 optimizable parameters (Table 1, section 4.2). To examine the sensitivity of the optimized model solution to the initial parameter choice, a series of new optimization experiments (n = 15) were conducted with a varying subset of the initial parameter values perturbed by ± 50% of those used in the original optimization experiment (Table B1). These experiments showed that the optimized model results (i.e., the reference case; Table 1) were not sensitive to the initial choice of the parameters. The 15 different initial parameter sets resulted in a range of initial model-observation misfits, some substantially larger than the reference case (14.25-28.24 versus 14.85 for the reference case). However, the total normalized optimized cost values of the 15 sensitivity experiments (5.79-7.19) were similar to that of the reference case of 6.42. In the sensitivity experiment #12, the initial model-observation misfit was ~2 times larger than that of the reference case, and there was up to 76% of the reduction in the model-observational misfit (versus 58% of the reduction in the reference case; Table B1, Table 1). These results suggest that no matter where in parameter space the optimization process starts from, the optimization scheme takes the model cost function to similar

local minima. Importantly, this was achieved by similar subsets of the optimized parameters: μ_{DA}, μ_{CR}, $r^{A}_{max,BAC}$, and μ_{MZ} were optimized in all cases, while α_{DA}, α_{CR}, Θ, μ_{BAC}, g_{CR}, g'_{DA}, g_{BAC}, μ_{KR}, g_{MZ}, and $remv_{KR}$ were optimized except for a few cases (Table B1). The uncertainties of the optimized parameters were also similar among different optimizations, with most of the relative errors < 0.5. Constrained parameter values and their uncertainty ranges averaged over the sensitivity experiments (Table B2) were comparable to those in the original optimization experiments (Table 1).

Overall, there was a good model-data fit with the largely decreased cost value for each data type after optimization (Table 2). Optimization yielded J_f close to 1 for all data types, compared to the initial model solution where three data types – diatom Chl, crypto Chl, and bacterial biomass – had particularly poor model fits to observations and underestimated error variances ($J >> 1$). Compared to the initial (unoptimized) model results, the average errors (ε_{bias}, Doney et al., 2009; Stow et al., 2009) in the optimized model results indicated that diatom Chl, cryptophyte Chl, bacterial biomass, BP, and POC had reduced model biases, while NO_3, PO_4, PP, SDOC, and PON had increased model biases (for both positive and negative biases, defined as $\varepsilon_{bias} > 0$ and $\varepsilon_{bias} < 0$, for model overestimation and underestimation of the observation, respectively). Optimization resulted in the negative model bias for PO_4, compared to the positive model bias in the initial model results. The point-to-point comparison plots showed that there were seasonally consistent, negative model biases for PP, POC, and PON (Figure B1). Model skill was further evaluated with a Taylor diagram (Taylor, 2001) summarizing the statistics of the correlation coefficient between model output and observations, normalized standard deviation (by the standard deviation of the observations), and centred (bias removed) root-mean-square difference (RMSD) for each data type, in which a better model skill is characterized by a higher correlation, a normalized standard deviation close to 1, and a lower RMSD (Figure 4). Optimization resulted in better model skills for cryptophyte Chl, PP, BP, and bacterial biomass via increased correlation coefficients and lowered RMSD (Figure 4B), compared to those in the unoptimized model results (Figure 4A). After optimization the normalized standard deviations of PP, BP, bacterial biomass, phosphate, POC, and PON were closer to 1 (Figure 4B). Direct comparisons with the observational data showed that the optimized model parameter set captured better the increases in diatom biomass early in the season, cryptophyte biomass in January, and bacterial biomass in mid-February, compared to the unoptimized model parameter set (Figures 5A-B, Figure B2).

4.2 Optimized parameters

The number of the optimized parameters in this study is small and comparable to those from other data-assimilative model focused on different marine environments (Friedrichs, 2001; Friedrichs et al., 2006, 2007; Luo et al., 2010). This is consistent with the general behaviour of marine plankton ecosystem models, in which well-posed model solutions would be found with only a subset of independent model parameters due to many cross-correlated parameters inherent in nonlinear model equations (Fennel et al., 2001; Harmon & Challenor, 1997; Matear, 1996; Prunet et al., 1996). Ecosystem models with a relatively large number of unconstrained parameters (i.e., equivalent to the optimized parameters with high uncertainties in the present study) might reduce total costs to a greater extent, but could possess low predictive skill as a result of being over-tuned to fit noise in the observations (Friedrichs et al., 2007). Also, there are several field and lab-based studies at the study site or in a similar polar environment that reported the values of the model parameters used in the WAP-1D-VAR model v1.0, including the bacterial growth rate of 0.82 d^{-1}, total phytoplankton (including large cells like diatoms) growth rate of 0.33-0.55 d^{-1}, nanophytoplankton (corresponding to cryptophytes) growth rate of 0.52-0.99 d^{-1} (Garzio et al., 2013), and the microzooplankton growth rate of up to 1.0 d^{-1} (Caron et al., 2000). The optimized values of the maximum bacterial, diatom, cryptophyte, and microzooplankton growth rates in this study were 1.06 d^{-1} (0.93-1.20 d^{-1}), 0.77 d^{-1} (0.68-0.88 d^{-1}), 0.72 d^{-1} (0.61-0.85 d^{-1}), and 1.18 d^{-1} (1.10-1.26 d^{-1}), falling in the ranges of those measured bacterial, total phytoplankton, nanophytoplankton, and microzooplankton growth rates, respectively.

The ensemble of the model state variables and flows obtained from the Monte Carlo experiments had generally small standard variations at each model time step and grid, suggesting the robustness of the modelled fields against the variations in

the optimized parameter values (Figures B3, B4). To examine the translation of the optimized parameter values to altered functioning of the WAP biogeochemical processes, we compared two different sets of the model simulation results – one based on the initial parameter values (Figures 5A, 6A, 7A) and the other based on the optimized parameter values (Figures 5B, 6B, 7B). However, due to the non-linearities in the model it is not straightforward to identify what causes the parameter variations, except for a few cases in which the changes in the parameter values are clearly linked to the difference in the model state variables and flows. The first case is the relation of the increased g_{BAC} value (bacterial half-saturation concentration in microzooplankton grazing, mmol C m^{-3}) to the elevated bacterial accumulations after optimization (Table 1, Figures 5, 7). The second case is the link between Θ (maximum Chl/N ratio, g Chl *a* (mol N)$^{-1}$) and the relative dominance of cryptophytes in total phytoplankton accumulations. It has been demonstrated that the variations of Θ are driven by an imbalance between the rate of light absorption and energy demands for photosynthesis and biosynthesis in phytoplankton cells (Geider et al., 1997). Θ can also change because of the variations in phytoplankton photo-acclimation or physiological differences across phytoplankton groups, from a lower Θ value for smaller species to a higher Θ value for larger diatom cells (Geider, 1987). Θ was optimized to a 30% lower value than the initial parameter value (Table 1), in order to simulate the relatively larger proportion of cryptophytes in total phytoplankton accumulations in the optimized model results compared to the unoptimized model results (Figures 5, 7). By contrast, the remaining cases are not as clear because the first-order impact of parameter variations on the model results is less direct and more nuanced. Compared to the unoptimized results, the decreases in μ_{DA} (diatom C-specific maximum growth rate, d^{-1}), μ_{CR} (cryptophytes C-specific maximum growth rate, d^{-1}), α_{DA} (initial slope of P-I curve of diatoms, mol C (g Chl)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$), and α_{CR} (initial slope of P-I curve of cryptophytes, mol C (g Chl)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$) did not lead to decreased diatom and cryptophyte accumulations, presumably due to decreased g_{MZ} (microzooplankton half-saturation concentration in krill grazing, mmol C m^{-3}) and increased rem_{VKR} (krill removal rate by higher-trophic levels, (mmol C m^{-3})$^{-1}$ d^{-1}) after optimization (Table 1, Figures 5, 7). Similarly, the decreased μ_{BAC} (maximum bacterial growth rate, d^{-1}) and the increased $r^4_{MAX,BAC}$ (bacterial maximum active respiration rate, d^{-1}) did not lead to decreased bacterial accumulations, presumably due to the increased g_{BAC} (bacterial half-saturation concentration in microzooplankton grazing, mmol C m^{-3}) and the decreased g_{MZ} (microzooplankton half-saturation concentration in krill grazing, mmol C m^{-3}).

4.3 Ecosystem indices

We calculated key ecosystem indices for the modelled growth season, including NPP (directly comparable to ^{14}C-PP observations), net community production (NCP; i.e., NCP = NPP – bacterial-, microzooplankton-, and krill respiration), BP, and POC export (sinking) flux (Figure 6). Setting an upper limit for lateral or vertical carbon export from the euphotic zone (Dugdale & Goering, 1967), over appropriate time and space scales NCP is quantitatively equivalent to new production that is supported via external sources of nitrogen (Ducklow & Doney, 2013). In both optimized and unoptimized model results, NPP increased after complete sea-ice retreat, but a brief ice-edge bloom was simulated under sea ice at the beginning of the growth season (Figures 3, 6). Seasonal patterns of NCP resembled those of NPP and occasionally fell below zero (i.e., the net heterotrophy) in subsurface waters for both optimized and unoptimized cases (Figure 6). The POC export flux increased over time and reached the maximum value at the end of the growth season in both model results, but there were two major POC flux events separated by weaker, in-between flux events in December in the optimized results that the initial model results did not capture (Figure 6b). After optimization, the correlation coefficients adjusted from 0.88 to 0.36, 0.89 to 0.68, and 0.45 to 0.73 for the NPP-vs.-NCP pair, the NPP-vs.-BP pair, and the NCP-vs.-POC export flux (lagged by 30 days) pair (all $p < 0.001$). In the optimized model results, the growth-season mean of the depth-integrated NPP, NCP, and BP in the 60 m water column, and the 30-day lagged POC export flux at 60 m were 19 ± 8 mmol C m^{-2} d^{-1}, 10 ± 3 mmol C m^{-2} d^{-1}, 1 ± 1 mmol C m^{-2} d^{-1}, and 2 ± 0.3 mmol C m^{-2} d^{-1} (uncertainties propagated from season-averaging in Figure 6b and Monte Carlo uncertainties in Figure B4), compared to 28 ± 6 mmol C m^{-2} d^{-1}, 13 ± 3 mmol C m^{-2} d^{-1}, 3 ± 1 mmol C m^{-2} d^{-1}, and 2 ± 0.2 mmol C m^{-2} d^{-1} in the unoptimized model results (uncertainties from season-averaging in Figure 6a).

The mean *e*-ratio, defined as the growth-season mean of the 30-day lagged POC export flux divided by the growth-season mean NPP (i.e., particle export efficiency), was 0.11 ± 0.05 (uncertainties propagated from season-averaging in Figure 6b and Monte Carlo uncertainties in Figure B4) in the optimized model results, compared to 0.07 ± 0.02 (uncertainties from season-averaging in Figure 6a) in the unoptimized model results. The mean *f*-ratio, defined as the amount of NO_3 uptake divided by the amount of NO_3 and NH_4 uptake both, was 0.88 ± 1.52 in the optimized model results, compared to 0.84 ± 0.19 in the unoptimized model results (not shown). The higher mean *f*-ratio relative to the mean *e*-ratio in this study implies an imbalance between production and export at the study site, at least during the modelled period. Excess new production relative to export production (as derived from sediment traps and ^{234}Th disequilibrium; Ducklow et al., 2018) was previously observed in the WAP, presumably due to diel vertical migration, DOM export, lateral export, and diffusive loss of PON via diapycnal mixing (Stukel et al., 2015). Stukel et al (2015) reported up to 5 times larger new production via NO_3 uptake than export production via Th-based N export along the coastal WAP. Several additional mechanisms might be responsible for driving the discrepancy between production and export. First, given that the assimilated pool of suspended POC in the model formulation is not a good indicator of a rapidly sinking detrital pool dominating particle export, the WAP-1D-VAR model v1.0 does not capture large, short-lived particle flux events (e.g., fecal pellets produced by a large swarm of krill), underestimating POC export flux. Second, the WAP-1D-VAR model v1.0 export scheme does not consider DOC export that would lower the production-export discrepancy. Finally, RDOC is not explicitly modelled in the model, due to its much longer time scale than the model time scale, so accumulated and not-exportable RDOC pool would contribute to the deviation of the modelled *e*-ratio from the modelled *f*-ratio. Indeed, the modelled mean *e*-ratios in our study, for both optimized and unoptimized cases, situate at the lower end of the range of the *e*-ratios measured or estimated in the WAP waters (Ducklow et al., 2018; Sailley et al., 2013; Stukel et al., 2015; Weston et al., 2013), but optimization increased the *e*-ratio by 60% and thus made it closer to the literature values.

The mean BP/NPP ratio was 0.05 ± 0.06 (uncertainties propagated from season-averaging in Figure 6b and Monte Carlo uncertainties in Figure B4) in the optimized model results, compared to 0.11 ± 0.04 (uncertainties from season-averaging in Figure 6a) in the unoptimized model results. The modelled mean BP/NPP ratio for both optimized and unoptimized cases correspond well to the estimates from other measurement- and observation-based studies (Ducklow et al., 2012; Kim & Ducklow, 2016). Relatively low bacterial activity in productive Antarctic waters, typically reflected as a low BP/PP ratio, has been attributed to low LDOM availability for bacterial growth (Kirchman et al., 2009), low temperature (Pomeroy & Wiebe, 2001), or top-down control via grazing and viral lysis (Bird & Karl, 1999).

4.4 Mean carbon stocks and flows

We summarized the growth-season means of the carbon stocks and flows in the entire food web (Figure 7). The WAP-1D-VAR model v1.0 captured several key WAP ecological and biogeochemical features, including the lack of macronutrient limitation (NO_3 and PO_4 drawdown by phytoplankton utilization but remaining well above their half-saturation constants, Table 1) and comparable values of the assimilated and non-assimilated model state variables (Ducklow et al., 2007, 2012, 2018; Kim et al., 2016; Moline et al., 2008; Smith et al., 2008), providing confidence in the model simulations. For instance, growth-season measurements in 2017-2018 at Palmer Station showed a strongly patchy krill distribution, with the mean biomass of 0.12 ± 0.04 mmol C m^{-3} and the maximum biomass of 0.57 mmol C m^{-3} when krill were present (unpublished data provided by D. Steinberg), falling in the range of the modelled krill biomass values (0.13 ± 0.03 mmol C m^{-3}; calculated from Figure 7b). The WAP-1D-VAR model v1.0 also simulated several important ecosystem metrics comparable to other statistical modelling studies. For instance, the modelled phytoplankton seasonal patterns in the present study are consistent with physicochemical attributes revealed by a distinct ecological seascape pattern in the coastal WAP (Bowman et al., 2018), including low Chl and high nutrients in the first half of the growth season followed by high Chl and low nutrients in the second

half of the growth season. A steady-state solution based, inverse modelling study quantified different food-web states using ecosystem network indices from Palmer LTER annual summer cruises along the WAP shelf region (Sailley et al., 2013). Their network indices include the ratio of C export to total PP (i.e., equivalent to *e*-ratio in our study) and the ratio of recycling (the sum of flows into respiration and DOC pool) to total PP, where more (less) recycling favourable microbial food-webs are characterized by greater (smaller) ratios of recycling to total PP and smaller (greater) ratios of total C export to total PP (Legendre & Rassoulzadegan, 1996). As discussed above, the modelled mean *e*-ratio in the present study is smaller than the estimates in the inverse modelling study for the WAP shelf region (Sailley et al., 2013), but consistent with their conclusion on the food-web status of the modelled growth season (2002-2003) positioned on the microbial food-web side. The discrepancy in the *e*-ratio values between the present study and Sailley et al. (2013) may be attributed to fundamentally different model formulation (i.e., time-evolving modelling for the WAP-1D-VAR model v1.0 versus steady-state modelling) and optimization approach, or due to relatively strong microbial food-web activity at our coastal site compared to the shelf region. Microbial food-web activity can be approximated by quantifying the amount of fixed carbon flowing through heterotrophic bacteria (Carlson et al., 1999; del Giorgio & Cole, 1998; Ducklow, 2000; Ducklow et al., 2012). According to this approach, microbial food-web activity from the optimized model results was around 38 ± 16%, calculated as the ratio of bacterial L- and SDOC uptake to PP (i.e., (arrow 13 + arrow 14)/arrow 1 in Figure 1, mean ± uncertainties from season-averaging and Monte Carlo uncertainties in Figure 7B). On average, SDOC supported 1 ± 2% of the total bacterial C uptake, or C demand (i.e., arrow 14/(arrow 13 + arrow 14) in Figure 1, mean ± uncertainties from season-averaging and Monte Carlo uncertainties in Figure 7b), but could be an important bacterial C source when LDOC became scarce as the growth season progressed (Figure 5b). Indeed, several observational studies speculated that the WAP bacteria utilize SDOM in short of LDOM (Ducklow et al., 2011; Kim & Ducklow, 2016; Luria et al., 2017).

5 Summary

We developed the WAP-1D-VAR model v1.0, a one-dimensional variational data assimilation model specific to the coastal WAP region, evaluated the model performance and robustness using a variety of quantitative metrics, and discussed the model applicability with regard to capturing the key WAP ecological and biogeochemical features using the data from the example growth season. The data assimilation scheme significantly reduced the model-observation misfits via the optimized model parameter set that adjusted the simulation to better match the Palmer LTER observations in 2002-2003. We also explored the nuanced question of how the observations influenced the data assimilation process, drove the variations in optimized parameter values relative to their corresponding initial parameter values, and affected the resulting model simulations. The WAP-1D-VAR model v1.0 successfully simulated the variables and flows not included in data assimilation, with the values consistent and comparable with other measurement- and observation-based studies in the coastal WAP. Importantly, the data assimilation scheme enabled the available observational data to constrain processes that were poorly understood, including the partitioning of NPP by different phytoplankton groups, the optimal Chl/C ratio of the WAP phytoplankton community, and the partitioning of DOC pools with different lability. Up to this point, a range of observational studies has provided snapshots of ecosystem and biogeochemical processes in the WAP. Yet, we have little understanding of the driving processes that underlie the connections between each component in complex food-web interactions. We used data-assimilative modelling to glue these snapshots together to explain better the observed dynamics and further understand the previously poorly constrained processes in the coastal WAP system.

Acknowledgements

We thank the many Palmer LTER team members and Palmer Station support staff for collecting, measuring, and analyzing the field samples and for providing the datasets used in the present study. We also thank Le Xie for providing

protocols for the parameter optimization and adjoint models, Cristina Schultz for providing the initial parameter values of the model parameters, Nicole Waite for compiling HPLC data, and Joe Cope and Lori Garzio for compiling the zooplankton data. Computing resources for model simulations were provided on the Rivanna high-performance computing system by the Advanced Research Computing Services of the University of Virginia and on the Poseidon high-performance computing system by Woods Hole Oceanographic Institution. Kim and Doney were supported by the National Aeronautics and Space Administration Ocean Biology and Biogeochemistry Program (grant NNX14AL86G) and US National Science Foundation Office of Polar Programs (grant PLR-1440435 to Hugh Ducklow at Columbia University; Palmer LTER). Kim was additionally supported by the Investment in Science Fund and the Reuben F. and Elizabeth B. Richards Endowed Fund from Woods Hole Oceanographic Institution. Schofield and Steinberg were supported by US NSF grant PLR-1440435. Luo was supported by National Natural Science Foundation of China project 41890802.

Data availability

Complete Palmer LTER time-series data used for data assimilation are available online (http://pal.lternet.edu/data). Surface downward solar radiation flux data used for physical forcing of the model simulations are found in the National Centers for Environmental Prediction website (https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.surface.html).

Code availability

The Tangent linear and Adjoint Model Compiler (TAPENADE) used to construct an adjoint model is available online (http://www-sop.inria.fr/tropics/). The current version of the WAP-1D-VAR model (v1.0) is available from the project website: https://zenodo.org/record/5041139 under the Creative Commons Attribution 4.0 International license. The exact version of the model used to produce the results, input data, and scripts to run the model and produce the plots for all the simulations presented in this paper are archived on Zenodo (Kim, H. Heather, Luo, Ya-Wei, Ducklow, Hugh W., Schofield, Oscar M., Steinberg, Deborah K., & Doney, Scott C. (2021, January 26). WAP-1D-VAR v1.0: A One-Dimensional Variational Data Assimilation Model for the West Antarctic Peninsula (Version v1.0). Zenodo (http://doi.org/10.5281/zenodo.5041139). A user manual is available as a separate supplement to this article.

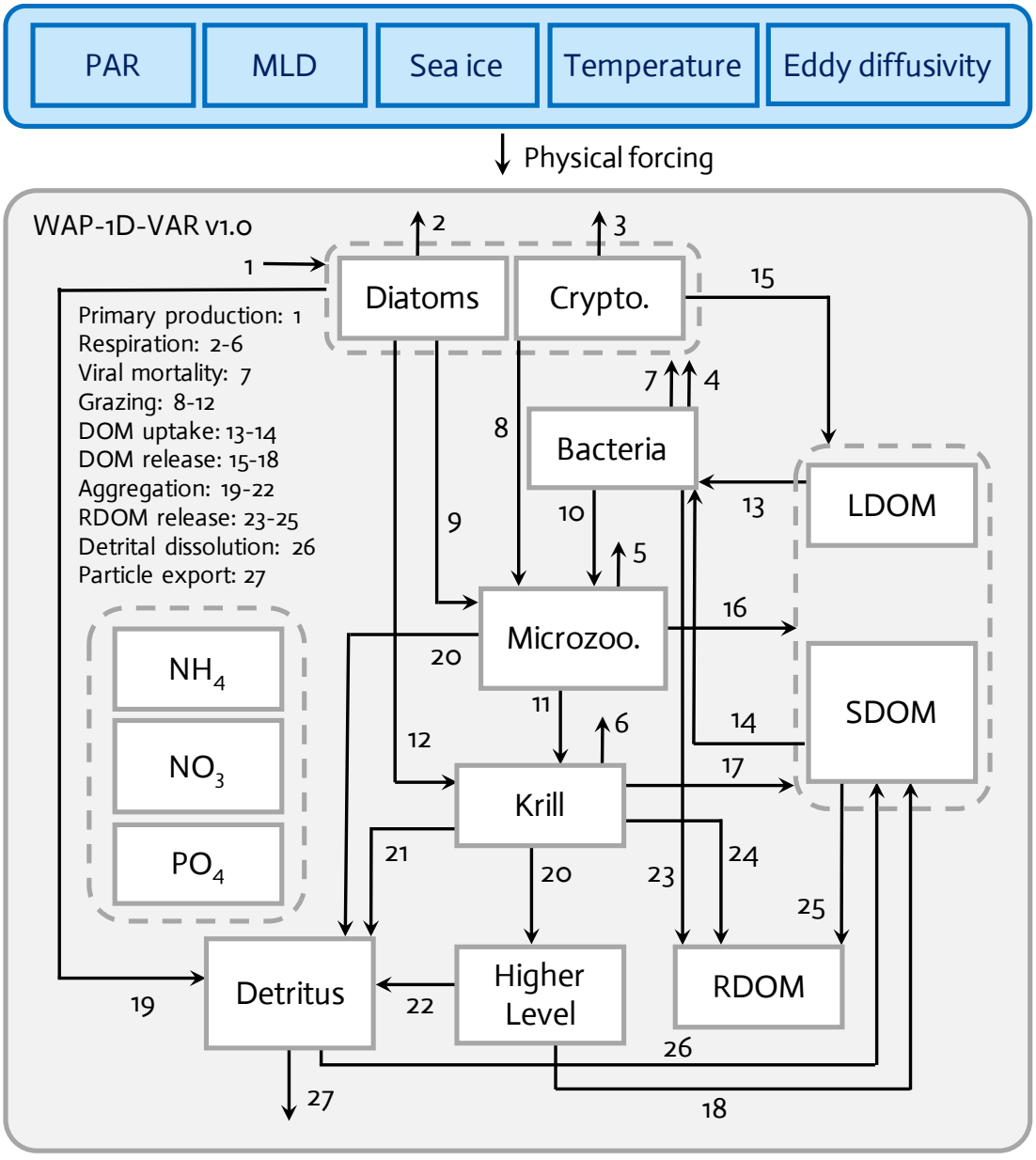
Author contributions

HHK developed the model, performed model simulations, and wrote the manuscript. YWL contributed to model simulations. HWD, OMS, and DKS provided observational data. SCD supervised the research and significantly the manuscript.

Competing interest

The authors declare that they have no conflict of interest.

Figure 1. Ecosystem model. The WAP-1D-VAR model v1.0 is forced by five different physical forcing, denoted as a horizontal row across the top of the schematic. The ecosystem component incorporates eleven different prognostic state variables. Higher level and refractory dissolved organic matter (RDOM) are represented implicitly.



PAR: photosynthetically active radiation, MLD: mixed layer depth, Crypto.: cryptophytes, Microzoo.: microzooplankton, LDOM: labile dissolved organic matter (DOM), SDOM: semi-labile DOM, RDOM: refractory DOM, Higher Level: higher-level zooplankton

Figure 2. Variational data assimilation. A variational adjoint scheme is employed for the parameter optimization and data assimilation processes (adapted from Glover et al., 2011). Gradient: the sensitivity of the total cost function with respect to model parameter from optimization.

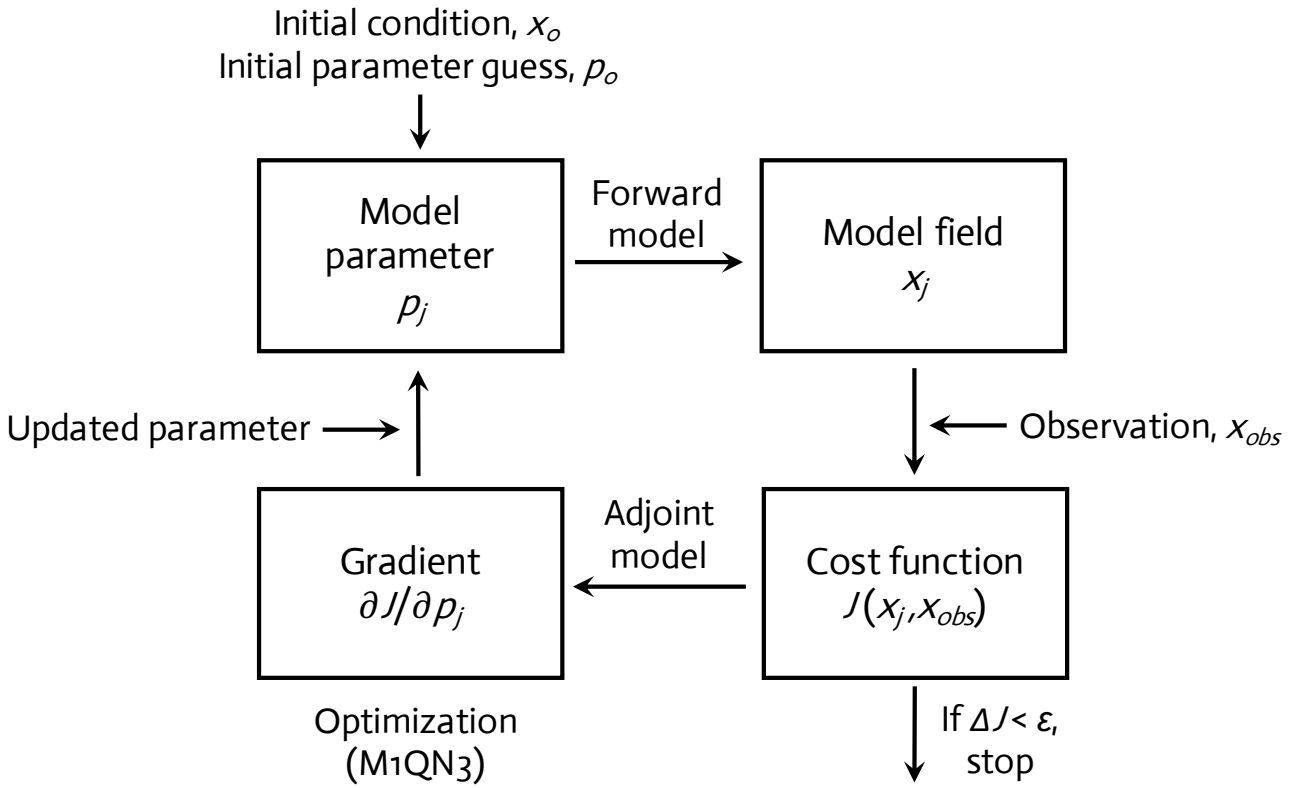


Figure 3. Physical forcing. Physical forcing used in the WAP-1D-VAR model v1.0, including surface photosynthetically active radiation (PAR) (a), sea-ice concentration (b), water temperature (c), and vertical eddy diffusivity (d) overlaid with mixed layer depth (MLD; dashed line) in the modelled growth season 2002-2003.

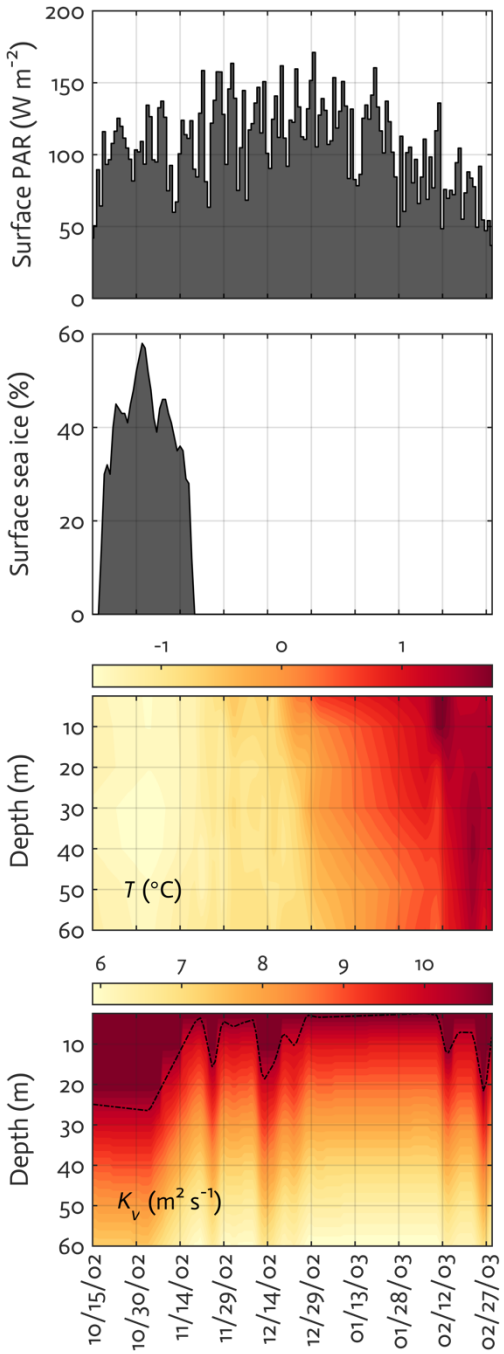


Figure 4. Model skill assessment. The Taylor diagrams using a polar-coordinate system summarizing the model-observational correspondence for each model stock and flow for the modelled growth season 2002-2003 before (a) and after optimization (b). The angular coordinate for the Pearson correlation coefficient (r), the distance from the origin for the standard deviation normalized by the standard deviation of the observation, and the distance from point (1,0), marked as REF on x-axis, for the centred (bias removed) root-mean-square difference (RMSD) between model results and observations.

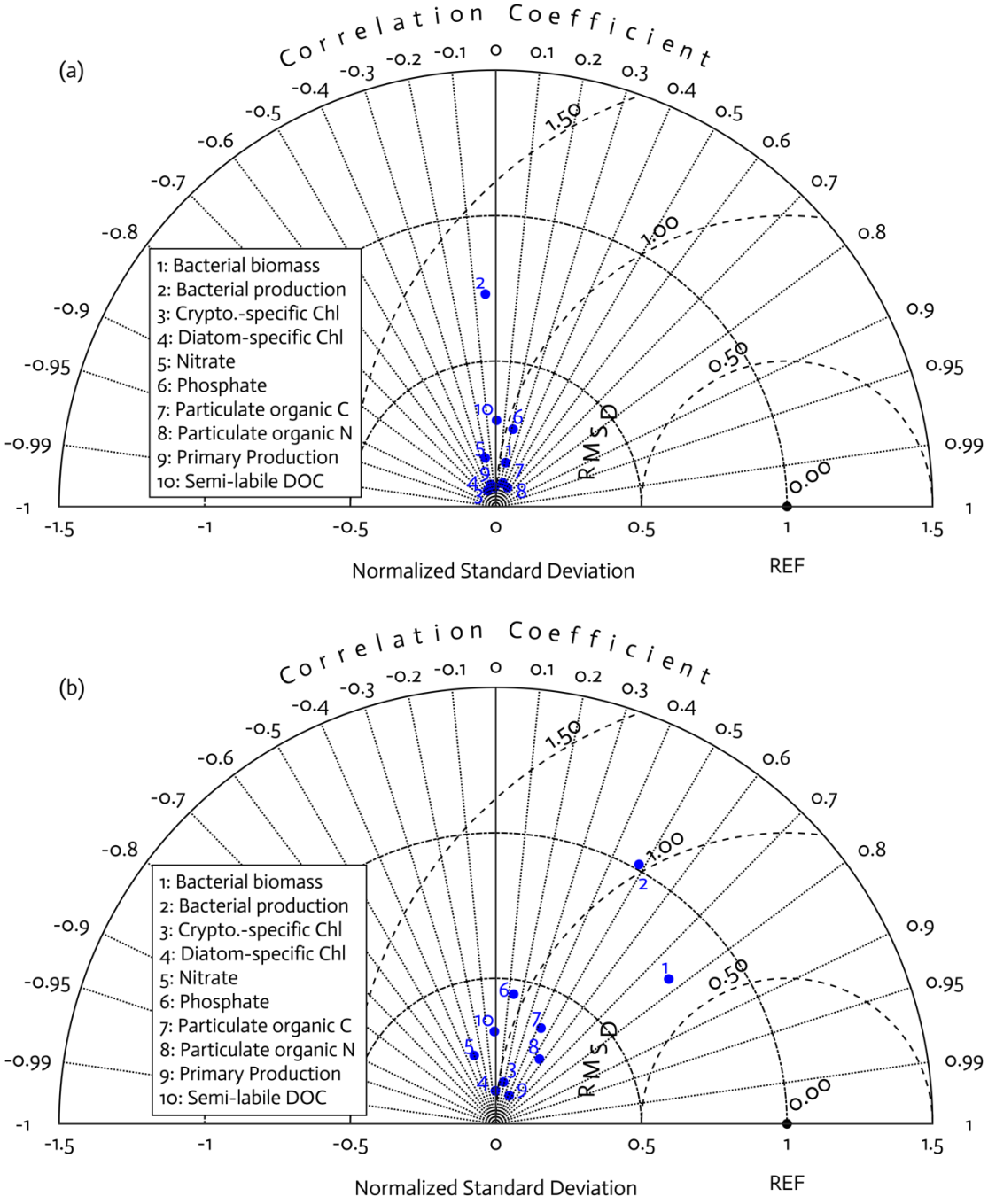


Figure 5. Model state variables before and after optimization. The model state variables for the modelled growth season 2002-2003 (x-axis; month/day) for before (a) and after optimization (b). init. as the initial (unoptimized) model results and opt. as the optimized model results. The error (standard deviation) of each model state variable from the Monte Carlo experiments (n = 1000) is available in Figure B3.

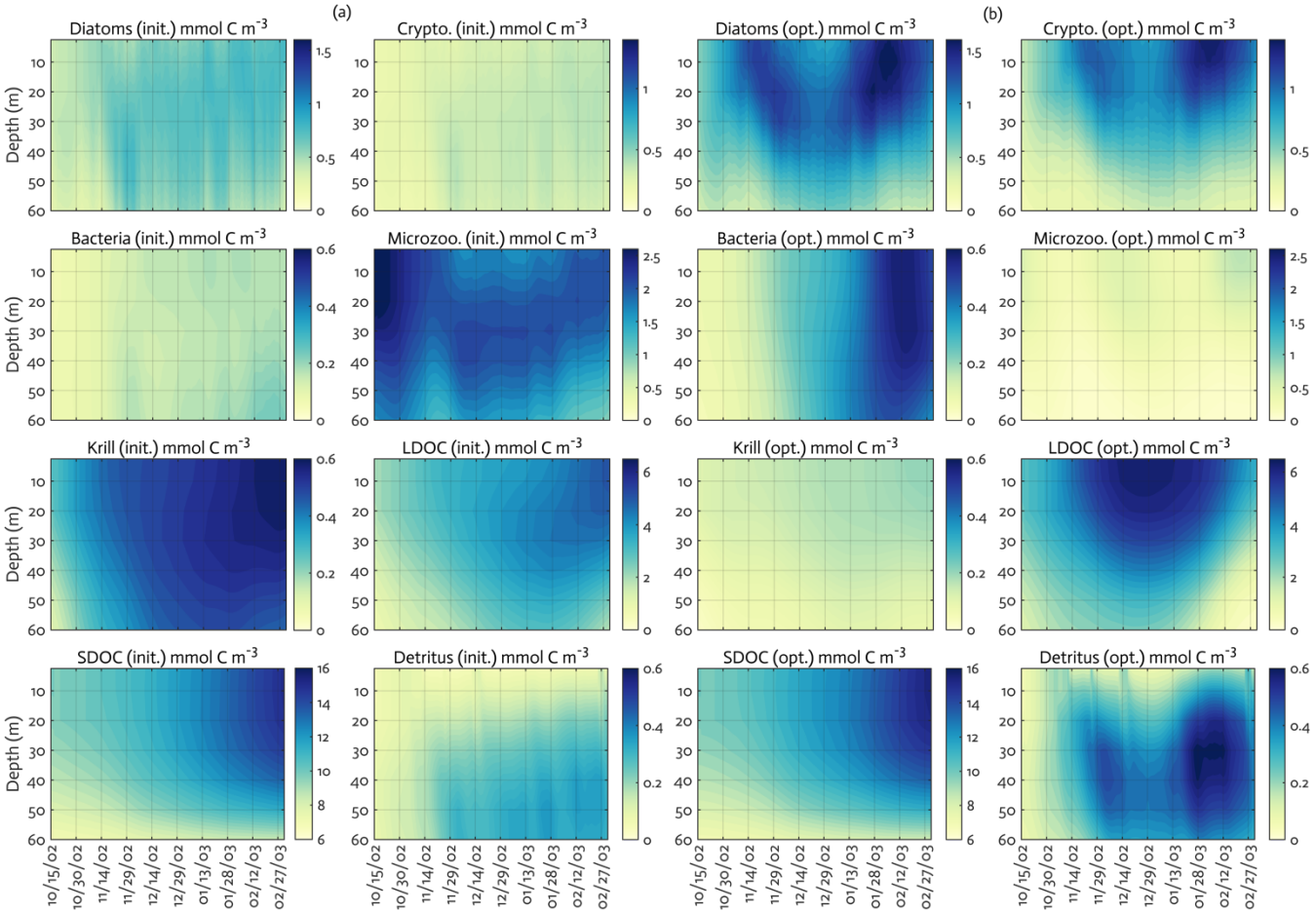


Figure 6. Model ecosystem indices before and after optimization. The key ecosystem indices for the modelled growth season 2002-2003 (x-axis; month/day). init. as the initial (unoptimized) model results and opt. as the optimized model results. NPP: net primary production, NCP: net community production, C export flux: particulate organic carbon (POC) export flux, and BP: bacterial production. The error (standard deviation) of each rate process from the Monte Carlo experiments (n = 1000) is available in Figure B4.

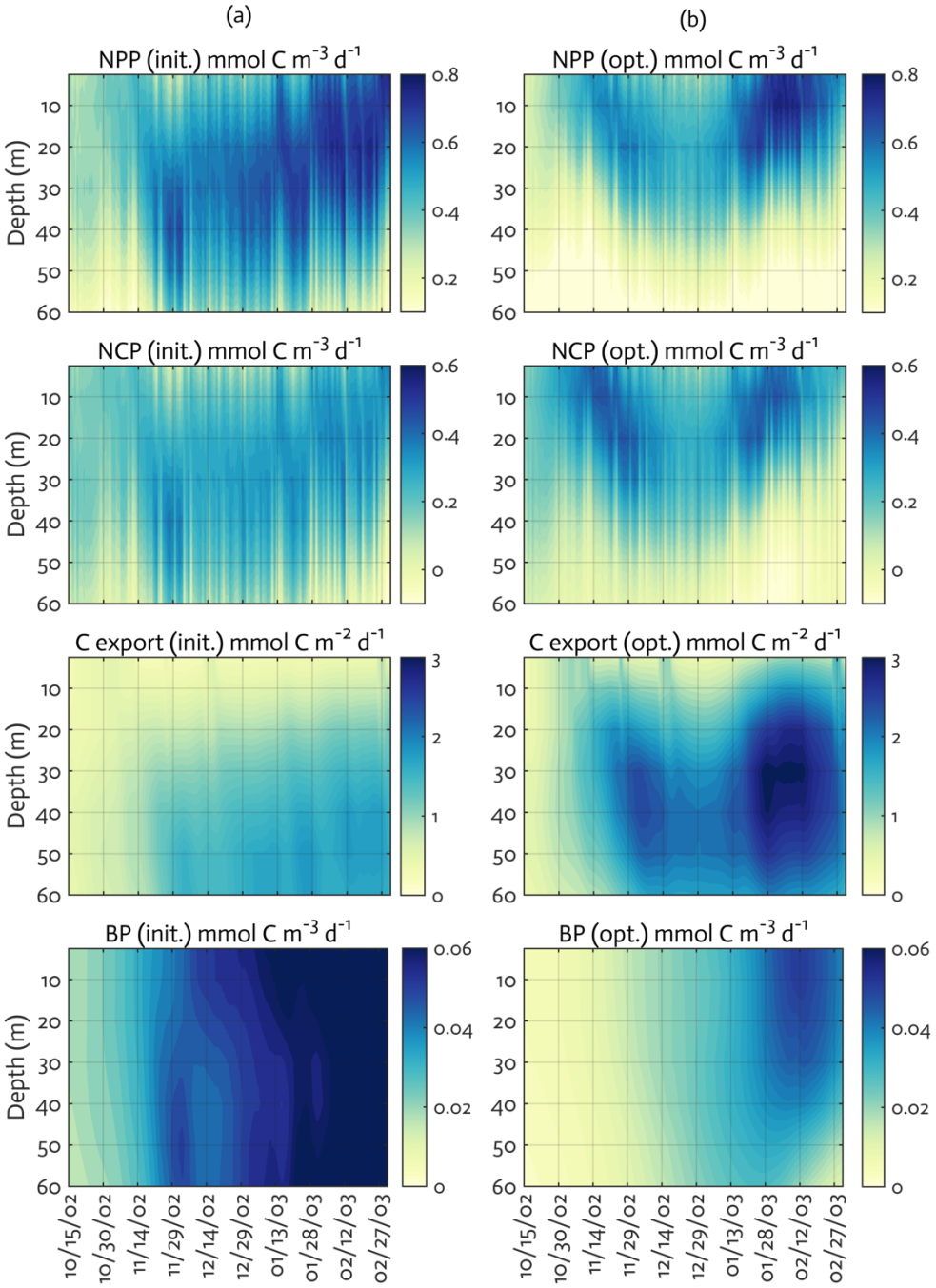


Figure 7. Mean carbon stocks and flows. Depth-integrated (0-60 m) carbon stocks (mmol C m^{-3}), flows (mmol C m^{-3} d^{-1}), and POC export flux at 60-m (mmol C m^{-2} d^{-1}) averaged over the modelled Austral growth season (October 2002 – March 2003) before (a) and after optimization (b). Values in parentheses as uncertainties from season-averaging (a) and as uncertainties propagated from season-averaging and depth-integration of the Monte Carlo errors for (b).

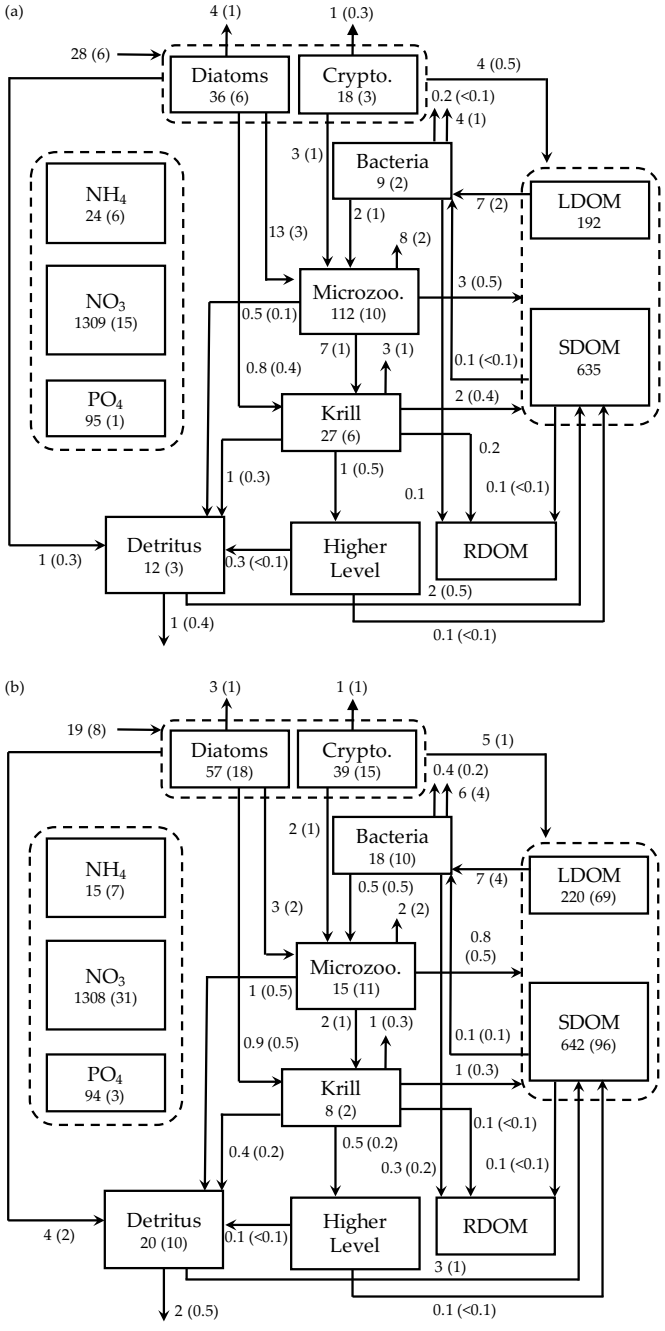


Table 1. Summary of model parameters. Summary of the model parameter symbol and definition, initial parameter values (p_0) and optimized values (p_f) for optimizable parameters, the cost function gradient with regard to the optimized parameter ($\partial J/\partial p$), and prescribed values for fixed model parameters over the course of simulations. The parameter with 'n/a' in the parenthesis is an optimized parameter with a large relative uncertainty, while the parameter with values in the parenthesis is a constrained parameter (optimized with a low relative uncertainty) with its upper and lower bounds. The uncertainties for these upper and lower bounds are calculated as: $p_f \times e^{\pm\sigma_f}$ where p_f is the value of the constrained parameter and σ_f is the square roots of diagonal elements of the inverse of the Hessian matrix. The cost function gradient with regard to the optimized parameter ($\partial J/\partial p$) after data assimilation is defined as: $\Delta J/e^{\Delta p}$ where $e^{\Delta p} \approx \Delta p$ for an infinitely small Δp. For example, a 10% change of a parameter (Δp = 10%) leads to a total cost change equivalent to 10% of the corresponding gradient.

Model parameter symbol and definition (optimizable)	p_0	p_f	$\partial J/\partial p$
A_E, Arrhenius parameter for temperature function	4000.00	-	-1.15
μ_{DA}, Diatom C-specific maximum growth rate, d^{-1}	2.00	0.77 (0.68-0.88)	-5.53×10^{-5}
μ_{CR}, Crypto. C-specific maximum growth rate, d^{-1}	1.00	0.72 (0.61-0.85)	2.51×10^{-4}
α_{DA}, Initial slope of P-I curve of diatoms, mol C (g Chl)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$	0.30	0.13 (0.10-0.19)	-1.55×10^{-4}
α_{CR}, Initial slope of P-I curve of crypto., mol C (g Chl)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$	0.20	3.89×10^{-2} (n/a)	0.45
β_{DA}, Light inhibition parameter for diatom photosynthesis (W m^{-2})$^{-1}$	5.00×10^{-3}	-	-1.10
β_{CR}, Light inhibition parameter for crypto. photosynthesis (W m^{-2})$^{-1}$	5.00×10^{-3}	-	0.32
$v^N_{REF,DA}$, Maximum N uptake rate per diatom C biomass, mol N (mol C)$^{-1}$ d^{-1}	0.50	-	-5.13×10^{-2}
$v^N_{REF,CR}$, Maximum N uptake rate per crypto. C biomass, mol N (mol C)$^{-1}$ d^{-1}	0.30	-	-3.07×10^{-2}
k^{NH4}_{DA}, NH_4 half-saturation concentration for diatom uptake, mmol m^{-3}	0.10	-	0.29
k^{NH4}_{CR}, NH_4 half-saturation concentration for crypto. uptake, mmol m^{-3}	0.10	-	0.14
k^{NO3}_{DA}, NO_3 half-saturation concentration for diatom uptake, mmol m^{-3}	1.00	-	-0.29
k^{NO3}_{CR}, NO_3 half-saturation concentration for crypto. uptake, mmol m^{-3}	0.60	-	-0.14
$v^P_{REF,DA}$, Maximum P uptake rate per diatom C biomass, mol P (mol C)$^{-1}$ d^{-1}	0.03	-	0.32
$v^P_{REF,CR}$, Maximum P uptake rate per crypto. C biomass, mol P (mol C)$^{-1}$ d^{-1}	0.03	-	0.15
k^{PO4}_{DA}, PO_4 half-saturation concentration for diatom uptake, mmol m^{-3}	0.05	-	-9.93×10^{-3}
k^{PO4}_{CR}, PO_4 half-saturation concentration for crypto. uptake, mmol m^{-3}	0.04	-	-3.68×10^{-3}
ζ^{NO3}, C requirement (respiration) to assimilate NO_3, mol C (mol N)$^{-1}$	2.00	-	-1.42
Θ, Maximum Chl/N ratio, g Chl *a* (mol N)$^{-1}$	2.90	2.27 (1.82-2.82)	5.95×10^{-5}
$ex_{PSV,DA}$, Diatom passive excretion rate per biomass, d^{-1}	0.05	-	0.86
$ex_{PSV,CR}$, Crypto. passive excretion rate per biomass, d^{-1}	0.05	-	2.17
$ex_{ACT,DA}$, Diatom active excretion rate per growth rate, d^{-1}	0.05	-	2.26×10^{-2}
$ex_{ACT,CR}$, Crypto. active excretion rate per growth rate, d^{-1}	0.05	-	4.06×10^{-3}
pom_{DA}, POM production rate by diatom aggregation, (mmol C m^{-3})$^{-1}d^{-1}$	0.04	-	1.99
pom_{CR}, POM production rate by crypto. aggregation, (mmol C m^{-3})$^{-1}d^{-1}$	0.03	-	0.61
k^{DOC}, DOC half-saturation concentration for bacterial uptake, mmol C m^{-3}	0.65	-	1.00
r_{SDOM}, Parameter controlling SDOM lability	5.00×10^{-3}	-	-0.64
μ_{BAC}, Maximum bacterial growth rate, d^{-1}	2.00	1.06 (0.93-1.20)	1.54×10^{-4}
$b_{R,BAC}$, Parameter control bacterial active respiration rate versus production, (mmol C m^{-3} d^{-1})$^{-1}$	1.50×10^{-2}	-	9.60×10^{-2}
$ex_{ADJ,BAC}$, Bacterial extra SDOC excretion rate, d^{-1}	2.00	-	0.00
$remi_{BAC}$, Bacterial nutrient regeneration rate, d^{-1}	6.00	-	-0.11
$ex_{REFR,BAC}$, Bacterial RDOC production rate, d^{-1}	1.70×10^{-2}	-	2.77
f_S, Bacterial selection strength on SDOM	0.25	-	-5.18×10^{-3}
r^B_{BAC}, Bacterial basal respiration rate, d^{-1}	1.27×10^{-2}	-	0.40
$r^A_{min,BAC}$, Bacterial minimum active respiration rate, d^{-1}	3.50×10^{-2}	-	-4.37×10^{-3}
$r^A_{max,BAC}$, Bacterial maximum active respiration rate, d^{-1}	0.58	0.80 (0.77-0.84)	-6.59×10^{-4}
$mort_{BAC}$, Bacterial mortality rate, d^{-1}	0.02	-	3.10
μ_{MZ}, Microzoo. C-specific maximum growth rate, d^{-1}	1.00	1.18 (1.10-1.26)	-7.41×10^{-4}

g_{DA}, Diatom half-saturation concentration in microzoo. grazing, mmol C m^{-3}	1.00	-	1.85
g'_{DA}, Diatom half-saturation concentration in krill grazing, mmol C m^{-3}	1.00	-	-1.15
g_{CR}, Crypto. half-saturation concentration in microzoo. grazing, mmol C m^{-3}	1.00	-	1.32
g_{BAC}, Bacterial half-saturation concentration in microzoo. grazing, mmol C m^{-3}	0.55	0.81 (0.64-1.03)	3.75×10^{-5}
ex_{MZ}, Total DOM excretion rate per microzoo. gross growth, d^{-1}	0.15	-	0.52
$f_{ex,MZ}$, Fraction of LDOC of total microzoo. DOC excretion	0.75	-	-0.37
r^{B}_{MZ}, Microzoo. basal respiration rate, d^{-1}	0.01	-	3.92×10^{-2}
r^{A}_{MZ}, Microzoo. active respiration rate, d^{-1}	0.42	-	-0.63
$ex_{ADJ,MZ}$, Microzoo. extra SDOM excretion rate, d^{-1}	2.00	-	0.00
$remi_{MZ}$, Microzoo. nutrient regeneration rate, d^{-1}	4.68	-	2.93×10^{-3}
pom_{MZ}, POM production rate per microzoo. gross growth, d^{-1}	2.70×10^{-2}	-	2.87×10^{-2}
μ_{KR}, Maximum krill C-specific growth rate, d^{-1}	0.80	1.02 (0.97-1.07)	2.17×10^{-4}
g_{MZ}, Microzoo. half-saturation concentration in krill grazing, mmol C m^{-3}	1.00	0.15 (n/a)	-0.95
ex_{KR}, Total DOM excretion rate per krill gross growth, d^{-1}	0.30	-	-0.29
$f_{ex,KR}$, Fraction of labile DOC of total krill DOC excretion	0.75	-	-0.45
r^{B}_{KR}, Krill basal respiration rate, d^{-1}	0.03	-	-0.50
r^{A}_{KR}, Krill active respiration rate, d^{-1}	0.30	-	-1.08
$ex_{ADJ,KR}$, Krill extra SDOM excretion rate, d^{-1}	2.00	-	0.00
$remi_{KR}$, Krill nutrient regeneration rate, d^{-1}	4.00	-	-2.81×10^{-2}
pom_{KR}, POM production rate per krill gross growth, d^{-1}	0.15	-	-0.38
$ex_{REFR,KR}$, Krill RDOC production rate, d^{-1}	0.02	-	-6.43×10^{-2}
$remv_{KR}$, Krill removal rate by higher-trophic levels, (mmol C m^{-3})$^{-1}$ d^{-1}	0.10	0.43 (n/a)	0.86
f_{KR}, Fraction of SDOM production by krill	0.10	-	5.11×10^{-2}
$f_{POM,HZ}$, Fraction of POM production by higher-trophic level	0.20	-	4.86×10^{-2}
$ex_{REFR,SDOM}$, Conversion rate of SDOM to RDOM, d^{-1}	9.00×10^{-4}	-	-5.17×10^{-2}
$q^{C}_{N,RDOM}$, RDOM N/C ratio, mol N (mol C)$^{-1}$	0.05	-	-0.10
$q^{C}_{P,RDOM}$, RDOM P/C ratio, mol P (mol C)$^{-1}$	6.50×10^{-4}	-	4.70×10^{-3}
$q^{C}_{N,POM}$, N/C ratio for POM production by microzoo. and krill, mol N (mol C)$^{-1}$	0.12	-	0.12
$q^{C}_{P,POM}$, P/C ratio for POM production by microzoo. and krill, mol P (mol C)$^{-1}$	4.50×10^{-3}	-	6.24×10^{-2}
r_{ntrf}, Nitrification rate (NH_4 to NO_3), d^{-1}	7.60×10^{-2}	-	-3.74×10^{-2}
prf_N, Preference for dissolving N content in POM	1.10	-	0.27
prf_P, Preference for dissolving P content in POM	4.00	-	1.67×10^{-4}
$wnsv$, Detritus vertical sinking velocity, m d^{-1}	5.00	-	0.26
$diss$, Detrital dissolution rate, d^{-1}	0.14	-	1.07
Model parameter symbol and definition (fixed)	***p***		
T_{ref}, Reference temperature in Arrhenius function, °C	15.00		
$q^{C}_{N,MIN,DA}$, Minimum N/C ratio of diatoms	3.40×10^{-2}		
$q^{C}_{N,MAX,DA}$, Maximum N/C ratio of diatoms	0.17		
$q^{C}_{N,RDF,DA}$, Reference (Redfield) N/C ratio of diatoms	0.15		
$q^{C}_{P,MIN,DA}$, Minimum P/C ratio of diatoms	1.90×10^{-3}		
$q^{C}_{P,MAX,DA}$, Maximum P/C ratio of diatoms	1.59×10^{-2}		
$q^{C}_{P,RDF,DA}$, Reference (Redfield) P/C ratio of diatoms	9.40×10^{-3}		
$q^{C}_{N,MIN,CR}$, Minimum N/C ratio of crypto.	3.40×10^{-2}		
$q^{C}_{N,MAX,CR}$, Maximum N/C ratio of crypto.	0.17		
$q^{C}_{N,RDF,CR}$, Reference (Redfield) N/C ratio of crypto.	0.15		
$q^{C}_{P,MIN,CR}$, Minimum P/C ratio of crypto.	1.90×10^{-3}		
$q^{C}_{P,MAX,CR}$, Maximum P/C ratio of crypto.	1.59×10^{-2}		
$q^{C}_{P,RDF,CR}$, Reference (Redfield) P/C ratio of crypto.	9.40×10^{-3}		
$q^{C}_{N,BAC}$, Reference (optimal) N/C ratio of bacteria	0.18		
$q^{C}_{P,BAC}$, Reference (optimal) P/C ratio of bacteria	0.02		
$q^{C}_{N,MZ}$, Reference (optimal) N/C ratio of microzoo.	0.20		

$q^{C}_{P,MZ}$, Reference (optimal) P/C ratio of microzoo.	2.20×10^{-2}
$q^{C}_{N,KR}$, Reference (optimal) N/C ratio of krill	0.20
$q^{C}_{P,KR}$, Reference (optimal) P/C ratio of krill	8.00×10^{-3}
ϵ_{DA}, Grazing limit to the amount of diatoms available for microzoo. grazing, mmol C m^{-3}	1.00×10^{-3}
ϵ_{CR}, Grazing limit to the amount of crypto. available for microzoo. grazing, mmol C m^{-3}	2.95

Table 2. Data types, observed means, coefficient of variation, target errors, and costs before and after optimization. The observed mean ($\bar{a}$), coefficient of variation (CV), and target error (σ) of each assimilated data type used for calculating the cost function before and after optimization. J_0 is the normalized cost function before optimization and J_f is the normalized cost function after optimization (Eq. 6). Data type units: mmol m^{-3} for NO_3, PO_4, diatom Chl, cryptophyte Chl, bacterial biomass, SDOC, and POC; mmol N m^{-3} for PON; and mmol C m^{-3} d^{-1} for PP and BP. The average error (ε_{bias}) of each data type (for non-transformed or raw Chl and PP) is calculated from Stow et al. (2009) before and after optimization where a positive value indicates the model overestimation of the observation and vice versa.

Data types	n	$\bar{a}$	CV	σ	J_0	J_f	$\varepsilon_{bias,0}$	$\varepsilon_{bias,f}$
NO_3	75	21.54	0.04	0.80	0.76	0.90	-0.59	-1.15
PO_4	75	1.43	0.03	0.05	0.40	0.47	0.03	-0.04
$\log_{10}$ diatom Chl	86	-0.07	0.20	0.08	2.29	1.30	-0.37	-0.21
$\log_{10}$ crypto. Chl	86	-0.27	0.24	0.10	3.09	0.56	-0.42	-0.13
$\log_{10}$PP	92	1.30	0.50	0.21	0.73	0.57	-16.6	-17.6
Bacterial biomass	55	0.44	0.08	0.04	4.61	0.58	-0.30	-0.02
BP	55	0.05	0.16	0.01	1.11	0.22	0.04	0.01
SDOC	55	11.39	0.20	2.30	0.66	0.67	0.04	0.10
POC	91	19.78	0.13	2.58	0.50	0.47	-7.43	-7.19
PON	91	2.66	0.12	0.32	0.71	0.68	-1.00	-1.01
Total cost function					14.85	6.42		

Appendix A

1. Temperature effect

$$T_f = \exp\{ -A_E \times (1/T - 1/T_{ref}) \} \quad \text{(A.1.1)}$$

2. Diatom processes

- Cellular quota (ratio):

$$Q^C_{N,DA} = N_{DA}/C_{DA} \quad \text{(A.2.1)}$$

$$Q^C_{P,DA} = P_{DA}/C_{DA} \quad \text{(A.2.2)}$$

$$Q^C_{CHL,DA} = CHL_{DA}/C_{DA} \quad \text{(A.2.3)}$$

- N and P limitation function:

$$N_{f,DA} = (Q^C_{N,DA} - q^C_{N,MIN,DA})/(q^C_{N,RDF,DA} - q^C_{N,MIN,DA}) \qquad 0 \le N_{f,DA} \le 1 \quad \text{(A.2.4)}$$

$$P_{f,DA} = (Q^C_{P,DA} - q^C_{P,MIN,DA})/(q^C_{P,RDF,DA} - q^C_{P,MIN,DA}) \qquad 0 \le P_{f,DA} \le 1 \quad \text{(A.2.5)}$$

- Maximum primary production rate:

$$P^C_{MAX} = \mu_{DA} \times T_f \times \min(N_{f,DA}, P_{f,DA}) \quad \text{(A.2.6)}$$

- C-specific gross primary production:

$$G^C_{DA} = C_{DA} \times P^C_{MAX} \times \{ 1 - \exp(-\alpha_{DA} \times Q^C_{CHL} \times PAR)/P^C_{MAX}) \} \times \exp(-\beta_{DA} \times PAR) \quad \text{(A.2.7)}$$

- Limitation on N and P uptake:

$$V^N_{MAX} = (q^C_{N,MAX} - Q^C_{N,DA})/(q^C_{N,MAX} - q^C_{N,RDF}) \qquad 0 \le V^N_{MAX} \le 1 \quad \text{(A.2.8)}$$

$$V^P_{MAX} = (q^C_{P,MAX} - Q^C_{P,DA})/(q^C_{P,MAX} - q^C_{P,RDF}) \qquad 0 \le V^P_{MAX} \le 1 \quad \text{(A.2.9)}$$

- N assimilation:

$$G^{NH4}_{DA} = C_{DA} \times V^N_{REF} \times T_f \times V^N_{MAX} \times \{ NH4/(NH4 + k^{NH4} + NO3 \times k^{NH4}/k^{NO3}) \} \quad \text{(A.2.10)}$$

$$G^{NO3}_{DA} = C_{DA} \times V^N_{REF} \times T_f \times V^N_{MAX} \times \{ NO3/(NO3 + k^{NO3} + NH4 \times k^{NO3}/k^{NH4}) \} \quad \text{(A.2.11)}$$

$$G^N_{DA} = G^{NH4}_{DA} + G^{NO3}_{DA} \quad \text{(A.2.12)}$$

- P assimilation:

$$G^{PO4}_{DA} = C_{DA} \times V^P_{REF} \times T_f \times V^P_{MAX} \times \{ PO4/(PO4 + k^{PO4}) \} \quad \text{(A.2.13)}$$

- Chlorophyll production:

$$G^{CHL}_{DA} = \theta \times (G^{NH4}_{DA} + G^{NO3}_{DA}) \times \{ G^C_{DA}/\alpha \times CHL_{DA} \times PAR \times \exp(-\beta \times PAR) \} \quad \text{(A.2.14)}$$

- Respiration:

$$R^C_{DA} = G^{NO3}_{DA} \times \zeta^{NO3} \quad \text{(A.2.15)}$$

- Passive excretion of LDOM:

$$E^C_{DA,LDOC,PSV} = ex_{DA,PSV} \times C_{DA} \quad \text{(A.2.16)}$$

$$E^N_{DA,LDON,PSV} = ex_{DA,PSV} \times N_{DA} \quad \text{(A.2.17)}$$

$$E^P_{DA,LDOP,PSV} = ex_{DA,PSV} \times P_{DA} \quad \text{(A.2.18)}$$

- Active excretion of LDOC:

$$E^C_{DA,LDOC,ACT} = ex_{DA,ACT} \times G^C_{DA} \quad \text{(A.2.19)}$$

- Active excretion of SDOC:

$$E^C_{DA,SDOC,ACT} = 0.5 \times C_{DA} \times \max(1 - Q^C_{N,DA}/q^C_{N,RDF,DA}, 1 - Q^C_{P,DA}/q^C_{P,RDF,DA}, 0) \quad \text{(A.2.20)}$$

- Active excretion of SDON and SDOP (if $EX^C_{DA,SDOC,ACT} > 0$, otherwise 0):

$$E^N_{DA,SDON,ACT} = 0.5 \times 0.25 \times N_{DA} \times \max(1 - Q^N_{P,DA}/q^C_{P,RDF,DA}/q^C_{N,RDF,DA}, 0) \quad \text{(A.2.21)}$$

$$E^P_{DA,SDOP,ACT} = 0.5 \times 0.25 \times P_{DA} \times \max(1 - Q^P_{N,DA}/q^C_{N,RDF,DA}/q^C_{P,RDF,DA}, 0) \quad \text{(A.2.22)}$$

- Partitioning between LDOM and SDOM:

$$E^{C}_{DA,LDOC} = E^{C}_{DA,LDOC,PSV} + 0.75 \times E^{C}_{DA,LDOC,ACT} \quad \text{(A.2.23)}$$

$$E^{N}_{DA,LDON} = E^{N}_{DA,LDON,PSV} \quad \text{(A.2.24)}$$

$$E^{P}_{DA,LDOP} = E^{P}_{DA,LDOP,PSV} \quad \text{(A.2.25)}$$

$$E^{C}_{DA,SDOC} = E^{C}_{DA,SDOC,ACT} + 0.25 \times E^{C}_{DA,LDOC,ACT} \quad \text{(A.2.26)}$$

$$E^{N}_{DA,SDON} = E^{N}_{DA,SDON,ACT} \quad \text{(A.2.27)}$$

$$E^{P}_{DA,SDOP} = E^{P}_{DA,SDOP,ACT} \quad \text{(A.2.28)}$$

- POM production by aggregation:

$$D^{C}_{DA} = pom_{DA} \times C_{DA} \times C_{DA} \quad \text{(A.2.29)}$$

$$D^{N}_{DA} = Q^{C}_{N,DA} \times D^{C}_{DA} \quad \text{(A.2.30)}$$

$$D^{P}_{DA} = Q^{C}_{P,DA} \times D^{C}_{DA} \quad \text{(A.2.31)}$$

$$D^{CHL}_{DA} = Q^{C}_{CHL,DA} \times D^{C}_{DA} \quad \text{(A.2.32)}$$

- Grazing by microzooplankton:

$$GZ^{C}_{DA,MZ} = T_f \times \mu_{MZ} \times C_{MZ} \times [\,(C_{DA} - \epsilon_{DA})^2 / \{(C_{DA} - \epsilon_{DA})^2 + g_{DA}^2 + (C_{CRYPTO} \times g_{DA}/g_{CR})^2 + (C_{BAC} \times g_{DA}/g_{BAC})^2\}\,] \quad \text{(A.2.33)}$$

$$GZ^{N}_{DA,MZ} = Q^{C}_{N,DA} \times GZ^{C}_{DA,MZ} \quad \text{(A.2.34)}$$

$$GZ^{P}_{DA,MZ} = Q^{C}_{P,DA} \times GZ^{C}_{DA,MZ} \quad \text{(A.2.35)}$$

$$GZ^{CHL}_{DA,MZ} = Q^{C}_{CHL,DA} \times GZ^{C}_{DA,MZ} \quad \text{(A.2.36)}$$

- Grazing by krill:

$$GZ^{C}_{DA,KR} = T_f \times \mu_{KR} \times C_{KR} \times [\,C_{DA}^2 / \{C_{DA}^2 + g'^{2}_{DA} + (C_{MZ} \times g'_{DA}/g_{MZ})^2\}\,] \quad \text{(A.2.37)}$$

$$GZ^{N}_{DA,KR} = Q^{C}_{N,DA} \times GZ^{C}_{DA,KR} \quad \text{(A.2.38)}$$

$$GZ^{P}_{DA,KR} = Q^{C}_{P,DA} \times GZ^{C}_{DA,KR} \quad \text{(A.2.39)}$$

$$GZ^{CHL}_{DA,KR} = Q^{C}_{CHL,DA} \times GZ^{C}_{DA,KR} \quad \text{(A.2.40)}$$

- The net growth rate equations:

$$\frac{dC_{DA}}{dt} = G^{C}_{DA} - E^{C}_{DA,LDOC} - E^{C}_{DA,SDOC} - D^{C}_{DA} - R^{C}_{DA} - GZ^{C}_{DA,MZ} - GZ^{C}_{DA,KR} \quad \text{(A.2.41)}$$

$$\frac{dN_{DA}}{dt} = G^{N}_{DA} - E^{N}_{DA,LDON} - E^{N}_{DA,SDON} - D^{N}_{DA} - GZ^{N}_{DA,MZ} - GZ^{N}_{DA,KR} \quad \text{(A.2.42)}$$

$$\frac{dP_{DA}}{dt} = G^{P}_{DA} - E^{P}_{DA,LDOP} - E^{P}_{DA,SDOP} - D^{N}_{DA} - GZ^{P}_{DA,MZ} - GZ^{P}_{DA,KR} \quad \text{(A.2.43)}$$

$$\frac{dCHL_{DA}}{dt} = G^{CHL}_{DA} - D^{CHL}_{DA} - GZ^{CHL}_{DA,MZ} - GZ^{CHL}_{DA,KR} \quad \text{(A.2.44)}$$

3. Cryptophyte processes

- Cellular quota (ratio):

$$Q^{C}_{N,CR} = N_{CR}/C_{CR} \quad \text{(A.3.1)}$$

$$Q^{C}_{P,CR} = P_{CR}/C_{CR} \quad \text{(A.3.2)}$$

$$Q^{C}_{CHL,CR} = CHL_{CR}/C_{CR} \quad \text{(A.3.3)}$$

- N and P limitation function:

$$N_{f,CR} = (Q^{C}_{N,CR} - q^{C}_{N,MIN,CR})/(q^{C}_{N,RDF,CR} - q^{C}_{N,MIN,CR}) \qquad 0 \le N_{f,CR} \le 1 \quad \text{(A.3.4)}$$

$$P_{f,CR} = (Q^C_{P,CR} - q^C_{P,MIN,CR})/(q^C_{P,RDF,CR} - q^C_{P,MIN,CR}) \qquad 0 \le P_{f,CR} \le 1$$ (A.3.5)

- Maximum primary production rate:

$$P^C_{MAX} = \mu_{CR} \times T_f \times \min(N_{f,CR}, P_{f,CR})$$ (A.3.6)

- C-specific gross primary production:

$$G^C_{CR} = C_{CR} \times P^C_{MAX} \times \{ 1 - \exp(-\alpha_{CR} \times Q^C_{CHL} \times PAR)/P^C_{MAX}) \} \times \exp(-\beta_{CR} \times PAR)$$ (A.3.7)

- Limitation on N and P uptake:

$$V^N_{MAX} = (q^C_{N,MAX} - Q^C_{N,CR})/(q^C_{N,MAX} - q^C_{N,RDF}) \quad 0 \le V^N_{MAX} \le 1$$ (A.3.8)

$$V^P_{MAX} = (q^C_{P,MAX} - Q^C_{P,CR})/(q^C_{P,MAX} - q^C_{P,RDF}) \qquad 0 \le V^P_{MAX} \le 1$$ (A.3.9)

- Nitrogen assimilation:

$$G^{NH4}_{CR} = C_{CR} \times V^N_{REF} \times T_f \times V^N_{MAX} \times \{ NH4/(NH4 + k^{NH4} + NO3 \times k^{NH4}/k^{NO3}) \}$$ (A.3.10)

$$G^{NO3}_{CR} = C_{CR} \times V^N_{REF} \times T_f \times V^N_{MAX} \times \{ NO3/(NO3 + k^{NO3} + NH4 \times k^{NO3}/k^{NH4}) \}$$ (A.3.11)

$$G^N_{CR} = G^{NH4}_{CR} + G^{NO3}_{CR}$$ (A.3.12)

- Phosphorus assimilation:

$$G^{PO4}_{CR} = C_{CR} \times V^P_{REF} \times T_f \times V^P_{MAX} \times \{ PO4/(PO4 + k^{PO4}) \}$$ (A.3.13)

- Chlorophyll production:

$$G^{CHL}_{CR} = \theta \times (G^{NH4}_{CR} + G^{NO3}_{CR}) \times \{ G^C_{CR}/\alpha \times CHL_{CR} \times PAR \times \exp(-\beta \times PAR) \}$$ (A.3.14)

- Respiration:

$$R^C_{CR} = G^{NO3}_{CR} \times \zeta^{NO3}$$ (A.3.15)

- Passive excretion of LDOM:

$$E^C_{CR,LDOC,PSV} = ex_{CR,PSV} \times C_{CR}$$ (A.3.16)

$$E^N_{CR,LDON,PSV} = ex_{CR,PSV} \times N_{CR}$$ (A.3.17)

$$E^P_{CR,LDOP,PSV} = ex_{CR,PSV} \times P_{CR}$$ (A.3.18)

- Active excretion of LDOC:

$$E^C_{CR,LDOC,ACT} = ex_{CR,ACT} \times G^C_{CR}$$ (A.3.19)

- Active excretion of SDOC:

$$E^C_{CR,SDOC,ACT} = 0.5 \times C_{CR} \times \max(1 - Q^C_{N,CR}/q^C_{N,RDF,CR}, 1 - Q^C_{P,CR}/q^C_{P,RDF,CR}, 0)$$ (A.3.20)

- Active excretion of SDON and SDOP (if $EX^C_{CR,SDOC,ACT} > 0$, otherwise 0):

$$E^N_{CR,SDON,ACT} = 0.5 \times 0.25 \times N_{CR} \times \max(1 - Q^N_{P,CR}/q^C_{P,RDF,CR}/q^C_{N,RDF,CR}, 0)$$ (A.3.21)

$$E^P_{CR,SDOP,ACT} = 0.5 \times 0.25 \times P_{CR} \times \max(1 - Q^P_{N,CR}/q^C_{N,RDF,CR}/q^C_{P,RDF,CR}, 0)$$ (A.3.22)

- Partitioning between LDOM and SDOM:

$$E^C_{CR,LDOC} = E^C_{CR,LDOC,PSV} + 0.75 \times E^C_{CR,LDOC,ACT}$$ (A.3.23)

$$E^N_{CR,LDON} = E^N_{CR,LDON,PSV}$$ (A.3.24)

$$E^P_{CR,LDOP} = E^P_{CR,LDOP,PSV}$$ (A.3.25)

$$E^C_{CR,SDOC} = E^C_{CR,SDOC,ACT} + 0.25 \times E^C_{CR,LDOC,ACT}$$ (A.3.26)

$$E^N_{CR,SDON} = E^N_{CR,SDON,ACT}$$ (A.3.27)

$$E^P_{CR,SDOP} = E^P_{CR,SDOP,ACT}$$ (A.3.28)

- POM production by aggregation:

$$D^C_{CR} = pom_{CR} \times C_{CR} \times C_{CR}$$ (A.3.29)

$$D^N_{CR} = Q^C_{N,CR} \times A^C_{CR}$$ (A.3.30)

$$D^P_{CR} = Q^C_{P,CR} \times A^C_{CR}$$ (A.3.31)

$$D^{CHL}_{CR} = Q^C_{CHL,CR} \times A^C_{CR}$$ (A.3.32)

- Grazing by microzooplankton:

$$GZ^{C}_{CR} = T_f \times \mu_{MZ} \times C_{MZ}$$
$$\times [\ (C_{CR} - \epsilon_{CR})^2 / \{(C_{CR} - \epsilon_{CR})^2 + g_{CR}^2 + (C_{DA} \times g_{CR}/g_{DA})^2 + (C_{BAC} \times g_{CR}/g_{BAC})^2\}\ ] \quad \text{(A.3.33)}$$

$$GZ^{N}_{CR} = Q^{C}_{N,CR} \times GZ^{C}_{CR,MZ} \quad \text{(A.3.34)}$$

$$GZ^{P}_{CR} = Q^{C}_{P,CR} \times GZ^{C}_{CR,MZ} \quad \text{(A.3.35)}$$

$$GZ^{CHL}_{CR} = Q^{C}_{CHL,CR} \times GZ^{C}_{CR,MZ} \quad \text{(A.3.36)}$$

- The net growth rate equations:

$$\frac{dC_{CR}}{dt} = G^{C}_{CR} - E^{C}_{CR,LDOC} - E^{C}_{CR,SDOC} - D^{C}_{CR} - R^{C}_{CR} - GZ^{C}_{CR} \quad \text{(A.3.37)}$$

$$\frac{dN_{CR}}{dt} = G^{N}_{CR} - E^{N}_{CR,LDON} - E^{N}_{CR,SDON} - D^{N}_{CR} \quad - GZ^{N}_{CR} \quad \text{(A.3.38)}$$

$$\frac{dP_{CR}}{dt} = G^{P}_{CR} - E^{P}_{CR,LDOP} - E^{P}_{CR,SDOP} - D^{N}_{CR} \quad - GZ^{P}_{CR} \quad \text{(A.3.39)}$$

$$\frac{dCHL_{CR}}{dt} = G^{CHL}_{CR} \quad - D^{CHL}_{CR} \quad - GZ^{CHL}_{CR} \quad \text{(A.3.40)}$$

4. Bacterial processes

- Cellular quota (ratio):

$$Q^{C}_{N,BAC} = N_{BAC}/C_{BAC} \quad \text{(A.4.1)}$$
$$Q^{C}_{P,BAC} = P_{BAC}/C_{BAC} \quad \text{(A.4.2)}$$
$$Q^{P}_{N,BAC} = N_{BAC}/P_{BAC} \quad \text{(A.4.3)}$$
$$Q^{C}_{N,LDOM} = N_{LDOM}/C_{LDOM} \quad \text{(A.4.4)}$$
$$Q^{C}_{P,LDOM} = P_{LDOM}/C_{LDOM} \quad \text{(A.4.5)}$$
$$Q^{C}_{N,SDOM} = N_{SDOM}/C_{SDOM} \quad \text{(A.4.6)}$$
$$Q^{C}_{P,SDOM} = P_{SDOM}/C_{SDOM} \quad \text{(A.4.7)}$$

- N and P limitation function:

$$N_{f,BAC} = Q^{C}_{N,BAC}/q^{C}_{N,BAC} \qquad 0 \le N_{f,BAC} \le 1 \quad \text{(A.4.8)}$$
$$P_{f,BAC} = Q^{C}_{P,BAC}/q^{C}_{P,BAC} \qquad 0 \le P_{f,BAC} \le 1 \quad \text{(A.4.9)}$$

- Maximum available LDOC and SDOC:

$$ALC = C_{LDOC} \quad \text{(A.4.10)}$$
$$ASC = r_{SDOC} \times C_{SDOC} \quad \text{(A.4.11)}$$

- Bacterial uptake of LDOC and SDOC (i.e., bacterial gross C growth):

$$G^{C}_{BAC,LDOC} = \mu_{BAC} \times T_f \times C_{BAC} \times \min(N_{f,BAC}, P_{f,BAC}) \times \{\ ALC/(ALC + k^{DOC} + ASC)\ \} \quad \text{(A.4.12)}$$
$$G^{C}_{BAC,SDOC} = \mu_{BAC} \times T_f \times C_{BAC} \times \min(N_{f,BAC}, P_{f,BAC}) \times \{\ ASC/(ASC + k^{DOC} + ALC)\ \} \quad \text{(A.4.13)}$$
$$G^{C}_{BAC,DOC} = G^{C}_{BAC,LDOC} + G^{C}_{BAC,SDOC} \quad \text{(A.4.14)}$$

- Bacterial N uptake:

$$G^{C}_{BAC,LDON} = G^{C}_{BAC,LDOC} \times Q^{C}_{N,LDOM} \quad \text{(A.4.15)}$$
$$G^{C}_{BAC,SDON} = G^{C}_{BAC,SDOC} \times \min\{\ q^{C}_{N,BAC},\ Q^{C}_{N,SDOM} + fs/N_{f,BAC} \times (q^{C}_{N,BAC} - Q^{C}_{N,SDOM})\ \} \quad \text{(A.4.16)}$$
$$G^{C}_{BAC,NH4} = G^{C}_{BAC,LDON} \times NH4/N_{LDOM}/\min(1, N_{f,BAC}) \quad \text{(A.4.17)}$$

if $N_{f,BAC} < 1$,

$$G^{C}_{BAC,NO3} = \min\{\ 0.1 \times NO3 \times 1/\min(1, N_{f,BAC}) \times (G^{C}_{BAC,LDON} + G^{C}_{BAC,SDON})/(N_{LDOM} + N_{SDOM}),$$
$$(NO3 + NH4) \times (G^{N}_{BAC,LDON} + G^{N}_{BAC,SDON})/(N_{LDOM} + N_{SDOM}) - G^{NH4}_{BAC}\ \} \quad \text{(A.4.18)}$$

else, $G^{C}_{BAC,NO3} = 0$ (A.4.19)

$$G^{C}_{BAC,N} = G^{C}_{BAC,LDON} + G^{C}_{BAC,SDON} + G^{C}_{BAC,NH4} + G^{C}_{BAC,NO3} \quad \text{(A.4.20)}$$

- Bacterial P uptake:

$$G^{C}_{BAC,LDOP} = G^{C}_{BAC,LDOC} \times Q^{C}_{P,LDOM} \quad \text{(A.4.21)}$$

$$G^{C}_{BAC,SDOP} = G^{C}_{BAC,SDOC} \times \min\{\, q^{C}_{P,BAC},\ Q^{C}_{P,SDOM} + f_S/P_{f,BAC} \times (q^{C}_{P,BAC} - Q^{C}_{P,SDOM}) \,\} \quad \text{(A.4.22)}$$

$$G^{C}_{BAC,PO4} = G^{C}_{BAC,LDON} \times PO4/P_{LDOM}/\min(1,\ P_{f,BAC}) \quad \text{(A.4.23)}$$

$$G^{C}_{BAC,P} = G^{C}_{BAC,LDOP} + G^{C}_{BAC,SDOP} + G^{C}_{BAC,PO4} \quad \text{(A.4.24)}$$

- Respiration:

$$R^{C}_{BAC} = \zeta^{NO3} \times G^{C}_{BAC,NO3} + r^{B}_{BAC} \times T_f \times C_{BAC} + \{\, r^{A}_{min,BAC} + (r^{A}_{max,BAC} - r^{A}_{min,BAC}) \times \exp(-b_{R,BAC} \times G^{C}_{BAC,DOC}) \,\} \times G^{C}_{BAC,DOC} \quad \text{(A.4.25)}$$

- RDOC release:

$$E^{C}_{BAC,RDOC} = refr_{BAC} \times C_{BAC} \quad \text{(A.4.26)}$$

$$E^{N}_{BAC,RDON} = E^{C}_{BAC,RDOC} \times q^{C}_{N,RDOM} \quad \text{(A.4.27)}$$

$$E^{P}_{BAC,RDOP} = E^{C}_{BAC,RDOC} \times q^{C}_{P,RDOM} \quad \text{(A.4.28)}$$

- Remineralization of inorganic nutrients:

if $Q^{C}_{N,BAC} > q^{C}_{N,BAC}$ and $Q^{C}_{P,BAC} > q^{C}_{P,BAC}$ (i.e., C in short)

$$REMI^{N}_{BAC} = remi_{BAC} \times (N_{BAC} - C_{BAC} \times q^{C}_{N,BAC}) \quad \text{(A.4.29)}$$

$$REMI^{P}_{BAC} = remi_{BAC} \times (P_{BAC} - C_{BAC} \times q^{C}_{P,BAC}) \quad \text{(A.4.30)}$$

$$remi_{BAC}$$

elseif $Q^{C}_{N,BAC} < q^{C}_{N,BAC}$ and $Q^{P}_{N,BAC} < q^{C}_{N,BAC}/q^{C}_{P,BAC}$ (i.e., N in short)

$$REMI^{N}_{BAC} = 0 \quad \text{(A.4.31)}$$

$$REMI^{P}_{BAC} = 0 \quad \text{(A.4.32)}$$

else (i.e., P in short)

$$REMI^{N}_{BAC} = 0 \quad \text{(A.4.33)}$$

$$REMI^{P}_{BAC} = 0 \quad \text{(A.4.34)}$$

- SDOM excretion to adjust stoichiometry:

if $Q^{C}_{N,BAC} > q^{C}_{N,BAC}$ and $Q^{C}_{P,BAC} > q^{C}_{P,BAC}$ (i.e., C in short)

$$E^{C}_{BAC,SDOC} = 0 \quad \text{(A.4.35)}$$

$$E^{N}_{BAC,SDON} = 0 \quad \text{(A.4.36)}$$

$$E^{P}_{BAC,SDOP} = 0 \quad \text{(A.4.37)}$$

elseif $Q^{C}_{N,BAC} < q^{C}_{N,BAC}$ and $Q^{P}_{N,BAC} < q^{C}_{N,BAC}/q^{C}_{P,BAC}$ (i.e., N in short)

$$E^{C}_{BAC,SDOC} = ex_{ADJ,BAC} \times (C_{BAC} - N_{BAC}/q^{C}_{N,BAC}) \quad \text{(A.4.38)}$$

$$E^{N}_{BAC,SDOC} = 0 \quad \text{(A.4.39)}$$

$$E^{P}_{BAC,SDOP} = ex_{ADJ,BAC} \times (P_{BAC} - N_{BAC}/q^{C}_{N,BAC} \times q^{C}_{P,BAC}) \quad \text{(A.4.40)}$$

else (i.e., P in short)

$$E^{C}_{BAC,SDOC} = ex_{ADJ,BAC} \times (C_{BAC} - P_{BAC}/q^{C}_{P,BAC}) \quad \text{(A.4.41)}$$

$$E^{N}_{BAC,SDON} = ex_{ADJ,BAC} \times (N_{BAC} - P_{BAC}/q^{C}_{P,BAC} \times q^{C}_{N,BAC}) \quad \text{(A.4.42)}$$

$$E^{N}_{BAC,SDOP} = 0 \quad \text{(A.4.43)}$$

- Grazing by microzooplankton:

$$GZ^{C}_{BAC} = T_f \times \mu_{MZ} \times C_{MZ} \times [\, C_{BAC}^2/\{C_{CR}^2 + g_{BAC}^2 + (C_{DA} \times g_{BAC}/g_{DA})^2 + (C_{CR} \times g_{BAC}/g_{CR})^2\} \,] \quad \text{(A.4.44)}$$

$$GZ^{N}_{BAC} = GZ^{C}_{BAC} \times Q^{C}_{N,BAC} \quad \text{(A.4.45)}$$

$$GZ^{P}_{BAC} = GZ^{C}_{BAC} \times Q^{C}_{P,BAC} \quad \text{(A.4.46)}$$

- Viral mortality:

$$M^{C}_{BAC} = m_{BAC} \times C_{BAC} \quad \text{(A.4.47)}$$

$$M^{N}_{BAC} = m_{BAC} \times N_{BAC} \quad \text{(A.4.48)}$$

$$M^{P}_{BAC} = m_{BAC} \times P_{BAC} \quad \text{(A.4.49)}$$

- Net flux of inorganic nutrients through bacteria:

$$FLUX^{NH4}_{BAC} = REMI^{N}_{BAC} - G^{C}_{BAC,NH4} \quad \text{(A.4.50)}$$

$$FLUX^{NO3}_{BAC} = -G^{C}_{BAC,NO3} \quad \text{(A.4.51)}$$

$$FLUX^{PO4}_{BAC} = REMI^{P}_{BAC} - G^{C}_{BAC,PO4} \quad \text{(A.4.52)}$$

- The net growth rate equations:

$$\frac{dC_{BAC}}{dt} = G^{C}_{BAC,DOC} - E^{C}_{BAC,SDOC} - E^{C}_{BAC,RDOC} - R^{C}_{BAC} - GZ^{C}_{BAC} - M^{C}_{BAC} \quad \text{(A.4.53)}$$

$$\frac{dN_{BAC}}{dt} = G^{N}_{BAC,DON} - E^{N}_{BAC,SDON} - E^{N}_{BAC,RDON} - R^{N}_{BAC} - GZ^{N}_{BAC} - M^{N}_{BAC} \quad \text{(A.4.54)}$$

$$\frac{dP_{BAC}}{dt} = G^{P}_{BAC,DOP} - E^{P}_{BAC,SDOP} - E^{P}_{BAC,RDOP} - R^{P}_{BAC} - GZ^{P}_{BAC} - M^{P}_{BAC} \quad \text{(A.4.55)}$$

5. Microzooplankton processes

- Cellular quota (ratio):

$$Q^{C}_{N,MZ} = N_{MZ}/C_{MZ} \quad \text{(A.5.1)}$$

$$Q^{C}_{P,MZ} = C_{MZ}/P_{MZ} \quad \text{(A.5.2)}$$

- Gross growth:

$$G^{C}_{MZ} = GZ^{C}_{CR} + GZ^{C}_{BAC} \quad \text{(A.5.3)}$$

$$G^{N}_{MZ} = GZ^{N}_{CR} + GZ^{N}_{BAC} \quad \text{(A.5.4)}$$

$$G^{P}_{MZ} = GZ^{P}_{CR} + GZ^{P}_{BAC} \quad \text{(A.5.5)}$$

- LDOM excretion:

$$E^{C}_{MZ,LDOC} = f_{ex,MZ} \times ex_{MZ} \times G^{C}_{MZ} \quad \text{(A.5.6)}$$

$$E^{N}_{MZ,LDON} = f_{ex,MZ} \times ex_{MZ} \times G^{N}_{MZ} \quad \text{(A.5.7)}$$

$$E^{P}_{MZ,LDOP} = f_{ex,MZ} \times ex_{MZ} \times G^{P}_{MZ} \quad \text{(A.5.8)}$$

- SDOM excretion:

$$E^{C}_{MZ,SDOC,1} = (1 - f_{ex,MZ}) \times ex_{MZ} \times G^{C}_{MZ} \quad \text{(A.5.9)}$$

$$E^{N}_{MZ,SDON,1} = (1 - f_{ex,MZ}) \times ex_{MZ} \times G^{N}_{MZ} \times Q^{C}_{N,MZ}/q^{C}_{N,MZ} \quad \text{(A.5.10)}$$

$$E^{P}_{MZ,SDOP,1} = (1 - f_{ex,MZ}) \times ex_{MZ} \times G^{P}_{MZ} \times Q^{C}_{P,MZ}/q^{C}_{P,MZ} \quad \text{(A.5.11)}$$

- SDOM excretion to adjust stoichiometry:

$$E^{C}_{MZ,SDOC,2} = ex_{ADJ,MZ} \times C_{MZ} \times \max(0, 1 - Q^{C}_{N,MZ}/q^{C}_{N,MZ}, 1 - Q^{C}_{P,MZ}/q^{C}_{P,MZ}) \quad \text{(A.5.12)}$$

$$E^{N}_{MZ,SDON,2} = 0.5 \times E^{C}_{MZ,SDOC,2} \times Q^{C}_{N,MZ} \quad \text{(A.5.13)}$$

$$E^{P}_{MZ,SDOP,2} = 0.5 \times E^{C}_{MZ,SDOC,2} \times Q^{C}_{P,MZ} \quad \text{(A.5.14)}$$

- Remineralization of inorganic nutrients:

$$REMI^{N}_{MZ} = remi_{MZ} \times \max(0, N_{MZ} - C_{MZ} \times q^{C}_{N,MZ}, N_{MZ} - q^{C}_{N,MZ}/P_{MZ} \times q^{C}_{P,MZ}) \quad \text{(A.5.15)}$$

$$REMI^{P}_{MZ} = remi_{MZ} \times \max(0, P_{MZ} - C_{MZ} \times q^{C}_{P,MZ}, P_{MZ} - q^{C}_{P,MZ}/N_{MZ} \times q^{C}_{N,MZ}) \quad \text{(A.5.16)}$$

- Respiration:

$$R^C_{MZ} = r^B_{MZ} \times T_f \times C_{MZ} + r^A_{MZ} \times G^C_{MZ} \quad \text{(A.5.17)}$$

- POM production:

$$P^C_{MZ} = pom_{MZ} \times G^C_{MZ} \quad \text{(A.5.18)}$$

$$P^N_{MZ} = q^C_{N,POM} \times G^C_{MZ} \quad \text{(A.5.19)}$$

$$P^P_{MZ} = q^C_{P,POM} \times G^C_{MZ} \quad \text{(A.5.20)}$$

- Grazing by krill:

$$GZ^C_{MZ} = T_f \times \mu_{MZ} \times C_{KR} \times [\, C_{MZ}^2 / \{C_{MZ}^2 + g_{MZ}^2 + (C_{DA} \times g_{MZ}/g_{DA})^2\} \,] \quad \text{(A.5.21)}$$

$$GZ^N_{MZ} = Q^C_{N,MZ} \times GZ^C_{MZ} \quad \text{(A.5.22)}$$

$$GZ^P_{MZ} = Q^C_{P,MZ} \times GZ^C_{MZ} \quad \text{(A.5.23)}$$

- The net growth rate equations:

$$\frac{dC_{MZ}}{dt} = G^C_{MZ} - E^C_{MZ,LDOC} - E^C_{MZ,SDOC,1} - E^C_{MZ,SDOC,2} - P^C_{MZ} - R^C_{MZ} - GZ^C_{MZ} \quad \text{(A.5.24)}$$

$$\frac{dN_{MZ}}{dt} = G^N_{MZ} - E^N_{MZ,LDON} - E^N_{MZ,SDON,1} - E^N_{MZ,SDON,2} - P^N_{MZ} - R^N_{MZ} - GZ^N_{MZ} \quad \text{(A.5.25)}$$

$$\frac{dP_{MZ}}{dt} = G^P_{MZ} - E^P_{MZ,LDOP} - E^P_{MZ,SDOP,1} - E^P_{MZ,SDOP,2} - P^P_{MZ} - R^P_{MZ} - GZ^P_{MZ} \quad \text{(A.5.26)}$$

6. Krill processes

- Cellular quota (ratio):

$$Q^C_{N,KR} = N_{KR}/C_{KR} \quad \text{(A.6.1)}$$

$$Q^C_{P,KR} = C_{KR}/P_{KR} \quad \text{(A.6.2)}$$

- Gross growth:

$$G^C_{KR} = GZ^C_{DA,KR} + GZ^C_{MZ} \quad \text{(A.6.3)}$$

$$G^N_{KR} = GZ^N_{DA,KR} + GZ^N_{MZ} \quad \text{(A.6.4)}$$

$$G^P_{KR} = GZ^P_{DA,KR} + GZ^P_{MZ} \quad \text{(A.6.5)}$$

- LDOM excretion:

$$E^C_{KR,LDOC} = f_{ex,KR} \times ex_{KR} \times G^C_{KR} \quad \text{(A.6.6)}$$

$$E^N_{KR,LDON} = f_{ex,KR} \times ex_{KR} \times G^N_{KR} \quad \text{(A.6.7)}$$

$$E^P_{KR,LDOP} = f_{ex,KR} \times ex_{KR} \times G^P_{KR} \quad \text{(A.6.8)}$$

- SDOM excretion:

$$E^C_{KR,SDOC,1} = (1 - f_{ex,KR}) \times ex_{KR} \times G^C_{KR} \quad \text{(A.6.9)}$$

$$E^N_{KR,SDON,1} = (1 - f_{ex,KR}) \times ex_{KR} \times G^N_{KR} \times Q^C_{N,KR}/q^C_{N,KR} \quad \text{(A.6.10)}$$

$$E^P_{KR,SDOP,1} = (1 - f_{ex,KR}) \times ex_{KR} \times G^P_{KR} \times Q^C_{P,KR}/q^C_{P,KR} \quad \text{(A.6.11)}$$

- SDOM excretion to adjust stoichiometry:

$$E^C_{KR,SDOC,2} = ex_{ADJ,KR} \times C_{KR} \times \max(0,\, 1 - Q^C_{N,KR}/q^C_{N,KR},\, 1 - Q^C_{P,KR}/q^C_{P,KR}) \quad \text{(A.6.12)}$$

$$E^N_{KR,SDON,2} = 0.5 \times E^C_{KR,SDOC,2} \times Q^C_{N,KR} \quad \text{(A.6.13)}$$

$$E^{P}_{KR,SDOP,2} = 0.5 \times E^{C}_{KR,SDOC,2} \times Q^{C}_{P,KR} \quad \text{(A.6.14)}$$

- Remineralization of inorganic nutrients:

$$REMI^{N}_{KR} = remi_{KR} \times \max(0, N_{KR} - C_{KR} \times q^{C}_{N,KR}, N_{KR} - q^{C}_{N,KR}/P_{KR} \times q^{C}_{P,KR}) \quad \text{(A.6.15)}$$

$$REMI^{P}_{KR} = remi_{KR} \times \max(0, P_{KR} - C_{KR} \times q^{C}_{P,KR}, P_{KR} - q^{C}_{P,KR}/N_{KR} \times q^{C}_{N,KR}) \quad \text{(A.6.16)}$$

- Respiration:

$$R^{C}_{KR} = r^{B}_{KR} \times T_f \times C_{KR} + r^{A}_{KR} \times G^{C}_{KR} \quad \text{(A.6.17)}$$

- POM production:

$$P^{C}_{KR} = pom_{KR} \times G^{C}_{KR} \quad \text{(A.6.18)}$$

$$P^{N}_{KR} = q^{C}_{N,POM} \times G^{N}_{KR} \quad \text{(A.6.19)}$$

$$P^{P}_{KR} = q^{C}_{P,POM} \times G^{P}_{KR} \quad \text{(A.6.20)}$$

- RDOC release:

$$E^{C}_{KR,RDOC} = refr_{KR} \times C_{KR} \quad \text{(A.6.21)}$$

$$E^{N}_{KR,RDON} = E^{C}_{KR,RDOC} \times q^{C}_{N,RDOM} \quad \text{(A.6.22)}$$

$$E^{P}_{KR,RDOP} = E^{C}_{KR,RDOC} \times q^{C}_{P,RDOM} \quad \text{(A.6.23)}$$

- Removal by higher trophic levels

$$M^{C}_{KR} = mort_{KR} \times C_{KR} \times C_{KR} \quad \text{(A.6.24)}$$

$$M^{N}_{KR} = M^{C}_{KR,RDOC} \times Q^{C}_{N,KR} \quad \text{(A.6.25)}$$

$$M^{P}_{KR} = M^{C}_{KR,RDOC} \times Q^{C}_{P,KR} \quad \text{(A.6.26)}$$

- The net growth rate equations:

$$\frac{dC_{KR}}{dt} = G^{C}_{KR} - E^{C}_{KR,LDOC} - E^{C}_{KR,SDOC,1} - E^{C}_{KR,SDOC,2} - E^{C}_{KR,RDOC} - P^{C}_{KR} - R^{C}_{KR} - M^{C}_{KR} \quad \text{(A.6.27)}$$

$$\frac{dN_{KR}}{dt} = G^{N}_{KR} - E^{N}_{KR,LDON} - E^{N}_{KR,SDON,1} - E^{N}_{KR,SDON,2} - E^{N}_{KR,RDON} - P^{N}_{KR} - R^{N}_{KR} - M^{N}_{KR} \quad \text{(A.6.28)}$$

$$\frac{dP_{KR}}{dt} = G^{P}_{KR} - E^{P}_{KR,LDOC} - E^{P}_{KR,SDOC,1} - E^{P}_{KR,SDOC,2} - E^{P}_{KR,RDOC} - P^{P}_{KR} - R^{P}_{KR} - M^{P}_{KR} \quad \text{(A.6.29)}$$

7. Detrital processes

- Dissolution:

$$DISS^{C}_{DET} = diss \times C_{DET} \quad \text{(A.7.1)}$$

$$DISS^{N}_{DET} = diss \times prf_N \times N_{DET} \quad \text{(A.7.2)}$$

$$DISS^{P}_{DET} = diss \times prf_P \times P_{DET} \quad \text{(A.7.3)}$$

- The net change equations:

$$\frac{dC_{DET}}{dt} = D^{C}_{DA} + D^{C}_{CR} + D^{C}_{MZ} + D^{C}_{KR} + DISS^{C}_{HZ} - DISS^{C}_{DET} \quad \text{(A.7.4)}$$

$$\frac{dN_{DET}}{dt} = D^{N}_{DA} + D^{N}_{CR} + D^{N}_{MZ} + D^{N}_{KR} + DISS^{N}_{HZ} - DISS^{N}_{DET} \quad \text{(A.7.5)}$$

$$\frac{dP_{DET}}{dt} = D^{P}_{DA} + D^{P}_{CR} + D^{P}_{MZ} + D^{P}_{KR} + DISS^{P}_{HZ} - DISS^{P}_{DET} \quad \text{(A.7.6)}$$

where $DISS^{C}_{HZ} = f_{POM,HZ} \times M^{C}_{KR}$

$DISS^{N}_{HZ} = f_{POM,HZ} \times M^{N}_{KR}$

$DISS^{P}_{HZ} = f_{POM,HZ} \times M^{P}_{KR}$

8. DOM processes

- Conversion of SDOM to RDOM:

$$REFR^{C}_{SDOM} = ex_{REFR,SDOM} \times C_{SDOM} \times \exp\{ 1 - \min(Q^{C}_{N,SDOM}/q^{C}_{N,RDOM}, Q^{C}_{P,SDOM}/q^{C}_{P,RDOM}) \} \quad \text{(A.8.1)}$$

$$REFR^{N}_{SDOM} = REFR^{C}_{SDOM} \times q^{C}_{N,RDOM} \quad \text{(A.8.2)}$$

$$REFR^{P}_{SDOM} = REFR^{C}_{SDOM} \times q^{C}_{P,RDOM} \quad \text{(A.8.3)}$$

- The net change equations:

$$\frac{dC_{LDOM}}{dt} = E^{C}_{DA,LDOC} + E^{C}_{CR,LDOC} + E^{C}_{MZ,LDOC} + E^{C}_{KR,LDOC} + M^{C}_{BAC} - G^{C}_{BAC,LDOC} \quad \text{(A.8.4)}$$

$$\frac{dN_{LDOM}}{dt} = E^{N}_{DA,LDON} + E^{N}_{CR,LDON} + E^{N}_{MZ,LDON} + E^{N}_{KR,LDON} + M^{N}_{BAC} - G^{N}_{BAC,LDON} \quad \text{(A.8.5)}$$

$$\frac{dP_{LDOM}}{dt} = E^{P}_{DA,LDOP} + E^{P}_{CR,LDOP} + E^{P}_{MZ,LDOP} + E^{P}_{KR,LDOP} + M^{P}_{BAC} - G^{P}_{BAC,LDOP} \quad \text{(A.8.6)}$$

$$\frac{dC_{SDOM}}{dt} = E^{C}_{DA,SDOC} + E^{C}_{CR,SDOC} + E^{C}_{BAC,SDOC} + E^{C}_{MZ,SDOC,1} + E^{C}_{MZ,SDOC,2} + E^{C}_{KR,SDOC,1} + E^{C}_{KR,SDOC,2} + E^{C}_{HZ,SDOC} + DISS^{C}_{DET} - REFR^{C}_{SDOM} - G^{C}_{BAC,SDOC} \quad \text{(A.8.7)}$$

$$\frac{dN_{SDOM}}{dt} = E^{N}_{DA,SDON} + E^{N}_{CR,SDON} + E^{N}_{BAC,SDON} + E^{N}_{MZ,SDON,1} + E^{N}_{MZ,SDON,2} + E^{N}_{KR,SDON,1} + E^{N}_{KR,SDON,2} + E^{N}_{HZ,SDON} + DISS^{N}_{DET} - REFR^{N}_{SDOM} - G^{N}_{BAC,SDON} \quad \text{(A.8.8)}$$

$$\frac{dP_{SDOM}}{dt} = E^{P}_{DA,SDOP} + E^{P}_{CR,SDOP} + E^{P}_{BAC,SDOP} + E^{P}_{MZ,SDOP,1} + E^{P}_{MZ,SDOP,2} + E^{P}_{KR,SDOP,1} + E^{P}_{KR,SDOP,2} + E^{P}_{HZ,SDOP} + DISS^{P}_{DET} - REFR^{P}_{SDOM} - G^{P}_{BAC,SDOP} \quad \text{(A.8.9)}$$

9. Dissolved inorganic nutrient processes

- Nitrification:

$$NTRF = r_{ntrf} \times NH4 \quad \text{(A.9.1)}$$

- The net change equations:

$$\frac{dNH4}{dt} = FLUX^{NH4}_{BAC} + REMI^{N}_{MZ} + REMI^{N}_{KR} + REMI^{N}_{HZ} - G^{NH4}_{DA} - G^{NH4}_{CR} - NTRF \quad \text{(A.9.2)}$$

$$\frac{dNO3}{dt} = FLUX^{NO3}_{BAC} - G^{NO3}_{DA} - G^{NO3}_{CR} + NTRF \quad \text{(A.9.3)}$$

$$\frac{dPO4}{dt} = FLUX^{PO4}_{BAC} + REMI^{P}_{MZ} + REMI^{P}_{KR} + REMI^{P}_{HZ} - G^{PO4}_{DA} - G^{PO4}_{CR} \quad \text{(A.9.4)}$$

where $REMI^{N}_{HZ} = M^{N}_{KR} - D^{N}_{HZ} - E^{SDON}_{HZ}$

$REMI^{P}_{HZ} = M^{N}_{KR} - D^{P}_{HZ} - E^{SDOP}_{HZ}$

Appendix B

Figure B1. Comparison of the observations to the initial (unoptimized) and optimized model results. The dot points in the second panels represent how much larger model output value is compared to the corresponding observational data (i.e., the model value minus the observational value). Normalized observation: observations normalized by the mean of each model state variable.

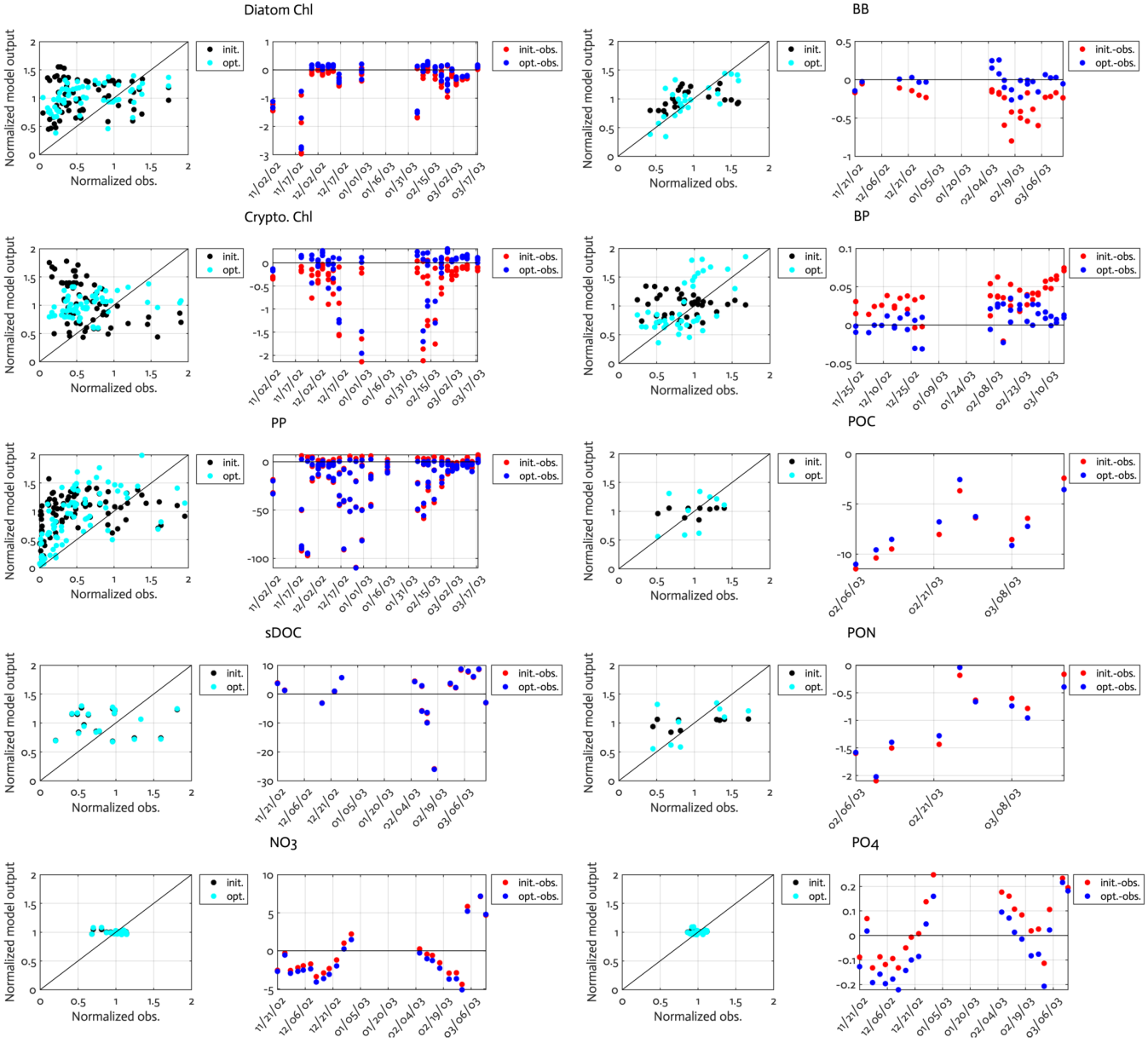


Figure B2. Observations assimilated for each data type. BAC: bacterial biomass, CR: cryptophyte Chl, DA: diatom Chl.

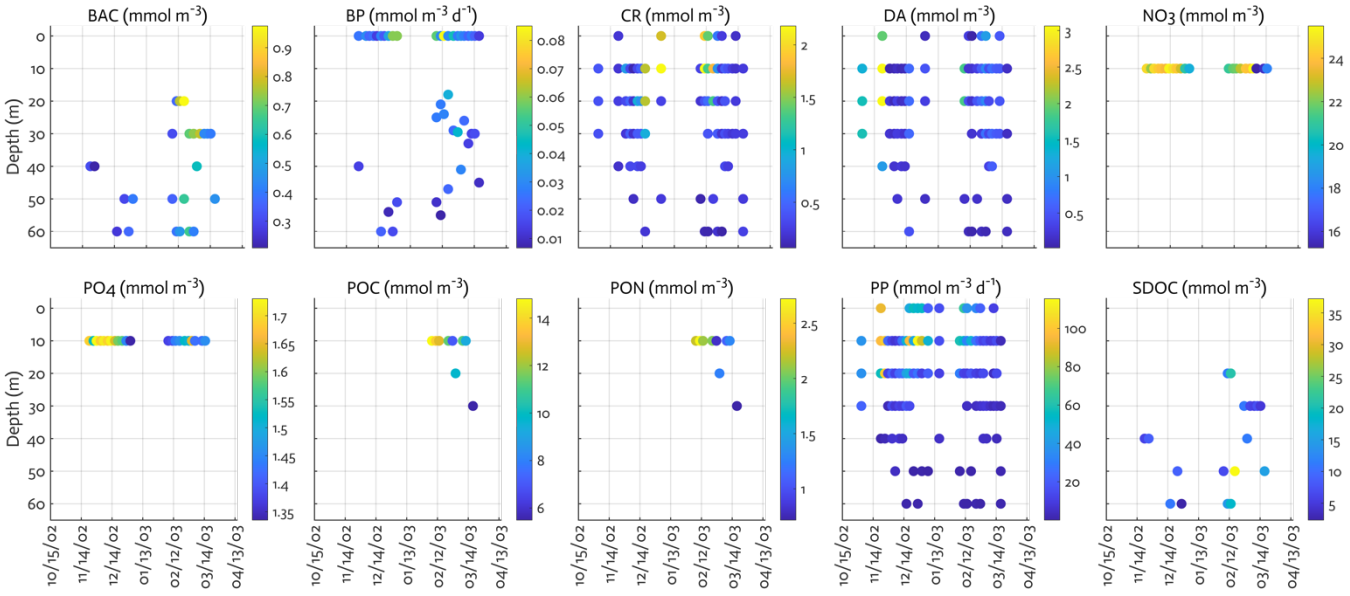


Figure B3. The uncertainties (standard deviation) of the model state variables for the modelled growth season 2002-2003 (x-axis; month/day) from the Monte Carlo experiments (n = 1000). Note different contour scales among panels.

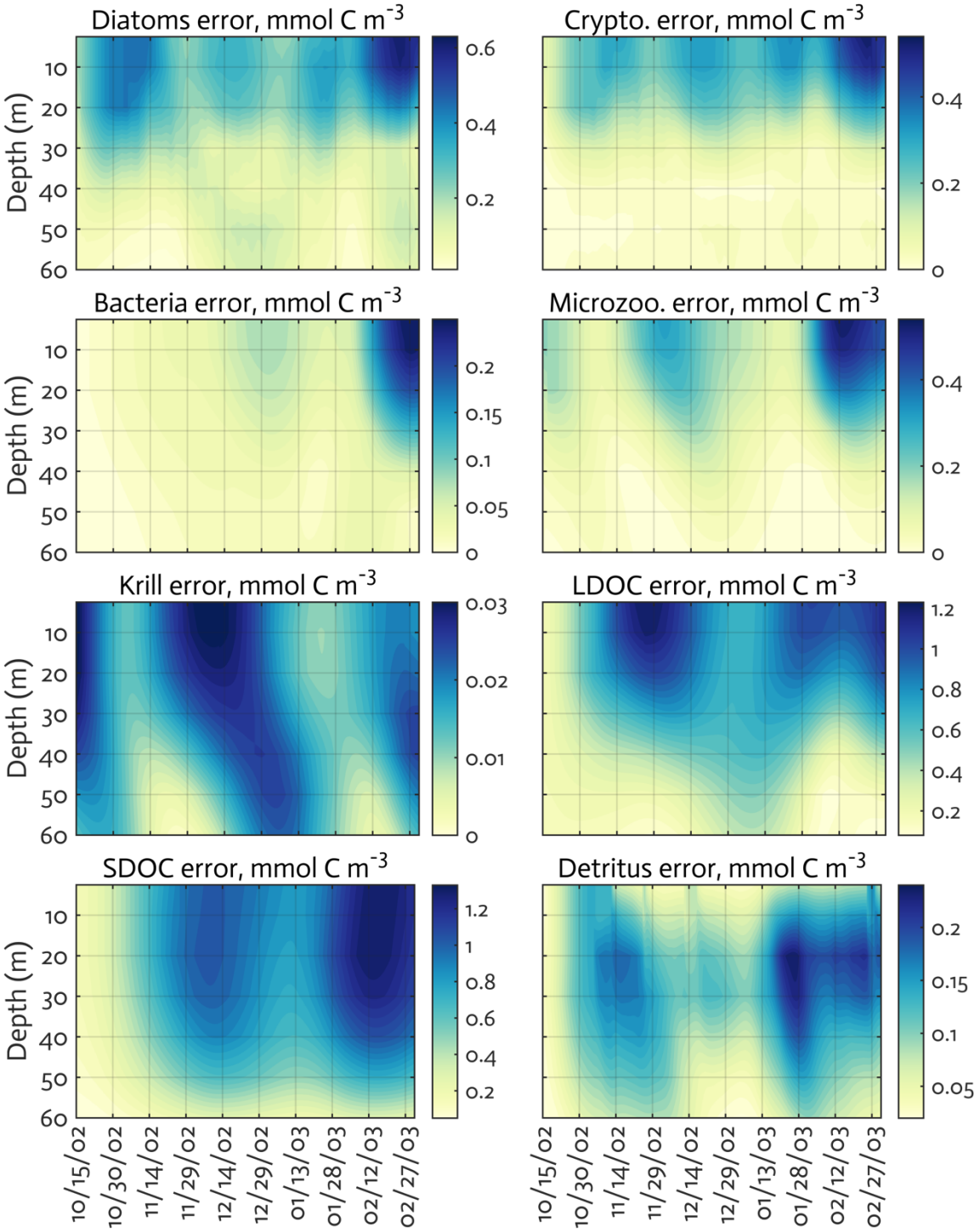


Figure B4. The uncertainties (standard deviation) of the ecosystem indices for the modelled growth season 2002-2003 (x-axis; month/day) from the Monte Carlo experiments (n = 1000). Note different contour scales among panels.

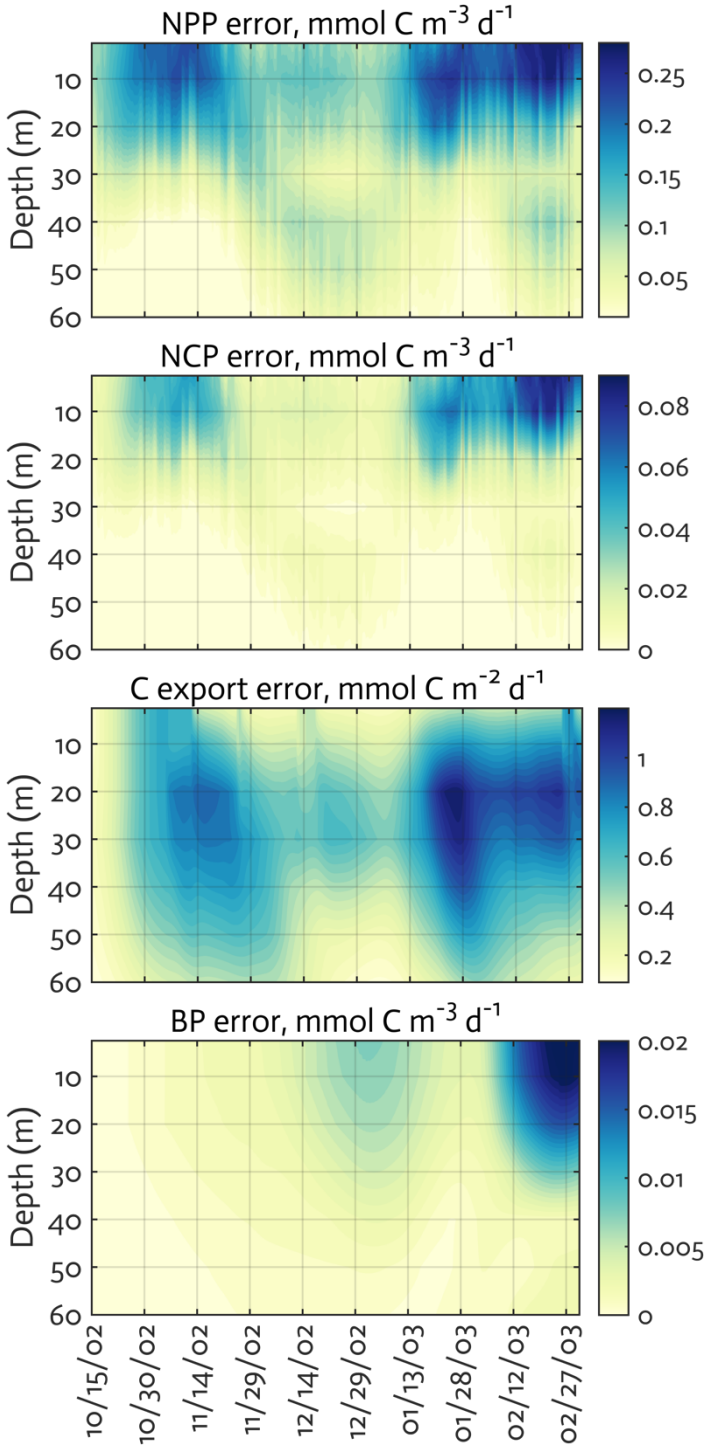


Table B1. Sensitivity tests to varying initial parameter values. The parameters with 'n/a' in the parentheses are optimized parameters with large relative uncertainties (i.e., 'optimized parameters'). 50% perturbations to A_E, *wnsv*, and *remin* (Table 1) are considered ecologically unrealistic and therefore excluded from these sensitivity experiments.

Perturbations to initial parameter values	J_0	J_f	Optimized parameters (optimized value, ranges)
Original optimization experiment: Initial parameter values (Table 1)	14.85	6.42	μ_{DA} (0.77, 0.68-0.88), μ_{CR} (0.72, 0.61-0.85), α_{DA} (0.13, 0.10-0.19), α_{CR} (3.89×10^{-2}, n/a), Θ (2.27, 1.82-2.82), μ_{BAC} (1.06, 0.93-1.20), $r^A_{max,BAC}$ (0.80, 0.77-0.84), μ_{MZ} (1.18, 1.10-1.26), g_{BAC} (0.81, 0.64-1.03), μ_{KR} (1.02, 0.97-1.07), g_{MZ} (0.15, n/a), $remv_{KR}$ (0.43, n/a)
Sensitivity experiment #1: μ_{DA} and μ_{CR} perturbed by +50%	17.02	5.79	μ_{DA} (1.15, 0.99-1.34), μ_{CR} (1.25, 0.85-1.85), α_{DA} (0.19, 0.16-0.23), α_{CR} (0.04, 0.03-0.06), Θ (2.34, 1.91-2.87), $r^A_{max,BAC}$ (0.85, 0.83-0.87), μ_{MZ} (1.50, n/a), g_{BAC} (1.75, n/a), μ_{KR} (0.87, 0.84-0.90), g_{MZ} (0.22, n/a), $remv_{KR}$ (0.14, n/a)
Sensitivity experiment #2: μ_{DA} and μ_{CR} perturbed by -50%	24.98	6.40	μ_{DA} (0.72, 0.61-0.84), μ_{CR} (0.97, 0.84-1.12), Θ (3.38, 2.46-4.64), μ_{BAC} (1.14, 0.96-1.35), $r^A_{max,BAC}$ (0.84, 0.80-0.88), μ_{MZ} (0.75, 0.65-0.86), g_{CR} (0.65, 0.47-0.90), g_{BAC} (1.23, 0.85-1.78), μ_{KR} (0.94, 0.87-1.02), g_{MZ} (0.40, 0.30-0.54), $remv_{KR}$ (0.24, n/a)
Sensitivity experiment #3: μ_{DA}, μ_{CR}, α_{DA}, and α_{CR} perturbed by +50%	14.87	6.61	μ_{DA} (0.71, 0.60-0.83), μ_{CR} (0.60, 0.54-0.67), α_{DA} (0.08, n/a), Θ (2.28, 1.83-2.84), μ_{BAC} (1.46, 1.32-1.61), $r^A_{max,BAC}$ (0.83, 0.80-0.86), μ_{MZ} (1.14, 1.03-1.27), g_{MZ} (0.17, 0.12-0.24), $remv_{KR}$ (0.21, 0.17-0.26)
Sensitivity experiment #4: μ_{DA}, μ_{CR}, α_{DA}, and α_{CR} perturbed by -50%	20.35	6.63	μ_{DA} (0.86, 0.77-0.96), μ_{CR} (0.77, 0.68-0.87), Θ (3.73, 2.75-5.06), $r^A_{max,BAC}$ (0.88, 0.86-0.90), μ_{MZ} (0.86, 0.76-0.98), g_{BAC} (2.19, 1.35-3.56), μ_{KR} (0.71, 0.68-10.75), g_{MZ} (0.41, 0.30-0.56)
Sensitivity experiment #5: μ_{DA}, μ_{CR}, α_{DA}, α_{CR}, Θ, and μ_{BAC} perturbed by +50%	16.72	6.72	μ_{DA} (0.75, 0.63-0.89), μ_{CR} (0.65, 0.59-0.71), α_{DA} (0.07, n/a), Θ (2.47, 1.99-3.07), μ_{BAC} (1.75, 1.60-1.91), $r^A_{max,BAC}$ (0.84, 0.82-0.86), μ_{MZ} (1.12, 1.01-1.24), μ_{KR} (0.65, 0.62-0.68), g_{MZ} (0.21, 0.16-0.28)
Sensitivity experiment #6: μ_{DA}, μ_{CR}, α_{DA}, α_{CR}, Θ, and μ_{BAC} perturbed by -50%	19.91	5.98	μ_{DA} (1.01, 0.87-1.17), μ_{CR} (0.98, 0.85-1.14), α_{CR} (0.11, n/a), Θ (2.78, 2.17-3.57), $r^A_{max,BAC}$ (0.86, 0.84-0.88), μ_{MZ} (1.55, 1.37-1.76), g_{BAC} (2.22, 1.47-3.35), g_{MZ} (0.26, 0.19-0.35), $remv_{KR}$ (0.13, n/a)
Sensitivity experiment #7: μ_{DA}, μ_{CR}, α_{DA}, α_{CR}, Θ, μ_{BAC}, g_{BAC}, and μ_{MZ} perturbed by +50%	16.68	6.40	μ_{DA} (1.14, 0.99-1.31), μ_{CR} (1.43, 1.15-1.78), α_{DA} (0.12, n/a), α_{CR} (0.04, n/a), μ_{BAC} (1.02, 0.92-1.13), $r^A_{max,BAC}$ (0.81, 0.77-0.85), μ_{MZ} (1.02, 0.91-1.15), g_{BAC} (1.64, 1.25-2.15), μ_{KR} (0.72, 0.68-0.76), g_{MZ} (0.38, 0.29-0.50)
Sensitivity experiment #8: μ_{DA}, μ_{CR}, α_{DA}, α_{CR}, Θ, μ_{BAC}, g_{BAC}, and μ_{MZ} perturbed by -50%	25.06	7.19	μ_{DA} (0.64, 0.57-0.72), μ_{CR} (0.47, 0.41-0.53), α_{CR} (0.08, n/a), Θ (3.13, 2.44-4.01), μ_{BAC} (1.14, 0.99-1.32), $r^A_{max,BAC}$ (0.79, 0.75-0.83), μ_{MZ} (0.40, 0.36-0.44), μ_{KR} (0.81, 0.71-0.77), g_{MZ} (0.55, 0.39-0.77), $remv_{KR}$ (0.12, n/a)
Sensitivity experiment #9: μ_{DA}, μ_{CR}, α_{DA}, α_{CR}, Θ, μ_{BAC}, g_{BAC}, μ_{MZ}, μ_{KR}, and g_{MZ} perturbed by +50%	16.36	6.22	μ_{DA} (0.88, 0.78-1.00), μ_{CR} (0.70, n/a), α_{DA} (0.27, 0.20-0.36), α_{CR} (0.04, n/a), Θ (2.40, 1.91-3.02), $r^A_{max,BAC}$ (0.87, 0.85-0.89), μ_{MZ} (0.95, 0.75-1.20), g_{BAC} (1.75, n/a), μ_{KR} (0.78, 0.74-0.82), g_{MZ} (0.21, 0.14-0.32), $remv_{KR}$ (0.18, n/a)
Sensitivity experiment #10: μ_{DA}, μ_{CR}, α_{DA}, α_{CR}, Θ, μ_{BAC}, g_{BAC}, μ_{MZ}, μ_{KR}, and g_{MZ} perturbed by -50%	18.27	6.75	μ_{DA} (0.91, 0.78-1.06), μ_{CR} (0.58, 0.51-0.66), Θ (2.71, 2.21-3.32), μ_{BAC} (1.70, n/a), $r^A_{max,BAC}$ (0.84, 0.82-0.86), μ_{MZ} (0.88, 0.71-1.09), g'_{DA} (2.5, n/a), g_{BAC} (1.0, n/a), μ_{KR} (1.89, 1.60-2.24), $remv_{KR}$ (0.04, n/a)
Sensitivity experiment #11: μ_{DA}, μ_{CR}, α_{DA}, α_{CR}, Θ, μ_{BAC}, g_{BAC}, μ_{MZ}, μ_{KR}, g_{MZ}, $r^A_{max,BAC}$, and $remv_{KR}$ perturbed by +50%	15.56	6.43	μ_{DA} (0.88, 0.79-0.98), μ_{CR} (0.75, 0.66-0.85), α_{DA} (0.30, n/a), α_{CR} (0.15, n/a), Θ (2.99, 2.33-3.83), $r^A_{max,BAC}$ (0.88, 0.87-0.89), μ_{MZ} (0.87, 0.75-1.01), g_{BAC} (1.62, n/a), μ_{KR} (0.73, 0.66-0.81), g_{MZ} (0.36, 0.26-0.50), $remv_{KR}$ (0.11, 0.08-0.15)
Sensitivity experiment #12: μ_{DA}, μ_{CR}, α_{DA}, α_{CR}, Θ, μ_{BAC}, g_{BAC}, μ_{MZ}, μ_{KR}, g_{MZ}, $r^A_{max,BAC}$, and $remv_{KR}$ perturbed by -50%	28.24	6.81	μ_{DA} (0.54, 0.49-0.60), μ_{CR} (0.41, 0.36-0.46), α_{DA} (0.12, n/a), Θ (1.96, 1.59-2.42), μ_{BAC} (1.40, 1.24-1.58), $r^A_{max,BAC}$ (0.81, 0.78-0.84), μ_{MZ} (0.32, 0.28-0.36), μ_{KR} (0.70, 0.65-0.75), g_{MZ} (0.20, n/a), $remv_{KR}$ (0.27, n/a)
Sensitivity experiment #13: A_E, *wnsv*, and *remin* perturbed by -50%	16.26	6.09	μ_{DA} (0.48, 0.44-0.53), μ_{CR} (0.57, 0.54-0.60), Θ (3.23, 2.54-4.11), $r^A_{max,BAC}$ (0.88, 0.86-0.90), μ_{MZ} (1.04, 1.00-1.09), g_{CR} (0.56, 0.47-0.67), g_{BAC} (1.27, 0.82-1.97), μ_{KR} (0.54, 0.53-0.55), g_{MZ} (0.18, n/a)
Sensitivity experiment #14: g_{DA}, g'_{DA}, and g_{CR} perturbed by +50%	14.25	5.95	μ_{DA} (0.65, 0.56-0.75), μ_{CR} (1.10, 0.98-1.24), α_{DA} (0.20, n/a), α_{CR} (0.12, n/a), Θ (3.09, 2.40-3.98), $r^A_{max,BAC}$ (0.87, 0.85-0.89), μ_{MZ} (1.75, 1.55-1.97), g_{CR} (0.77, 0.63-0.95), g_{BAC} (2.60, 1.70-3.97), μ_{KR} (0.82, 0.78-0.86), g_{MZ} (0.24,0.18-0.32)
Sensitivity experiment #15: g_{DA}, g'_{DA}, and g_{CR} perturbed by -50%	24.17	6.51	μ_{DA} (1.22, 1.03-1.45), μ_{CR} (0.86, 0.75-0.99), α_{DA} (0.10, n/a), α_{CR} (0.15, n/a), Θ (4.53, 3.36-6.11), μ_{BAC} (1.24, n/a), $r^A_{max,BAC}$ (0.85, 0.83-0.87), μ_{MZ} (0.86, 0.73-1.01), g_{CR} (0.53, 0.42-0.67), g_{BAC} (0.89, 0.55-1.43), μ_{KR} (1.02, 0.94-1.10), g_{MZ} (0.37, 0.25-0.55), $remv_{KR}$ (0.19, n/a)

Table B2. Summary of model parameters from initial parameter perturbation experiments. Summary of the model parameter symbol and definition, initial parameter values ($p_{0,\,ref}$) and optimized values ($p_{f,\,ref}$) for the original (reference) experiments (Table 1) and initial parameter values (p_0) and optimized values (p_f) averaged from the sensitivity experiments (n = 15). $p_{0,\,ref}$ and the mean p_0 are the same for most parameters because of the perturbations by ±50% their original initial parameter values (with standard deviation in parentheses), while $p_{0,\,ref}$ and the mean p_0 are different for A_E and *wnsv* because of the perturbations only by -50% their original initial parameter values (with standard deviation in parentheses). Numbers in parentheses for p_f are the uncertainty ranges (lower and upper bounds) averaged across the sensitivity experiments as follows. First, for each sensitivity experiment lower and upper bounds of the constrained parameter are calculated as $p_f \times e^{+\sigma f}$ and $p_f \times e^{-\sigma f}$ (where p_f is the value of the constrained parameter and σ_f is the square roots of diagonal elements of the inverse of the Hessian matrix), respectively. Then we form the "lower (upper) bound parameter set" that only consists of the lower (higher) bounds of the constrained parameters from each experiment, and average those across the sensitivity experiments (n = 15) to calculate the lower (upper) bound listed in parentheses.

Model parameter symbol and definition (optimizable)	$p_{0,\,ref}$	p_0	$p_{f,\,ref}$	p_f
A_E, Arrhenius parameter for temperature function	4000	3867 (516)	-	-
μ_{DA}, Diatom C-specific maximum growth rate, d^{-1}	2.00	2.00 (0.93)	0.77 (0.68-0.88)	0.84 (0.73-0.96)
μ_{CR}, Crypto. C-specific maximum growth rate, d^{-1}	1.00	1.00 (0.46)	0.72 (0.61-0.85)	0.81 (0.69-0.95)
α_{DA}, Initial slope of P-I curve of diatoms, mol C (g Chl)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$	0.30	0.30 (0.13)	0.13 (0.10-0.19)	0.18 (0.17-0.19)
α_{CR}, Initial slope of P-I curve of crypto., mol C (g Chl)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$	0.20	0.20 (0.08)	3.89×10^{-2} (n/a)	0.13 (0.13-0.13)
Θ, Maximum Chl/N ratio, g Chl *a* (mol N)$^{-1}$	2.90	2.90 (1.10)	2.27 (1.82-2.82)	3.03 (2.42-3.82)
μ_{BAC}, Maximum bacterial growth rate, d^{-1}	2.00	2.00 (0.76)	1.06 (0.93-1.20)	1.86 (1.80-1.92)
$r^A_{max,BAC}$, Bacterial maximum active respiration rate, d^{-1}	0.58	0.58 (0.11)	0.80 (0.77-0.84)	0.85 (0.83-0.87)
μ_{MZ}, Microzoo. C-specific maximum growth rate, d^{-1}	1.00	1.00 (0.33)	1.18 (1.10-1.26)	1.00 (0.89-1.13)
g_{DA}, Diatom half-saturation concentration in microzoo. grazing, mmol C m^{-3}	1.00	1.00 (0.19)	-	-
g'_{DA}, Diatom half-saturation concentration in krill grazing, mmol C m^{-3}	1.00	1.00 (0.19)	-	1.10 (1.10-1.10)
g_{CR}, Crypto. half-saturation concentration in microzoo. grazing, mmol C m^{-3}	1.00	1.00 (0.19)	-	0.90 (0.87-0.95)
g_{BAC}, Bacterial half-saturation concentration in microzoo. grazing, mmol C m^{-3}	0.55	0.55 (0.18)	0.81 (0.64-1.03)	1.34 (1.07-1.75)
μ_{KR}, Maximum krill C-specific growth rate, d^{-1}	0.80	0.80 (0.21)	1.02 (0.97-1.07)	0.86 (0.80-0.93)
g_{MZ}, Microzoo. half-saturation concentration in krill grazing, mmol C m^{-3}	1.00	1.00 (0.27)	0.15 (n/a)	0.34 (0.28-0.43)
$remv_{KR}$, Krill removal rate by higher-trophic levels, (mmol C m^{-3})$^{-1}$ d^{-1}	0.10	0.10 (0.02)	0.43 (n/a)	0.14 (0.14-0.15)
wnsv, Detritus vertical sinking velocity, m d^{-1}	5.00	4.83 (0.65)	-	-
diss, Detrital dissolution rate, d^{-1}	0.14	0.14 (0.02)	-	-

References

Bertilsson, S., Berglund, O., Karl, D. M., & Chisholm, S. W. (2003). Elemental composition of marine Prochlorococcus and Synechococcus: Implications for the ecological stoichiometry of the sea. *Limnology and Oceanography*, *48*(5), 1721–1731. https://doi.org/10.4319/lo.2003.48.5.1721

Bird, D. F., & Karl, D. M. (1999). Uncoupling of bacteria and phytoplankton during the austral spring bloom in Gerlache Strait, Antarctic Peninsula. *Aquatic Microbial Ecology*, *19*(1), 13–27. https://doi.org/10.3354/ame019013

Bowman, J. S., & Ducklow, H. W. (2015). Microbial Communities Can Be Described by Metabolic Structure: A General Framework and Application to a Seasonally Variable, Depth-Stratified Microbial Community from the Coastal West Antarctic Peninsula. *PLOS ONE*, *10*(8), e0135868. https://doi.org/10.1371/journal.pone.0135868

Bowman, J. S., Kavanaugh, M. T., Doney, S. C., & Ducklow, H. W. (2018). Recurrent seascape units identify key ecological processes along the western Antarctic Peninsula. *Global Change Biology*, *24*(7), 3065–3078. https://doi.org/10.1111/gcb.14161

Caron D.A., Dennett M.R., Lonsdale D.J., Moran D.M., Shalapyonok L. Microzooplankton herbivory in the Ross sea, Antarctica. Deep Sea Research Part II: Topical Studies in Oceanography. 2000 (15-16):3249-72

Campbell, J. W. (1995). The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research: Oceans*, *100*(C7), 13237–13254. https://doi.org/10.1029/95JC00458

Carlson, C. A., Bates, N. R., Ducklow, H. W., & Hansell, D. A. (1999). Estimation of bacterial respiration and growth efficiency in the Ross Sea, Antarctica. *Aquatic Microbial Ecology*, *19*(3), 229–244.

Carvalho, F., Kohut, J., Oliver, M. J., Sherrell, R. M., Schofield, O. 2016. Mixing and phytoplankton dynamics in a submarine canyon in the West Antarctic Peninsula. Journal of Geophysical Research DOI: 10.1002/2016JC011650

Clarke, A., Griffiths, H. J., Barnes, D. K. A., Meredith, M. P., & Grant, S. M. (2009). Spatial variation in seabed temperatures in the Southern Ocean: Implications for benthic ecology and biogeography. *Journal of Geophysical Research: Biogeosciences*, *114*(G3). https://doi.org/10.1029/2008JG000886

Cook, A. J., Fox, A. J., Vaughan, D. G., & Ferrigno, J. G. (2005). Retreating Glacier Fronts on the Antarctic Peninsula over the Past Half-Century. *Science*, *308*(5721), 541–544. https://doi.org/10.1126/science.1104235

del Giorgio, P. A., & Cole, J. J. (1998). Bacterial Growth Efficiency in Natural Aquatic Systems. *Annual Review of Ecology and Systematics*, *29*(1), 503–541. https://doi.org/10.1146/annurev.ecolsys.29.1.503

Doney, S. C., Glover, D. M., McCue, S. J., & Fuentes, M. (2003). Mesoscale variability of Sea-viewing Wide Field-of-view Sensor (SeaWiFS) satellite ocean color: Global patterns and spatial scales. *Journal of Geophysical Research: Oceans*, *108*(C2). https://doi.org/10.1029/2001JC000843

Doney, S. C., Lima, I., Moore, J. K., Lindsay, K., Behrenfeld, M. J., Westberry, T. K., Mahowald, N., Glover, D. M., & Takahashi, T. (2009). Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data. *Journal of Marine Systems*, *76*(1), 95–112. https://doi.org/10.1016/j.jmarsys.2008.05.015

Droop, M. R. (1974). The nutrient status of algal cells in continuous culture. *Journal of the Marine Biological Association of the United Kingdom*, *54*(4), 825–855. https://doi.org/10.1017/S002531540005760X

Droop, M. R. (1983). 25 years of algal growth kinetics. A personal view. *Botanica Marina*. http://agris.fao.org/agris-search/search.do?recordID=US201302597810

Ducklow, H. W. (2000). Bacterial production and biomass in the ocean. In *Microbial Ecology of the Oceans, Second Edition* (pp. 85–120). John Wiley & Sons, Inc.

Ducklow, H. W., Doney, S. C., & Steinberg, D. K. (2008). Contributions of Long-Term Research and Time-Series Observations to Marine Ecology and Biogeochemistry. *Annual Review of Marine Science*, *1*, 279–302.

Ducklow, Hugh W, Baker, K., Martinson, D. G., Quetin, L. B., Ross, R. M., Smith, R. C., Stammerjohn, S. E., Vernet, M., & Fraser, W. (2007). Marine pelagic ecosystems: The West Antarctic Peninsula. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1477), 67–94. https://doi.org/10.1098/rstb.2006.1955

Ducklow, Hugh W., & Doney, S. C. (2013). What is the metabolic state of the oligotrophic ocean? A debate. *Annual Review of Marine Science*, *5*, 525–533. https://doi.org/10.1146/annurev-marine-121211-172331

Ducklow, Hugh W., Myers, K. M. S., Erickson, M., Ghiglione, J.-F., & Murray, A. E. (2011). *Response of a summertime Antarctic marine -bacterial community to glucose and ammonium enrichment*. http://agris.fao.org/agris-search/search.do?recordID=AV2012072112

Ducklow, Hugh W., Schofield, O., Vernet, M., Stammerjohn, S., & Erickson, M. (2012). Multiscale control of bacterial production by phytoplankton dynamics and sea ice along the western Antarctic Peninsula: A regional and decadal investigation. *Journal of Marine Systems*, *98–99*, 26–39. https://doi.org/10.1016/j.jmarsys.2012.03.003

Ducklow, Hugh W., Stukel, M. R., Eveleth, R., Doney, S. C., Jickells, T., Schofield, O., Baker, A. R., Brindle, J., Chance, R., & Cassar, N. (2018). Spring–summer net community production, new production, particle export and related water column biogeochemical processes in the marginal sea ice zone of the Western Antarctic Peninsula 2012–2014. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2122), 20170177. https://doi.org/10.1098/rsta.2017.0177

Dugdale, R. C., & Goering, J. J. (1967). Uptake of New and Regenerated Forms of Nitrogen in Primary Productivity1. *Limnology and Oceanography*, *12*(2), 196–206. https://doi.org/10.4319/lo.1967.12.2.0196

Fennel, K., Losch, M., Schröter, J., & Wenzel, M. (2001). Testing a marine ecosystem model: Sensitivity analysis and parameter optimization. *Journal of Marine Systems*, *28*(1), 45–63. https://doi.org/10.1016/S0924-7963(00)00083-X

Friedrichs, M. A. M. (2001). Assimilation of JGOFS EqPac and SeaWiFS data into a marine ecosystem model of the Central Equatorial Pacific Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*(1), 289–319.

Friedrichs, M. A. M., Dusenberry, J. A., Anderson, L. A., Armstrong, R. A., Chai, F., Christian, J. R., Doney, S. C., Dunne, J., Fujii, M., Hood, R., McGillicuddy, D. J., Moore, J. K., Schartau, M., Spitz, Y. H., & Wiggert, J. D. (2007). Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. *Journal of Geophysical Research: Oceans*, *112*(C8).

Friedrichs, M. A. M., Hood, R. R., & Wiggert, J. D. (2006). Ecosystem model complexity versus physical forcing: Quantification of their relative impact with assimilated Arabian Sea data. *Deep Sea Research Part II: Topical Studies in Oceanography*, *53*(5), 576–600.

Fukuda, R., Ogawa, H., Nagata, T., & Koike, I. (1998). Direct Determination of Carbon and Nitrogen Contents of Natural Bacterial Assemblages in Marine Environments. *Applied and Environmental Microbiology*, *64*(9), 3352–3358.

Garzio, L. M., Steinberg, D. K., Erickson, M., & Ducklow, H. W. (2013). *Microzooplankton grazing along the Western Antarctic Peninsula*. https://doi.org/10.3354/ame01655

Garzio, L., & Steinberg, D. (2013). Microzooplankton community composition along the Western Antarctic Peninsula. *Deep Sea Research Part I: Oceanographic Research Papers*, *77*, 36–49. https://doi.org/10.1016/j.dsr.2013.03.001

Geider, R. J., MacIntyre, H. L., & Kana, T. M. (1997). 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. *Marine Ecology Progress Series*, *148*(1/3), 187–200. JSTOR.

Geider, Richard J. (1987). Light and Temperature Dependence of the Carbon to Chlorophyll a Ratio in Microalgae and Cyanobacteria: Implications for Physiology and Growth of Phytoplankton. *The New Phytologist*, *106*(1), 1–34. JSTOR.

Gilbert, J. C., & Lemaréchal, C. (1989). Some numerical experiments with variable-storage quasi-Newton algorithms. *Mathematical Programming*, *45*(1), 407–435.

Glover, D. M., Doney, S. C., Oestreich, W. K., & Tullo, A. W. (2018). Geostatistical Analysis of Mesoscale Spatial Variability and Error in SeaWiFS and MODIS/Aqua Global Ocean Color Data. *Journal of Geophysical Research: Oceans*, *123*(1), 22–39. https://doi.org/10.1002/2017JC013023

Glover, D. M., Jenkins, W. J., & Doney, S. C. (2011). 10. Model analysis and optimization. In *Modeling Methods for Marine Science*. Cambridge University Press.

Harmon, R., & Challenor, P. (1997). A Markov chain Monte Carlo method for estimation and assimilation into models. *Ecological Modelling*, *101*(1), 41–59. https://doi.org/10.1016/S0304-3800(97)01947-9

Henley, S. F., Schofield, O. M., Hendry, K. R., Schloss, I. R., Steinberg, D. K., Moffat, C., Peck, L. S., Costa, D. P., Bakker, D. C. E., Hughes, C., Rozema, P. D., Ducklow, H. W., Abele, D., Stefels, J., Van Leeuwe, M. A., Brussaard, C. P. D., Buma, A. G. J., Kohut, J., Sahade, R., … Meredith, M. P. (2019). Variability and change in the west Antarctic Peninsula marine system: Research priorities and opportunities. *Progress in Oceanography*, *173*, 208–237. https://doi.org/10.1016/j.pocean.2019.03.003

Kim, H., Doney, S. C., Iannuzzi, R. A., Meredith, M. P., Martinson, D. G., & Ducklow, H. W. (2016). Climate forcing for dynamics of dissolved inorganic nutrients at Palmer Station, Antarctica: An interdecadal (1993–2013) analysis. *Journal of Geophysical Research: Biogeosciences*, *121*(9), 2369–2389.

Kim, H., & Ducklow, H. W. (2016). A decadal (2002–2014) analysis for dynamics of heterotrophic bacteria in an Antarctic coastal ecosystem: Variability and physical and biogeochemical Forcings. *Frontiers in Marine Science*, *3*. https://doi.org/10.3389/fmars.2016.00214

King, J. C. (1994). Recent climate variability in the vicinity of the antarctic peninsula. *International Journal of Climatology*, *14*(4), 357–369. https://doi.org/10.1002/joc.3370140402

Kirchman, D. L., Morán, X. A. G., & Ducklow, H. (2009). Microbial growth in the polar oceans—Role of temperature and potential impact of climate change. *Nature Reviews Microbiology*, *7*(6), 451–459.

Klinck, J. M. (1998). Heat and salt changes on the continental shelf west of the Antarctic Peninsula between January 1993 and January 1994. *Journal of Geophysical Research: Oceans*, *103*(C4), 7617–7636. https://doi.org/10.1029/98JC00369

Lawson, L. M., Spitz, Y. H., Hofmann, E. E., & Long, R. B. (1995). A data assimilation technique applied to a predator-prey model. *Bulletin of Mathematical Biology*, *57*(4), 593–617.

Legendre, L., & Rassoulzadegan, F. (1996). Food-web mediated export of biogenic carbon in oceans:hydrodynamic control. *Marine Ecology Progress Series*, *145*, 179–193. https://doi.org/10.3354/meps145179

Luo, Y.-W., Friedrichs, M. A. M., Doney, S. C., Church, M. J., & Ducklow, H. W. (2010). Oceanic heterotrophic bacterial nutrition by semilabile DOM as revealed by data assimilative modeling. *Aquatic Microbial Ecology*, *60*(3), 273–287.

Luria, C. M., Amaral-Zettler, L. A., Ducklow, H. W., Repeta, D. J., Rhyne, A. L., & Rich, J. J. (2017). Seasonal Shifts in Bacterial Community Responses to Phytoplankton-Derived Dissolved Organic Matter in the Western Antarctic Peninsula. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.02117

Luria, C. M., Ducklow, H. W., & Amaral-Zettler, L. A. (2014). *Marine bacterial, archaeal and eukaryotic diversity and community structure on the continental shelf of the western Antarctic Peninsula*. https://doi.org/10.3354/ame01703

Matear, R. J. (1996). Parameter optimization and analysis of ecosystem models using simulated annealing: A case study at Station P. *Oceanographic Literature Review*, *43*(6). https://doi.org/info:doi/10.1357/0022240953213098

McCarthy, J. (1980). *Nitrogen. In: Morris I (ed) The physiological ecology of phytoplankton.* Blackwell, Oxford.

Meredith, M. P., & King, J. C. (2005). Rapid climate change in the ocean west of the Antarctic Peninsula during the second half of the 20th century. *Geophysical Research Letters*, *32*(19). https://doi.org/10.1029/2005GL024042

Moline, M., Karnovsky, N., Brown, Z., Divoky, G., Frazer, T., Jacoby, C., Torres, J., & Fraser, W. (2008). High Latitude Changes in Ice Dynamics and Their Impact on Polar Marine Ecosystems. *Annals of the New York Academy of Sciences*, *1134*, 267–319. https://doi.org/10.1196/annals.1439.010

Montégut, C. de B., Madec, G., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research: Oceans*, *109*(C12). https://doi.org/10.1029/2004JC002378

Montes-Hugo, M., Doney, S. C., Ducklow, H. W., Fraser, W., Martinson, D., Stammerjohn, S. E., & Schofield, O. (2009). Recent Changes in Phytoplankton Communities Associated with Rapid Regional Climate Change Along the Western Antarctic Peninsula. *Science*, *323*(5920), 1470–1473. https://doi.org/10.1126/science.1164533

Murphy, E. J., Cavanagh, R. D., Hofmann, E. E., Hill, S. L., Constable, A. J., Costa, D. P., Pinkerton, M. H., Johnston, N. M., Trathan, P. N., Klinck, J. M., Wolf-Gladrow, D. A., Daly, K. L., Maury, O., & Doney, S. C. (2012). Developing integrated models of Southern Ocean food webs: Including ecological complexity, accounting for uncertainty and the importance of scale. *Progress in Oceanography*, *102*, 74–92. https://doi.org/Murphy, E.J. ORCID: https://orcid.org/0000-0002-7369-9196 <https://orcid.org/0000-0002-7369-9196>; Cavanagh, R.D. ORCID: https://orcid.org/0000-0002-2474-9716 <https://orcid.org/0000-0002-2474-9716>; Hofmann, E.E.; Hill, S.L.; Constable, A.J.; Costa, D.P.; Pinkerton, M.H.; Johnston, N.M.; Trathan, P.N.; Klinck, J.M.; Wolf-Gladrow, D.A.; Daly, K.L.; Maury, O.; Doney, S.C.. 2012 Developing integrated models of Southern Ocean food webs: including ecological complexity, accounting for uncertainty and the importance of scale. Progress in Oceanography, 102. 74-92. https://doi.org/10.1016/j.pocean.2012.03.006 <https://doi.org/10.1016/j.pocean.2012.03.006>

Pomeroy, L. R., & Wiebe, W. J. (2001). Temperature and substrates as interactive limiting factors for marine heterotrophic bacteria. *Aquatic Microbial Ecology*, *23*(2), 187–204. https://doi.org/10.3354/ame023187

Prunet, P., Minster, J.-F., Echevin, V., & Dadou, I. (1996). Assimilation of surface data in a one-dimensional physical-biogeochemical model of the surface ocean: 2. Adjusting a simple trophic model to chlorophyll, temperature, nitrate, and pCO2 data. *Global Biogeochemical Cycles*, *10*(1), 139–158. https://doi.org/10.1029/95GB03435

Prunet, P., Minster, J.-F., Ruiz-Pino, D., & Dadou, I. (1996). Assimilation of surface data in a one-dimensional physical-biogeochemical model of the surface ocean: 1. Method and preliminary results. *Global Biogeochemical Cycles*, *10*(1), 111–138. https://doi.org/10.1029/95GB03436

Saba, G. K., Fraser, W. R., Saba, V. S., Iannuzzi, R. A., Coleman, K. E., Doney, S. C., Ducklow, H. W., Martinson, D. G., Miles, T. N., Patterson-Fraser, D. L., Stammerjohn, S. E., Steinberg, D. K., & Schofield, O. M. (2014). Winter and spring controls on the summer food web of the coastal West Antarctic Peninsula. *Nature Communications*, *5*(1), 1–8. https://doi.org/10.1038/ncomms5318

Sailley, S. F., Ducklow, H. W., Moeller, H. V., Fraser, W. R., Schofield, O. M., Steinberg, D. K., Garzio, L. M., & Doney, S. C. (2013). Carbon fluxes and pelagic ecosystem dynamics near two western Antarctic Peninsula Adélie penguin colonies: An inverse model approach. *Marine Ecology Progress Series*, *492*, 253–272. https://doi.org/10.3354/meps10534

Schofield, O., Saba, G., Coleman, K., Carvalho, F., Couto, N., Ducklow, H., Finkel, Z., Irwin, A., Kahl, A., Miles, T., Montes-Hugo, M., Stammerjohn, S., & Waite, N. (2017). Decadal variability in coastal phytoplankton community composition in a changing West Antarctic Peninsula. *Deep Sea Research Part I: Oceanographic Research Papers*, *124*, 42–54. https://doi.org/10.1016/j.dsr.2017.04.014

Sherrell, R. M., Annett, A. L., Fitzsimmons, J. N., Roccanova, V. J., & Meredith, M. P. (2018). A "shallow bathtub ring" of local sedimentary iron input maintains the Palmer Deep biological hotspot on the West Antarctic Peninsula shelf. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *376*(2122). https://doi.org/10.1098/rsta.2017.0171

Smith, D. A., Hofmann, E. E., Klinck, J. M., & Lascara, C. M. (1999). Hydrography and circulation of the West Antarctic Peninsula Continental Shelf. *Deep Sea Research Part I: Oceanographic Research Papers*, *46*(6), 925–949. https://doi.org/10.1016/S0967-0637(98)00103-4

Smith, R. C., Martinson, D. G., Stammerjohn, S. E., Iannuzzi, R. A., & Ireson, K. (2008). Bellingshausen and western Antarctic Peninsula region: Pigment biomass and sea-ice spatial/temporal distributions and interannual variabilty. *Deep Sea Research Part II: Topical Studies in Oceanography*, *55*(18), 1949–1963. https://doi.org/10.1016/j.dsr2.2008.04.027

Spitz, Y. H., Moisan, J. R., & Abbott, M. R. (2001). Configuring an ecosystem model using data from the Bermuda Atlantic Time Series (BATS). *Deep Sea Research Part II: Topical Studies in Oceanography*, *48*(8–9), 1733–1768.

Stammerjohn, S. E., Martinson, D. G., Smith, R. C., Yuan, X., & Rind, D. (2008). Trends in Antarctic annual sea ice retreat and advance and their relation to El Niño–Southern Oscillation and Southern Annular Mode variability. *Journal of Geophysical Research: Oceans*, *113*(C3). https://doi.org/10.1029/2007JC004269

Steinberg, D. K., Ruck, K. E., Gleiber, M. R., Garzio, L. M., Cope, J. S., Bernard, K. S., Stammerjohn, S. E., Schofield, O. M. E., Quetin, L. B., & Ross, R. M. (2015). Long-term (1993–2013) changes in macrozooplankton off the Western Antarctic Peninsula. *Deep Sea Research Part I: Oceanographic Research Papers*, *101*, 54–70. https://doi.org/10.1016/j.dsr.2015.02.009

Stow, C. A., Jolliff, J., McGillicuddy, D. J., Doney, S. C., Allen, J. I., Friedrichs, M. A. M., Rose, K. A., & Wallhead, P. (2009). Skill assessment for coupled biological/physical models of marine systems. *Journal of Marine Systems*, *76*(1), 4–15. https://doi.org/10.1016/j.jmarsys.2008.03.011

Stukel, M. R., Asher, E., Couto, N., Schofield, O., Strebel, S., Tortell, P., & Ducklow, H. W. (2015). The imbalance of new and export production in the western Antarctic Peninsula, a potentially "leaky" ecosystem. *Global Biogeochemical Cycles*, *29*(9), 1400–1420. https://doi.org/10.1002/2015GB005211

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183–7192.

Thibodeau, P. S., Steinberg, D. K., Stammerjohn, S. E., & Hauri, C. (2019). Environmental controls on pteropod biogeography along the Western Antarctic Peninsula. *Limnology and Oceanography*, *64*(S1), S240–S256. https://doi.org/10.1002/lno.11041

Tziperman, E., & Thacker, W. C. (1989). An Optimal-Control/Adjoint-Equations Approach to Studying the Oceanic General Circulation. *Journal of Physical Oceanography*, *19*(10), 1471–1485.

Vaughan, D. G. (2006). Recent Trends in Melting Conditions on the Antarctic Peninsula and Their Implications for Ice-sheet Mass Balance and Sea Level. *Arctic, Antarctic, and Alpine Research*, *38*(1), 147–152. https://doi.org/10.1657/1523-0430(2006)038[0147:RTIMCO]2.0.CO;2

Vaughan, D., Marshall, G., Connolley, W., Parkinson, C., Mulvaney, R., Hodgson, D., King, J., Pudsey, C., & Turner, J. (2003). Recent Rapid Regional Climate Warming on the Antarctic Peninsula. *Climatic Change*, *60*, 243–274. https://doi.org/10.1023/A:1026021217991

Ward, B. A., Friedrichs, M. A. M., Anderson, T. R., & Oschlies, A. (2010). Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models. *Journal of Marine Systems*, *81*(1), 34–43.

Weston, K., Jickells, T. D., Carson, D. S., Clarke, A., Meredith, M. P., Brandon, M. A., Wallace, M. I., Ussher, S. J., & Hendry, K. R. (2013). Primary production export flux in Marguerite Bay (Antarctic Peninsula): Linking upper water-column production to sediment trap flux. *Deep Sea Research Part I: Oceanographic Research Papers*, *75*, 52–66. https://doi.org/10.1016/j.dsr.2013.02.001

Whitehouse, M. J., Meredith, M. P., Rothery, P., Atkinson, A., Ward, P., & Korb, R. E. (2008). Rapid warming of the ocean around South Georgia, Southern Ocean, during the 20th century: Forcings, characteristics and implications for lower trophic levels. *Deep Sea Research Part I: Oceanographic Research Papers*, *55*(10), 1218–1228. https://doi.org/10.1016/j.dsr.2008.06.002